Asset Analytics
Performance and Safety Management
Series Editors: Ajit Kumar Verma · P. K. Kapur · Uday Kumar

Manoj Kumar Tiwari
Madhu Ranjan Kumar
Rofin T. M.
Rony Mitra *Editors*

Applications of Emerging Technologies and AI/ML Algorithms

International Conference on Data Analytics in Public Procurement and Supply Chain (ICDAPS2022)

Springer

Asset Analytics

Performance and Safety Management

The main aim of this book series is to provide a floor for researchers, industries, asset managers, government policy makers and infrastructure operators to cooperate and collaborate among themselves to improve the performance and safety of the assets with maximum return on assets and improved utilization for the benefit of society and the environment.

Assets can be defined as any resource that will create value to the business. Assets include physical (railway, road, buildings, industrial etc.), human, and intangible assets (software, data etc.). The scope of the book series will be but not limited to:

- Optimization, modelling and analysis of assets
- Application of RAMS to the system of systems
- Interdisciplinary and multidisciplinary research to deal with sustainability issues
- Application of advanced analytics for improvement of systems
- Application of computational intelligence, IT and software systems for decisions
- Interdisciplinary approach to performance management
- Integrated approach to system efficiency and effectiveness
- Life cycle management of the assets
- Integrated risk, hazard, vulnerability analysis and assurance management
- Adaptability of the systems to the usage and environment
- Integration of data-information-knowledge for decision support
- Production rate enhancement with best practices
- Optimization of renewable and non-renewable energy resources

Review Process

The proposal for each volume follows multi-pronged review process. The first level of review (single blind) is conducted by the series editors who may or may not decide to enlist the help of editorial board members for a second level of review. Proposals may also undergo a third level of peer review (double blind) if recommended by the Series Editors.

The series follows Ethics Statement found in the Springer standard guidelines here. https://www.springer.com/us/authors-editors/journal-author/journal-author-helpdesk/before-you-start/before-you-start/1330#c14214

Manoj Kumar Tiwari · Madhu Ranjan Kumar ·
Rofin T. M. · Rony Mitra
Editors

Applications of Emerging Technologies and AI/ML Algorithms

International Conference on Data Analytics in Public Procurement and Supply Chain (ICDAPS2022)

Springer

Editors
Manoj Kumar Tiwari
National Institute of Industrial Engineering (NITIE)
Mumbai, Maharashtra, India

Madhu Ranjan Kumar
Arun Jaitley National Institute of Financial Management
Faridabad, Haryana, India

Rofin T. M.
National Institute of Industrial Engineering (NITIE)
Mumbai, Maharashtra, India

Rony Mitra
Indian Institute of Technology Kharagpur
Kharagpur, India

ISSN 2522-5162 ISSN 2522-5170 (electronic)
Asset Analytics
ISBN 978-981-99-1018-2 ISBN 978-981-99-1019-9 (eBook)
https://doi.org/10.1007/978-981-99-1019-9

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Foreword

Governments procure large amounts of goods and services from the private sector that can enable them to implement policies and deliver public services. Across the OECD, public procurement expenditure is amounted to over 13% of GDP in 2020. As such, public procurement has the potential to be a potent economic and social lever that generates savings while supporting government goals for integrity, economic growth, inclusiveness, and sustainability.

An extensive examination of procurement data enables the identification of performance gaps and their underlying causes, the retention of focused initiatives, the methodical justification of alternative interventions and tactics, and the formulation of policy recommendations. Businesses should use the collected data to extract value from it and make significant business decisions. Data analytics is used to drive this purpose.

The supply chain management and public procurement applications of data analytics and machine learning are covered in this book volume. International Conference on Data Analytics in Public Procurement and Supply Chain (ICDAPS 2022) brought together leading international experts on procurement, service, supply, and logistics systems from academia, industry, and government to discuss pressing issues and research opportunities, mostly the applications of emerging technologies and AI/ML algorithms in E-procurement and Blockchains in Government Tendering, Supplier Prioritization & Risk Management in Procurement, Reverse Auction, Fintech-Based Digital Transactions Solutions, Perishable Food Supply Chain, and Vaccine Supply Chain Digital Twin.

Los Angeles, USA

Prof. Christopher S. Tang
UCLA Distinguished Professor
and Edward W. Carter Chair
in Business Administration

Contents

About the Editors

Prof. Manoj Kumar Tiwari (FNAE, FNASc, FIIIE, FIISE) has been Director, National Institute of Industrial Engineering (NITIE) Mumbai since November 2019. He is a Professor with Higher Academic Grade (HAG) in the Department of Industrial and Systems Engineering at the Indian Institute of Technology, Kharagpur. He is the winner of The David F. Baker Distinguished Research Award 2023. He is actively involved in research relevant to the application of optimization, modeling, decision support systems, and data mining in logistics, supply chain management, and manufacturing research domains. He is associate/senior editor of several renowned journals like IJPR, POMS, IISE, ESWA, JIM, IEEE-SMCA, INS, and CAIE. Prof. Tiwari has led numerous projects and consultancy assignments with prominent industry and government organizations in India, including Indian Air Force, Procter and Gamble, TATA Hitachi, and ports located in eastern India. In addition, he extensively engaged in collaborative research with partners from Loughborough University UK, Institute for Manufacturing (IfM) at the University of Cambridge UK, Warwick Manufacturing Group UK, Durham University UK, and the University of Connecticut USA. Prof. Tiwari is an author of more than 400 articles in leading international journals with H index-77.

Prof. Madhu Ranjan Kumar IRSS (Retd) heads the Procurement Research Centre (PRC) at Arun Jaitley National Institute of Financial Management (AJNIFM), Faridabad, India. The recent research areas of Prof. Kumar are use of data analytics in Reverse Auction and in pricing of bids. He is a Fellow of Indian Institute of Materials Management and a Fellow of Institution of Engineers. He is a Mechanical Engineer from IIT BHU and doctorate from Southern Cross University, Australia.

Dr. Rofin T. M. is currently working as an Assistant Professor in the area of Operations and Supply Chain Management at National Institute of Industrial Engineering (NITIE), Mumbai. He obtained his PhD from the Department of Industrial and Systems Engineering, Indian Institute of Technology Kharagpur. His research work is concerned with game-theoretic modelling of emerging multi-channel supply chain configurations, EPC contracts, Electric Vehicle Charging Supply Chain and Supply

Chain Risk and Resilience. He has more than 6 years of teaching experience at the post graduate level. He is a lifetime member of Operational Research Society of India.

Rony Mitra is a research scholar at the Indian Institute of Technology (IIT) Kharagpur with the Department of Mathematics. He completed his B.Sc. in Mathematics and M.Tech. in Computer Science and Data Processing. His areas of research interest are financial mathematics, financial modeling, stochastic processes, asset pricing, risk management, time series, portfolio management, supply chain management, data structures, etc. He developed innovative solutions for various industrial challenges, including predictive maintenance for baggage handling systems, sales forecasting for a multinational retail company, turn around time reduction for an Indian Port.

Chapter 1
Q-learning Approach to Mitigate Bacterial Contamination in Food Supply Chain

Meghna Maity, Ashesh Kumar Sinha, and Shing Chang

Abstract Due to globalization of markets, products are moving throughout the country and exported. Quality monitoring and traceability in a supply chain is hence essential. Blockchain technology ensures the privacy or resistance to data immutability of a decentralized database in case an attacker tries to alter or delete data inside the blockchain. The purpose of our research is to develop a holistic approach to make a perishable food supply chain resilient by ensuring a secure way of storing and monitoring supply chain data and identifying and optimizing the relevant factors that lead to bacterial contamination in food supply chain under uncertainty. We developed a blockchain interface to add, validate, and send data related to detailed wheat supply chain from farmers growing crops to harvesting, storing in the silos or elevators, then processors for milling and sifting to develop end products (bread, biscuit, buns, tortillas, etc.) and distribute to retailers who finally sell them to consumers. After the data is securely stored in the decentralized ledger, in our next step, we came up with a data-driven technology such as machine learning (Q-learning model) to filter relevant parameters impacting food quality at all stages of the supply chain and optimize them to ensure good quality.

Keywords Q-learning · Blockchain · Bacterial contamination · Wheat supply chain

M. Maity · A. K. Sinha (✉) · S. Chang
Industrial and Manufacturing Systems Engineering, Kansas State University, Manhattan, USA
e-mail: sinhaa@ksu.edu; ashesh.iitian@gmail.com

M. Maity
e-mail: mmaity@ksu.edu

S. Chang
e-mail: changs@ksu.edu

M. K. Tiwari et al. (eds.), *Applications of Emerging Technologies and AI/ML Algorithms*, Asset Analytics, https://doi.org/10.1007/978-981-99-1019-9_1

1.1 Introduction

With a predicted increment in population all over the globe, there is also an increase in requirement of food availability. Apart from population rise, unfortunate weather events, pandemic situations like COVID-19, or bacterial contamination of products places unprecedented disruptions in front of perishable food supply chains—bottlenecks in transport and logistics of perishable products, unexpected consumer demand, and food loss to name a few. An immense pressure to make a robust supply chain motivated our research work.

About 6–7% of the world's total wheat is produced by the United States alone and it is a major exporter (11%) of wheat too [1]. Wheat is the third most grown product after corn and soybeans in the United States. It is the principal source of grain consumption in the country and hence our point of interest. We focus on developing a dynamic monitoring system for a wheat supply chain to reduce bacterial contamination and maintain food quality. Research has proceeded profoundly in terms of traceability of products-RFID tags used for temperature tracking [2] to blockchain technology by Walmart and Nestle to monitor quality [3] and satisfy consumer expectations. The insights from the model reveal that the use of decentralized database could make these data accessible to peers (retailers, suppliers, and manufacturers) while mitigating data tampering by virtue of blockchain technology. This not only ensures transparency among the peers but also privacy and immutability of the data present in the decentralized database of different supply chain networks. Once the data is securely stored in the decentralized ledger, our next step is to identify parameters that affect the quality of a perishable item (wheat in our case) due to bacterial contamination. Such parameters can be temperature, humidity, pH of wheat in silos, etc., leading to contamination or bacterial growth at any point along the chain—production, processing, preparation, and distribution. According to CDC, more than 100 people were infected by a Salmonella outbreak in packaged seafood in 2021 and 1040 people were affected due to onions containing bacterial growth in 2020 [4]. Data was collected conventionally by reviewing records from restaurants and stores. Since past work mostly focused on how to trace or track batches that were already contaminated, we came up with a data-driven technology such as machine learning (Q-learning model [5]) to filter relevant parameters impacting food quality at all stages of the supply chain and optimize them to ensure good quality.

1.2 Traceability in Supply Chain Data—Blockchain Technology

With the invention of blockchain, centralized way of data storage is shifting to decentralized data. Supply chain systems involve complex entities to carry out the production and transportation of products through various stages [6]. We exclude export and

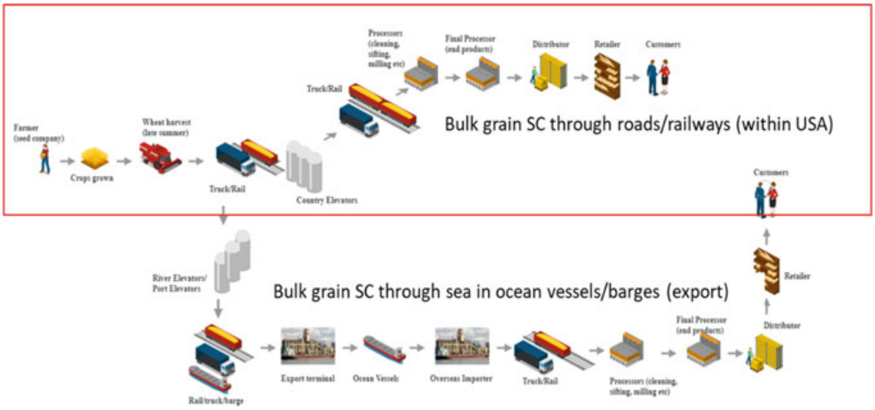


Fig. 1.1 A detailed wheat supply chain

consider a detailed wheat supply chain network inside the country for our purpose of study (see highlighted in the red box in Fig. 1.1).

To attain traceability and transparency, we developed a blockchain interface to store supply chain data. This platform enables us to dynamically monitor stored data related to farming, harvesting, transportation, storage, and parametric conditions like temperature, humidity, etc. We can add, send, validate, and update new block in the chain.

1.3 Monitoring Parameters and Optimization: Q-learning Model

In the upper half of Fig. 1.1, wheat is grown by farmers during the crop season and harvested at peak time which is late summer. Harvested wheat is transported by truck or rail logistics and stored in country elevators. During bumper harvest, the biggest challenge faced is storing wheat with keeping its quality intact in silos or elevators. Wheat, mostly in the storage areas is exposed to various parameters like temperature, CO_2 level, humidity, and pressure leading to bacterial or fungal growth affecting the quality of the product. From the country, it is further transported for cleaning, sifting, and milling in processing units. Distributors collect the processed wheat in the form of bread, biscuit, buns, tortillas, pasta, noodles, etc., and send it to retailers like Kroger, Walmart, Aldi, etc. Customers then can collect the final item which is ready to consume. Our point of investigation concerns areas where managing the quality of wheat is quintessential. This section outlines the Q-learning model of the supply chain network to monitor the quality of wheat at potential risk areas. We consider five levels of bacterial contamination from very high quality of wheat that has the lowest contamination to very low quality of wheat that has the highest contamination. For simplicity, we take two parameters temperature and humidity under study that affect the growth of bacteria in wheat.

Following are the key elements in Q-learning model formulation in determining the best course of action that would ensure good quality of wheat as it travels through the entire supply chain.

Environment: It is the abstract world through which the agent moves. The environment takes the current state and action of the agent as input and returns its next state and appropriate reward as the output.

States: The specific place at which an agent is present is called a state. This can either be the current situation of the agent in the environment or any of the future situations. A set of states S_k is defined as a tuple (k, q_k), where k is each step of supply chain starting from farming to harvesting, truck/rail, country elevators, processing units to finally customers. Quality of wheat in each step k ranges from the lowest quality of wheat, value represented by 1, to the highest quality of wheat, represented by 5, such that quality level $q_k =$ (very low, low, medium, high, very high). Table 1.1 shows the possible states of our model. Here, $S_3 = (3, 4)$ represents the state truck/rail and quality of wheat inside is high. Note that in our model there are 8 stages and 5 quality levels giving a possibility of 40 states that wheat can be in.

Actions: We define the action space as a tuple of temperature and humidity as parameters affecting bacterial contamination/quality of a state represented as $A_m = (t, h)$, where m is the action number. Temperature value and humidity values are represented by t and h, respectively. Values of t and h vary from 10 °C to 50 °C and 10% to 50%, respectively, with a step size of 1. Performing action $A_m = (t, h)$, the quality of wheat transitions from S_k to state S_l as shown in Fig. 1.2. Table 1.2 shows 5 possible actions for our model. Here, action $A_1 = (10, 30)$ represents the first action where the temperature and humidity at a stage of supply chain are 10 °C and 30%, respectively. Note that a combination of 41 values of temperature and humidity makes 41 * 41 = 1681 possible actions.

Reward: The system is in S_k and decision-maker chooses any action A_m available in state S_k resulting in a new state S_l in the next time period and giving the decision-maker a corresponding reward $R_{A_m}(S_k)$.

The Q-learning algorithm calculates the quality of a state–action combination given by the standard Eq. (1.1) [5].

Table 1.1 Possible set state space

S_k	k	q_k
S_1	1	1
S_2	1	2
S_3	3	4
S_4	5	4
S_5	5	5

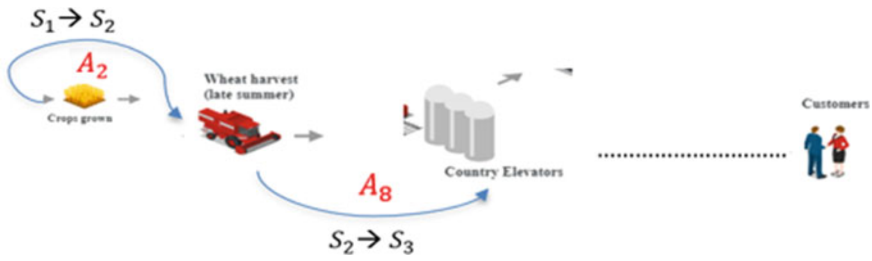


Fig. 1.2 Transition of one state to another by taking actions A_2 and A_8

Table 1.2 Possible action space

A_m	t	h
A_1	10	30
A_2	20	29
A_3	35	10
A_4	17	20
A_5	50	32

$$Q^{new}(S_k, A_m) \leftarrow Q(S_k, A_m) + \alpha\left[R_{A_m}(S_k) + \gamma * \max_{\forall A} Q(S_l, A) - Q(S_k, A_m)\right] \tag{1.1}$$

Note that $Q{:}S \times A \rightarrow R$, where S, A, and R are states, actions, and rewards, respectively. Q-learning works when we don't know which states are good or bad, or by taking an action which state we land on, in other words, the value of transition probability. It works by estimating the values of Q^{new} by exploring and exploiting. The model must try out state–action pairs from the data to figure out the best course of action that yields the maximum expected discounted reward. γ is the discount factor and α is the rate at which the model learns. Both values of γ and α range from 0 to 1.

1.4 Experiments and Analysis

We assume the performed an experiment at the elevator stage of the wheat supply chain to show the best and worst possible values of temperature and humidity at each quality level of the wheat. For country elevators, best and worst temperature values explored according to the Q function are represented in Fig. 1.3.

If elevator has a very low quality of wheat, the best decision would be to maintain temperature at 39 °C. If elevator has a high quality of wheat, the worst decision would be to maintain temperature at 33 °C. Similarly, for country elevators, the best and worst humidity values explored according to the Q function are represented in Fig. 1.4. If elevator has a medium quality of wheat, the best decision would be to

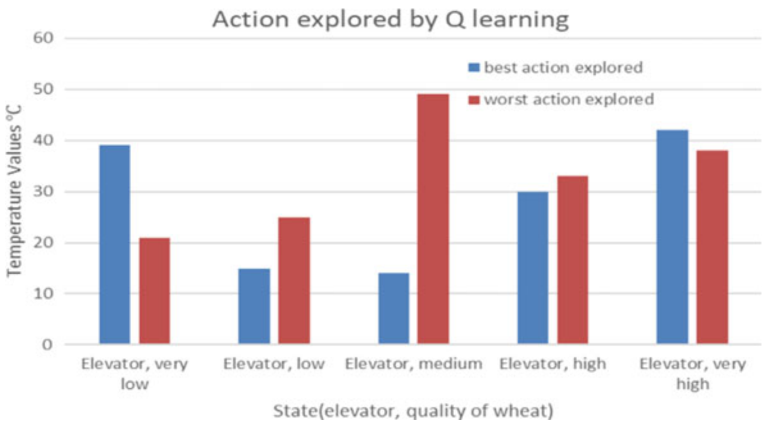


Fig. 1.3 Temperature values to select/avoid

maintain humidity at 49%. If elevator has a very high quality of wheat, the worst decision would be humidity at 36%.

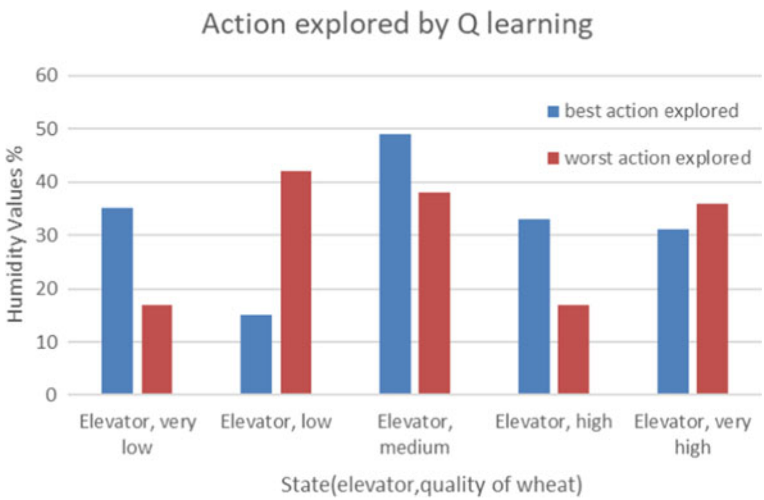


Fig. 1.4 Humidity values to select/avoid

1.5 Conclusion

We consider a detailed wheat supply chain where temperature and humidity are the two main parameters responsible for bacterial contamination at any point in the network. We implement a Q-learning algorithm because in our model transition probability is not known. This method determines the best values of temperature and humidity that are to be maintained to ensure good quality of wheat at every stage of the supply chain. A decision-maker can have an idea of what could be the best and worst decisions in determining the quality of wheat. The algorithm in our model is as good as the data. In terms of future research, we intend to explore interactions between several parameters other than temperature and humidity and validate the model with realistic data.

References

1. https://www.ers.usda.gov/topics/crops/wheat/wheat-sector-at-a-glance/
2. Amador C, Emond JP (2010) Evaluation of sensor readability and thermal relevance for RFID temperature tracking. Comput Electron Agric 73(1):84–90
3. Caro MP, Ali MS, Vecchio M, Giaffreda R (2018) Blockchain-based traceability in agri-food supply chain management: a practical implementation. In: 2018 IoT vertical and topical summit on agriculture-Tuscany (IOT Tuscany). IEEE, pp 1–4
4. https://www.cdc.gov/salmonella/outbreaks.html
5. Sutton RS, Barto AG (2018) Reinforcement learning: an introduction. MIT Press
6. Shahid A, Almogren A, Javaid N, Al-Zahrani FA, Zuair M, Alam M (2020) Blockchain-based agri-food supply chain: a complete solution. IEEE Access 8:69230–69243; Watkins CJ, Dayan P (1992) Q-learning. Mach Learn 8(3):279–292

Chapter 2
Optimization of Network Planning in a Real-Life Vehicle Logistics Distribution System

Vipul Kumar Mishra, Chandrasekharan Rajendran, Franz Lenher, A S. R. Suryanarayana Murty, A. S. Balakrishnan, Artie Jina, and Hinoj Pallath

Abstract The present work has been carried out in the logistics distribution of an automobile manufacturer in South Africa for improving their outbound vehicle logistics distribution efficiency. As far as we know, the literature on the distribution and logistics systems used in the automotive industry does not currently contain a Mixed Integer Linear Programming (MILP) model for implementing First In First Out (FIFO). This study presents the development of a comprehensive shipment plan for the automobile manufacturer to (i) deliver finished cars from the manufacturing plant to other facilities and (ii) adhere to a range of loading, routing, and capacity constraints while meeting the vessel-loading timetable and shipment plan with the minimum transportation and inventory costs. The model accounts for dynamic parameters such as train and trailer availability, production volume, ship timetable, and demand at the ports of South Africa. An MILP model is developed to capture the business constraints and the time-dependent storage costs at the facilities, which vary depending on the number of days cars are stored in the facilities. The proposed MILP model uses FIFO strategy to ensure that finished cars leave port warehouses in the same order that they arrived. The model offers a cost-effective

V. K. Mishra (✉) · C. Rajendran · A. S. R. Suryanarayana Murty
Indian Institute of Technology Madras, Chennai, India
e-mail: vipulmishra5659@gmail.com

C. Rajendran
e-mail: craj@iitm.ac.in

F. Lenher
University of Passau, Passau, Germany
e-mail: franz.lehner@uni-passau.de

A. S. Balakrishnan · H. Pallath
Ford Motor Private Limited, Chennai, India

A. Jina
Ford of Southern Africa, Pretoria, South Africa

M. K. Tiwari et al. (eds.), *Applications of Emerging Technologies and AI/ML Algorithms*, Asset Analytics, https://doi.org/10.1007/978-981-99-1019-9_2

shipment plan to meet the vessel demands and the daily transportation mode requirements. These results will benefit the decision-maker in automating the shipment plan.

2.1 Introduction

South Africa has long been Africa's automotive leader, producing over half a million automobiles of various varieties every year. With the advancement of technology, the automobile manufacturing business is rapidly expanding. The automotive manufacturing supply chain is one of the most complex supply chains in the world [1].

Automobile distribution is a critical component of the logistics of the automotive industry, involving the shipping of finished automobiles from hundreds of manufacturing factories across multiple continents. The shipping of completed automobiles is a crucial aspect of the entire automobile manufacturing sector.

The efficient distribution and transportation of vehicles require efficient supply chain management. The automotive supply chain is becoming increasingly complex because of competition and technological advancements, necessitating practical solutions from vehicle manufacturers and brands. The primary purpose of the current work is to map the hurdles in the outbound logistics process and develop an efficient shipment plan. Outbound logistics focuses on the demand side of the supply–demand equation. The procedure entails storing and transporting items to the customers and final users. The inventory and transportation decisions are the two most important elements in outbound logistics. The proper inventory and transportation strategies can help a company save money and succeed in their business.

With increasing knowledge on the supply chain and the introduction of new technologies, there is a need for further contribution from research studies on real-world settings to understand better the impact of accurate shipment plans on the supply chain, their outcomes, and effectiveness.

2.2 Problem Description

The major challenge for the automobile manufacturer in this study is to develop a detailed shipment plan. The manufacturer faced problems in determining the quantity and schedule of shipment of finished cars from one site to another. The automaker has the manufacturing facility and have a storage compound. The finished cars need to be transported to the Port Elizabeth and Port Durban to meet the demand and deadline of the ships. They also have rented facilities near the ports for the storage of cars. The storage of cars at the 3PL warehouses and ports involves storage costs. The storage cost at the ports is differential in nature, such that it varies with the number of days cars are stored in the port. The shipment of cars can be made via trailers or trains. There are several constraints on the availability of trailers and train capacity.

The manufactured cars can be stored at the Silverton manufacturing plant and the Westwood storage compound, and the 3PL warehouses near the ports. The storage at Silverton and Westwood is free of cost as these facilities are owned by the car manufacturer. The storage at the 3PL warehouses is rented based on the need, and the storage of cars is charged on a per-day basis. When the cars are stored at the ports, the storage for the first 14 days is free at the port, and after that port authority charges a very high storage cost, which increases with every single day car stay at the port.

The arrows in Fig. 2.1 indicate the possibility of transportation of cars. The blue arrow indicates that the cars can be transported via road only, and the red arrow indicates that the cars can be transported using rail as well as road. The trailers are available on all days of the week, but their capacity is limited, whereas trains are available only on a few days of the week. The transportation cost by road is cheaper than by rail. The finished cars must be available at the ports when the ship arrives. To tackle these complexities, a Mixed Integer Linear Programming (MILP) model is developed. First In First Out (FIFO) logic is implemented at the ports to reduce the penalty cost at the ports. The proposed MILP model accommodates the FIFO logic, which is not accounted for, in most similar works in the available literature.

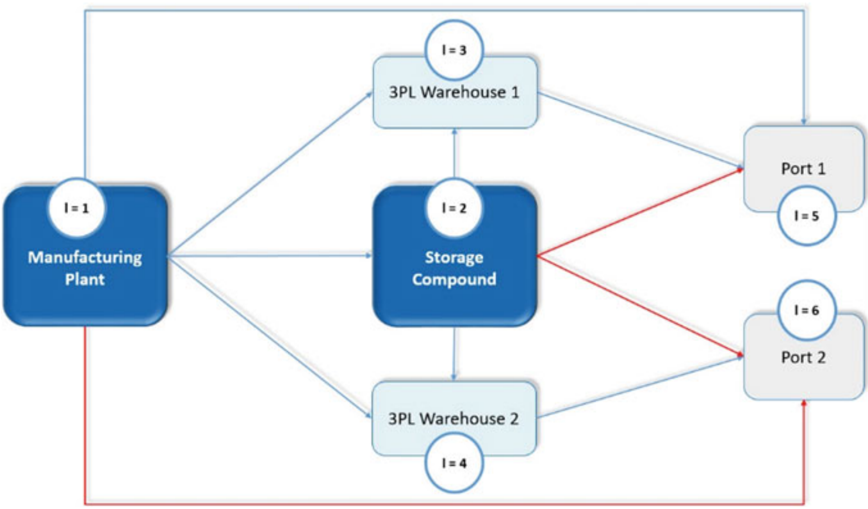


Fig. 2.1 Network diagram

2.3 Related Research

Optimization of network planning involves optimizing the storage cost at different locations and the transportation cost by different modes of transport. The optimization process involves three significant decisions, including (1) Loading of finished cars in different auto carriers, (2) Transportation route of auto carriers, and (3) Inventory level at each of the locations. In our paper, a mathematical model is proposed to solve the above-listed problems in an integrated manner. To the best of our knowledge, there is no MILP model for implementing FIFO available to date in the literature on the logistics and distribution systems of the automotive industry.

The first formal method to solve the loading problem is reported in 1998 [2]. The authors characterized the auto carrier as a fixed set of slots, and they created a loading network to express LIFO precedence between slots. The loading network and any additional pairwise incompatibilities between slots and vehicles are considered when assigning a slot to each vehicle. They used a branch-and-bound strategy to solve a nonlinear NP-hard assignment problem.

Researchers have used different methods to solve the loading and routing problem of the automotive industry. A genetic algorithm has been used to resolve the mode of transportation allocation to minimize the logistics cost [3]. The MILP model has also been used in the automotive industry for the optimization of transportation cost where multimodal transportation is used [4]. Also, some researchers used iterated local search algorithm for the routing part and mathematical model and enumeration techniques for the loading part [5].

With increasing competition in the automotive sector, multi-modal transportation for outbound logistics is vital to increase supply chain efficiencies. To optimize the multimodal transportation cost, a MILP model is proposed to provide a real-life automobile manufacturer with an effective outbound transportation network plan [3]. Road and sea modes are considered for transporting the finished cars in their model with the objective to minimize the total transportation cost. They did not consider the strategy of shipment in the objective function of their MILP model.

The automobile manufacturer uses two modes of transportation, and the inventory carrying costs are different at all the locations, including the differential carrying cost at the ports. So, this paper proposes an MILP model that accounts for all the scenarios mentioned above for an automobile manufacturer with the objective of minimizing the transportation and storage costs. FIFO strategy is also implemented mathematically at the ports to minimize the storage costs.

2.4 Methodology

This section presents a mathematical model developed to obtain the shipment plan. An MILP model is built, capturing all the business constraints with an objective function of minimizing the total cost that includes the inventory carrying cost at all the locations and transportation cost by different modes of transportation. MILP is a mathematical modeling technique for determining the best solution for a system with constraints. It is widely used in a variety of optimization fields, including production planning, transportation and network design, to name a few. MILP models are significantly more challenging to solve than linear programming models.

Following are the notations used in the mathematical model:

Indices:

m index of model (m = 1, 2, …, 7)
t,t' index of days (t,t' ∈ {1,2,-----,28})
l index for locations (1, 2, …, 6)
l' index for locations (1, 2, …, 6),
SD = {(l,l') = (1,2), (1,3), (1,4), (2,3), (2,4), (2,5), (2,6), (2,5), (3,5), (4,6)}

Parameters:

T_c capacity of a trailerl (in terms of the number of units).
T_r capacity of a trainl (in terms of the number of units).
C_l Storage capacity of the facility at location l l(in terms of the number of units).
$P_{m,t}$ Number of units of model m produced on day t.
$CT_{l,l'}$ Per unit cost of transportation by trailer from l to l'.
$CR_{l,l'}$ Per unit cost of transportation by rail from l to l'.
$SC_{t,l}$ Storage cost for day t at location.
$T^A_{l,l',t}$ Number of trailers available on day t from l to l'.
$R^A_{l,l',t}$ Number of trailers available on day t from l to l'.
$D_{m,l,t}$ Number of units of model *m*required on day l at location l.

Decision Variables:

$N_{m,l,t}$ Number of units of model m entering location l on day t.
$N'_{m,l,t}$ Number of units of model m leaving location l on day t.
$S^T_{m,l,l',t}$ Number of units of model m shipped from l to l' by Trailer on day t.
$S^R_{m,l,l',t}$ Number of units of model m shipped from l to l' by Train on day t.
$T_{l,l',t}$ Number of trailers used on day t from l to l'.
$R_{l,l',t}$ Number of trains used on day t from l to l'.
$N1_{m,t',t,l}$ Number of units of model m leaving location l on day t' which came on day t.
$IN_{m,t,l}$ Inventory for model m on day t at location l.

Objective function:

Min Cost

$$
\begin{aligned}
&= \sum_{l,l\prime \in SD}\left(\sum_{m}\sum_{t} CT_{l,l\prime} * S^{T}_{m,l,l\prime,t}\right) + \sum_{l,l\prime \in SD}\left(\sum_{m}\sum_{t} CR_{l,l\prime} * S^{R}_{m,l,l\prime,t}\right) \\
&\quad + \sum_{t}\sum_{l=1}^{4}\left(\sum_{m} SC_{t,l} * IN_{m,l,t}\right) + \sum_{t,t\prime}\left(\sum_{l=5}^{6}\left(SC_{(t-t\prime),l} * \sum_{m} N1_{m,t\prime,t,l}\right)\right) \qquad (2.1)
\end{aligned}
$$

Constraints:

$$N_{m,t,1} = P_{m,t}, \forall m, t \qquad (2.2)$$

$$N_{m,l',t} = \sum_{l} S^{T}_{m,l,l',t-2} + \sum_{l} S^{R}_{m,l,l',t-2}, \forall l', m, t, (l, l\prime \in SD) \qquad (2.3)$$

$$N'_{m,l,t} = \sum_{l'} S^{T}_{m,l,l',t} + \sum_{l'} S^{R}_{m,l,l',t}, \forall m, t, (l, l\prime \in SD) \qquad (2.4)$$

$$\sum_{t'=1}^{t} N_{m,l,t'} \geq \sum_{t'=1}^{t} N\prime_{m,l,t'}, \forall m, t, l \qquad (2.5)$$

$$C_{l} \geq \sum_{m=1}^{7}\sum_{t'=1}^{t} (N_{m,l,t'} - N'_{m,l,t'}), \forall t, l \qquad (2.6)$$

$$T^{A}_{l,l',t} \geq T_{l,l',t} \geq \frac{\sum_{m=1}^{7} S^{T}_{m,l,l',t}}{T_{c}}, \forall t, (l, l' \in SD) \qquad (2.7)$$

$$R^{A}_{l,l',t} \geq R_{l,l',t} \geq \frac{\sum_{m=1}^{7} S^{R}_{m,l,l',t}}{R_{c}}, \forall t, (l, l' \in SD) \qquad (2.8)$$

$$IN_{m,l,t} = \sum_{t'=1}^{t} (N_{m,l,t'} - N'_{m,l,t'}), \forall m, t, l \qquad (2.9)$$

$$N_{m,l,t} = \sum_{t'=t}^{\min(28,t+28)} N1_{m,t,t',l}, \forall m, t; l \in (5, 6), t\prime \geq t \qquad (2.10)$$

$$N'_{m,l,t} = \sum_{t'=max(t-28,1)}^{t} N1_{m,t',t,l}, \forall m, t; l \in (5, 6), t \geq t\prime \qquad (2.11)$$

$$\sum_{t'=1}^{t} N'_{m,l,t'} = \sum_{t'=1}^{t} D_{m,l,t'}, \forall m, t; l \in (5, 6) \qquad (2.12)$$

The model's objective function is depicted in equation number (2.1). The objective function includes the transportation cost by train and trailers and the storage cost at 3PL warehouses and ports.The storage cost at ports follows an incremental structure, meaning that the cost increases as the duration of storage extends. For example, if cars are stored for one day, the cost per unit is \$50. However, if the storage period extends to two days, the cost per unit per day becomes \$150, and for the third day, it rises to \$350 per unit per day. This incremental cost is represented by the last term in the objective function. By incorporating this term, the implementation of FIFO (First In, First Out) logic is ensured at ports, thereby minimizing the storage cost incurred. It's important to note that these numbers are used for illustration purposes and may not reflect the actual costs. Equation (2.2) is used to ensure that whatever is getting produced at the manufacturing plant is entering the storage at the manufacturing plant. Equation (2.3) ensures that the inflow at the location equals the sum of outflows from all other associated locations. Equation (2.4) ensures the outflow at a location is the sum of cars transported by trains and trailers to all other associated locations. Equation (2.5) maintains the conservation of flow . Equation (2.6) is for ensuring that the capacity at a location is not exceeded. Equations (2.7, 2.8) are for the capacity constraints on the trailer and train capacity. Equation (2.9) gives the indication about the inventory level at a facility. Equations (2.10–2.12) implement the FIFO logic at the ports.

The above model has been solved using the CPLEX solver in GAMS software based on the real-life situation.

2.5 Results

The model is tested on several dummy datasets to understand its working in different scenarios. The model is run for the 28 days cycle. On the first day of the cycle, the model is run using the forecasted values of production and demand data. From the second day onwards, the actual shipment is fed as yesterday's production data to obtain the revised shipment plan for the remaining days of the cycle.

The model's output is converted into a detailed shipment plan that indicates the number of units of finished cars for a particular market that needs to be transported from which origin to which destination and by which mode of transport. The requirement of train and trailer capacity is obtained as the model's output and the inventory level at each of the locations.

Currently, shipments are made based on a subjective system based on the daily availability of trains and trailers.

The exact data and the developed shipment plan cannot be presented due to the data privacy concerning the automobile manufacturer.

2.6 Conclusions

Mixed integer linear programming model is used in this study to identify the best shipment schedule for a real-world vehicle manufacturing company in South Africa. Working on a real-world situation makes the research more practical, expands knowledge in the area of supply chain management, logistics, and transportation, and allows our model to be implemented using real-time scenarios. The presented transportation and inventory carrying cost-optimized model with FIFO strategy of shipment and its computational result can be used as part of outbound logistics planning, assisting tactical decision-making on the mode of transport, the need for different modes of transport, delivery quantity, and delivery schedule from each of the locations while also meeting demand at ports in terms of quantity and time.

References

1. B. Global (2022) Top 4 Automotive supply shain challenges and solutions. https://www.blumeglobal.com/learning/automotive-supply-chain/
2. Agbegha GY, Ballou RH, Mathur K (1998) Optimizing auto-carrier loading. Transp Sci 32(2):174–188
3. Ngọc HGB, Phong HT Multimodal transportation optimization for an automobile production in Vietnam
4. Zeng T, Hu D, Huang G (2013) The transportation mode distribution of multimodal transportation in automotive logistics. Procedia Soc Behav Sci 96:405–417
5. Dell'Amico M, Falavigna S, Iori M (2015) Optimization of a real-world auto-carrier transportation problem. Transp Sci 49(2):402–419

Chapter 3
Efficient Supplier Selection in e-Procurement Using Graph-Based Model

Puja Sarkar, Rony Mitra, Vivekanand B. Khanapuri, and Manoj Kumar Tiwari

Abstract Public procurement is described as a dynamic, real-time negotiation between a purchasing association and several pre-qualified suppliers contesting against one another to achieve the opportunity to deliver goods or services to the purchasing organization. A significant number of enterprises use Reverse auction (RA) for supplying products. This paper investigates aspects of the RA system using the optimization model to keep up the work on comprehending the characteristics of the impact of performing RA on the bidding results for multiple products. In particular, a graph-based model has been developed for the RA bidding framework to evaluate the consequences of the bids collected in public procurement. Our method aims to confirm the determination, whether procurement is profitable based on the previous auction datasets. The analysis involves a scenario consisting of companies that will dominate the market. This study also consists in determining the decision to occur an auction.

Keywords Auctions/bidding · Reverse auctions · e-procurement · Strategic procurement · Multiunit auctions

P. Sarkar (✉) · R. Mitra · V. B. Khanapuri · M. K. Tiwari
National Institute of Industrial Engineering Mumbai, Mumbai, India
e-mail: puja.921013@gmail.com

V. B. Khanapuri
e-mail: vbkhanapuri@nitie.ac.in

M. K. Tiwari
e-mail: mkt09@hotmail.com

R. Mitra · M. K. Tiwari
Indian Institute of Technology Kharagpur, Kharagpur, India

M. K. Tiwari et al. (eds.), *Applications of Emerging Technologies and AI/ML Algorithms*, Asset Analytics, https://doi.org/10.1007/978-981-99-1019-9_3

3.1 Introduction

In the modern world of technology, online reverse auctions (RA) have developed swiftly due to the rage of digitization. Prices can be proposed indefinitely for each round of a RA before the termination of the bidding duration. The cut-price can be seen in real time as credentials (but it is unknown who submits the cut-price), or the bidder can see only their own position in the bidding stack at the bidding time. After the bidding is over, the organizer declares the supplier as winner who has won the bid.

In traditional RAs, buyers need to precisely describe product features like origination, descriptions, models, and additional details, which are the premise of RA. In point of fact, buyers often are not familiar with the attributes when buying products. Chen et al. [2] demonstrated a structural depiction of optimal lot-sizing policies using super modularity arguments, concavity, and monotonicity for expected revenue function on the single-auction. Huang et al. [4] framed an online auction framework based on the limited quantity of items and established a profitable online auction for buyers and sellers. Instead, buyers should exhaustively deal with the opportunities and obstacles of individual characteristic and select a seller with the most incredible capability to supply goods [8]. This procedure is essential to depict and pertain to a new RA mechanism called the multi-attribute RA.

In the case of competitive procurement [9], the supplier (company) that will sign a contract is determined via online auction. Procurement requirements (specification of goods/services/works) are published before the auction and any company can apply online by downloading the necessary documents and becoming a participant. During the auction, participants are competing by lowering price that they ask for fulfilling the contract. After three rounds of auction, participant that offered the lowest price is announced a winner. After this, organizer checks the documents submitted by the winner before the auction. If everything is alright, organizer signs the contract with the winner and this company becomes a supplier. If there are problems with documents, organizer disqualifies such company and proceeds to check the documents of company that offered the second lowest price. This process is continued until the supplier is determined or there are no more non-disqualified participants left. Li et al. [6] studied the mixed-strategy equilibrium in which the buyer designed an optimal procurement policy with an unobservable supplier. Khamesi et al. [5] introduced a mechanism for reverse auction to facilitate an incentive-based demand response auction. Engelbrecht et al. [3] demonstrated the dominant strategy of the bidder, which is a part of a hybrid mechanism. To determine the real-time bidding [1], an auto pricing strategy can be used by the advertising agencies.

The initial value of public procurement auction corresponds to the reserve price from the auction theory. It is the highest acceptable value of contract that is published by organizer as a part of the procurement announcement. During the auction, the first bid of a participant should be lower or equal to the initial value. The final value is the lowest value of a contract offered as a result of an auction.

The rest of this paper is structured as follows. Section 2 describes the problem related to reverse auction. In Sect. 3, we illustrate the solution methodology. In Sect. 4, we explain the results and solutions. Finally, Sect. 5 concludes the paper and provides some future research directions.

3.2 Problem Description

Our main objective is to find out whether an online RA is fruitful for an organizer or not. Also, to find out expedient participants for participating in RA specifying the suppliers (i.e., companies or vendors) which will dominate in the market winning the highest number of tenders and to decide the time interval for successfully occurring RA.

3.3 Methodology

The auction process typically solves multiple item selections from sellers' perspectives using the online RA process. This section instigates the detail explanation of the components of auction model. This approach is trying to find out the effectiveness of public procurement, which is held yearly. Also, it is analyzed whether there is some evidence of increased competition, and whether that always leads to increased public savings or not. For this purpose, first procuring entities are distinctly identified from past available RA datasets and determine the total number of participants who took part in a reverse auction for a particular object. The participants who win the contract become suppliers. Then, the graph-based algorithm is used to find the companies that will dominate in the market, winning maximum agreements. This will help organizers to get information on which companies are more active in procurement. The algorithm consists of creating graphs out of data with configurable data points as nodes of the graph. A graph $G(V, E)$ is constructed from available datasets where $\boldsymbol{V}$ is the set of unique vertices or nodes derived from the dataset features "*participant_code*", "*organizer_code*", "*lot_cpv*" and $\boldsymbol{E}$ is the set of edges. Edges are defined between two nodes if one participant participated in a RA conducted by an organizer for a particular product. The out-degree for each node of "*participant_code*" will exactly decide the active participation of companies in the auction. For example, in Fig. 3.1, $B = \{B_1, B_2, B_3\}$ is the set of organizer codes, $P = \{P_1, P_2\}$ is the set of procured products and $C = \{C_1, C_2, C_3, C_4\}$ is the set of participant codes in reverse auction and each element of the sets B, P, C are considered as nodes of the graph $\boldsymbol{G}$. The

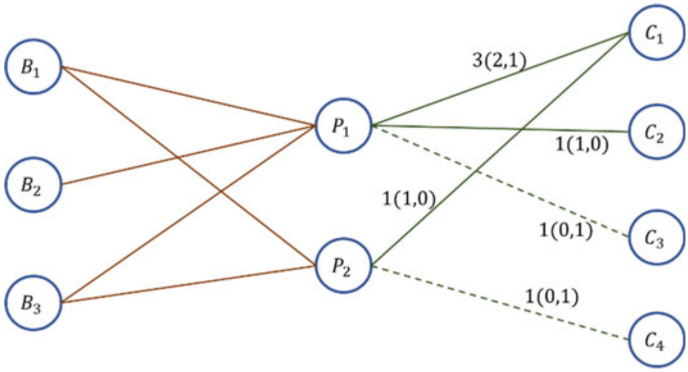


Fig. 3.1 Graph-based visualization of multi-item reverse auction

edge weight is given in the format $a(b, c)$, where a is the total number of times a participant participated in the auction, b is the count of winning the auction, and c is the count of not winning the auction. The final graph considers all the records similar to Fig. 3.1.

Considering the participants to become suppliers winning in multiple reverse auctions is usually demonstrated as the supplier determination problem (SDP). Organizers and participants are auctioneers and bidders, respectively. The organizer's goal is to pick out the top bidding portfolio to maximize their demands to satisfy all the commodity requirements based on the final bid proposed by the participant. The participant company who has concluded the transaction and submits the selected price is the bidder. Our goal is to find the company or supplier that will dominate the market, winning the tenders in RA. Then the SDP can be expressed as an IP problem (1).

$$MaxZ = \sum_{i \in N, j \in M} w_i x_{ij} \tag{3.1}$$

$$Subject\ to,$$

$$\sum_{i=1}^{N} w_i = 1 \tag{3.2}$$

$$w_i = \{0, 1\}, \forall i \in N \tag{3.3}$$

$$x_{ij} \leq \alpha d_{ij}, \forall i \in N, \forall j \in M \tag{3.4}$$

where variable parameter α is the probability of winning for a participant in multiple reverse auctions. In the objective function, N contains the list of distinct procured objects, and M is the unique list of participants who participated in the RA. x_{ij} is the total count of winning contracts for a participant in RA for a particular object. d_{ij} is the sum of participants for a selective object or product in procurement. w_i is a binary variable. When the participant i is selected, $w_i = 1$; otherwise, $w_i = 0$. If there have not sufficiently strong, competitive, and qualified suppliers who will participate, RA should not organize by the purchaser. Also, the organizers have to check the profitability of RA concerning the time, cost, and other relevant factors for conducting a successful RA. The organizer will calculate the cost for conducting the auction and the percentage of price reduction in the RA and will come to a conclusion defining a user-defined threshold (δ) for coming to a decision. If the percentage of price reduction exceeds the threshold (δ), we have to check the percentage of common suppliers per year. Then it can be decided to perform an auction if most of the suppliers are not repetitive. The minimum percentage rate of common suppliers is also user desirable threshold (γ). This proposed approach is analysed and illustrated in Fig. 3.2.

3.3.1 Data Collection

For the purposes of our analysis, we obtained 5 years (2015–2019) of detailed transactional data from Ukrainian's online reverse auction system provided by the Parliament of Ukraine "On public procurement," which launched the electronic procurement system. Each year was provided with the following data sets: an item-level data set containing detailed information about each auction with one procurement and one participant (company), representing a total of 2,313,168 observations. Thus, one procurement is related to multiple rows if more than one company competes to become a supplier. This dataset includes all companies that supplied goods or services to ulfil a public procurement consisting of both competitive and non-competitive procurement contracts. It provides procurement objects or products, organizer code, a description of the procurement entity, purchased goods/services/works, and a code of the supplier company. The data preparation phase consists of two main steps. The first step is to summarize the data to obtain a final dataset in which each observation represents an auction conducted for a specific item. The second step is to clean the data to ensure its consistency as well as the coherence of the final sample used for the analysis. This involves removing observations with inconsistent dates and times, duplicated identifiers, or missing values [7]. The main variables are generated during the data preparation phase including measures of (i) potential savings of price from the initial fixed price; and (ii) common suppliers of each year.

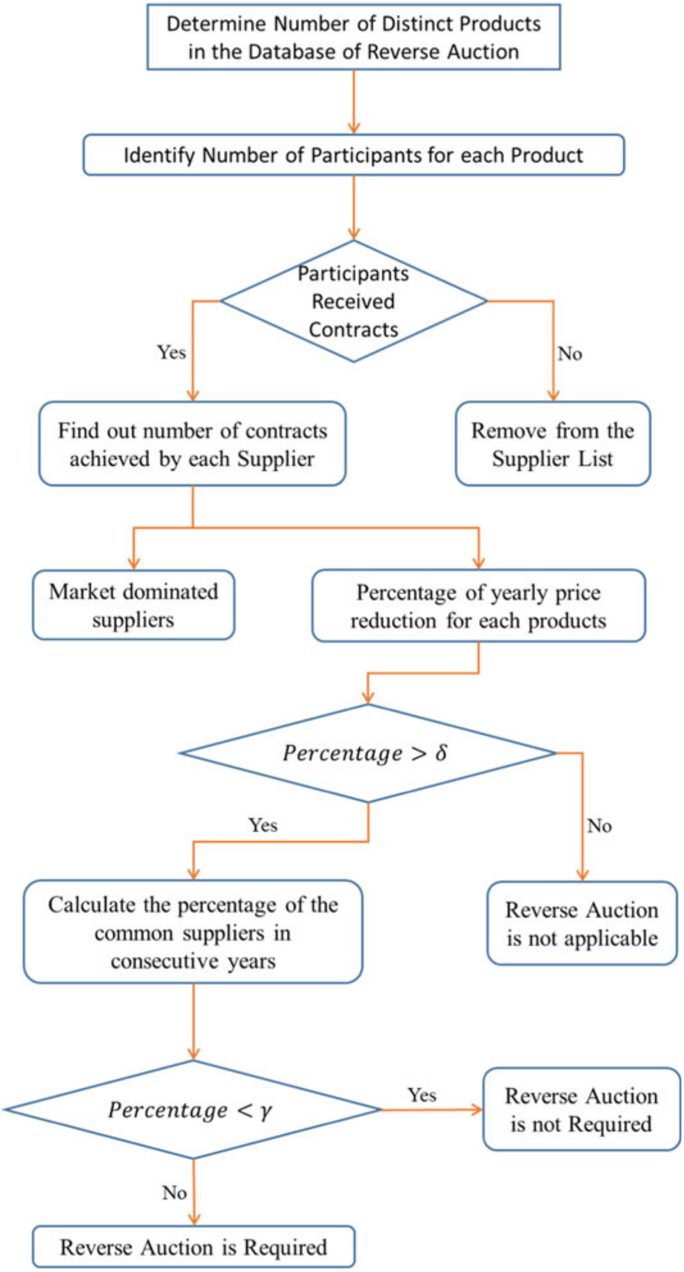


Fig. 3.2 Sequence of the reverse auction validation process for multiple organizations and products

3.4 Results and Analysis

To test the hypotheses mentioned above, we use a number of control variables, such as the year of auction, participant name & code, lot_cpv, and organizer name & code.

The final sample used in our analysis consists of 2,313,078 observations, with each row in the dataset representing an auction process for the purchase of an item. The data reveal various products and services through Ukrainian public procurement e-reverse auctions. Between 2015 and 2019, the government purchased more than 984,859 different products. As shown in Fig. 3.3, the number of organizers and the number of participants stabilized over the last 3 years.

Taking into account, the relatively stable number of participants and organizers, it can be caused by the higher participation rate per company (i.e., each company participates in more RAs). Consequently, it increases competition and helps an organizer to get the best price for a particular product purchasing. Consequently, companies can enjoy an increase in short-term profit, long-term savings, and investments. Figure 3.4 shows the relative proportion of the top 20 suppliers winning maximum contracts in procurement, listed in order from *"least"* to *"most"*. Through the analysis, it is vividly clear that the supplier or company with *code number* 25394112 will dominate the market.

We found that the most frequently procured object is "33,600,000−6_ *Pharmaceutical products*". The percentage of reduction price is almost the same throughout the 5 years (Fig. 3.5). So, we cannot decide whether RAs should be performed yearly. If we check the common suppliers for consecutive 2 years, we can see in Fig. 3.6, there are most of the suppliers are not common. Hence, it is profitable if the RA is performed yearly, and the product is kept in public procurement.

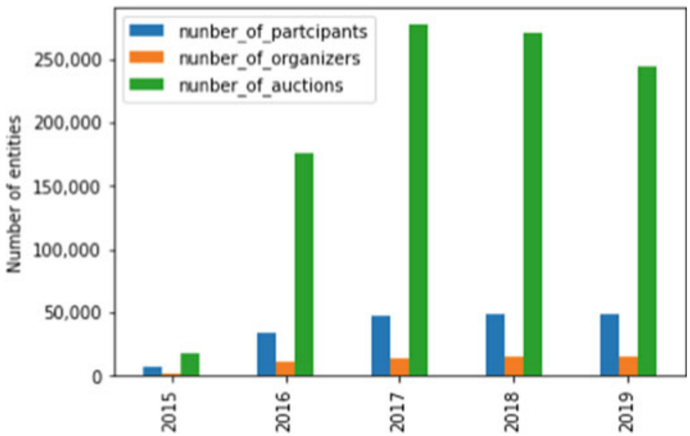


Fig. 3.3 Numbers of organizations, products, and companies in different years

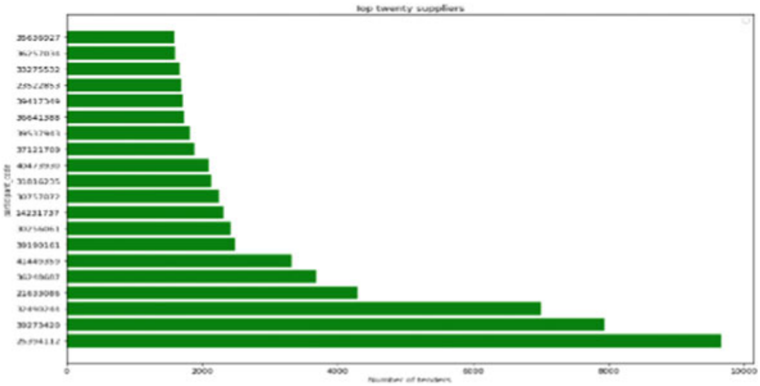


Fig. 3.4 Number of contracts achieved by participants

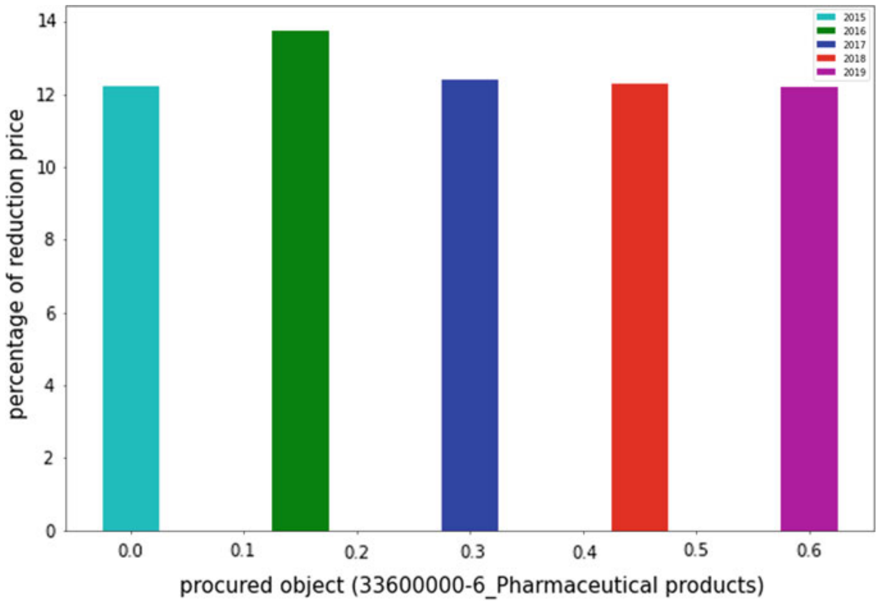


Fig. 3.5 Percentage of price reduction with respect to produced pharmaceutical products

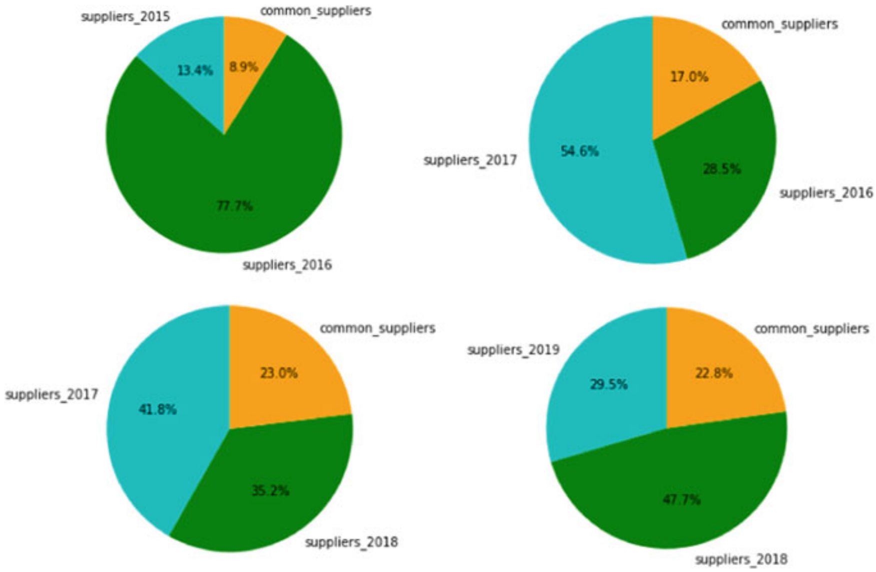


Fig. 3.6 Percentage of common supplier in consecutive years

3.5 Conclusion

In this research, the graph-based method is used as the modeling technique to analyze reverse auction mechanisms. This concept of analyzing the dataset as a collection of nodes and edges can be extended to any data type. Choosing the features of data according to the domain of application will make the algorithm work for any real world problems. The method also provides the flexibility of choosing the column and having the structure of the graph required for any specific domain of application. Additionally, this approach can be applied to solve numerous applications in the industrial sector, including Supply Chain Management, Process Optimization, Fault Detection and Diagnosis, Network Optimization, Asset Management, Quality Control, and Risk Assessment.

References

1. Adikari S, Dutta K (2019) A new approach to real-time bidding in online advertisements: Auto pricing strategy. INFORMS J Comput 31(1):66–82. https://doi.org/10.1287/ijoc.2018.0812
2. Chen X, Ghate A, Tripathi A (2011) Dynamic lot-sizing in sequential online retail auctions. Eur J Oper Res 215(1):257–267. https://doi.org/10.1016/j.ejor.2011.05.051
3. Engelbrecht-Wiggans R, Katok E (2006) E-sourcing in procurement: Theory and behavior in reverse auctions with noncompetitive contracts. Manage Sci 52(4):581–596. https://doi.org/10.1287/mnsc.1050.0474

4. Huang YS, Yang MS, Ho JW (2022) Bundling Decisions for Selling Multiple Items in Online Auctions. Decis Anal 19(1):44–62. https://doi.org/10.1287/deca.2021.0440
5. Khamesi AR, Silvestri S (2020) Reverse auction-based demand response program: A truthful mutually beneficial mechanism. In: Proceeding—2020 IEEE 17th International Conference on Mobile Ad Hoc and Smart Systems, MASS 2020, 427–436. https://doi.org/10.1109/MASS50613.2020.00059
6. Li C (2019. Manufacturing and service operations management supplier competition and cost reduction with endogenous information asymmetry supplier competition and cost reduction with endogenous information asymmetry. Manuf & Serv Oper Manag Publ, Dec
7. Liu A, Xiao Y, Li Z, Wang R (2022) An agent-based multiattribute reverse auction approach for online secondhand commodities. Manag Decis Econ 43(1):129–145. https://doi.org/10.1002/mde.3363
8. Schoenherr T, Mabert VA (2007) Online reverse auctions: Common myths versus evolving reality. Bus Horiz 50(5):373–384. https://doi.org/10.1016/j.bushor.2007.03.003
9. Shokr I, Torabi SA (2017) An enhanced reverse auction framework for relief procurement management. Int J Disaster Risk Reduct, *24*(August 2016), 66–80. https://doi.org/10.1016/j.ijdrr.2017.05.019

Chapter 4
Understanding the Role of e-Procurement and Blockchains in Government Tendering—A Case Evidence from China

Kai Sun, Vikas Kumar, and Meng Wang

Abstract Digital transformation is changing the business landscape; hence Digital technology integrates suppliers, customers, regulatory agencies, stakeholders and other related parties to provide solutions to the problems faced by supply chain information. This research, therefore, aims to analyse the impact of e-Procurement and blockchain on government public bidding in China and further explores the obstacles faced in implementing and integrating technology into their systems. Through interviews with government employees and suppliers, this research analyses the application performance and challenges of e-Procurement and blockchain in procurement and the benefits to downstream supply chain finance. The results show that e-Procurement and blockchain are handy tools in improving the process of government public bidding and enhancing the transparency of the procedures and the speed of information flow. They help downstream suppliers to get more financial support and help the government procurement industry to develop better. A transparent market among customers, retailers, suppliers and manufacturers in the supply chain allows companies and governments to achieve a win–win situation.

K. Sun · M. Wang
Faculty of Business, Law and Social Sciences, Birmingham City University, Birmingham, UK
e-mail: sun-kai.sun@warwick.ac.uk

M. Wang
e-mail: Meng.Wang@bath.edu

V. Kumar (✉)
Warwick Manufacturing Group, University of Warwick, Coventry, UK
e-mail: Vikas.Kumar@bcu.ac.uk

M. K. Tiwari et al. (eds.), *Applications of Emerging Technologies and AI/ML Algorithms*, Asset Analytics, https://doi.org/10.1007/978-981-99-1019-9_4

4.1 Introduction

Early research emphasized that public procurement is the government's main economic activity because it accounts for 15%–20% of global and national revenue. However, in developing economies, these numbers are between 20 and 70% [1]. As digital transformation takes place, government organizations worldwide are also adapting to new technologies to help improve their efficiency, especially in the government's public bidding and procurement. Government public bidding and procurement include several major steps, from setting technical indicators to experts discussing the plan's feasibility and finally to public bidding for social companies, including a series of complete industrial supply chains. Procurement is very important in all areas of supply chain management. Procurement is a more strategic way of acquiring goods and services and includes a series of behaviours, including the acquisition of displayed goods and services and the management after the contract. Electronic procurement refers to the procurement and sale of materials, jobs and services through the Internet and other information and network systems, such as electronic data exchange, for B2B (business to business) or B2C (business to consumer) or B2G (business to government). This is not a new technology, but it has affected many industries' supply chain management aspects and has been applied in many countries. However, blockchain is the latest technology tool in supply chain management. Presently, the number of newly initiated and applied blockchain projects by governments and public management agencies of various countries is rapidly increasing [2]. It provides customers with a platform for decentralized transactions and anonymous purchases. The technology integrates suppliers, customers, regulatory agencies, stakeholders and other related parties to provide solutions to the problems faced by supply chain information. This study explores the application of these new technologies in government public bidding.

However, the inefficiency and risks, such as severe environmental damage and socio-economic impact, are because most governments often fail to conduct due diligence in this process [3, 4], corruption [5] and legal losses. Blockchain and e-Procurement can reduce the excessively lengthy process for the government and its supply chain companies in the bidding process, provide tamper-resistant and verified transaction supply chain activities under the government-led private chain, and improve the work of both parties' efficiency and increase information transparency. Blockchain is expected to have an important impact on society, politics and the environment when it is used in the provision of public services [6]. The technological process of blockchain, also known as distributed ledger technology, provides users with a guarantee that stored information will not be disturbed. This principle guarantees the authenticity between different actors, and these actors may lack trust in each other to some extent. For many years, the government's public procurement management has relied on outdated and labour-intensive paper contracts, and offline meetings and interviews have been the most effective means of communication. The empowerment of these two technologies can reduce paper-based contracts and the potential for human tampering with contracts, enhance the efficiency of online communication and promote environmental sustainability. Blockchain is being used to solve the complexity of logistics. It breaks information islands, realizes the automation of transactions and bureaucratic procedures, improves transparency and ensures the authenticity of the supply chain [6]. At the same time, with the government-led private chain, government suppliers are equivalent to government credit endorsements, making it easier for suppliers to obtain bank loans.

4.1.1 Challenges of Chinese Government Open Tendering

China is the world's second largest economy and the world's largest developing country. Government procurement has played a significant role in boosting the economy. China's government procurement has increased 19.3 times from 165.94 billion Chinese yuan in 2013 to 3211.4 billion Chinese yuan in 2017 [7]. However, under such a huge volume of government procurement transactions every year, many local government departments still adopt the old-fashioned government procurement system. Moreover, only some economically developed areas have begun to explore the application of electronic procurement and blockchain in government procurement [8]. Pointed out that information blocking exists in the information transmission mechanism of government procurement. Currently, the transmission mechanism of government procurement information in China is mainly established on the Internet platform, i.e. the government procurement website. The release speed, saving effect and transmission way of government procurement information all affect the quality of government procurement. The government procurement information platform mainly relies on Internet technology. However, the technology type and information transmission mode cannot be updated in time during the construction and maintenance of the platform. In that case, problems such as unimpeded information communication and untimely information sharing between procurement departments and suppliers will occur [7]. These problems will greatly impact the procurement project bidding efficiency and the quality of project performance. Currently, the standardization of China's government procurement industry is still at the content level, such as standardizing the content of bidding announcements, prequalification announcements and result announcements [7]. The current related trading platforms in China, such as centralized procurement agency platforms, local public resource trading platforms and various professional website platforms, are increasingly scattered and cannot fully realize information sharing [9]. This leads to an increase in the management cost of information managers and the search cost of information users. The decentralization of trading platforms and the diversification of the construction forms of trading platforms in different regions not only increase the difficulty of finding information on suppliers but also bring some trouble to other participants in the government procurement system. This is not conducive to the public's understanding and supervision of government procurement behaviour [10]. Additionally, in all kinds of government procurement procedures, many costs and resources are needed for information verification [11]. Government procurement contract financing refers to the use of government procurement as a platform to help small and medium-sized enterprises raise funds (Huang et al., 2020). Because the information in the current market comes from the business systems of various enterprises, there is a lack of effective supervision and management, and the authenticity of the documents is not easy to grasp, resulting in opaque and inaccurate information in the trading market [13]. However, the key tools to solve these problems are using e-Procurement and blockchain.

4.1.2 Blockchain and e-Procurement in Government Tendering

Highlight that many public services are provided by non-governmental organizations and private organizations or through mixed ownership partnerships such as companies, municipal partnerships, third sectors or public–private partnerships [14]. Although blockchain is still in the early stages of implementation and development in the public sector, breakthrough research is needed regarding its effectiveness, efficiency and usefulness in public management [15]. From the big picture, at the international level, establish a global public infrastructure based on blockchain, seeking to improve coordination and information sharing between governments, enterprises and citizens of different countries. Similarly, at the national level, such improvement is needed as well. Blockchain is looking forward to positively impacting government transparency, accountability and trust [16, 17]. In the field of government procurement, there is a huge amount of transaction data and information to be shared and mined every day, showing the characteristics of multiple departments, multiple types and multiple levels, which coincides with the characteristics of blockchain technology [18]. Pointed out that blockchain and e-Procurement technology can achieve close to the highest level of transparency, such as openness and integrity, confidentiality, accountability, fairness and efficiency at the initial stage of the procedure by shortening process time and automating procedures. At the same time, more and more companies are concerned about the transparency of supply chain information.

The problems faced by the Chinese government's public bidding and procurement are exactly the motivations for applying blockchain and e-Procurement. This study will explore whether these technologies will change processes and overall operations by improving efficiency and their role in the government procurement industry. This research will follow a series of steps to in-depth research and analysis to examine the role of blockchain and e-Procurement technology in the Chinese government's public bidding and procurement.

4.2 Methodology

This research explores the role of electronic contracts and blockchain in government procurement bidding. This study mainly discusses the role of electronic contracts and blockchain in government bidding and whether it can help improve the administrative efficiency of government bidding and accelerate the information flow rate between suppliers and the government. Furthermore, when analysing the role of electronic contracts and blockchain in Chinese government bidding, the application performance and challenges of e-Procurement and blockchain technology are analysed through interviews with government employees, and suppliers focused on the procurement field. In this research, primary data was collected through semi-structured interviews with six government procurement officers and six people from suppliers focused on government tendering (chosen from six different companies). This study adopted a purposive sampling approach. All 12 interviewees were engaged in government bidding procurement projects for more than 15 years on average, as shown in Table 4.1. As a result, both parties were very familiar with every step and process of government procurement and its related plans, strategies, goals and overall planning. The study did not pose any known risks to the participants' emotional, physical or mental health or well-being. This research complies with the ethical guidelines outlined in the Belmont Report (1978), which is 'respect for humanity, kindness and justice' for all people involved in the research. Adequate ethical approval was obtained before conducting the interviews.

4.3 Findings and Discussions

Face-to-face interviews were conducted with 12 government procurement specialists selected to take part in the study, of which 6 participants were government-side workers, and 6 were supplier-side workers willing to be interviewed to provide information for the research. The participants had many years of experience in government procurement and had participated in at least hundreds of government public tenders. These participants also worked in different departments and thus brought different views on e-Procurement and blockchain technology. The bidding, bid evaluation and uploading of bids have enabled employees of these departments to use electronic procurement daily. The study explored whether there is an application model that can simplify and secure the entire process. Through interviews with 12 government officials and suppliers, the study identified the Chinese government's bidding process. A full-fledged process diagram is shown in Fig. 4.1, whereas Fig. 4.2 presents a simplified bidding process.

Figure 4.1 appears to be complex and detailed, in which bidding management and supplier management corresponds to the tenderer and bidder in this study. Figure 4.2 is a simplified version which shows four direct participants in the bidding process: the tenderer, the bidder, the regulatory agency and the trading centre. The tenderer

Table 4.1 Interview participants and their job function

Participant	Position	Working experiences in the government procurement field (years)	Number of government procurement projects participated in (Since work)
A	Government procurement staff	10 years	Over 500
B	Government procurement experts	25 years	Over 2000
C	Government procurement staff	7 years	Over 700
D	Government procurement staff	17 years	Over 1300
E	Government procurement experts	23 years	Over 3000
F	Government procurement staff	9 years	Over 1000
G	Company staff	4 years	Over 300
H	Company senior manager	15 years	Over 1800
I	Company CEO	21 years	Over 2000
J	Company senior manager	13 years	Over 1600
K	Company staff	7 years	Over 800
L	Company staff	6 years	Over 300

Source Authors

mainly submits the qualification information about the company or the institution at the initial stage, and the tenderer shall upload the items it needs to buy, the parameters and amount, as well as the credit rating of the institution. The bidder uploads the parameters of the corresponding goods, the quotation and the credit rating of the company according to the announcement issued by the tenderer. Then the regulatory agency reviews the documents submitted by both sides, including the tender documents of the tenderer and the tender documents of the bidder. And comprehensively evaluate whether both parties have the due qualification, including the legal compliance information about this open tender. After the regulatory approval and the completion of the open tender, the trading centre will start to review the legality of the open tender transaction and whether the contract is legal and compliant and record the transaction information and details as a memorandum for future review.

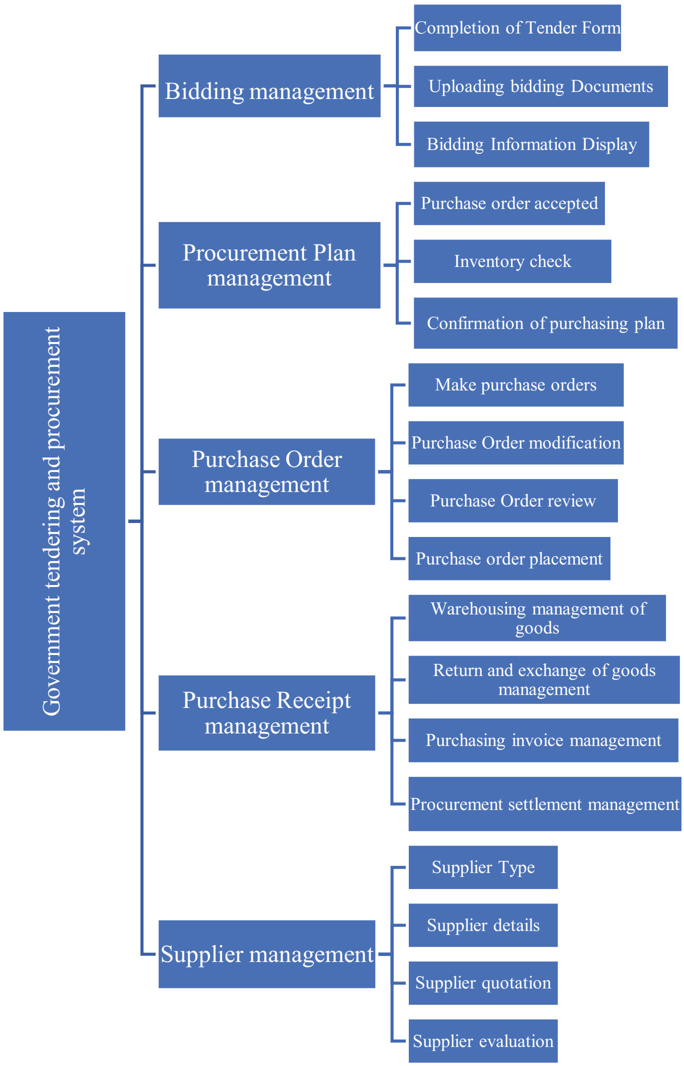


Fig. 4.1 Government tendering and procurement system *Source* Authors

The processes described in Fig. 4.2 can be optimized entirely by e-Procurement and blockchain. According to the interviews with 12 interviewees, all these processes involved manual filling and paper. For each public tender, four copies of the same documents, ranging from 300 to 1,000 pages, are prepared for each of the four departments, costing over ten thousand Chinese Yuan to print each year. And e-Procurement and blockchain can make the whole process completely paperless and electronic. In today's world, environmental protection can reduce paper usage, be environmentally friendly and reduce annual fixed printing costs.

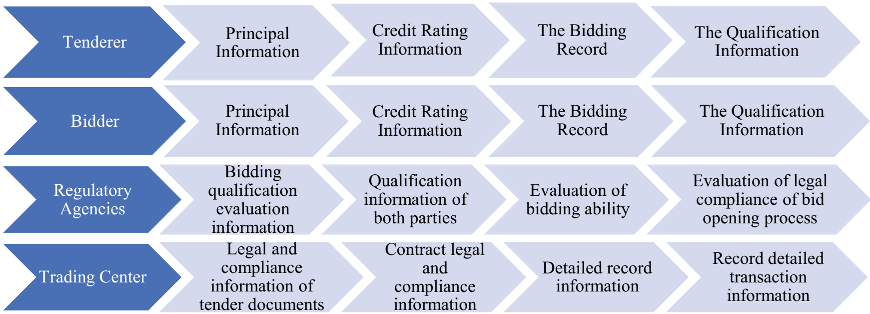


Fig. 4.2 Simplified Bidding Process *Source* Authors

According to the interviewees, the government's most time-consuming and tedious part is making bidding documents, reviewing bidding documents and reviewing the qualification information of bidding companies. For suppliers, bidding preparation, qualification information certification and document processing before bid opening are the most time-consuming. In addition, there are some obstacles in transmitting and updating information between the two parties, and they cannot deliver the information needed by each other in real time. All 12 interviewees said the entire bidding process took an average of 15–30 days from start to finish. Participant C (government) stated, '*For the government, the production of bidding documents, review and qualification verification is more troublesome and time-consuming. Sometimes, because the update of enterprise qualification is not timely, government staff need to call the supplier constantly to urge the update of information*'.

The issues that governments have been paying serious attention to when using public funds are transparency and efficiency in bidding, especially the transparency and efficiency in public infrastructure projects [20]. However, due to the heavy management work, some errors are prone to occur: file preparation errors, incomplete information, file confusion, unnecessary copying and possible leakage of valuable information, which ultimately affects the timely implementation and cost of the project. [20] pointed out that e-Procurement can reduce these errors to a certain extent. Reducing or avoiding these unnecessary management errors benefits the project and all participants. For government public bidding, e-Procurement can improve the speed of information flow and reduce the average time required by government procurement. It also enhances the security of data and the effectiveness of long-term storage (electronic archiving). It also saves time and energy for the daily repeated bidding management work and saves management costs [21, 22]. In the past, if each department under the government conducted an open tender, the information was only posted on the department's own official website. As a result, suppliers need to constantly search for bidding information on the websites of various government departments and make a bidding document that meets the special requirements of this department according to the requirements of the bidding document format of this department. After the implementation of e-Procurement, the bidding information of each department will be uniformly published on the government's e-Procurement platform, and the bidding information of all departments can be easily obtained on the platform. This was evident from the responses from both government and suppliers. For example, Participant F (supplier) said, *'With the electronic procurement platform, the unification of the bidding document format makes us feel a lot easier. In the past, the company would appoint a bidding document production professional to produce a document that meets the requirements. Now there is no need for specialized personnel. This spare person can find more projects for the company, which indirectly reduces the employment cost of our enterprise. At the same time, we do not need to appoint special personnel to search for appropriate projects on the official websites of various government departments and then personally send the documents to different departments. Generally speaking, the efficiency of information transmission is much higher than before, and it also saves many labour and transportation costs'*.

e-Procurement has been widely used in the early stage of government public bidding. However, e-Procurement has limitations in the latter stages of expert reasoning, blind evaluation, delivery and payment. Therefore, blockchain can be the underlying architecture for the entire e-Procurement system. Due to the tamper-proof characteristics of the blockchain, the competent unit and regulatory unit can directly review and supervise every link on the chain, breaking the information island created by the previous multi-department collaborative work. As a result, it reduces the personnel and time cost of supervision and verification of multiple departments and systems. For example, Participant B (government) suggested, *'In the past, many companies falsely declared their qualifications in order to qualify for the bidding process. In my department, I might find out and disqualify these companies. But he will go to another department to continue to apply for new projects. If the inspectors do not carefully review, the company may be shortlisted in the final bid evaluation stage. After the blockchain system is established, if the information can be shared between various departments, once this incompletely qualified company is on the chain, the reviewers of other projects in the future can quickly discover the company's problems. Until the company's qualifications reach the standard, it can be eligible for selection again. I think this may be a more practical aspect of blockchain technology'*.

The findings from the interviews clearly indicate that a blockchain-based solution can improve the existing efficiency situation, and it is possible to reduce or completely eliminate fraud and other behaviours. Findings also show that due to the anonymous nature of blockchain, all bidding companies are anonymous to the counterparties in the initial review stage of the project, which fundamentally eliminates bidding companies from conspiring to win the project. As Participant F (supplier) stated, *'The anonymity and immutability of blockchain can make the bidding process fairer. Because the reviews are all anonymous, experts can only use hard metrics to determine which company won'*. Blockchain also brings a strong state of trust between different stakeholders when it involves various market participants involving multiple transactions between them. For example, this technology is very useful in the government's public bidding, especially for promoting the value capture of the two parties' transactions [23]. The contract data after the successful bidding is uploaded to the blockchain. It includes the legal compliance of the entire bidding process and the completion of the subsequent contract execution, combined with the complete audit information. And through the use of big data, achieve multiple authorities, resolving the credit issues of the participants in the bidding process. At the same time, complete data, and other information traceability for subsequent historical projects to promote rating and authenticity verification. One of the government employees (Participant B) advised, *'For the contract execution stage, blockchain can faithfully record the problems during the contract execution so that the tenderer can trace to the source and check at any time after the project ends. Suppose there is any problem with the goods, or the supplier fails to fulfil the obligations according to the contract after the project is over. In that case, we can quickly find it in the government's public bidding in the future'*.

In China, the bidding process from beginning to end involves payment in the following aspects: first, when the public bidding begins, each supplier participating in the bidding shall pay a security deposit to the tenderer, which is about 5% of the total contract amount. Second, when the tender is over, the deposit paid to the tenderer will be returned to each supplier. In contrast, the tenderer will pay a deposit of about 10% of the total contract amount to the supplier who wins the project. Third, after all the purchased goods are delivered and accepted, the tenderer will pay 80%–85% of the contract's total value to the supplier, leaving the final 5%–10% as the warranty money. Finally, the final warranty money will be paid to the supplier after the warranty expiration if there is no quality problem with the goods during the warranty period. At this point, all payments regarding the open tender are completed. Through this series of links, it can be found that the supplier has to advance the money needed to purchase goods.

In some cases, a supplier may have more than one project at the same time, which may result in a lack of cash flow for the supplier in the case of advance payment. At this time, suppliers need to borrow money from banks to solve the problem, but many suppliers are small and medium-sized enterprises, and banks are unwilling to issue loans. The e-Procurement and blockchain systems are likely to address this challenge. As one of the interviewees (Participant I (supplier)) pointed, *'At present, based on electronic procurement, we have the relevant services landed. After the bidding and entering the execution stage, if our cash flow is insufficient and we need loans, we can take the contracts signed with the government departments and go to the banks cooperating with the government to obtain loans with the contracts as vouchers. For example, we have insufficient cash flow and submitted relevant evidence, but the government's auditor failed to pass'*. At the same time, the open and transparent data before and after the deposit is combined with the blockchain and the smart contract automatically repays the deposit to ensure the interests of bidders.

With the smart contract in the blockchain, the system will automatically record all these payment behaviours. All parties can use smart contracts to design auto-executed contract relationships without additional monitoring or execution costs [24]. The smart contract is used to enforce the efficient implementation of the bidder's bid bond process, combined with a part of the chain penalty mechanism to solve the practical problem of bidders getting back the bid bond interest. And the entire information is on the chain, clear and transparent, to prevent funds from being misappropriated for other purposes or delayed in returning to bidders. The smart contract in the blockchain allows triggering under some individual conditions and automatically executes certain tasks under pre-agreed conditions, such as ordering, sending invoices, and making payments. With the existence of the smart contract, it is equivalent to the supplier getting a contract with government credit endorsement. Without the need for any intermediary to review and approve the situation, the supplier can directly obtain a loan from the bank based on the contract generated by the smart contract, reducing human intervention in supply chain finance. After the project is completed, if the supplier has a bank loan, the smart contract can directly transfer the payment paid by the government to the bank, reducing the possibility of the supplier defaulting or embezzling funds for other purposes. One of the participants (Participant I (supplier)) also captured these benefits, who mentioned, *'Although this service is currently available under the framework of e-Procurement. A system with smart contracts based on the blockchain will be more convenient. It will completely reduce the possibility of human intervention and makes loans more convenient than before*'.

Although e-Procurement and blockchain technologies can address many challenges, the interviewees identified four difficulties. First is the cost of learning because many experienced government procurement specialists are middle-aged and above. The electronic procurement system built by blockchain means they have to learn repeatedly, which is undoubtedly a relatively big resistance. Secondly, although there is a willingness to share data among government departments, not every department is willing to share all its data with other departments, which involves interests and regulatory uncertainty. Chang, Iakovou and Shi (2019) stated that, for the government, blockchain's most important costs and risks are related to regulatory uncertainty, and interoperability is one of these risks. The ability to easily share information, operations and transactions across different departments is an important feature of interoperability. Third, the cost of building and maintaining the whole system. Because to build a system that connects almost all departments to achieve information sharing, the requirements for system construction, upgrade and maintenance will be very high. Another cost of blockchain comes from the fact that, in the early stages of development, many professionals need to be invested in construction, and it lacks sufficient flexibility to adapt to different situations [26]. Point out the maintenance cost is relatively high. The final difficulty is the obstacles of vested interests within the government to the promotion of the system. Because this system will make many intermediaries useless, they will inevitably have a counterproductive effect on promoting the new system. Our findings thus identify both the benefits and challenges of e-Procurement and blockchain adoption in the Chinese government tendering process.

4.4 Conclusions

The results show that e-Procurement and blockchain are very useful tools to improve the process of government public bidding and enhance the transparency of the procedures and the speed of information flow. The application of e-Procurement makes the early stages of the bidding process from offline all things completed by people to online fixed procedures, thereby simplifying the process and improving efficiency. Blockchain is a viable solution, not an idea, and has now been gradually promoted and implemented. The government began to understand transformation as a keyword. Transformation can shorten the procurement process and stimulate the vitality of government and private economic activities. e-Procurement and blockchain have been confirmed as suitable solutions to make bidding activities traceable, non-tamperable and improve organizational transparency.

Government open procurement accounts for the largest part of economic activities in developing countries. As the leading party, the government should explore more cutting-edge technologies such as blockchain, big data, artificial intelligence and other tools that can improve productivity. At present, the government's application of cutting-edge technology is not enough. Under the special system of China, the government, as the relatively strong leader in the market, should first conduct research and application of cutting-edge technology, form a system and summarize the experience. Then, the system and experience will be gradually extended to economic activities between small and medium-sized enterprises, and the new technology reform will be carried out from top to bottom to improve the society's overall efficiency and economic vitality. As for the new technology of blockchain and e-Procurement, only the government has a relatively large volume, with scale cost and strong fault tolerance ability to improve the application degree of technology. e-Procurement has been implemented to improve the administrative efficiency of government bidding in the past 2 years. The bidding documents of a unified format have been adopted among government departments, and information has also been integrated. Suppliers can have more time to win more projects by reducing the pressure of material reporting.

One of the limitations of this study is the small sample size. The study's sample size is 12 people, accounting for only a small part of the public procurement professional population. Furthermore, all participants were from only one province, and hence the geographical limitation is also reflected. Therefore, although each participant gave their opinion on the industry, the result is not representative of the entire population. Therefore, future research should involve more respondents specializing in public procurement by the governments of different provinces to minimize the deviation of the sample. Another limitation of this study is the focus on the external application of blockchain; hence, the underlying technology of blockchain is not introduced in detail. In addition, few local governments in China have carried out pilot projects because blockchain is too advanced for government procurement. Therefore, all respondents' opinions on blockchain application in government procurement are based on their years of experience in the industry and detailed introduction of blockchain by researchers. Therefore, for the future research direction, research can start from designing the underlying architecture, with blockchain as the bottom layer and e-Procurement as the form of expression, to make a complete operational system. And adopt the opinions of government procurement departments of different regions in line with the region's characteristics.

References

1. World Bank (2012) Why reform public procurement? Washington DC: World Bank. https://www.worldbank.org/content/dam/Worldbank/document/MNA/Why_Reform_Public_Procurement_English.pdf. Accessed 11 October 2022
2. Consensys (2020) Which Governments are Using Blockchain right now?, Available at: https://media.consensys.net/which-governments-are-using-blockchain-right-now-327c0e7fb8a1. Accessed 20 October 2022
3. Walker H, Brammer S (2012) The relationship between sustainable procurement and e-procurement in the public sector. Int J Prod Econ 140(1):256–268. https://doi.org/10.1016/j.ijpe.2012.01.008
4. Preuss L (2009) Addressing sustainable development through public procurement: the case of local government. Supply Chain Manag: Int J 14(3):213–223. https://doi.org/10.1108/13598540910954557
5. Lio M-C, Liu MC, Ou YP (2011) Can the internet reduce corruption? A cross-country study based on dynamic panel data models. Gov Inf Q 28(1):47–53. https://doi.org/10.1016/j.giq.2010.01.005
6. Cagigas D, Clifton J, Diaz-Fuentes D, Fernandez-Gutierrez M (2021) Blockchain for public services: a systematic literature review. IEEE Access 9:13904–13921. https://doi.org/10.1109/ACCESS.2021.3052019
7. Peng Z (2019) Brief discussion on the application prospect of blockchain technology in government procurement. China GovMent Procure Rep, 5
8. Li Z, Guo H, Wang WM, Guan Y, Barenji AV, Huang GQ, McFall KS, Chen X (2019) A blockchain and automl approach for open and automated customer service. IEEE Trans Industr Inf 15(6):3642–3651. https://doi.org/10.1109/TII.2019.2900987
9. Xia Z, Liu D (2019) The application and challenge of innovative techniques in government procurement. J Chongqing Three Gorges Univ 35(179):83–90
10. Zhu B, Jia P, Sun N (2020) Probe into the influence of new information environment on university procurement and bidding. China Manag InfIzation 23(3):224–226
11. Zheng Z, Xie S, Dai H, Chen X, Wang H (2017) An overview of blockchain technology: architecture, consensus, and future trends. IEEE Int Congr Big Data (BigData Congress) 2017:557–564. https://doi.org/10.1109/BigDataCongress.2017.85
12. Huang H, Liu Q, Ye C, Chen X (2021) Game analysis of government procurement contract financing based on blockchain technology. J Syst Simul, 33(8):1947–1958. https://doi.org/10.16182/j.issn1004731x.joss.20-0233
13. Long Y, Zhang J, Ai R (2019) Research on the optimization of supply chain financial system based on blockchain technology. Southwest Finance 450(1):72–79
14. Clifton J, Warner ME, Gradus R, Bel G (2019) Re-municipalization of public services: trend or hype? J Econ Policy Reform 24(3):293–304. https://doi.org/10.1080/17487870.2019.1691344
15. Rodríguez Bolívar MP, Scholl HJ (2019) Mapping potential impact areas of Blockchain use in the public sector. Information Polity 24(4):359–378. https://doi.org/10.3233/IP-190184
16. Berryhill J, Bourgery T, Hanson A (2018) Blockchains unchained: Blockchain technology and its use in the public sector. OECD Work Pap Public Govance, 28:1–53. https://doi.org/10.1787/19934351

17. Nugent T, Upton D, Cimpoesu M (2016) Improving data transparency in clinical trials using blockchain smart contracts. F1000Research, 5:2541. https://doi.org/10.12688/f1000research.9756.1
18. Nin Sánchez S (2019) The implementation of decentralised ledger technologies for public procurement. Eur Procure & Public PrivE PartnShip Law Rev, 14(3):180–196. https://www.jstor.org/stable/26895828
19. The Belmont report (1978) National commission for the protection of human subjects of biomedical and behavioral research. Office of the Secretary Ethical Principles and Guidelines for the Protection of Human Subjects of Research, Accessed on 31 July 2022 https://www.hhs.gov/ohrp/regulations-and-policy/belmont-report/read-the-belmont-report/index.html
20. Santoso DS, Bourpanus N (2019) Moving to e-bidding: Examining the changes in the bidding process and the bid mark-up decisions of Thai contractors. J Financ Manag Prop Constr 24(1):2–18. https://doi.org/10.1108/JFMPC-05-2018-0025
21. Nguyen TA, Chovichien V, Tran NTB (2012) The differences between the perception and practical application of eBidding in construction industry. ASEAN Eng J Part C, 1(1):73–87. https://doi.org/10.11113/aej.v1.15330
22. Lindsley G, Stephenson P (2008) E-tendering tendering process within construction: A UK perspective. Tsinghua Sci Technol. 13(S1):273–278. https://doi.org/10.1016/S1007-0214(08)70161-5
23. Matyskevic J, Kremer-Matyskevic I (2020) The economic advantages of blockchain technology in E-procurement. RegNal Form & Dev Stud, 31(2): 17–27. http://dx.doi.org/https://doi.org/10.15181/rfds.v31i2.2094
24. Kiviat T (2015) Beyond bitcoin: issues in regulating blockchain transactions. Duke Law Journal, 65:569–608. Available at: https://scholarship.law.duke.edu/dlj/vol65/iss3/4
25. Chang Y, Iakovou E, Shi W (2020) Blockchain in global supply chains and cross border trade: a critical synthesis of the state-of-the-art, challenges and opportunities. Int J Prod Res 58(7):2082–2099. https://doi.org/10.1080/00207543.2019.1651946
26. Filippi PD, Hassan S (2016) Blockchain technology as a regulatory technology: From code is law to law is code. First Monday, 21(12). https://doi.org/10.5210/fm.v21i12.7113

Chapter 5
Omni-Channel Distribution Network Design for Fresh Food Procurement Considering Freshness-Keeping Effort and Food Quality Loss

Indira Roy and Lohithaksha M. Maiyar

Abstract This study aims at designing a distribution network for public procurement of fresh fruits and vegetables in India. The present model is conceptualized in an omni-channel environment to offer a seamless buying experience to consumers through government-regulated markets. It addresses the key challenges of fresh food loss due to quality degradation and freshness-keeping effort along with transportation network design during fresh food distribution. Hence, a cost optimization model is developed which minimizes the cost associated with transportation, quality loss, and freshness-keeping effort using a mixed-integer nonlinear programming model. The proposed model is solved using an exact solution approach in a pyomo environment with the CPLEX solver for different problem sizes. The effects of quality loss cost and freshness-keeping effort on the total cost and freshness-keeping effort decisions were also analyzed.

Keywords Fresh food distribution · Omni-channel · Freshness-keeping effort · MINLP · Manakuragayalu

5.1 Introduction and State of the Art

With 26.4% of the world population (~2 billion people) living under moderate to severe levels of food insecurity, global food security is an enormous challenge today [1]. Around 1.3 billion tons of food is lost or wasted every year which accounts for nearly one-third of the produced food globally. India is the second largest producer of fresh fruits and vegetables in the world. According to the study conducted by Jha. et al., (2015), 0.6% of GDP (INR 926.51 billion) was lost due to post-harvest losses

I. Roy · L. M. Maiyar (✉)
Department of Entrepreneurship and Management, Indian Institute of Technology Hyderabad-502284, Hyderabad, India
e-mail: l.maiyar@em.iith.ac.in

M. K. Tiwari et al. (eds.), *Applications of Emerging Technologies and AI/ML Algorithms*, Asset Analytics, https://doi.org/10.1007/978-981-99-1019-9_5

in India during FY14. This is the result of a lack of proper storage and handling techniques; lack of refrigeration; and spoilage in transport and processing [3].

Fresh food supply chain networks differ from the traditional food supply chain because of the distinction in the food with respect to shorter shelf life, requirement of shorter delivery time, and ambient storage facilities. Especially, with consumers now being more digitally active than before, owing to personal convenience and health safety reasons are increasingly looking for door-step deliveries [4]. Hence, omni-channel retailing has become increasingly popular among consumers as it is a consumer-centric approach and aims at providing a seamless shopping experience at every step. In the Indian context, although many large fresh food retailers like Spar-Hypermarket, Spencer's, STAR Market, and others have become truly omni-channel, the government-regulated markets operate only in traditional brick-and-mortar channels. However, with digital India campaigns and initiatives like e-NAM, the government-aided supply chain is speculative of becoming omni-channel in the coming years. There are substantial studies with omni-channel distribution, yet its penetration into fresh food supply chain has not taken an upward trend. Our work is focused on deducing practical implications for omni-channel distribution network setting for transport of fresh produce in Indian scenario.

One of the major challenges with operating in omni-channel environment is to maintain the quality and freshness of the produce throughout the network across different channels. Cold chain systems include a series of freshness-keeping activities such as low-temperature storage, distribution by reefer vehicles, thermal insulated packaging, transportation with coolant refrigeration systems, and so on [5]. In the context of freshness-keeping activities, [6] introduced freshness-keeping effort as a decision variable which describes the resources and capital invested to prolong and maintain freshness level throughout the distribution network. Few researchers have emphasized on considering freshness-keeping effort as one of the major variables for shipment consolidation [7], comparison of different sales modes [8], and cap and trade regulation in cold chain systems [9]. However, the effect of freshness-keeping effort on transportation cost and quality loss cost considerations have not been studied yet. Also, to make decisions on investments towards freshness-keeping effort in an omni-channel scenario keeping the transportation cost and purchase probability at the demand points optimum, a cost trade-off model must be formulated. Hence, in the current study, we focus on a fresh food transportation network design incorporating the decisions regarding the freshness-keeping investments and transportation network decisions. It also addresses the key challenge of maintaining optimum transportation cost with cost tradeoff between quality loss cost and freshness-keeping effort in an omni-channel distribution network.

5.2 Problem Description and Formulation

This study considers a multi-echelon three-tier fresh food distribution network in an omni-channel environment as shown in Fig. 5.1. For the present study, the focus is on the downstream flow of fresh produce from collection centers till consumers in the presence of three different types of channels—Brick and Mortars, Buy online deliver at home (BODH), and Buy online pickup at store (BOPS). To maintain the quality and freshness of the produce throughout the network in the most economical manner, a mixed-integer nonlinear programming model is formulated based on the following assumptions:

- Potential locations of collection centers, distribution centers, retail outlets, fulfillment centers, and pick-up points are fixed.
- The model considers one distinct type of fresh produce transportation across all channels of demand.
- All demands and availability of fresh produce at various echelons are known and deterministic.
- All fresh produce is transported with a full truckload.
- The transportation takes place within a single time period.

The notations for decision variables and parameters used in problem formulation are described in Tables 5.1, 5.2, respectively.

Objective function:

Total Transportation cost,

$$TTC = \sum_{i\in I}\sum_{j\in J} f_{ij}y_{ij} + \sum_{j\in J}\sum_{k\in K} f_{jk}y_{jk} + \sum_{j\in j}\sum_{l\in L} f_{jl}y_{jl} + \sum_{j\in J}\sum_{m\in M} f_{jm}y_{jm} + \sum_{l\in L}\sum_{o\in O} f_{lo}y_{lo} + \sum_{i\in I}\sum_{j\in J} c_{ij}x_{ij} + \sum_{j\in J}\sum_{k\in K} c_{jk}x_{jk} + \sum_{j\in J}\sum_{l\in L} c_{jl}x_{jl} + \sum_{j\in J}\sum_{m\in M} c_{jm}x_{jm} + \sum_{l\in L}\sum_{o\in O} c_{lo}x_{lo} \quad (5.1)$$

Total freshness-keeping cost, {TFC} =

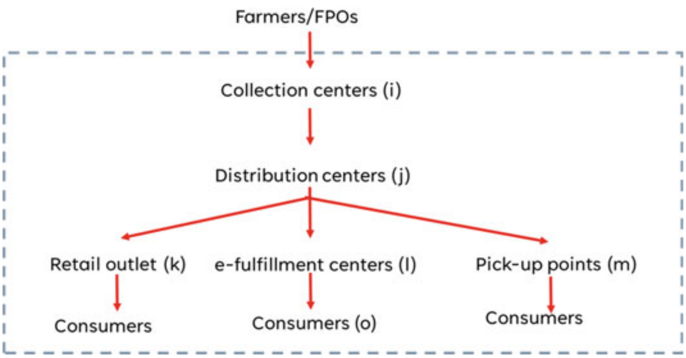


Fig. 5.1 Schematic representation of proposed distribution network for fresh food procurement

Table 5.1 Notation for the decision variables used in this study.

	Decision variables
y_{ij}	1 if the flow between i to j is positive, 0 otherwise
y_{jk}	1 if the flow between j to k is positive, 0 otherwise
y_{jl}	1 if the flow between j to l is positive, 0 otherwise
y_{jm}	1 if the flow between j to m is positive, 0 otherwise
y_{lo}	1 if the flow between l to o is positive, 0 otherwise
z_i	1 if collection center makes an investment towards freshness-keeping effort,0 otherwise
z_j	1 if distribution center makes an investment towards freshness-keeping effort between, 0 otherwise
z_l	1 if fulfillment center makes an investment towards freshness-keeping effort, 0 otherwise
x_{ij}	Quantity(kg.) of fresh food transported from collection centers to distribution center
x_{jk}	Quantity(kg.) of fresh food transported from distribution center to retail outlets
x_{jl}	Quantity (kg.) of fresh food transported from distribution center to fulfillment center
x_{jm}	Quantity (kg.) of fresh food transported from distribution center to fulfillment center
x_{lo}	Quantity (kg.) of fresh food transported from fulfillment center to consumers

$$TFC = \sum_{i\in I}\sum_{j\in J}\lambda_{ij}z_i + \sum_{j\in J}\sum_{k\in K}\lambda_{jk}z_j + \sum_{j\in J}\sum_{l\in L}\lambda_{jl}z_j + \sum_{j\in J}\sum_{m\in M}\lambda_{jm}z_j + \sum_{l\in L}\sum_{o\in O}\lambda_{lo}z_l + \sum_{i\in I}\sum_{j\in J}\rho_{ij}z_i x_{ij} + \sum_{j\in J}\sum_{k\in K}\rho_{jk}z_j x_{jk} + \sum_{j\in J}\sum_{l\in L}\rho_{jl}z_j x_{jl} + \sum_{j\in J}\sum_{m\in M}\rho_{jm}z_j x_{jm} + \sum_{l\in L}\sum_{o\in L}\rho_{lo}z_l x_{lo} \tag{5.2}$$

Total cost incurred due to quality loss,

$$TCQ = \sum_{i\in I}\sum_{j\in J}P_j(1-p_{ij})(1-\tau_{ij}z_i)x_{ij} + \sum_{j\in J}\sum_{k\in K}P_k(1-p_{jk})(1-\tau_{jk}z_j)x_{jk} + \sum_{j\in J}\sum_{l\in L}P_l(1-p_{jl})(1-\tau_{jl}z_j)x_{jl} + \sum_{j\in J}\sum_{m\in M}P_m(1-p_{jm})(1-\tau_{jm}z_j)x_{jm} + \sum_{l\in L}\sum_{o\in O}P_o(1-p_{lo})(1-\tau_{lo}z_l)x_{lo} \tag{5.3}$$

$$Minimize\ TC = TTC + TFC + TCQ \tag{5.4}$$

The model considers the total transportation cost, the total freshness-keeping effort cost, and the total cost incurred due to quality loss in the objective function, Eq. 5.1 represents the total transportation cost consisting of fixed and variable transportation cost as a function of the quantity of fresh produce transported between each echelon. Equation 5.2. represents the total investment required towards freshness-keeping effort. The freshness-keeping cost, $\rho = a + b(\tau^2)$ is calculated as a quadratic function of freshness-keeping effort, where $\tau (0 \le \tau \le 1)$ represents the amount of

Table 5.2 Notation for the parameters used in this study

Parameters			
$\boldsymbol{I}$	Set of collection centers $\{i_1, i_{2,\ldots} i_n\}$	ρ_{jm}	Variable cost associated with freshness-keeping effort per kg. between j to m
$\boldsymbol{J}$	Set of distribution centers $\{j_1, j_{2,\ldots} j_n\}$	ρ_{lo}	Variable cost associated with freshness-keeping effort. between l to o
$\boldsymbol{K}$	Set of retail stores $\{k_1, k_{2,} \ldots k_n\}$	λ_{ij}	Fixed cost associated with freshness-keeping effort per kg. between i to j
$\boldsymbol{L}$	Set of fulfillment centers $\{l_1, l_{2,} \ldots l_n\}$	λ_{jk}	Fixed cost associated with freshness-keeping effort per kg. between j to k
$\boldsymbol{M}$	Set of pickup points $\{m_1, m_{2\ldots} m_n\}$	λ_{jl}	Fixed cost associated with freshness-keeping effort per kg. between j to l
$\boldsymbol{O}$	Set of consumers for BODH mode $\{o_1, o_2, \ldots, o_n\}$	λ_{jm}	Fixed cost associated with freshness-keeping effort per kg. between j to l
f_{ij}	Fixed cost associated with flow i to j	λ_{lo}	Fixed cost associated with freshness-keeping effort per kg. between j to m
f_{jk}	Fixed cost associated with flow j to k	P_k	Quality loss cost per unit kg of fresh food at the retail stores
f_{jl}	Fixed cost associated with flow j to l	P_l	Quality loss cost per unit kg. of fresh food at the fulfillment center
f_{jm}	Fixed cost associated with flow j to m	P_m	Quality loss cost per unit kg. of fresh food at the pickup point
f_{lo}	Fixed cost associated with flow l to o	P_o	Quality loss cost per unit kg. at the BODH mode
c_{ij}	Unit variable cost per unit ton of quantity transported from i to j	p_{lj}	Purchase probability distribution center
c_{jk}	Unit variable cost per unit ton of quantity transported from j to k	p_{jk}	Purchase probability at retail outlets
c_{jl}	Unit variable cost per unit ton of quantity transported from j to l	p_{jl}	Purchase probability at fulfillment centers
c_{jm}	Unit variable cost per unit ton of quantity transported from j to m	p_{jm}	Purchase probability at pick-up points
c_{lo}	Unit variable cost per unit ton of quantity transported from l to o	p_{lo}	Purchase probability at BODH points
ρ_{ij}	Variable cost associated with freshness-keeping effort per kg. between i to j	D_{ck}	Consumer demand at the retail store

(continued)

Table 5.2 (continued)

Parameters			
I	Set of collection centers $\{i_1, i_{2,\ldots} i_n\}$	ρ_{jm}	Variable cost associated with freshness-keeping effort per kg. between j to m
ρ_{jk}	Variable cost associated with freshness-keeping effort per kg. between j to k	D_{lo}	Consumer demand at the BODH point
ρ_{jl}	Variable cost associated with freshness-keeping effort per kg. between j to l	D_{cm}	Consumer demand at the pick-up point

Table 5.3 Summary of the model parameters under two different problem instances

Problem size	Low	High
No. of collection centers	2	5
No. of distribution center	1	1
No. of fulfillment center	1	1
No. of Retail outlets	2	5
No. of PUPs	2	5
No.of BODH points	2	3
Number of variables	28	58
Number of constraints	20	40
Objective function value (Rs.)	6423.40	6562.5

resources and capital investment required to prolong or maintain the freshness of the produce throughout the network. The variable cost of freshness keeping is a function of decision made on whether to invest on freshness-keeping effort or not and the quantity of fresh produce transported between each echelon and the fixed cost of freshness keeping only depends on the choice of decision to invest in freshness keeping. Equation 5.3 represents the quality loss cost evaluated by penalizing loss in sales of consumable food due to loss of food quality which is captured by the value $(1 - p)$, where p represents purchase probability. $(1 - \tau z)$ represents loss of freshness for quantity of food that cannot be purchased. $(1 - \tau)$ takes value 1 when there is no freshness-keeping effort taken (when z = 0) and takes a value between 0 and 1, otherwise, thereby reducing the probability of fresh food not purchased by a proportionate amount to the value of freshness-keeping effort. According to Bortolini, M. et al., (2015), the purchase probability p, is the function of transportation time and is calculated as

$$p = \min\left(\frac{1 - \frac{tt}{ss}}{1 - QRP}, 1\right) \tag{5.5}$$

The objective function as described in Eq. 5.4 is subjected to various constraints which are discussed below.

Subjected to constraints

$$\sum_{j \in J} x_{ij} \leq A_i \forall i \tag{5.6}$$

$$\sum_{i \in I} x_{ij} \geq \sum_{k \in K} x_{jk} + \sum_{l \in L} x_{jl} + \sum_{m \in M} x_{jm} \forall j \tag{5.7}$$

$$\sum_{o \in O} x_{lo} \leq \sum_{j \in J} x_{jl} \forall l \tag{5.8}$$

$$\sum_{j \in J} x_{jk} \geq D_{ck} where, D_{ck} = \alpha D_c \forall k \tag{5.9}$$

$$\sum_{l \in L} x_{lo} \geq D_{co} where, D_{co} = \beta D_c \forall o \tag{5.10}$$

$$\sum_{j \in J} x_{jm} \geq D_{cm} where, D_{cm} = (1 - \alpha - \beta) D_c \forall m \tag{5.11}$$

$$x_{ij} \leq M y_{ij} \forall i, \forall j \tag{5.12}$$

$$x_{jk} \leq M y_{jk} \forall j, \forall k \tag{5.13}$$

$$x_{jl} \leq M y_{jl} \forall j, \forall l \tag{5.14}$$

$$x_{jm} \leq M y_{jm} \forall j, \forall m \tag{5.15}$$

$$x_{lo} \leq M y_{lo} \forall l, \forall o \tag{5.16}$$

$$y_{ij,} y_{jk,} y_{jl,} y_{jm}, y_{lo}, z_i, z_j, z_l \in \{0, 1\} \forall i, \forall j, \forall k, \forall l, \forall m, \forall o \tag{5.17}$$

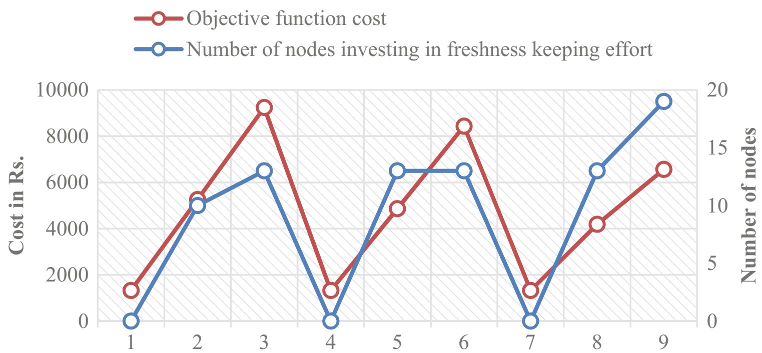
$$x_{ij}, x_{jk}, x_{jl}, x_{jm}, x_{lo} \geq 0 \forall i, \forall j, \forall k, \forall l, \forall m, \forall o \tag{5.18}$$

Equation 5.6 ensures that quantity transported from collection center to distribution center is within the availability of fresh produce at the collection center A_i. Equations 5.7, 5.8 are the transshipment constraints to ensure consistent flow from collection centers to retail outlets, pickup points, fulfillment centers, and fulfillment center-BODH points through the distribution center. Equations 5.9–5.11 represent demand constraints at the retail, pickup points, and BODH points where α and β are demand constant for the total demand D_C. Equations 5.12–5.16 ensure that there will be positive flow in a given route between the echelons if and only if there is a positive route selected. Lastly, Eqs. 5.17, 5.18 represent the integrality and non-negativity constraints of the decision variables.

5.3 Numerical Study and Analysis

In India, traditionally two different routes of fresh food distribution network exist through government-regulated markets. First is the Mandi route which is regulated under APMC Act and starts from farmers to APMCs /Mandis to commission agents, traders, wholesalers, retailers, and consumers. While the second follows the farm-to-fork model where the government provides infrastructure (Rythu Bazaar) for the farmers to sell their produce directly to consumers. Both routes have their own challenges. On one hand, Mandi route lacks traceability due to the long chain. While on the other hand, although the farmers in the farm-to-fork model are at the top of the value chain yet the reach of distant farmers to avail of such facilities is limited. To overcome this, in the year 2016, the Telangana government started a pilot project of Manakuragayalu to support distant farmers of semi-urban and rural areas to sell their produce to Hyderabad and the Secunderabad region of Telangana. They set up various collection centers in major districts like Medak, Rangareddy, a distribution center at Bowenpally market, and Manakuragayalu retail outlets at various locations to serve the consumers of Hyderabad and Secunderabad. The work presented in this paper is further conceptualized one step ahead of the Manakuragalyalu project in an omni-channel environment.

We consider a set of five collection centers (Medak, Siddipet, Rangareddy, Sangareddy, Mahabubnagar); one distribution center (Bowenpally); one Fulfillment center (Hyderabad); five Retail outlets; five Pick-up points, and three BODH points. The mixed-integer nonlinear programming model formulated in Sect. 2 is solved using the exact method in Python. The framework used is pyomo with IBM CPLEX solver. The problem is solved for two instances with varied sizes at freshness-keeping effort value $\tau = 0.5$, Penalty cost, P = Rs.5, Demand Dc = 5000 kg ($\alpha = 0.5$, $\beta = 0.2$), shelf life,ss = 24 h, and QRP = 0.3, as described in Table 5.3. Theeffect of freshness-keeping effort and penalty cost for quality loss on decision variable to whether invest on freshness-keeping effort and corresponding objective function value was analyzed as shown in Fig. 5.2. It was observed that with lower values of freshness-keeping effort and penalty cost, total cost is minimum with no requirement to investment towards freshness-keeping effort. As the penalty cost increases, the need for investment towards freshness-keeping effort is realized. With high penalty cost, table he trade-off between quality loss cost due to no freshness-keeping and freshness-keeping cost due to freshness is more clearly seen and it becomes imperative to invest for freshness-keeping efforts with higher value of freshness-keeping effort to attain optimal total cost of operation.



τ	0.1	0.1	0.1	0.25	0.25	0.25	0.5	0.5	0.5
P	0.125	2.5	5	0.125	2.5	5	0.125	2.5	5

Fig. 5.2 Effect of freshness-keeping effort and penalty cost on total cost and freshness-keeping decision

5.4 Conclusions

This work captured the trade-off between freshness-keeping cost and quality loss cost in combination with transportation cost for omni-channel fresh food distribution network while simultaneously evaluating freshness-keeping decisions and transportation network decisions for multi-echelon fresh food supply chains in the Indian context. It was observed that with the increase in quality loss cost, it becomes essential to deploy freshness-keeping effort into the distribution network to attain optimized total cost. While with penalty cost being low, decision on freshness-keeping effort depends on the purchase probability at each node. The integrated cost-effective decision support model would be helpful to decision-makers and stakeholders in designing and choosing the best choice of the omni-channel framework based on the scale and nature of the supply chain. The model would also assist in reducing fresh food wastage by its ability to support food quality monitoring and suggest the right levels of freshness-keeping effort to be maintained. The present work holds immense potential to be extended further by incorporating multiple products, multiple time periods, decision to choose between the type of omni-channel point of sale, and environmental concerns with single as well as multi-objective modeling approaches.

References

1. FAO (2019) The state of food andagriculture (2019). Moving forward on food loss and waste reduction, Rome
2. Jha SN, Vishwakarma RK, Ahmad T, Rai A, Dixit AK (2015). Assessment of quantitative harvest and Post-harvest losses of major crops and commodities in India. Ludhiana: ICAR-CIPHET

3. Poore J, Nemecek T (2018) Reducing food's environmental impacts through producers and consumers. Science 360(6392):987–992
4. Industry 4.0 in Food Industry India Food Report. Deloitte 2018
5. Chao L, Weidong C, Qian Z, Jing M (2021) Modelling dynamic freshness keeping effort over finite time horizon in a two-echelon online fresh product supply chain. Eur J Oper Res 293:511–528
6. Cai X, Chen Y, Xu X (2010) Optimization coordination of fresh product supply chains with freshness-keeping effort. Prod Oper Manag 19(3):261–278
7. Jing C, Ming D, Lei X (2018) A perishable product shipment consolidation model considering freshness-keeping effort. Transp Res Part-E 115:56–86
8. Yang L, Ruihong T (2019) Comparison of sales modes for a fresh product supply chain with freshness-keeping effort. Transportation Research Part-E 125(425–448):2019
9. Xueli M, Jian W, Qingguo B, Shuyun W (2020) Optimization of three-echelon cold chain considering freshness keeping effort inder cap-and-trade regulation in Industry 4.0. Int J Prod Econ, 220
10. Marco B, Maurizio F, Emilio F, Mauro G, Francesco P (2015). Fresh food sustainable distribution: cost, delivery time and carbon footprint three-objective optimization. J Food Eng

Chapter 6
Redesigning of Procurement Distribution System Network: An Application of Clustering Algorithms

M. R. Ganesh and T. Radharamanan

Abstract With the objective of minimizing the total distance travelled, the work proposes a novel procurement-distribution design network of the supply chain for the fair price shops (FPS) in the public distribution system of India. The work is a two-step process. The first step is to understand the existing distribution system. This starts from identifying the existing location of the FPS and converting them to its geographical locations in terms of longitudes and latitudes which are converted to distances to estimate the total distance travelled. The study, as a case, is conducted for one of the districts of Kerala, which is the rice bowl of Kerala which has 14 depots that serves the total 941 FPS in the district. The total distance travelled by the current distribution network is identified in the first step. The second step of the work is to determine the optimal number of depots that would be required to serve the existing 941 FPS of the district. To determine the optimum number of depots K-mean clustering applying the centroid approach is proposed to be adopted. We are proposing a novel distribution cum procurement supply chain. The trucks on delivery of the goods can collect the grains procured at the FPS from the farmers. The second step would primarily optimize the number of depots required to serve the FPS, and then identify the FPS that could be distributed from the given depot. The objective is to provide a scientific solution to minimize the total transportation distance by identifying the geographical location and adopting the k-mean clustering procedure. Finally, the benefits of implementing the proposed model will be quantified and discussed in detail by comparing the existing distribution system.

M. R. Ganesh
Mechanical Engineering Department, National Institute of Technology Calicut, Kozhikode, Kerala, India

T. Radharamanan (✉)
School of Management Studies, National Institute of Technology Calicut, Kozhikode, Kerala, India
e-mail: radha_ramanan@nitc.ac.in

M. K. Tiwari et al. (eds.), *Applications of Emerging Technologies and AI/ML Algorithms*, Asset Analytics, https://doi.org/10.1007/978-981-99-1019-9_6

6.1 Introduction

For feeding the enormous population of 1.35 billion population of our country, a high and efficient food grain production and distribution system is very important. For providing basic food and non-food commodities to the economically vulnerable sections of society, the Public Distribution System (PDS) distributes it through government-sponsored Fair Price Shops (FPSs) at cheap prices in every part of the country. Central government holds the responsibility of procurement, storage, transportation and bulk allocation of essential commodities to the states, while state governments take operational responsibility of allocation within state, identification of eligible families, issue of ration cards and supervision of the functioning of FPSs.

This work focuses on the possibility of the procurement of food grains from the farmers at the FPSs which will reduce the supply chain cost of PDS [1]. Transportation cost being the major contributor to the total cost, effort is done for reducing the same. In an attempt for that, the FPSs are clustered so that the total distance travelled is reduced thereby reducing the transportation cost.

6.2 Literature Review

According to Deb et al. [2], k-mean clustering offers good performance in globular cluster distinguishing. Implementation of the elbow method helps to overcome the weakness of dependence on the initial selection of the number of clusters. The method is a visual method of testing the consistency of the best number of clusters by comparing the difference of sum of squares of each cluster, the extreme difference forming the angle of elbow showing the best cluster according to Umargono et al. [3]. Frohlich and Westbrook [4] showed that cluster analysis helps in the betterment of supply chain management. Keeney [5] recommended using clustering for assigning customer areas to different facilities after which many researches have been done in exact and non-exact methods in this direction. It has been found that for customer segmentation studies k-mean clustering model has been employed effectively [6].

For routing and spare parts distribution problem, Feliu [7] developed a two-step algorithm, in which the customer clustering and distribution center determination is performed first, and in the second step, distribution channels are determined. The problem discussed in the problem is very much similar but in a different area of application.

From the literature, it was found that k-mean clustering can be efficiently employed for the improvement of supply chain efficiency. Hence, the algorithm is used for redistribution of the FPSs under various depots from they are served by the food grains for public distribution.

6.3 Model Formulation

K-means has a high clustering speed and performs well in large datasets. The objective of the classical K-means clustering method is to find the set $J(V)$ of C clusters C_J with cluster mean v_j for the sake of decreasing the amount of squared errors [7].

$$J(V) = \sum_{i=1}^{c} \sum_{j=1}^{c_i} \left(\|x_i - v_j\| \right)^2 \tag{6.1}$$

where $\|x_i - v_j\|$ is the Euclidean distance between x_i and v_j, c_j is the number of data points in the ith cluster, c is the number of cluster centres and v is the set of centres given by

$$v_i = (1/c_i) \sum_{j=1}^{c_i} x_i \tag{6.2}$$

In this problem, the data points are the set of FPS locations and the centroid points are depot locations.

6.4 Solution Methodology

The test problem for the model consists of 941 FPSs and 14 depots from the district of Palakkad in Kerala. The location of FPSs and the depots are obtained from Aadhaar-enabled Public Distribution System—AePDS website (https://epos.kerala.gov.in/dfso_fps_details) and google maps. The location in latitude–longitude format is utilized for the processing of data. After obtaining the locations k-mean clustering is performed for the obtained data in Google Colab (https://colab.research.google.com/notebooks).

Initially, the current system is evaluated for the total distance travelled using the data obtained. The FPSs are allotted to the depots on a taluk basis. All the FPSs in a particular taluk are assigned to one or more depots in the particular taluk. If the net demand of FPSs is not satisfied by the depots in the taluk, then the remaining FPSs may be assigned to the nearby depots. This administrative restriction limits the possibility of supplying food grains to the nearest FPSs from the depots resulting in increasing the total distance travelled by trucks for distribution. To overcome this issue, possibility of reducing the distance travelled by trucks is investigated in this work.

The problem is solved in two stages. In the first stage, the FPS shops are rearranged into the existing depots without considering the administrative boundaries. Hence, the depots will serve the FPSs irrespective of whether the FPSs lie in the same taluk or not. Here, only the distance is taken as a criterion. In the second stage, clustering of FPSs is done and the optimal locations for the depots are found out. A significant reduction in transportation is expected by optimizing the depot locations.

6.5 Results and Discussions

Initially, 941 FPSs are currently served from 14 depots on the basis of administrative boundaries of taluks. The total distance travelled in the distribution system is found to be 8879.12 km.

In the first stage of the problem, the FPSs are clustered into these existing depots without considering the administrative boundaries. The total distance travelled by the trucks for distribution is obtained as 5724.75. A reduction of 35.5% in total distance compared to the current system is observed in this case. In the second case, the FPSs are clustered using the k-mean cluster algorithm with $k = 14$ (which is the number of depots in the district). K-mean algorithm successfully provides the location of new depot points, and the total distance travelled from these depots to the FPS locations is obtained as 5297.28 km. The reduction in total distance travelled is 40.34% compared to the present distribution network.

The clusters for both of the stages are shown in Fig. 6.1. In Fig. 6.1, the result from Python is illustrated. The clusters and the corresponding cluster points are shown by various coloured dots. The red crossed points show the location of existing depots. The blue symbols represent the location of new depots.

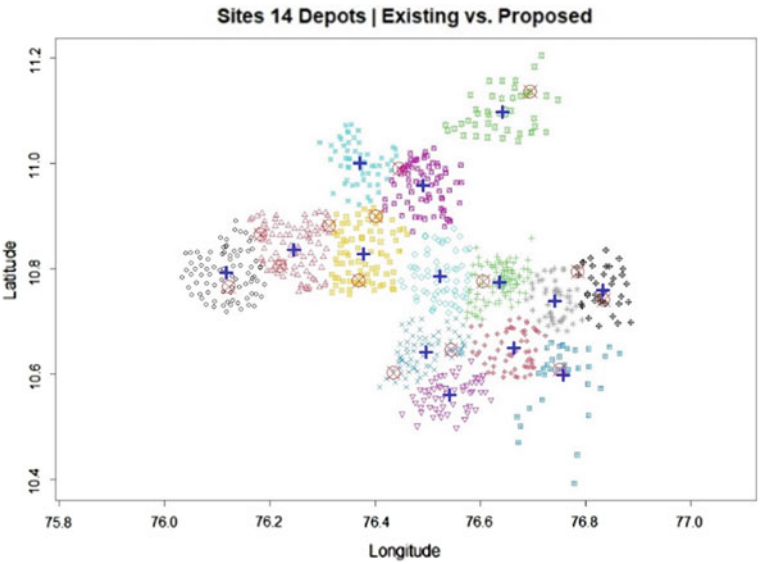


Fig. 6.1 Existing and new location of PDS depots. Red circles show the existing locations whereas the blue marks indicate the new locations

Table 6.1 Table showing number of FPS being served from the depots in existing system and after redistribution of FPS

Depot No.	No. FPS in the existing system	No. of FPS after redistribution
1	49	70
2	44	103
3	53	47
4	170	51
5	39	142
6	35	53
7	44	40
8	77	83
9	58	94
10	113	42
11	56	61
12	106	40
13	53	58
14	44	57
Total no. depots	941	941
Distance (in km)	8897.12	5297.28
Sum of squared error		3.5

Table 6.1 shows the number of FPS being served from the depots in the existing system and the number of FPS served from the depots after the redistribution of FPS to existing depot locations. As mentioned earlier, a reduction of 35.5% of the total distance travelled is found in this case.

Table 6.2 shows the number of FPS being served from the relocated depots. The k-mean clustering is performed, and 14 clusters are obtained with new centroid points (depot locations). The total distance is also shown which is reduced up to 40.37%. The sum of squared error values is also indicated in the tables.

6.6 Conclusions

The paper discusses the redistribution of FPSs on the basis of distance travelled from the depots so that the transportation cost of the food supply chain network is reduced. Implementing the k-mean clustering to the problem it was found that with redistribution of FPSs alone to the existing depots, a 35.5% reduction in the total distance travelled is noted. When the depots are relocated, about 40.37% of reduction is observed. Hence, it is inferred that the redistribution of FPSs will help in reducing the total distance travelled in the food distribution network thereby reducing the total

Table 6.2 Table showing number of FPS being served from the relocated depots

Depot No.	No. FPS
I	57
II	38
III	75
IV	69
V	50
VI	40
VII	105
VIII	67
IX	85
X	84
XI	93
XII	77
XIII	57
XIV	44
Total no. depots	941
Distance (in km)	5262.65
Sum of squared error	2.8

supply chain cost. Determining the distribution network for the clusters is a work that can be performed as an extension of this work. It will further improve the efficiency of the supply chain in terms of cost and time of procurement and delivery.

References

1. Ganesh MR, Rijas BM, Radharamanan T (2019) Modeling and analysis of food grains procurement: a case study. In: Proceedings of IMEC 2019 international mechanical engineering congress (IMEC-2019), NIT Tiruchirappalli, India
2. Deb S, Tian Z, Fong S, Wong R, Millham R, Wong KK (2018) Elephant search algorithm applied to data clustering. Soft Comput 22(18):6035–6046
3. Umargono E, Suseno JE, Gunawan SV (2020) K-means clustering optimization using the elbow method and early centroid determination based on mean and median formula. In: The 2nd international seminar on science and technology (ISSTEC 2019), pp 121–129
4. Frohlich MT, Westbrook R (2001) Arcs of integration: an international study of supply chain strategies. J Oper Manag 19(2):185–200
5. Keeney RL (1972) A method for districting among facilities. Oper Res 20(3):613–618
6. Shin HW, Sohn SY (2004) Segmentation of stock trading customers according to potential value. Expert Syst Appl 27(1):27–33
7. Gonzalez-Feliu J (2013) Vehicle routing in multi-echelon distribution systems with cross-docking: A systematic lexical-metanarrative analysis. Comp Inform Sci 6(3). ISSN 1913-8989
8. Zeebaree DQ, Haron H, Abdulazeez AM, Zeebaree SR (2017) Combination of K-means clustering with genetic algorithm: a review. Int J Appl Eng Res 12(24):14238–14245

Chapter 7
Big Data Analytics and Its Applications in Supply Chain Management: A Literature Review Using SCOR Model

Abhijeet Ghadge and D. G. Mogale

Abstract As we move further into the twenty-first century, technology has played an essential role in people's lives inevitably, and generated massive information called 'Big Data (BD)'. With the ability to manage massive dataset, big data analytics can usefully extract the insight from big data and support the firm to leverage decision-making. Hence, the interest in big data applications has spread to comprehend many areas of study, including Supply Chain Management (SCM). However, the academic articles studying the employment of Big Data Analytics (BDA) in SCM are limited. Besides, most of those academic papers offer less interest in the entire SC systems. Most of them prefer to study in an individual SC area. Thus, this study aims to investigate state of the art in this domain through Systematic Literature Review (SLR) and discuss future research opportunities. We found that optimisation, simulation and visualisation tend to be the most appropriate BD tools to apply in SCM. Also, linear programming, statistics, association rule mining, fuzzy logic and decision tree are likely to be the most suitable BDA techniques for SC operations.

Keywords Big data analytics · Supply chain management · Big data tools and techniques · SCOR model · And Systematic literature review

7.1 Introduction

Big data (BD) is mostly interpreted as the enormous collection of data that has rapidly increased ahead of the analytical capability of traditional applications. Consequently, Dey et al. [6] described that big data is the massive store of information kept on any

A. Ghadge
Centre for Logistics and Supply Chain Management, School of Management, Cranfield University, Cranfield MK43 0AL, UK
e-mail: Abhijeet.Ghadge@cranfield.ac.uk

D. G. Mogale (✉)
Logistics and Operations Management Section, Cardiff Business School, Cardiff University, Cardiff, UK
e-mail: MogaleD@cardiff.ac.uk

M. K. Tiwari et al. (eds.), *Applications of Emerging Technologies and AI/ML Algorithms*, Asset Analytics, https://doi.org/10.1007/978-981-99-1019-9_7

servers that is difficult to analyse by conventional tools. Even definitions of BD generally concentrate on data size, other characteristics comprehensively explain the word 'big data' known as the four Vs: Volume—the size of data, Velocity—the speed of data, Variety—the structure form of data and Veracity—the accuracy of data. BD presents characteristically as large volume, high velocity, less structured or unstructured forms, and less accuracy.

Big data analytics (BDA) has a high capability to manage the complexity of BD. Bolstorff and Rosenbaum [3] explained that BDA includes 'big data' and 'advanced analytics' which can discover valuable business trends. Arunachalam et al. [1] classified data analytics techniques into Descriptive, Predictive and Prescriptive analytics. Supply Chain Operations Reference (SCOR), established by the Supply Chain Council (SCC), is a reference model for SC activities to facilitate the SC organisations in case of exchanging the knowledge or information. The SCOR model includes five compositions (Plan, Source, Make, Deliver and Return) [7]. Presently, limited authors discussed the collaboration of BDA and SCM and they explained specific operation like production, planning and procurement [12]. Hence, there is a need of holistic analysis of application of BDA in SC. Following this background, this study attempts to answer the research question: which are BDA tools and techniques that appropriate for each stage of the SCOR model? The aim of this study is to analyse the tools and techniques of BDA used in SCM and their benefits through a Systematic Literature Review (SLR).

The reminder of the paper is arranged as follows. Section 7.2 'Research Methodology' provides the methods of research delimitation. Then, the paper describes the key findings separated into 'descriptive findings' performed in Sect. 7.3 and 'thematic findings' described in Sect. 7.4. Lastly, Sect. 7.5 'conclusion' represents the research limitations and concluding statements.

7.2 Research Methodology

A literature review offers an essential tool for consolidating, improving, and synthesising several papers to create a research theory of study field [16]. According to Bearman et al. [2], the literature review can be classified into a traditional narrative review and a systematic review. The first one, the traditional review performs a specific literature review through the viewpoint of the reviewer. Secondly, the systematic literature review (SLR) is a descriptive approach that employs reproducible scientific procedures, such as the developed search strategies and a well-organised presentation of the research findings. Owing to Tranfield et al. [19], the procedure of SLR has been claimed to be the most efficient method that contributes to the high-quality results of the literature review. Based on Ghadge et al. [8], the following are three significant steps of the SLR approach: 1. Identification of sources, 2. Data screening and synthesis and 3. Data analysis and dissemination.

In the first process, search strings were listed through the exploration of the primary studies about the application of BDA in SCM. This procedure was operated based on the 'Boolean method'. Accordingly, the words, 'Big Data Analytics' and 'Supply Chain Management' were the main keywords. Also, the other words that frequently occur in the BDA and SCM literature were included in the searching words to cover all relevant fields of BDA and SCM. Then, two renowned academic databases, Scopus and Web of Science, were chosen for data extraction. Both data sources are commonly used for literature review and proven to offer inclusive outcomes [10].

There are five criteria applied for this study. The first one is the publication year. Secondly, language is identified to be one of the criteria. Only the literature in English were considered in this paper because English is a universal language for academic articles. Thirdly, only academic journals were selected for the study to ensure the quality of the included literature. Other academic and non-academic documents like letters, textbooks, book chapters, conference papers and editorial were excluded from the study. Fourthly, the comprehended articles of this research were peer-reviewed because these peer-reviewed literature have been guaranteed the journal quality by the experts. Finally, the full-text assessment may be significant for the literature review. It can confirm that the author can read all contents in the paper and synthesise the research method and findings comfortably. Then, all included papers of the study were full-text articles. In second stage, the inclusion and exclusion criteria were employed to screen the articles to obtain the high quality of the evidence-based approach. Only literature that met all the pre-defined criteria were included in the literature review [19]. Accordingly, a framework for data screening and synthesis was created based on inclusion and exclusion criteria. After the data extraction and the criteria examination, the authors read the title, keywords and abstracts of each paper to approve that their contents involve the application of BDA in SCM. The studies that cover only BDA or SCM individually were excluded from the lists of selected journals.

In the last stage, the method of data analysis and dissemination is recognised as the most meaningful section of SLR because this part presents the summary and evaluation of article findings that developed from the collective literature to the readers [19]. Overall, data analysis is split into the descriptive and thematic analysis. The descriptive analysis applies statistical knowledge to explain the results. For the thematic analysis, the information from the included literature review will be developed to achieve the research objective and answer the research question.

7.3 Descriptive Analysis

A descriptive finding of previous studies obtained from Scopus and Web of science following the SLR process in Sect. 7.2 is presented here. The results of a systematic review are statistically described owing to publication year. There are 20

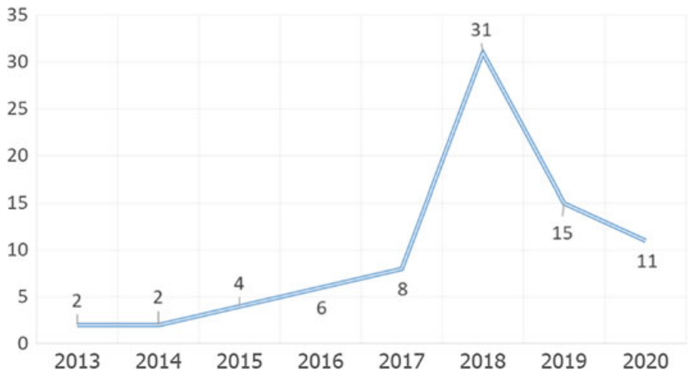


Fig. 7.1 The yearly number of BDA articles in the context of SCM

published journals distributed from the included works of literature. Further, the articles published in the International Journal of Production Research and Computer & Industrial Engineering mostly appear in the literature review, eight papers of each journal.

7.3.1 Year Wise Publications

Figure 7.1 shows the yearly number of studies increases continuously, and peak in 2018, 31 articles. Then, the quantities of papers decrease to 15 in 2019, and 11 in the half-year of 2020.

7.3.2 Methodological Distribution of Articles

Based on Fig. 7.2, the quantitative method, 49%, is the most favoured approach used by the researchers to conduct the academic studies related to BDA and SCM, followed by the qualitative method, 42%. Also, a mixed method which employs both qualitative and quantitative methodology offers the least popular practice among the authors, 9%.

7.3.3 Distribution of BDA Types

According to Fig. 7.3, predictive analytics that supports future forecasting is the most prevalent BDA type when applying BDA in SCM, 43%. This statistical result agrees with the descriptive conclusion of supply chain function and SCOR activities that

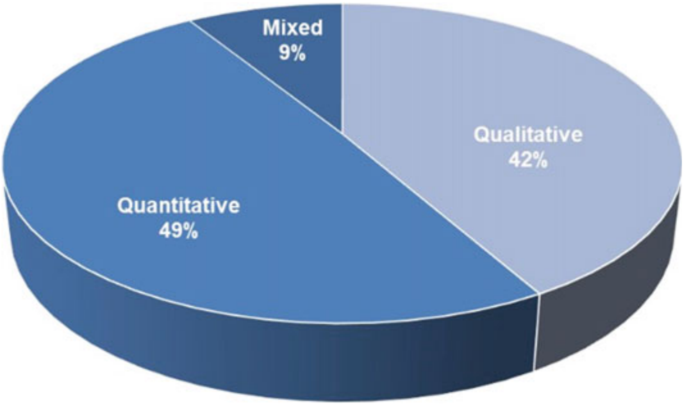


Fig. 7.2 The percentage of articles based on the methodologies

retail/customer and plan process are the preferable areas to conduct BDA. Further, prescriptive places as the second rank of BDA types that usually applies in SCM research with 31%, while descriptive/diagnostic shows the least percentage, 26%.

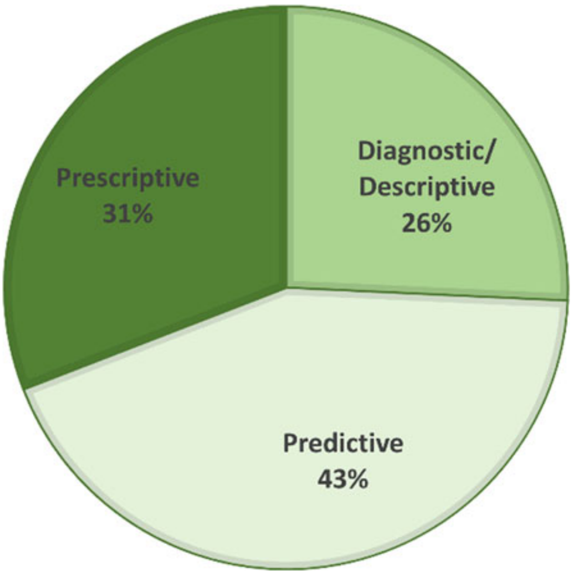


Fig. 7.3 The percentage of articles based on BDA types

7.3.4 Distribution of BDA Tools

Owing to Fig. 7.4, mixed technique, especially data mining, machine learning and statistical modelling, tend to be the most popular tools, 19.67%, used to solve SCM issues. An example of an academic article that suggests the data mining approach is 'A data mining-based framework for supply chain risk management [13]'. Also, the instance research that recommends the machine learning technique is 'Supply chain management and Industry 4.0: conducting research in the digital age [11]'.

7.4 Thematic Analysis

Thematic analysis aims to answer the research question.

SCOR model activities: Theme to describe the elements of the SCOR model, which consists of plan, source, make, deliver, and return.

Supply chain functions: Theme to understand the core events of logistics and supply chain management comprising procurement, manufacturing, logistics and transportation, warehousing and distribution, and retail and customer. The theme is established based on Nguyen et al. [15] to explore which SCM functions relate to each SCOR activity.

Types of BDA analytics: Theme to provide the classification of big data models is categorised into descriptive and diagnostics analytics, predictive analytics and prescriptive analytics. This theme is developed by referencing Arunachalam et al. [1] to explain the level of data analytics applied in a conceptual model.

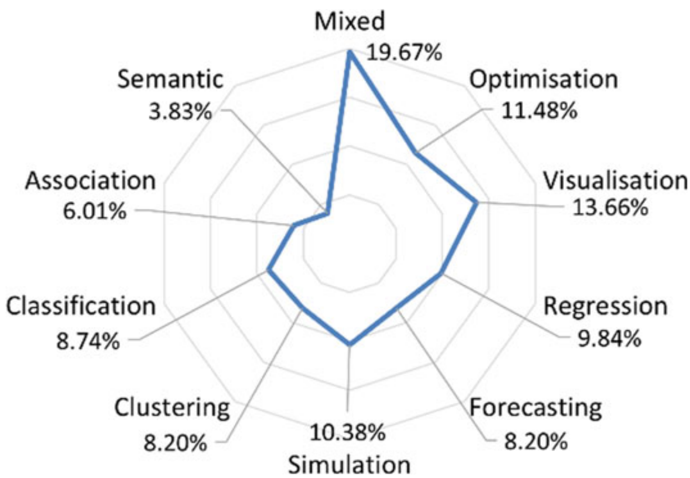


Fig. 7.4 The comparison of BDA tools applied in SCM

BDA tools: Theme to indicate the BDA tools is employed in SCM. The subcategories of this theme are identified according to Nguyen et al. [15] which include association, clustering, classification, regression, forecasting, semantic, optimisation, simulation, visualisation and mixed tools. For mixed algorithms, the author discovers from the descriptive analysis that data mining, machine learning and basic statistical modelling mostly present in the selected articles. Then, these three algorithms are added to the BDA tools to increase the accuracy of data analysis. The appropriate BDA tools are chosen based on the frequency of appearance in the literature review.

BDA techniques: After the literature review, the theme of BDA techniques is created to explore the most suitable BDA techniques used in SCM. The subcategories of this theme are essentially identified according to the findings from SLR, Nguyen et al. [15] and Chehbi-Gamoura et al. [5] including association rule mining (ARM), spatial-temporal visualisation (STV), decision tree (DT), support vector machine (SVM), linear programming (LP), linear regression (LIR), logistics regression (LOG), genetic algorithm (GA), neural network (NN), text mining (TM), k-mean clustering (KMC), time-series forecasting (TSF), Bayesian network (BN), fuzzy logic (FL), sequential analysis (SEG), sentiment analysis (SEN), natural language processing (NLP) and statistics (STA). The proposed BDA techniques are selected due to the frequency of presence in the SLR.

7.4.1 Plan Process

According to Wang et al. [20], BDA provides a significant role in the planning stage of SCM. It encourages organisations to enhance their decision making on demand planning, sourcing, production planning, inventory management, transportation, and the supply chain network design. Then, it is noticeable that the use of BDA in the planning phase is likely to cover all critical events in the supply chain.

Tiwari et al. [18] state that big data applications can notice demand signal, which facilitates the companies in terms of demand forecasting. Hence, organisations are likely to caption the new trends of markets and prepare for the change in the market and the fluctuation of demand more effectively. Consequently, BDA can support planners through the tracking of cash-cycle management, inventory replenishment and forecasting accuracy by tracking day-to-day sales information. Thereby, organisations can prevent missing products which lead to lost sales and good preparation for the peak season of products Hahn and Packowski [9]. Thus, BDA tends to offer improvements to supply chain planning significantly. In order to provide more specific guidance on the BDA employment in SCM, this study has explored the most appropriate BDA tools and techniques for the planning stage through the literature review.

7.4.2 Source Process

According to Nguyen et al. [15] and Tiwari et al. [18], the use of BDA in procurement frequently comprehends supplier selection, sourcing improvement, supply risk management and supplier performance. With the aim of leveraging supply risk management and supplier performance, big data applications can differentiate risks into the risk that should avoid and risk that should take to monitor market trends. Thereby, the firms can update and capture the alterations in sourcing or supplier markets, helping the company to respond to the supply change capably [20]. For supplier selection and sourcing improvement, BDA can efficiently investigate and analyse supplier performance [20]. Further, BDA can offer in-depth data analysis of the return on investment (ROI). Hence, the organisations may improve strategic sourcing through the worthy of financial information Tiwari et al. [18]. In an attempt to observe more BDA utilisation in the sourcing supply chain, this research has examined the most suitable BDA tools and techniques that mostly apply in sourcing activity.

7.4.3 Make Process

Due to Nguyen et al. [15] production control currently gains the interest of applying BDA. Wang et al. [20] state that big data tools can describe the production costs to the manufacturers. Moreover, BDA can provide the impact of manufacturing profits when there are any changes in production costs. Accordingly, the big data techniques can ensure the right production parts go to the right assembly by extracting the insight of production capacity. Further, BDA can be used to examine the amount of material waste and seek the proper manufacturing tools to decrease the level of material waste.

As well as that, inventory management seems to be a significant area for applying BDA. Big data applications assist the companies in leveraging the design of current stock operations by using the optimisation system. It is likely to help the firms to decrease the challenges of multi-channel inventories [20]. Also, Tiwari et al. [18] suggest that the sharing of information on internal and external production system would enhance the benefits of big data on inventory management. Accordingly, this study has examined the most advisable big data tools and techniques that mostly apply in the production process.

7.4.4 Deliver Process

The logistics function relates to the distribution of goods from the supply sites to the distribution centres, and the end customers. Also, logistics data can be massively generated by many players in the delivery route, such as carriers, transportation service providers, and shippers. Thereby, the development of BDA in distributing area seems significant to leverage SCM. The employment of BDA in logistics and transportation focuses on the fundamental logistics functions which consist of supply chain network design, route optimisation, and proactive safety management [17, 21]. With the purpose of enhancing the efficiency of supply chain network design, big data tools can calculate the optimal locations of the distribution centres by analysing the massive data generated from warehousing, transportation, and consumer demand [21].

7.4.5 Return Process

According to Chehbi-Gamoura et al. [5], the return process significantly applies big data applications in the areas of the green supply chain which specially concerns with ecology and cost improvement, reverse logistics management and the closed-loop supply chain. The green supply chain essentially focusses on the decrease in waste. Then, the returns of goods for remanufacturing, recycling and reuse can reduce the number of wastes spread in the environment. To support the green supply chain, BDA can be applied in the process control to restrict the pollution emission and manage the natural resources sustainably [21].

Furthermore, big data applications have a high potential to employ in reverse logistics and closed-loop supply chains. Nguyen et al. [15] claim that BDA has been used in product lifecycle design which is useful for enterprises to forecast the number of returned products. This information is significantly beneficial for the planning of production capacity and remanufacturing scheduling in reverse logistics. Consequently, big data enables organisations to adopt customer complaints to identify the root causes of defects. It enables the company to improve its product quality [4]. Optimisation and visualisation are likely to be the most appropriate BDA tools for raising the performance of the return process.

7.5 Conclusion and Future Work

Following the SLR process, this study aims to examine the state of the art in the domain of big data applications in SCM based on the SCOR model. Consequently, the study demonstrates BDA tools and techniques that seem appropriate to apply in each stage of the SCOR model. With an attempt to achieve the goal, this study aims to answer the research question: which are big data analytical tools and techniques that appropriate for each of the SCOR models?

In order to address research question, the SLR framework is conducted to obtain a high quality of academic articles for the literature review. Accordingly, papers related to the application of BDA in the context of supply chain operations are selected to collect and analyse critical information. The descriptive analysis shows that the researchers have significantly gained an interest in the utilisation of BDA in SCM during the last four years. Consequently, most of the researchers seem to concentrate on the strategic level of SCM. As well as that, demand management is related to retail, and customer tends to be the most favoured area for applying big data technology.

According to the thematic analysis, it has discovered the appropriate BDA tools and techniques for each stage of the SCOR model, which answer the research question potentially. The results show that LP optimisation, and ARM, DT, and FL simulations have a high potential to improve planning performance. Also, STA visualisation, LP optimisation, and ARM and DT simulations tend to support the operations in procurement usefully. To better the manufacturing practice, STA visualisation, LP and STA optimisations, and FL and STA simulations can be employed in this field. Consequently, LP optimisation, STA visualisation and STA simulation can be applied in logistics and transportation management to develop the delivery functions. For return activity, STA optimisation and STA visualisation are likely to leverage the reverse logistics performance.

The study offers significant implications for the theoretical research overview. It introduces the potential BDA algorithms that have a high capacity to improve the entire SCM with further academic evidence. The findings also provide the most suitable BDA tools and techniques to apply in the individual stage of SCM owing to the SCOR model and level of management. These findings tend to support the statement of Chehbi-Gamoura et al. [5] that the study of the employment of BDA across the whole supply chain at one research is essentially needed to manage the massive information generated in the SCM simultaneously.

The results of this study provide several practical implications for the practitioners. First, this study offers the guidance of what BDA tools and techniques that the organisations should consider applying in their supply chain practice to leverage the performance. The integrated big data applications from the beginning to the ending stages of SCM are presented through the conceptual model. It would support the company on the decision of big data investment across the entire chain.

Even this study is likely to offer useful insights, there are some limitations to concern. Firstly, the author only extracted the academic literature from SCOPUS and WOS databases to explore the research implications. Hence, some relevant articles published on other sources can be missed from the literature review. Also, the results are determined based on the author perspective only. Therefore, some analysis may be presented in different ways compared to the other papers. Thus, future research should employ the cross-checking process to minimise bias and errors []. A conceptual model integrating different BDA tools and techniques could be developed for guidance for the researchers or practitioners in case of BDA investment in SCM.

References

1. Arunachalam D, Kumar N, Kawalek JP (2018) Understanding big data analytics capabilities in supply chain management: unravelling the issues, challenges and implications for practice. Transp Res Part E Logist Transp Rev 114:416–436
2. Bearman M, Smith CD, Carbone A, Slade S, Baik C, Hughes-Warrington M, Neumann DL (2012) Systematic review methodology in higher education. High Educ Res Dev 31(5):625–640
3. Bolstorff P, Rosenbaum R (2011) Supply chain excellence: a handbook for dramatic improvement using the SCOR model, 3rd edn.
4. Brinch M, Stentoft J, Jensen JK, Rajkumar C (2018) Practitioners understanding of big data and its applications in supply chain management. Int J Logist Manage.
5. Chehbi-Gamoura S, Derrouiche R, Damand D, Barth M (2020) Insights from big data analytics in supply chain management: an all-inclusive literature review using the SCOR model. Prod Plan Control 31(5):355–382
6. Dey N, Das H, Naik B, Behera HS (eds) (2019) Big data analytics for intelligent healthcare management. Academic Press
7. Georgise FB, Thoben KD, Seifert M (2012) Adapting the SCOR model to suit the different scenarios: a literature review & research agenda. Int J Bus Manage 7(6):2
8. Ghadge A, Wurtmann H, Seuring S (2020) Managing climate change risks in global supply chains: a review and research agenda. Int J Prod Res 58(1):44–64
9. Hahn GJ, Packowski J (2015) A perspective on applications of in-memory analytics in supply chain management. Decis Support Syst 76:45–52
10. Harzing AW, Alakangas S (2016) Google Scholar, Scopus and the Web of Science: a longitudinal and cross- disciplinary comparison. Scientometrics 106(2):787–804
11. Hofmann E, Sternberg H, Chen H, Pflaum A, Prockl G (2019) Supply chain management and Industry 4.0: conducting research in the digital age. Int J Phys Distrib Logistics Manage
12. Kache F, Seuring S (2017) Challenges and opportunities of digital information at the intersection of Big Data Analytics and supply chain management. Int J Oper Prod Manage
13. Kara ME, Fırat SÜO, Ghadge A (2020) A data mining-based framework for supply chain risk management. Comput Ind Eng 139:105570
14. Mishra D, Gunasekaran A, Papadopoulos T, Childe SJ (2018) Big Data and supply chain management: a review and bibliometric analysis. Ann Oper Res 270(1):313–336
15. Nguyen T, Li ZHOU, Spiegler V, Ieromonachou P, Lin Y (2018) Big data analytics in supply chain management: a state-of-the-art literature review. Comput Oper Res 98:254–264
16. Seuring S, Gold S (2012) Conducting content-analysis based literature reviews in supply chain management. Supply Chain Manage Int J
17. Singh A, Kumari S, Malekpoor H, Mishra N (2018) Big data cloud computing framework for low carbon supplier selection in the beef supply chain. J Clean Prod 202:139–149
18. Tiwari S, Wee HM, Daryanto Y (2018) Big data analytics in supply chain management between 2010 and 2016: insights to industries. Comput Ind Eng 115:319–330
19. Tranfield D, Denyer D, Smart P (2003) Towards a methodology for developing evidence-informed management knowledge by means of systematic review. Br J Manag 14(3):207–222
20. Wang G, Gunasekaran A, Ngai EW, Papadopoulos T (2016) Big data analytics in logistics and supply chain management: certain investigations for research and applications. Int J Prod Econ 176:98–110
21. Zhao R, Liu Y, Zhang N, Huang T (2017) An optimization model for green supply chain management by using a big data analytic approach. J Clean Prod 142:1085–1097

Chapter 8
Leveraging Machine Learning of Indian Railways Public Procurement Data for Managerial Insights

Samir Maity, Bodhibrata Nag, and Sushovan Khatua

Abstract This paper proposes and demonstrates (a) a supervised machine learning methodology to predict the expenditure incurred on procurement, repair, and reconditioning of components as well as expenditure incurred on procurement of fuel based on the performance of the Indian Railways and (b) an unsupervised machine learning methodology to classify the good and poor performing administrative zones, using the data of expenditure incurred on procurement, repair, and reconditioning of components as well as expenditure incurred on procurement of fuel and the performance. The first methodology will aid managers in determining whether the expenditure incurred is more than what should be incurred. Further, it may also benefit the managers in fine-tuning the frequency of replacement of components. The second methodology will assist managers in searching for best practices of maintenance in good-performing zones, which can be propagated in the poor-performing zones to lower the overall expenditure on maintenance.

Keywords Public procurement · Indian Railways · Machine learning · Performance prediction · Classification

8.1 Introduction

Public procurement is the term given for procurement made by the public sector. Public procurement includes (a) goods for maintenance and repair of equipment, plant, and machinery, (b) goods for the operation of assets, and (c) works and

S. Maity
Data Science Department, University of Kalyani, Kalyani 741235, India

B. Nag (✉)
Operations Management Group, IIM Calcutta, Kolkata 700104, India
e-mail: bnag@iimcal.ac.in

S. Khatua
Computer Science & Engineering Department, Moulana Abul Kalam Azad University of Technology, Haringhata 741249, India

M. K. Tiwari et al. (eds.), *Applications of Emerging Technologies and AI/ML Algorithms*, Asset Analytics, https://doi.org/10.1007/978-981-99-1019-9_8

services for the construction of buildings and assets, providing security to premises, maintaining computer assets, etc. [3].

Public procurement involves humongous expenditure. The annual expenditure incurred on public procurement in India is estimated at USD 500 billion, which is approximately 20–22% of India's GDP [2]. The bulk of the public procurement expenditure is incurred by organizations such as Indian Railways, Defence, Ministry of Road Transport, Central Public Works Department, and BSNL.

We examine the case of Indian Railways for this paper. Indian Railways is the world's fourth-largest railway. It comprises 17 administrative zones. It carries twenty million passengers and three million tons of freight every day by operating more than twenty thousand trains daily. Its assets include a network of 68 thousand kilometers, 155 thousand bridges, 6200 stations, 61 thousand kilometers of optical fiber cable network for communication and train operations, 12 thousand diesel and electric locomotives, 8 thousand passenger coaches, 300 thousand freight wagons, 300 depots for repair and maintenance of locomotives, coaches and wagons, 1.2 million quarters for employee housing, 700 hospitals and health units with around 14 thousand beds, 100 schools and colleges and seven factories for the production of locomotives and carriages. Its annual expenditure is USD 23 billion. The expenditure incurred on wages and employee benefits is approximately 75%, the fuel bill is approximately 15%, and the remaining is incurred towards the procurement of goods and services for repairs and maintenance.

There must be minimal failures of assets to ensure the safe and reliable operation of twenty thousand trains daily on the Indian Railways network. However, every asset is subject to enormous wear and tear due to the continuous operation of train services, resulting in failure. For example, the track may fail due to corrosion and rusting, change in geometry of rail table, corrugation, cracks, gauge changes, bridge foundation or superstructure or bearing failures, overhead catenary wire defects, signaling system defects; locomotives may fail due to problems with its transformers, motors, relays, overhead current collection system or superstructure, bogies and bearings; passenger carriages may fail due to problems with batteries, air conditioners, lighting system, superstructure, bogies, and bearings; freight wagons may fail due to problems with superstructure, bogies, and bearings. Indian Railways thus has to carry out both scheduled and unscheduled maintenance of all its assets so that their failure rate is as low as possible. During maintenance, failure-prone components are replaced with new or reconditioned components so that the asset is not prone to failure while in service. The frequency of replacement of failure-prone components is decided by trial and error based on the mean time between failure data collected continuously by Indian Railways. Further, failed components are replaced with new or reconditioned components during the repair of failed assets. While the replacement of failed components of the track is done in situ, the replacement of failed components of locomotives, carriages, and wagons may be done either in repair depots or elsewhere as required during operation. The cost incurred in procurement, repair, and reconditioning of components accounts for a sizeable share of Indian Railways expenses, as indicated earlier.

Indian Railways is part of the Ministry of Railways of the Government of India. Therefore, it is required to seek approval of its annual budget from the Indian Parliament through a budget document termed "Demand for Grants." This budget document details the actual expenditure incurred under various heads of expenses and the projected expenditure in the financial year for which the approval is sought from Parliament. Further, the Indian Railways publishes an annual statistical document wherein the actual expenditure incurred in the various administrative zones under multiple heads of expenses. The annual budget documents and the annual statistical documents are available in the public domain, and their data have been used for this paper.

The expenditure incurred for procurement, repair, and reconditioning of components is captured in the following five heads of expenses:

(a) Primary Unit 4: cover the expenditure incurred for procurement, repair, and reconditioning of components of the track.
(b) Primary Unit 5: cover the expenditure incurred for procurement, repair, and reconditioning of components of locomotives in repair depots.
(c) Primary Unit 6: cover the expenditure incurred for procurement, repair, and reconditioning of components of carriages and wagons in repair depots.
(d) Primary Unit 8: cover the expenditure incurred for procurement, repair, and reconditioning of components of locomotives, carriages, and wagons during operation.

Further, another Primary Unit 10 captures the expenditure incurred in the procurement of fuel for operations.

The expenditure incurred on procurement, repair, and reconditioning of components is monitored by the Indian Railways management for percentage changes on a year-to-year basis. However, there is no methodology currently available to monitor the expenditure in relation to the performance parameters of the Railways. This paper proposes and demonstrates (a) a machine learning methodology to predict the expenditure incurred on procurement, repair, and reconditioning of components as well as expenditure incurred on procurement of fuel based on the performance of the Indian Railways and (b) a machine learning methodology to classify the good and poor performing administrative zones, using the data of expenditure incurred on procurement, repair, and reconditioning of components as well as expenditure incurred on procurement of fuel and the performance. The first methodology will aid managers in determining whether the expenditure incurred is more than what should be incurred. Further, it may also assist the managers in fine-tuning the frequency of replacement of components. The second methodology will aid managers in searching for best practices of maintenance in good-performing zones, which can be propagated in the poor-performing zones to lower the overall expenditure on maintenance.

The performance parameters of Indian Railways used for the purpose of the study discussed in this paper are (a) Annual Net ton-km (NTKM) of freight carried by Indian Railways (obtained by multiplying the weight in tons of each freight consignment by the distance of transportation and adding the result for all freight consignments carried by Indian Railways over the year) and (b) Annual Passenger km (PKM) carried by Indian Railways (obtained by adding the distance of transportation of each passenger consignments carried by Indian Railways over the year).

To achieve the above-proposed goal, machine learning (ML) is the best available methodological tool to apply to the available data for predicting and classification of the performance of the Railways. ML-based approaches can be used to determine the interdependencies of the different railway expenditures to the performance parameters stated above. Since clustering is the key methodology in discovering the hidden pattern of the data, this paper proposes a k-means clustering strategy to classify different administrative zones of the Indian Railways based on the procurement expenditures in various areas.

The rest of the paper is structured as follows: a concise literature review is given in Sect. 8.2; a methodological description is given in Sect. 8.3; a discussion on managerial insights is given in Sect. 8.4; conclusions and further directions of research is given in Sect. 8.5.

8.2 Literature Survey

There is very little research in the areas of application of machine learning in procurement. Rodríguez et al. [4] paper proposes an award price estimator based on bidding price using the random forest regression method. It was trained and tested on a dataset of 58,337 tenders for public procurement in Spain from 2012 to 2018. Domingos et al. [1] paper proposed a CRISP-DM data mining methodology to predict anomalies in IT purchases in the Brazilian Federal Procurement System. This paper will add to the growing repertoire of machine learning applications in public procurement.

8.3 Model Development

8.3.1 Feature Selection and Data Preparation

The expenditure incurred for procurement, repair, and reconditioning of components is captured in the following five heads of expenses. As described above, these have been chosen as features for this machine learning study to study the relationship of the expenditure in relation to the performance parameters of the Railways. The rationale for their choice is that the expenses incurred under these five heads vary with the utilization of railway assets (such as track, locomotives, passenger coaches, and freight wagons). The remaining heads of expenses deal with expenditure incurred on offices, staff benefits, and amenities for passengers and staff and thus do not vary with the utilization of railway assets.

The expenditure data for the period 2002–03 to 2020–21 has been used for the purpose of this study. Few missing data were imputed using means of data in the adjacent years.

8.3.2 *Model Building*

The data set was split to obtain training and testing sets with properties as given in Table 8.1.

Multiple linear regression, logistic regression, decision trees, random forest, and Naïve Bayes supervised machine learning algorithms were systematically applied with varying attribute subsets (with or without fuel cost, track repair and maintenance cost, etc.) to obtain a set of heterogeneous models for the model selection step to evaluate and choose the final model for identifying performance of the system. The performance of the developed models for different algorithms is given n Table 8.2. Based on the performance parameters indicated in Table 8.2, the Extra Tree Regressor algorithm model was chosen for this study.

Further, we used the unsupervised machine learning k-means algorithm to classify the good and poor performing administrative zones, using the data of expenditure

Table 8.1 Statistical properties of training and testing sets

Data points	Training set	Testing set
Min	14	12
Max	18	16
Mean	15.45	13.45

Table 8.2 Performance of the developed models

Models	Adjusted R-squared	R-squared	RMSE
KernelRidge	516.42	−256.71	860,837.86
MLPRegressor	479.89	−238.44	888,812.91
RANSACRegressor	5.85	−1.42	199,232.01
LinearRegression	4.76	−0.88	200,582.81
RidgeCV	3.12	−0.06	112,789.76
KNeighborsRegressor	2.98	0.01	129,523.91
Ridge	2.39	0.30	81,753.11
BaggingRegressor	1.49	0.75	43,366.67
ExtraTreesRegressor	1.49	0.75	44,404.03
RandomForestRegressor	1.44	0.78	21,716.34
ExtraTreeRegressor	1.37	0.81	26,530.33
DecisionTreeRegressor	1.37	0.82	20,634.53

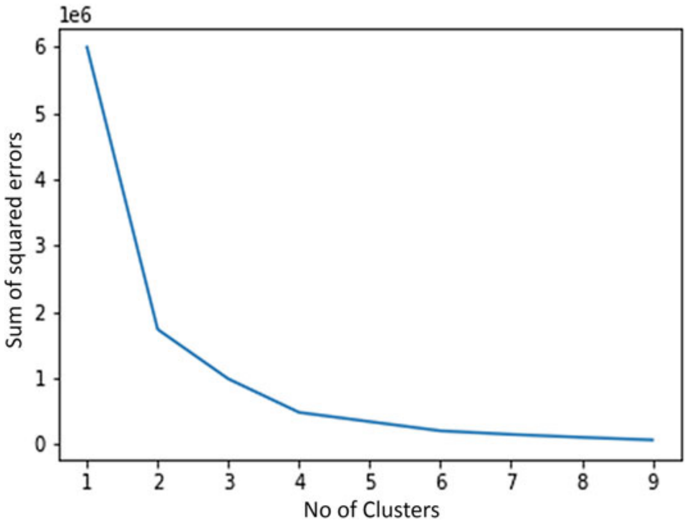


Fig. 8.1 Determination of optimal number of clusters using elbow method

incurred on procurement, repair, and reconditioning of components as well as expenditure incurred on procurement of fuel and the performance. For this purpose, the optimal number of clusters was determined as four using the Elbow method (as shown in Fig. 8.1).

8.4 Model Results

8.4.1 Influential Features for Performance Parameters

Application of the ExtraTreeRegressor algorithm model (discussed in Sect. 8.2) to the data corresponding to expenditure incurred on procurement, repair, and reconditioning of components as well as expenditure incurred on procurement of fuel and the performance parameters shows that fuel expenses are the highest influential feature for both passenger and freight transportation performance parameters as seen from Figs. 8.2 and 8.3.

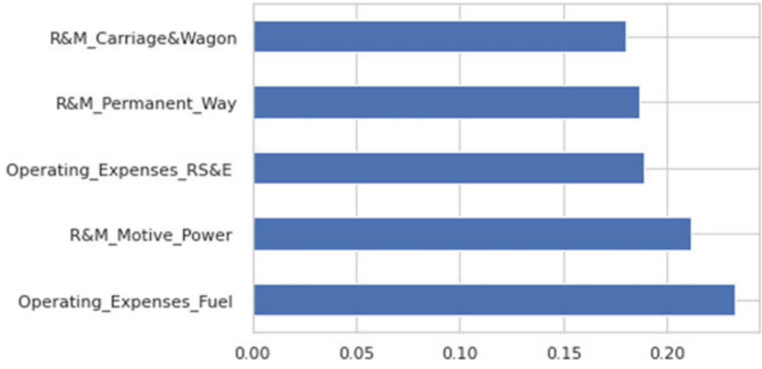


Fig. 8.2 Feature relationship with Freight Transportation Performance (NTKM)

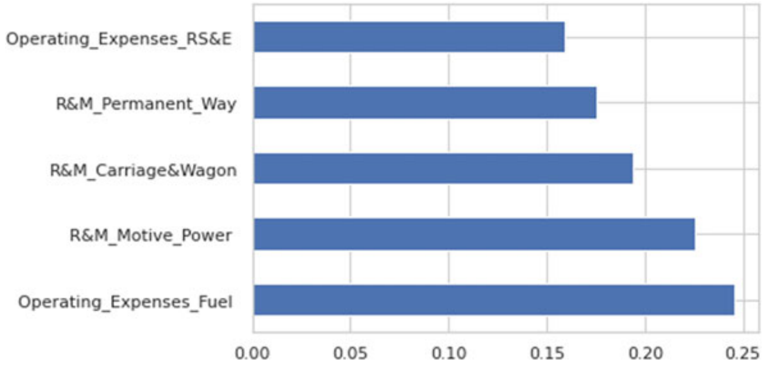


Fig. 8.3 Feature relationship with Passenger Transportation Performance (NTKM)

8.4.2 Comparative Performance Analysis of Administrative Zones

K-Means Clustering Algorithm was used to determine the classification of performance of administrative zones in four categories: (i) "Very High," (ii) "High," (iii) "Medium," and (iv) "Low." The classification of administrative zones is shown in Fig. 8.4.

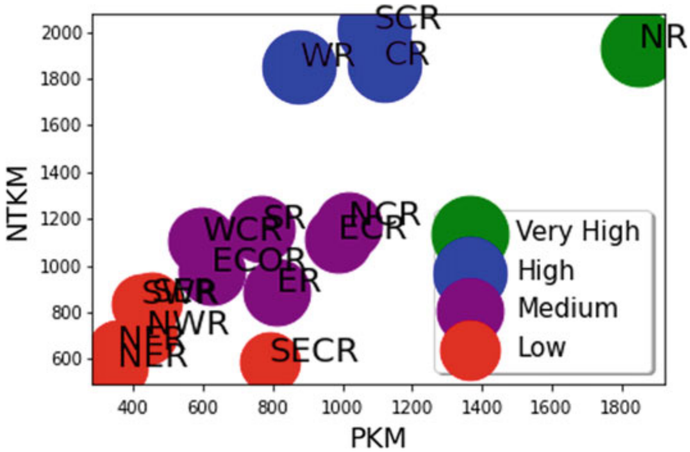


Fig. 8.4 Comparison of performance of zones

8.5 Conclusion

It will be seen from the above section the machine learning methodology enables the prediction of expenditure incurred on procurement, repair, and reconditioning of components as well as expenditure incurred on procurement of fuel based on the performance of the Indian Railways. Further, the machine learning methodology enables the classification of administrative zones in different performance categories, using the data of expenditure incurred on procurement, repair, and reconditioning of components, and expenditure incurred on procurement of fuel and the performance. These methodologies can be extended to apply more granular data in terms of the geography of operation of railway services, make of asset, age of the asset, and intensity of use of assets. It will thereby provide opportunities for the manager to decide the choice of make of asset, age of replacement of asset or its components, and maintenance schedules of assets.

References

1. Domingos SL, Carvalho RN, Carvalho RS, Ramos GN (2016) Identifying IT purchases anomalies in the Brazilian Government Procurement system using deep learning. In: Proceedings of the 15th IEEE international conference on machine learning and applications. IEEE, Anaheim CA, pp 722–727
2. Ministry of Finance, Govt of India (2019) FM reviews capital expenditure & payments of maharatnas and navratnas CPSEs. Retrieved from Press Information Bureau: https://pib.gov.in/PressReleasePage.aspx?PRID=1586546#:~:text=Public%20procurement%20as%20a%20percentage,of%20works%2C%20goods%20and%20services

3. Nag B (2015) Combating corruption in Indian public procurement—some exploratory case studies. J Inst Public Enterp 1–34
4. Rodríguez MJ, Montequín VR, Fernández FO, Balsera JM (2019) Public procurement announcements in Spain: regulations, data analysis, and award price estimator using machine learning. Complexity 1–20

Chapter 9
Blockchain Technology and Its Application in 3D Parts Procurement: A Case Study

Chiranjib Biswas, Uday Venkatadri, and Cenk Şahin

Abstract In this paper, we present a review of Blockchain based technologies and their application in supply chain procurement. The blockchain has two main characteristics that may be leveraged for adoption in supply chain networks and logistics. First, the exchange of information in blockchain is real time, secure, verifiable, trustable and these are accessible to all members of the network. 3D Printing is an excellent application for Blockchain based procurement. In Blockchain based 3D printing, the aim is to secure contracts from suppliers that may be used in assemblies or for spare parts. An OEM can issue a call for proposals in real time from 3D printing suppliers. The bids can be received securely using the blockchain with partial/complete order quantities. In this paper, we show how the blockchain can be used to integrate the four stages of 3D printing procurement: Scan and Design, Build and Monitor, Test and Validate, and Deliver and Manage. Finally, we present a case study for scheduling 3D printers in manufacturing network. We examine a 3D printer scheduling problem by considering multiple printing alternatives in which various types of products can be printed in different forms using Mixed-Integer linear programming (MILP) under scenario-based experiments.

Keywords Block chain technology · 3D printing · Procurement · Supply chain

C. Biswas · U. Venkatadri (✉)
Dalhousie University, 5269 Morris Street, Halifax, NS B3H-4R2, Canada
e-mail: uday.venkatadri@dal.ca

C. Biswas
e-mail: chiranjib.biswas@dal.ca

C. Şahin
Cukurova University, Balcali Mahallesi, Saricam, Adana 01250, Turkey
e-mail: cenksahin@cu.edu.tr

M. K. Tiwari et al. (eds.), *Applications of Emerging Technologies and AI/ML Algorithms*, Asset Analytics, https://doi.org/10.1007/978-981-99-1019-9_9

9.1 Introduction

3D printers have ushered in new opportunities in manufacturing. Over the last few decades, technology savvy companies have been using 3D printers to produce prototypes. However, in recent years, the cost of these printers has reduced drastically, thanks to the expiration of a few key patents used in this technology. This has helped the exponential growth and use of 3D printers. 3D printers are now being used to manufacture spare parts, mission critical components, and to create an xtremely agile supply chain, thanks to the ability to print these parts on site, where they are needed. The entire printing can take place in proximity to customers, reducing transportation costs and lead time. But for the technology to gain widespread acceptance, there must be seamless integration between the OEM and 3D printing service providers so that printing files can be securely transmitted online to printers. There is a need to ensure that the quantity and quality of the products printed exactly match the contracts the OEM, 3D printing service providers and customers have agreed upon. To maintain the authenticity of these products, these need to be certified. To fulfill all the services associated with 3D printing described above, there must be a right technological platform which will support the end-to-end supply chain and procurement process. The 3D printing process needs to be executed within a highly secured and verifiable environment. The blockchain is a disruptive technology which can be explored to this end because of its indelible, immutable, and tamper proof attributes.

In this paper, we show how the blockchain can be used to integrate the four stages of 3D printing procurement: Scan and Design, Build and Monitor, Test and Validate, and Deliver and Manage. Finally, we present a case study for scheduling 3D printers in a manufacturing network. The literature presents the problem of scheduling 3D printers assuming identical machines. Products can be printed in whole or in parts. These parts which can be assembled later usually have different shapes. Due to the decentralization of 3D printers, thanks to blockchain technology, different products can be manufactured in different locations, close to customers or even end users. These products can even be manufactured in parts instead of the total product and assembled on site.

The remainder of the paper is organized as follows: in Sect. 9.2, we present a literature review of blockchain technologies with a focus in supply chain procurement application. In Sect. 9.3, we present additive manufacture networks and show how block chain technology can be exploited to drive procurement through secure contracts in this industry. In Sect. 9.4, we present a case study for additive manufacturing parts and component scheduling that can potentially work within a block chain framework. In this section, we look at the problem of scheduling 3D printers using mixed-integer linear programming under scenario-based experiments to evaluate multiple printing alternatives. Section 9.5 concludes the paper and offers directions for future research.

9.2 Literature Review of Blockchain with Application in Supply Chain

Blockchain is a distributed ledger technology wherein transaction between users (who are part of the network) are stored in a secure, verifiable, and permanent way. All these data are saved cryptographically inside blocks and these blocks are arranged hierarchically as per sequence. Thus, an endless chain of data blocks is created. It allows all the transactions be traceable and verifiable at any time by the participating users in the network. Once a transaction is certified and saved in chain, it can neither be tampered with nor modified. The process of adding a new transaction to the blockchain involves relaying the new transaction to all the stakeholders in the network for review and audit. If a majority of stakeholders consent to this transaction, then this addition is made to the chain, as a new block, in accordance with the consensus agreement (Fig. 9.1). All these stakeholders are connected through computers in the network, each computer is called a node. Security is ensured by storing a record of the transaction across multiple distributed nodes. As the name suggests, smart contracts are a key element of the blockchain, capable of making sure that credible transactions take place without the involvement of third parties.

Blockchaintechnology has two main characteristics which can be leveraged for adoption in supply chain /supply networks and logistics [3]:

- The exchange of information in the blockchain is real time, secured, verifiable and trustworthy. They are accessible to all members of the network or to others (depending on security privileges and the type of blockchain).
- When defined requirements are met, automatic verification and execution of agreed transactions through smart contracts become a possibility.

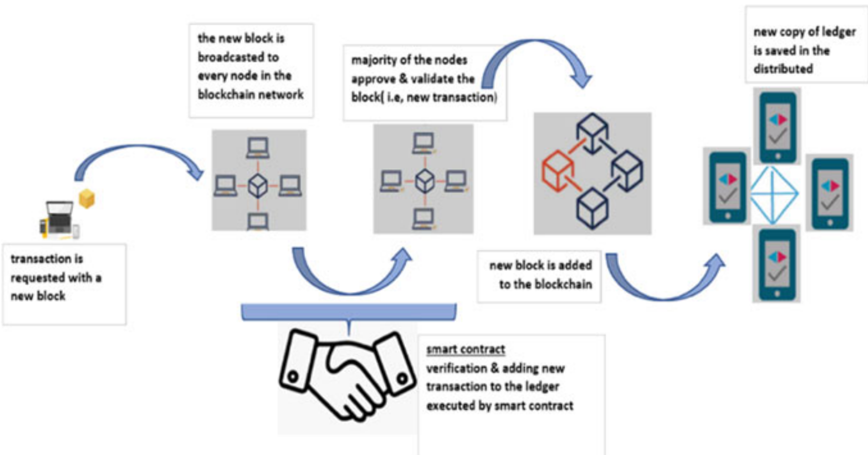


Fig. 9.1 Steps involved in a BC transaction adapted from [1] and [2]

In supply chain & procurement, block chains have been proposed to be used in the following areas: tracking goods (traceability), goods flow (visibility), forecasting, open access to information who are part of the supply chain network, prevention of fraud as the blocks are indelible, and automated transactions through smart contract. This makes the entire supply chain sustainable [3].

9.3 Additive Manufacturing & Blockchain-Based AM Decentralized Supply Networks

Additive manufacturing (AM) is also known as 3D printing or additive layer manufacturing (ALM); it is an industrial production term for a computer-controlled process that deposits material in layers to create three-dimensional objects. Starting from the 3D computer aided design (CAD) model, a layered digital version of the model is created by software integrated into the machine control system, or through online services. In the subsequent deposition or sintering of the material, the pattern of the resulting layers guides the printer [3].

Decentralized AM supply networks incorporating blockchain technology provide an array of benefits, including shorter lead times, savings in transportation costs, reduced inventory cost and increased transparency and communication [5]. AM has four phases which blockchain technology can be implemented into: Scan and Design, Build and Monitor, Test and Validate, and Deliver and Manage [4].

The Scan and Design phase consists of several hand offs, for example, from a CAD system to tools analysis and from there on, to production software. This happens in an environment where many engineers could be working on different components, which can be a significantly complex process. Blockchain facilitates append only timestamp tracking of these changes, which can be distributed over multiple participating departments/organizations accounting of all the changes in design acting as a ledger. This phase of digital thread for Additive Manufacturing (DTAM) comes with several challenges related to file formats and data standardization. While Blockchain does not provide any explicit solution to these challenges, the blockchain can successfully manage to create a format-agnostic accounting of changes to files, aid in performing analyses, and transmit information from software tool to software tool. Blockchain provides an opportunity to track file signatures across platforms in this design and analysis process [4] to ensure authenticity, traceability, and digital security.

The Build and Monitor phase transforms a digital model (from the previous phase) into a physical component. This phase deals with large amounts of data generated during the build process, for example, while capturing sensor data during a build. These data are extremely important while certifying components. This phase involves challenges stemming from the potential distributed nature of AM across a supply chain involving builds in multiple locations. These need adequate systems and infrastructure for tracking, controlling, process feedback, and collecting data. Blockchain can be a potential solution to these challenges. From the perspective of AM being distributed for the point of manufacturing, Blockchain can serve as a tamper-resistant transaction layer, which is stored in a decentralized and distributed manner participating stakeholders. In particular, for complex component manufacturing that needs audit trails for certification, such as the aerospace industry, Blockchain can become extremely valuable [4].

In the Test and Validate phase, records are tied up with digital models of serialized unit component. These records are created during the inspection process. They are vital from for quality assurance, certification, process tuning and continuous improvement. A full digital representation of the component and its test and validation reports are important. Blockchain facilitates maintenance of individual component history, which includes transaction records of test data sets [4].

In the Deliver and Manage phase, products are delivered for their intended use. This creates opportunity to collect data relating to a component's performance in real life, by integrating sensors and monitors and continuously feeding back data to component design. This is a classic example of using blockchain with integrated IoT connected devices [4].

Figure 9.2 schematizes the framework of our proposed Blockchain-Based AM decentralized procurement network technology and shows how the blockchain can be used to integrate the four stages of 3D printing procurement. The parties here involve an OEM, an Engineering Service Provider, an AM service provider and a Life Cycle Agent to sustainably dispose of products. In the proposed decentralized blockchain-based AM supply network, public and private information are stored in each block together with corresponding documents, which are then stored in the cloud repository.

In the first step of the proposed Blockchain-Based AM decentralized supply networks, the OEM shares files of assembly components with descriptions and gets the design from an engineering service provider (Step 1 and Step 2). OEM uploads CAD design of the component in blockchain, seeking quotations from additive manufacturing (AM) service providers (Step 3). At this point, blockchain calculates the capacity available with suppliers and accordingly selects the eligible suppliers who

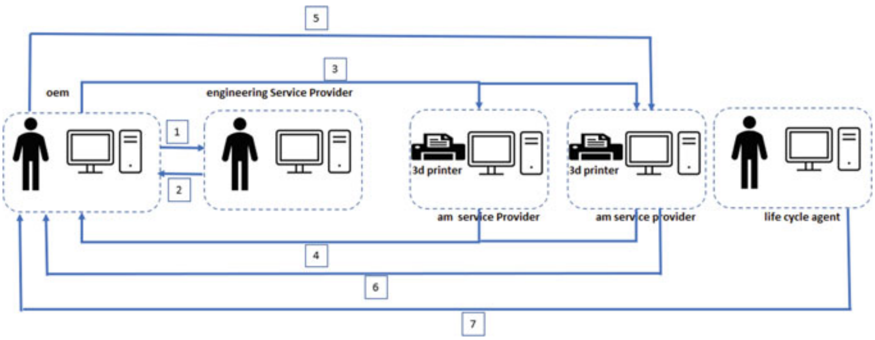


Fig. 9.2 Proposed Blockchain-Based AM decentralized supply networks (adapted from [6])

in turn get access to view the RFP. AM Service providers reply with bids for the RFP, which includes process time, delivery time, cost, alternative process architecture etc. (Step 4). Smart contract capability in the block chain helps the OEM decide the best proposals based on several parameters such as cost, delivery, reliability, quality, and location, especially if the OEM needs to deliver the component to its customer who is situated far away from its normal operational location. After the selection is over, OEM transfers the design files in a secured environment to the selected AM service provider, including license data, such as print count, technical specifications, delivery due date and receiver keys (Step 5). At the AM service provider's end, printer set up, printing and post printing data are captured in the blockchain. The AM service provider sends details of an order dispatch transaction to OEMs, which include information regarding package weight and dimensions, assembly details, material details, and recycling or dismantling instructions (Step 6). At the end of the product life cycle, the transaction requests instructions for disassembling and recycling the product in such a way that it can be recycled or disposed of sustainably (Step 7). In the above-described process, production is heavily reliant on knowledge, communication, and information and not necessarily on material characteristics [7]. Customers track the detailed information of products which in turn increase their trust attached with product [8].

9.4 Scheduling 3D Printers

Due to the decentralization of 3D printers, thanks to blockchain technology, different products can be manufactured in different locations, close to the consumers or even end users. Even these products can be manufactured in parts instead of the total product in one single set up. As a starting point of this study, we focused on a 3D printer scheduling problem by considering multiple printing options, which opens the possibility of printing a variety of types of products in different formats. When printing products with 3D printers, products can either be printed as full products or as smaller constituent parts of larger products. Smaller constituent parts are of varied shapes and can be assembled later to produce the larger products. Therefore, we can say that each product can be printed differently via a 3D printer because of the different printing options.

9.4.1 Mathematical Model

This paper considers the problem of finding an optimal job assignment in terms of parts to be produced to 3D printers considering printing alternatives and having constraints on availability of printers due to scheduling efficiency. Kim et al. (2017) [9] developed a mixed-integer linear programming (MILP) model for this problem and used a metaheuristic approach to solve it. The objective is to minimize the period of time it takes to complete all processing operations, i.e., the makespan. Production is assumed to take place in a fixed time period.

Suppose there are M identical 3D printers with same speed and f number of products to be printed from M printers. For each product, there are l alternatives. A product can be manufactured as a whole part or in different parts. Based on the alternatives, the parts that need to be produced are determined. A 3D printer is assumed to be able to process one item at a time with non-preemptive processing. Additionally, setup times are considered to be independent of sequence and reflected in the printing time. The notation for the parameters and decision variables used in the model is as follows:

Indices	
f	index of products
i	index of machines
j	index of parts
k	index of order
l	index of printing alternatives
Parameters	
a_{flj}	adjacency matrix representing the parts to be manufactured Depending on products and their printing alternatives
p_j	printing time of part j
Decision variables	
C_{max}	makespan
x_{ijk}	$\begin{cases} 1, & \text{If part } j \text{ is processed on machine } i \text{ for the } kth \text{ order} \\ 0, & \text{otherwise} \end{cases}$
y_{fl}	$\begin{cases} 1, & \text{If product } f \text{ takes production alternative } l \\ 0, & \text{otherwise} \end{cases}$

Table 9.1 is the adjacency matrix showing two products with three alternatives each. Alternative 1 represents the whole part, alternative 2 comprises two consecutive parts and alternative 3 has three consecutive parts. In the table, if product 2 is produced with alternative 3, part 10, 11 and 12 should be produced because $\boldsymbol{a_{2310}}$, $\boldsymbol{a_{2311}}$ and $\boldsymbol{a_{2312}}$ are 1.

The first print alternative ($l = 1$) for each product consists of only one part and requires the shortest total print time. Based on this notation, the mathematical model is as follows:

$$Minimize\, C_{\max} \tag{9.1}$$

Subject to:

$$\sum_i \sum_k x_{ijk} \geq \sum_l y_{fl} a_{flj} \forall f, l \tag{9.2}$$

Table 9.1 Adjacency matrix with two products having three alternatives

Product	Alternative	Part											
		1	2	3	4	5	6	7	8	9	10	11	12
1	1	1	0	0	0	0	0	0	0	0	0	0	0
	2	0	1	1	0	0	0	0	0	0	0	0	0
	3	0	0	0	1	1	1	0	0	0	0	0	0
2	1	0	0	0	0	0	0	1	0	0	0	0	0
	2	0	0	0	0	0	0	0	1	1	0	0	0
	3	0	0	0	0	0	0	0	0	0	1	1	1

$$\sum_j x_{ijk} \leq 1 \, \forall i, k \tag{9.3}$$

$$\sum_k \sum_j p_j x_{ijk} \leq C_{\max} \, \forall i \tag{9.4}$$

$$\sum_l y_{fl} = 1 \, \forall f \tag{9.5}$$

$$x_{ijk} \in \{0, 1\} \forall i, j, k, y_{fl} \in \{0, 1\} \, \forall f, l \tag{9.6}$$

The objective function (1) aims to minimize the completion time of the last part. Constraint (2) ensures that once a printing alternative is selected for product f, the related parts are produced. Constraint (3) is imposed to ensure that only one production alternative for each product will be chosen. Constraint (4) requires the total printing time for all parts processed by machines to be less than the last part's completion time. Each printer can print one part at a time as indicated by constraint (5). Finally, constraints (6) express the domain of the variables.

9.5 A Case Study and Results

Numerical experiments are carried out for combinations of 4, 5 and 10 machines, 10 and 20 products and 3, 5 alternatives. A product is split into parts according to the number of production alternatives. For example, when the number of production alternatives for a product is 2, there are 2 options: a product is produced as a whole structure or produced as 2 parts. The experiments were conducted under the assumption that the total printing time increases randomly between 5 and 15% as the number of parts consisting of a product increase. For production alternative 1, the printing time is generated randomly within a range of 10 s to 300 s, and for parts separated by alternative, it is also generated randomly. The printing time (p_j) is given in Table 9.2. This problem is solved by Kim et al. (2017) [9] using the genetic algorithm (GA). In our approach, the MILP model was solved using the problem set (Table 9.2) under different scenario-based experiments. The MILP model was coded in GUSEK, a well known open source software available under the GNU General Public License and then converted into a.lp file which was subsequently run in Gurobi © 9.5.0 on a computer running Windows 11 on an Intel(R) Core(TM) i5 CPU 650 @ 3.20GHz 3 with 4.00 GB RAM. Table 9.3 shows the results of the experiments.

We can see from the experiments that, as we increase the printing time (p_j), the makespan ($C_{\max}$) also increases. If we reduce the number of machines, makespan ($C_{\max}$) increases. Conversely, when we increase the number of machines, makespan time ($C_{\max}$) decreases. Clearly, the approach seems viable for problems of these sizes, i.e., 20 products (f) with about 5 alternatives (l).

Table 9.2 The printing times associated with the production alternatives for 20 products

f	Printing time (p_j)										
	1	2 Parts		3 Parts			5 Parts				
1	261	294	6	80.9	140.2	111.9	91.6	75.8	21.8	92.7	71.1
2	12	1	13	0.008	9.9	5.1	2.9	3.1	0.1	3.7	7.2
3	82	18	68	40.3	26.2	29.5	18.9	43.6	14.1	19.7	6.7
4	250	32	236	103.3	98.3	106.3	69.5	92.0	71.2	65.6	42.7
5	225	150	93	48.8	115.4	100.8	70.4	23.4	101.2	7.0	95.0
6	38	12	29	18.1	8.8	19.1	4.2	11.9	17.1	12.9	5.9
7	85	89	9	43.3	0.0	64.7	28.2	35.7	3.9	53.5	2.8
8	128	11	131	69.2	62.6	30.2	45.9	40.6	0.2	46.6	41.7
9	75	9	74	33.1	24.0	34.9	8.4	2.4	42.4	42.9	4.9
10	244	46	227	131.3	46.6	112.1	90.2	54.1	17.9	93.4	66.4
11	105	23	87	7.7	48.8	69.5	20.7	3.1	0.6	109.8	3.7
12	254	221	46	15.4	119.4	150.2	54.7	96.3	93.2	31.7	29.2
13	28	14	17	22.5	8.2	3.3	6.8	0.4	8.1	16.3	4.3
14	160	83	88	76.3	16.3	88.5	73.4	48.6	45.6	1.1	32.3
15	113	110	14	47.5	68.5	22.0	5.6	27.6	45.8	24.4	51.7
16	118	35	94	59.7	19.2	68.1	10.0	42.6	32.0	46.2	36.2
17	236	180	89	21.9	189.5	95.6	7.5	49.8	107.7	82.2	80.8
18	115	50	82	3.2	59.9	83.9	43.3	27.7	16.6	38.6	36.7
19	223	200	45	119.4	119.8	27.8	103.3	68.0	16.9	84.7	21.2
20	297	111	207	95.3	124.3	139.4	94.6	58.8	109.4	96.5	21.8

Table 9.3 Scenario-based Experiments Results

Experiment	Machine (i)	Product (f)	Alternative (l)	$C_{\max}$ (sec)	Part (j)	Description
1	5	10	3	286.1	15	
2	5	10	3	314.8	16	10% increment in printing time (p_j)
3	5	10	3	372	15	30% increment in printing time (p_j)
4	4	10	3	353	14	No# of machines reduced by 1, with respect to experiment #1
5	10	10	3	167.5	26	No # of machines doubled, with respect to experiment #1
6	5	20	3	610	20	
7	5	20	5			Unsolved in Gurobi after an hour, the MIPGap was 0.14

9.6 Conclusions

In this paper, we discussed how block chain in supply chain management is real time, secure, verifiable and trustworthy for all members of the network, We then discussed how it can be used in the development of smart contracts and end to end data acquisition and storage for traceability purposes through the life cycle of products.

We also showed how the blockchain can be used to integrate the four stages of 3D printing procurement: Scan and Design, Build and Monitor, Test and Validate, and Deliver and Manage.

We presented a case study for scheduling 3D printers in a manufacturing network. The literature presents the problem of scheduling 3D printers assuming identical machines. Products can be printed in whole or in parts. Due to the decentralization of 3D printers, thanks to blockchain technology, different products can be manufactured in different locations, close to customers or even end users. We presented some feasibility results using MILP.

In the future, this type of problem can be extended to take both the OEM and supplier's perspectives depending on fully open (trust), fully confidential (no trust), and intermediate network settings. Using a negotiation-based approach, a Multi Agent System operating blockchain technology where the bids can be received securely can be developed and employed for simulation analysis to explore the further benefits of cooperation at the perspectives of both the OEM and supplier in the AM industry.

References

1. Saberi S, Kouhizadeh M, Sarkis J, Shen L (2019) Blockchain technology and its relationships to sustainable supply chain management. Int J Prod Res 57(7):2117–2135
2. Crosby M, Pattanayak P, Verma S, Kalyanaraman V (2016) Blockchain technology: beyond bitcoin. Appl Innov 2:6–9
3. Dujak D, Sajter D (2019) Blockchain applications in supply chain. In: Kawa A, Maryniak A (eds) SMART supply network. EcoProduction. Springer, Cham. https://doi.org/10.1007/978-3-319-91668-2_2
4. Mandolla C, Petruzzelli AM, Percoco G, Urbinati A (2019) Building a digital twin for additive manufacturing through the exploitation of blockchain: a case analysis of the aircraft industry. Comput Ind 109:134–152
5. Alkhader W, Albaaki N, Salah K, Jayaraman R, Arshad J, Omar M (2020) Blockchain-based traceability and management for additive manufacturing. IEEE-Access 8:188363–188377
6. Papakostas N, Newell A, Hargaden V (2019) A novel paradigm for managing the product development process utilising blockchain technology principles. CIRP Ann Manuf Technol 68(1):137–140
7. Pazaitis A, De Filippi P, Kostakis V (2017) Blockchain and value systems in the sharing economy: the illustrative case of backfeed. Technol Forecast Soc Chang 125:105–115
8. Tian F (2016) An agri-food supply chain traceability system for China based on RFID & blockchain technology. In: 13th international conference on service systems and service management (ICSSSM)
9. Kim J, Park S, Kim H (2017) Scheduling 3D printers with multiple printing alternatives. In: 13th IEEE conference on automation science and engineering (CASE)

Chapter 10
A Mathematical Model-Based Heuristic for Clustering, Logistics, and Order Pickup in the Constrained In-Bound Multi-Period Multi-Part Inventory Routing Problem with Heterogeneous Vehicles

Anushee Jain and Chandrasekharan Rajendran

Abstract A typical problem faced in any supply chain is the simultaneous planning of inventory and logistics. In this work, we consider a representative problem of clustering, lot size optimization, and transport planning faced in many logistics industries. It is a problem of an inbound logistics system for which demand is generated through the outbound logistics. The demand for parts for a finite time period is given and is deterministic. The demand occurring at a warehouse is met by procuring raw material from the suppliers that are spread out geographically. The order pickup is done using heterogeneous trucks which start from the warehouse, pick up the order from the supplier, and return to the warehouse. The entire trip should be completed within the given lead time without any permissible backorder or delay. The problem can be described as Constrained In-Bound Multi-Period Multi-Part Inventory Routing Problem with Heterogeneous Vehicles (CIBIRPHV). The problem is distinct and different from the Inventory Routing Problem present in the literature. The objective is to minimize the total costs, comprising of transportation costs, inventory carrying costs, and warehouse storage costs. We propose a mathematical model to first cluster the suppliers, followed by another mathematical model for order planning and pickup.

A. Jain · C. Rajendran (✉)
Indian Institute of Technology Madras, Chennai, India
e-mail: craj@iitm.ac.in

A. Jain
e-mail: anushee27@gmail.com

M. K. Tiwari et al. (eds.), *Applications of Emerging Technologies and AI/ML Algorithms*, Asset Analytics, https://doi.org/10.1007/978-981-99-1019-9_10

10.1 Introduction

Procurement in a supply chain is an important function which refers to all the tasks involved in obtaining the optimal product from the optimal vendor on optimal terms. Procurement covers the following list of tasks: purchase planning, determining the standard of quality, identifying suitable suppliers, price negotiation, financing purchases, acquiring required goods and services and inventory control, and disposal of waste production. Hence, Supply Chain Management is considered one of the many responsibilities faced by a procurement function. Procurement logistics involves the flow of raw material and parts required for manufacturing which are procured from suppliers. The shift to Just-in-Time production has led to many companies carrying out production by procuring the required materials in only the required amount at the required times, which also leads to reducing inventory costs. In order to meet customer demands, prevent stock-outs, and reduce bottlenecks in production, manufacturers must ensure appropriate quantities of storage and must ensure optimal distribution and procurement plans that include routing of the vehicles. This coordination of inventory management and transportation has given rise to a class of problems known as Inventory Routing Problem.

10.1.1 Inventory Routing Problem

Inventory Routing Problem (IRP) is a class of problems arising from the combination of inventory management and routing [8, 15]. IRP seeks to find a trade-off between keeping inventories as low as possible without affecting the transportation costs. IRPs can be found with both in-bound and out-bound distribution. In-bound IRPs are common in industries that have a large number of low to medium volume suppliers where consolidation of order quantity provides significant savings in the total cost. Out-bound IRPs are most common in the Vendor Managed Inventory (VMI) setting where a manufacturer/distributer manages the inventory at the retailer. The inventory position and demand are known across the different entities in the supply chain. This information sharing leads to improved efficiency in supply chain operations. In general, IRP comprises procuring/distributing goods via pre-determined routes to a group of customers, each of whom has varied daily demand, as well as inventory management to ensure that no customer runs out of stock at any given time, at minimum total cost. Applications of IRP in real-life systems have been found in the distribution of ammonia and groceries [1, 9] studied the problem of periodic pickup of auto parts. In this paper, we discuss a representative problem in a manufacturing company. There are various suppliers scattered geographically. The suppliers are managed centrally by a warehouse, and all decisions related to planning and routing are made at the warehouse. The aim is to develop a lot size optimization tool, dynamic order placement, and transport planning, to optimize various supply chain costs, namely, Inventory carrying cost, Transportation cost, Warehouse space cost, and recommend order quantity/lot size and order frequency coupled with vehicle movement and transport planning having the least landed cost. We develop a mathematical model for the same, respecting the business constraints.

10.2 Literature Review

In the literature review, we discuss a few studies related to different variants of IRP. [8] defined Inventory Routing Problem as,

"a problem that involves a set of customers, where each customer has a different demand each day…The objective is to minimize the annual delivery costs, while attempting to ensure that no customer runs out of the commodity at any time."

The name 'Inventory Routing Problem' (IRP) has synonymously been used for combined inventory management and routing problems in various studies. The literature on IRP can be categorized based on different criteria [5]: time horizon (finite, infinite), structure (one to one, one to many, many to many), routing (direct, multiple, and continuous), inventory policy (maximum level (ML), order up to level (OU)), inventory decisions (lost sales, backorder, non-negative), fleet composition (homogeneous, heterogeneous), fleet size (single, multiple, unconstrained). General Integrated Transportation Inventory models (GITI) have been used for making strategic decisions for long-term planning compared to specific models which are used for decisions pertaining to medium time horizons [12].

Another scheme in which the papers can be classified is related to the availability of information on demand (deterministic or stochastic). In this paper, IRPs with deterministic known demand have been discussed.

10.2.1 Literature Review on Multi-Product Multi-Period IRP

Multi-product and multi-period IRPs very closely resemble real-life problems found in many industries. Most of the MILP models for multi-product IRPs found in the literature are np-hard and can handle only small instances. Speranza & Ukovich, (1994) proposed integer (IP) and MILP for the problem of determining frequencies at which multiple products for a single customer need to be delivered on a common link to minimize the sum of transportation and inventory costs [3]. Extended this problem for multiple products and multiple customers. The heuristics proposed in the literature for multi-period IRP seem to decompose the problem into hierarchical sub-problems, with the result of each sub-problem used as a starting point for the next one [4]. Developed a two-phase heuristic to solve the allocation-routing problem arising in the grocery industries with backorder costs involved and with a limited fleet of homogeneous vehicles.

Studied an IRP with many to one structure, multiple products, and multi-periods in a finite time horizon [11]. They proposed mathematical models to get lower and upper bounds followed by Genetic Algorithms, which gave improved upper bounds [10]. Improved their solutions by developing a variable neighborhood search heuristic.

Studied the many-to-many case and proposed a MILP formulation for a multi-product multi-depot IRP [7, 14]. Developed a branch and cut algorithm for the multi-item multi-vehicle version of the problem. They also incorporated several features like partial driving consistency by allowing some of the customers to be served by more than one truck driver at a time. An exact MILP was also proposed by Coelho & Laporte, (2013b) to solve a multi-vehicle multi-product multi-vehicle problem.

10.2.2 Complexity of IRP

The complexity of IRP arises from the fact that multiple decisions related to order planning and routing need to be taken simultaneously. This leads to numerous feasible solutions and solving for them becomes computationally difficult [13].

10.2.3 Research Gaps Identified from the Literature Review

- Most of the existing literature studies discuss IRPs with out-bound distribution.
- "Literature on handling multiple products by a central warehouse is less." [14].
- No study seems to consider the warehouse as a part of the cluster of suppliers/ customers.
- No study seems to give focus to the clusters formed by the consolidation of deliveries during the milk-run.
- No prior study seems to consider the case of minimum order quantity (MOQ) and order quantity in the multiples of lot size required by any supplier/customer in the order planning phase. Packaging order quantity in multiples of lot size leads to optimization of the warehouse storage space.

10.3 Proposed Methodology for CIBIRPHV

10.3.1 An Introduction to the Problem Statement

We consider a two-echelon supply chain consisting of N suppliers geographically dispersed and managed centrally by a warehouse W. A set of mutually exclusive parts $P = \{P_1, P_2, .., Pp\}$ needs to be procured from each supplier. The exclusive set of parts $P(s)$, $s \in \{1, 2, .., N\}$, is a subset of the parts supplied by supplier N from the whole product set P, with $\bigcap_{s=1}^{N} P(s) = 0$ i.e., a part supplied by a supplier is not supplied by any other supplier.

The problem consists of designing an order and routing plan for procuring multiple parts from a number of suppliers and delivers them at the warehouse. The demand of parts is deterministic but dynamic in nature and varies daily. There is no shortage or backorder allowed. The two types of vehicles involved in the problem consist of truck type 1 which provides direct pickup (LTL type) and truck type 2 (TL) which provides consolidation of deliveries using milk-run. A TL truck can visit a maximum of, say, 3 suppliers at a time and from the first pickup supplier visited, the other suppliers must be within a radius of, say, 200 kms. Each type of truck can travel a maximum distance of, say, 450 kms per day. A supplier can be visited by either LTL truck or TL truck, but not both on a given day. The trucks must start at the warehouse, traverse through the suppliers, and come back at the warehouse within a lead time of, say, 4 days. There is no restriction on the number of vehicles to be used at a given point of time. The problem is referred to as constrained in-bound multi-period multi-part Inventory Routing Problem with heterogeneous vehicles (CIBIRPHV) in the study. This variant of IRP is motivated by the real-life problem existing in many industries. The representation of the methodology has been done using a hypothetical dataset. The uniqueness of the problem lies in its consideration of many practical constraints, e.g., constraints on the distance ravelled, minimum order quantity, order quantity in multiples of lot size, etc., which are relevant to industries. A three-phase heuristic has been proposed in the study which decomposes the problem into inventory management, allocation followed by routing. A mathematical model (MIP) for clustering is proposed in the study. It is done to divide the supplier data into small and dense clusters for TSP. A MILP model is then proposed for the order pickup and transport planning which is executed on the sub-clusters obtained from the previous phase. In the third phase, routes are defined and routing cost is calculated. The proposed heuristic focuses to deal with the lot size optimization and minimization of overall cost, e.g., transportation, inventory, and warehouse storage cost in the supply chain. A numerical illustration is used to show the proposed model at the end of the chapter.

One of the variants of this one-warehouse, multiple supplier problems is when there is more than one warehouse managing the suppliers. The multi-warehouse problem is computationally expensive in nature and requires the one-warehouse multiple supplier problem to be solved individually for each of the warehouses [14].

10.3.2 The Proposed Methodology

We propose a three-phase approach to solve this problem. We divide the entire time horizon into 14-day time periods or cycles.

1. Use of clustering techniques to break the problem into small clusters: Distance-based clustering of suppliers is done to divide the large dataset into small clusters. This is done for computational feasibility and to get dense clusters for TSP.
2. Solving for order pickup quantity and transport planning for each sub-cluster separately using the MILP model: The solution to the MILP model gives the

pickup quantity, order pickup schedule, and allocations of suppliers and trucks as the result.

3. Design of actual routes for the allocated suppliers and costs adjustment: The approach is similar to the continuous review systems in the inventory management, where the inventory is replenished when the items reach a particular inventory position. In our case, the orders are picked up whenever there is a demand for the part; but the frequency of picking up the order varies.

Phase 1: Clustering of suppliers. The need for a separate model for clustering arises due to the complex nature of the problem. The multi-period multi-part IRP with heterogeneous vehicles having various business requirements is computationally difficult to solve in a single step. The proposed MIP model for clustering is used as a preprocessing step to split our data into small clusters before the actual execution of the MILP model for order pickup and transport planning. The proposed MIP model for clustering considers a special case, with every cluster having the warehouse as its member. It is done to ensure that logical clusters formed are oriented toward the warehouse, and any feasible clustering far away is not considered. We also add a constraint to limit the number of parts in a cluster as limiting both parts and suppliers yielded infeasible solutions. Primarily, we follow the K-medoid approach.

Proposed Model for clustering

Parameters

N	Number of suppliers
C	Number of clusters (experimented with different values of C to get a feasible solution)
$dist_{i,j}$	Distance between supplier i and supplier j.
$dist_{w,j}$	Distance between warehouse (w) and supplier j.
$dist_{j,w}$	Distance between supplier j and warehouse (w)
$parts_i$	Number of parts supplied by supplier i.
P_{max}	Maximum parts allowed in a cluster

Decision Variables

x_{jj} Binary variable which takes value 1, if supplier j acts as a focal node (i.e., as a medoid), 0 otherwise.

x_{ij} Binary variable which takes value 1, if supplier i is associated with focal node supplier j, 0 otherwise.

Objective

$$\text{Minimize Z} = \sum_{i=1}^{N}\sum_{j=1}^{N}\left(dist_{ij} \times x_{ij}\right) + \sum_{j=1}^{N}\left(\left(dist_{w,j} + dist_{j,w}\right) \times x_{jj}\right) \tag{10.1}$$

subject to the following constraints.

$$\sum_{j=1}^{N} x_{ij} = 1 \forall i \in 1, .., N \tag{10.2}$$

$$\sum_{j=1}^{N} x_{jj} = C \forall j \in 1, .., N. \tag{10.3}$$

$$x_{ij} \leq x_{jj} \forall j \in 1, .., N and i \in 1, .., N. \tag{10.4}$$

$$\sum_{i=1}^{N} (parts_i \times x_{ij}) \leq P_{max} \times x_{jj} \forall j \in 1, .., N. \tag{10.5}$$

The first term in the objective function minimizes the distance between the focal node j and supplier i. The focal node j is used to compute the total distance as a surrogate for density and hence the TSP distance. The second term in the objective function tries to bring the focal nodes closer to the warehouse, i.e., minimize the distance between the focal node and the warehouse. This helps to form logical clusters that are centered around the warehouse. Constraint (10.2) restricts that a supplier can be assigned to only one focal node j. Constraint (10.3) states that the number of clusters is equal to C. Constraint (10.4) states that a supplier i can be assigned to supplier j only when j acts as a focal node. Constraint (10.5) is to ensure that the maximum number of parts in a cluster does not exceed P_{max}. This is done for the computational tractability of the MILP model for order pickup and transport planning.

Phase 2: The MILP model for Order pickup and Transport Planning. After splitting the data into smaller clusters, we develop a mathematical model for the order pickup and transport planning problem. We split the entire time horizon of demand into, say, 14 days of time period or cycles. For the purpose of presentation, we consider a demand up to 28 days with 2 cycles of 14 days each in our study.

The following assumptions are considered in our model.

- Transportation Cost is proportional to distance.
- Vehicles are heterogeneous.
- A lead time of 4 days is permissible.
- Order quantity calculations are done from 1 to t days. Demand for the first 4 days is taken as 0 ; $D_{p,s,\tau} = 0 \forall \tau = 1, 2, 3,$ and 4 (since our demand realization is assumed to start from $t = 5$ day, in order to receive the order quantity at the beginning of $t = 5$th day by accounting for the transportation lead time).
- $D_{p,s,\tau}$ is realized at the beginning of the day. The order quantity, whose order has been placed $\tau - 4$ days before, is received on the beginning of day τ, before the demand for that day is realized. The demand is met out of the net inventory present at the beginning of the day.

- At the beginning of the planning period, there is some inventory available with us, known as beginning inventory (BI). $BI_{p,s,1}$=$SS_{p,s}$ = Average demand for part p over lead time.
- The first pickup supplier visited acts as a medoid in the cluster so that the business constraint of 200 kms radius from the first pickup supplier is satisfied. It is done for the sake of computational simplicity and for the approximate computation of distance. A medoid supplier may vary in different time periods.

  The objective function has the following cost components:

1. Transportation costs between nodes, based on distance and weight, for suppliers;
2. Fixed Charge for the trucks;
3. Inventory Carrying Cost at the warehouse; and
4. Warehouse Storage Cost.

Terminology

p	Index to denote a part
s	Index to denote a supplier
τ	Index to denote a time unit (of one day)
t'	Index to denote a time cycle, $t' \in \{0, 14\}$. When $t' = 0$, the MILP model is run for first 14 days demand
S	Set of suppliers who are supplying various parts in $P(s)$ (number of suppliers = n $\{S\}$ = N)
$P(s)$	Set of products supplied by a supplier (total number of parts = $\sum_{s \epsilon S}$n$\{P(s)\}$.

Parameters

$D_{p,s,\tau}$	Daily demand (number of units) of part p required from supplier s in time period τ
$BI_{p,s,1}$	Beginning Inventory available at the beginning of day 1. It is equal to average demand for part p over one week for $t' = 0$ corresponding to $t' \in \{0, 14\}$.
$SS_{p,s}$	Safety stock of part p supplied by supplier s. It is equal to average demand for part p over one week for $t' = 0$.
$w_{p,s}$	Per unit weight of part p from supplier s
$d_{s,s'}$	Distance between supplier s and the first pickup (notional medoid) supplier s'. From the geo codes provided, distance matrix indicating the distance between the supplier s and s' is obtained
$d'_{w,s'}$	Distance between warehouse w and the first pickup (notional medoid) supplier s'
$d'_{s',w}$	Distance between first pickup (notional medoid) supplier s' and warehouse w
$c1$	Unit transportation cost with respect to truck type 1 (LTL)
$c2$	Unit transportation cost with respect to truck type 2 (TL)

(continued)

(continued)

$D_{p,s,\tau}$	Daily demand (number of units) of part p required from supplier s in time period τ
$C1$	Capacity of truck type 1 (LTL)
$C2$	Capacity of truck type 2 (TL)
$OH1$	Fixed charge of using truck type 1 (LTL)
$OH2$	Fixed charge of using truck type 2 (TL)
M	A very large number (an integer ~ 99,999,999) greater than the truck capacity
T	Total time in days for which the model is run ($T = 28$ days) or 2 cycles of 14 days each (cycles commencing from $t^{'}$+1, with $t^{'} \in \{0, 14\}$)
$MOQp, s$	Minimum Order Quantity to be placed for part p by supplier s
$pallet^{'}_{p,s}$	Multiples of order quantity for part p by supplier s (or Supplier Enforced Order Multiple)
$Vol/unitp, s$	Volume occupied by one unit of part p by supplier s
$HC_{p,s}$	Inventory holding cost per day (on material cost)
$Wh.C$	Warehouse space cost charge
L	Transportation lead time (L= 4 days)

Decision Variables

$Q_{p,s,\tau}$	Quantity (number of units) of part p ordered from supplier s in time period τ.
$Inv_{p,s,t}$	End inventory of part p from supplier s at the end of time t
$BI_{p,s,t'+1}$	Beginning inventory available of part p supplied by supplier s at the beginning of planning period of 14 days corresponding to $t^{'}$ where $t^{'} \in \{0, 14\}$.
$Q^{'}_{s,t}$	Total weight of all products transported from supplier s in time t
$\delta_{s',s',t}$	Binary decision variable to represent if a supplier acts as a first pickup (notional medoid) supplier in time t
$\delta_{s,s',t}$	Binary decision variable to represent if a supplier is associated with first pickup (notional medoid) supplier with respect to its cluster of suppliers in time t
k_p	It is an integer, which indicates multiple of $pallet^{'}_{p,s}$ (Supplier Enforced Order Multiple) of units of part p corresponding to $Q_{p,s,t}$
$del1_{s',t}$	Binary decision variable to represent if truck type 1 (LTL) is chosen for the first pickup (notional medoid) supplier and the set of suppliers associated with the medoid including itself in given time t
$del2_{s',t}$	Binary decision variable to represent if truck type 2 (TL) is chosen for the first pickup (notional medoid) supplier and the set of suppliers associated with the medoid including itself in given time t
$D_{s',t}$	Distance traveled by the truck allocated to the first pickup (notional medoid) supplier in time t. It is a dependent variable, which depends on the value of $\delta_{s',s',t}$ in time t
$Q^{''}_{s,s',t}$	It is a variable which stores the weight of order pickup quantity supplied by supplier s if supplier is associated with a first pickup supplier in time t

(continued)

(continued)

$Q_{p,s,\tau}$	Quantity (number of units) of part p ordered from supplier s in time period τ.
$Q_{s',t}'''$	It is a variable which stores the total weight of parts carried from all suppliers including the first pickup supplier in time t. It is the summation of $Q''_{s,s',t}$ over all the suppliers S
$Cost_{s',t}$	Total Transportation cost of the truck allotted to notional medoid supplier s', in time period t
$Cost1_{s',t}$	Total charge per day incurred on using LTL type of truck in time period t
$Cost2_{s',t}$	Total charge per day incurred on using TL type of truck in time period t
$V_{p,s,t}$	Space required in the warehouse for the end inventory of parts. It is measured in m^3

Objective

Minimize

$$Z = \sum_{t=t'+1}^{t'+14} \sum_{s \in S} (Cost_{s',t}) + \sum_{t=t'+1}^{t'+14} \sum_{s \in S} (Cost1_{s',t})$$
$$+ \sum_{t=t'+1}^{t'+14} \sum_{s \in S} (Cost2_{s',t})$$
$$+ \sum_{t=t'+5}^{t'+18} \sum_{s \in S} \sum_{p \in P(s)} (Inv_{p,s,t} \times HC_{p,s}) + \sum_{t=t'+5}^{t'+18} \sum_{s \in S} \sum_{p \in P(s)} (V_{p,s,t} \times Wh.C) \tag{10.6}$$

subject to all constraints, and

for all, $t' \in \{0, 14\}$ (e.g., when $t' = 0$, the time window is from t=1 day to t=14 days and so on).

Constraint (10.7) ensures that in order to meet the demand in the warehouse (no shortage for the part) the quantity to be ordered for given part p from the given supplier s for time period t should be greater than or equal to the sum of the demand for the same part p from the supplier s over the lead time (L) of 4 days, with Safety Stock ($SS_{p,\mathrm{s}}$) for that part being maintained. Safety stock ($SS_{p,\mathrm{s}}$) is the minimum level of stock which is maintained in the inventory. At the beginning of day 1, i.e., at $t' = 0$, BI is entered by the user as an input. For subsequent periods, e.g., with $t' = 14$, BI on the 15th day is given by the inventory leftover at the end of 14th day.

$$BI_{p,s,t'+1} + \sum_{\tau=t'+1}^{t} Q_{p,s,\tau} \geq \sum_{\tau=t'+5}^{t+4} D_{p,s,\tau} + SS_{p,\mathrm{s}}$$

$$\forall s \in S;\ p \in P(s);\ t \in \left\{t' + 1, t' + 2, \ldots, t' + 14\right\} \quad (10.7)$$

Constraint (10.8) is the inventory balance equation for part p from the supplier s in time period t.

Inventory calculations are done from the (t'+5) day to (t +4) days as demand for the first four days is taken as 0. Inventory costs are also calculated from day (t'+5) to day (t +4) in the objective function.

Inventory at the end of a day is given as follows: Beginning inventory present at the beginning of a given cycle + cumulative order quantity received up to that day–cumulative demand satisfied by the orders received till that day.

$$Inv_{p,s,t} = BI_{p,s,t'+1} + \sum_{\tau=t'+1}^{t-4} Q_{p,s,\tau} - \sum_{\tau=t'+5}^{t} D_{p,s,\tau}$$

$$\forall s \in S;\ p \in P(s);\ t \in \left\{t' + 5, t' + 6, \ldots, t' + 18\right\} \quad (10.8)$$

Constraint (10.9) is used to compute the total weight of all products transported from supplier s in time period t is obtained by the product of the ordered quantity of each part from the supplier and the weight of the part from the supplier.

$$Q'_{s,t} = \sum_{p \in P(s)} \left(Q_{p,s,t} \times w_{p,s}\right)$$

$$\forall s \in S;\ p \in P(s);\ t \in \left\{t' + 1, t' + 2, \ldots, t' + 14\right\} \quad (10.9)$$

Constraints for Supplier Allocation: Constraint (10.10) ensures that a supplier can be assigned to a medoid or first pickup supplier s' only when supplier s' acts as a medoid pickup supplier.

$$\delta_{s,s',t} \le \delta_{s',s',t}$$

$$\forall s \in S;\ p \in P(s);\ t \in \left\{t' + 1, t' + 2, \ldots, t' + 14\right\} \quad (10.10)$$

Constraint (10.11) ensures that a supplier can have at most one medoid pickup supplier in a given time t.

$$\sum_{s' \in S} \delta_{s,s',t} \le 1$$

$$\forall s \in S;\ p \in P(s);\ t \in \left\{t' + 1, t' + 2, \ldots, t' + 14\right\} \quad (10.11)$$

Constraint (10.12) ensures that the supplier can act as a medoid pickup supplier only if there is some order quantity for that supplier in that time period t.

$$\delta_{s',s',t} \leq \sum_{p \in P(s)} Q_{p,s',t}$$

$$\forall s \in S;\ p \in P(s);\ t \in \left\{t^{'}+1, t^{'}+2, \ldots, t^{'}+14\right\} \tag{10.12}$$

Constraint (10.13) ensures that a supplier can be assigned to a medoid pickup supplier if and only if there is some order quantity for that supplier in that time period t.

$$\sum_{s' \in S} \delta_{s,s',t} \leq \sum_{p \in P(s)} Q_{p,s,t}$$

$$\forall s \in S;\ p \in P(s);\ t \in \left\{t^{'}+1, t^{'}+2, \ldots, t^{'}+14\right\} \tag{10.13}$$

Constraint (10.14) forces a supplier to get assigned to a medoid pickup supplier if order quantity of all the parts supplied by the supplier is greater than 0.

$$Q^{'}_{s,t} \leq M \sum_{s' \in S} \delta_{s,s',t} \forall s \in S;\ t \in \left\{t^{'}+1, t^{'}+2, \ldots, t^{'}+14\right\} \tag{10.14}$$

Constraint (10.15) is to ensure Minimum Order Quantity of a part p from a supplier s is given as a parameter so the order quantity for a part p from supplier s should be greater than or equal to the MOQ of that part p from supplier s.

$$Q_{p,s,t} \geq MOQ_{p,s} \times \sum_{s' \in S} \delta_{s,s',t}$$

$$\forall s \in S;\ p \in P(s);\ t \in \left\{t^{'}+1, t^{'}+2, \ldots, t^{'}+14\right\} \tag{10.15}$$

Constraints (10.16) and (10.17) ensure that order quantity $Q_{p,s,t}$ for a part p by supplier s in time t is in multiples of $pallet^{'}_{p,s}$ (which is a parameter) of part p by supplier s, only when $s^{'}$ acts as its first pickup supplier.

k_p is a variable which takes the value of the nearest integer corresponding to the value of $Q_{p,s,t}$.. Constraint (10.18) ensures that order quantity is 0 if $s^{'}$ does not acts as a first pickup supplier. Also, if $s^{'}$ acts as a first pickup supplier, i.e., $\sum_{s' \in S} \delta_{s,s',t} = 1$ then by virtue of constraints (10.16) and (10.17), $Q_{p,s,t}$ is forced to take a positive value ($Q_{p,s,t} > 0$).

Constraint (10.19) ensures that from Constraint (10.18), if order quantity $Q_{p,s,t}$ is 0, and then k_p is 0.

$$Q_{p,s,t} \geq pallet'_{p,s} \times k_p - M\left(1 - \sum_{s' \in S} \delta_{s,s',t}\right) \tag{10.16}$$

$$Q_{p,s,t} \leq pallet'_{p,s} \times k_p + M\left(1 - \sum_{s' \in S} \delta_{s,s',t}\right) \tag{10.17}$$

$$Q_{p,s,t} \leq M \sum_{s' \in S} \delta_{s,s',t} \tag{10.18}$$

$$k_p \leq Q_{p,s,t} \tag{10.19}$$

$$\forall s \in S;\ p \in P(s);\ t \in \left\{t' + 1, t' + 2, \ldots, t' + 14\right\}$$

k_p-multiples of order quantity.

Constraint (10.20) is used to compute distance (approximate) traveled by the truck allocated to a first pickup supplier s' and thereafter other suppliers associated with s' in time period .

$$D_{s',t} = \sum_{\substack{s \in S \\ s' \neq s}} \left(d_{s,s'} \times \delta_{s',s',t}\right) + \left((d'_{w,s'} + d'_{s',w}) \times \delta_{s',s',t}\right)$$

$$\forall s' \in S;\ t \in \left\{t' + 1, t' + 2, \ldots, t' + 14\right\} \tag{10.20}$$

Constraint (10.21) ensures that supplier s associated with its first pickup supplier s' should be within radius of 200 kms.

$$\sum_{\substack{s' \in S \\ s' \neq s}} \left(d_{s,s'} \times \delta_{s,s',t}\right) \leq 200$$

$$\forall s \in S;\ t \in \left\{t' + 1, t' + 2, \ldots, t' + 14\right\} \tag{10.21}$$

Constraint (10.22) ensures that only one type of truck can be assigned to the first pickup supplier in time t. It means in time t, either $del1_{s',t} = 1 or del2_{s',t} = 1$, only when supplier s' acts as a medoid supplier.

$$del1_{s',t} + del2_{s',t} = \delta_{s',s',t}$$

$$\forall s \in S;\ t \in \left\{t' + 1, t' + 2, \ldots, t' + 14\right\} \tag{10.22}$$

Constraint (10.23) is used for the distance traveled by the truck allocated to a first pickup supplier $s^{'}$ in time period t should be less than the maximum distance that can be traveled by the truck in its entire travel. A buffer distance is adjusted with the distance so that more number of suppliers can be accommodated in a single trip of the truck considering the 200 kms radius constraint.

$$D_{s',t} \le \left(del1_{s',t} \times 450\right) + \left(del2_{s',t} \times 1600\right)$$

$$\forall s^{'} \in S; t \in \left\{t^{'}+1, t^{'}+2, \ldots, t^{'}+14\right\} \tag{10.23}$$

Constraint (10.24) ensures the maximum number of pickup suppliers for the truck type 1 and truck type 2, i.e., to ensure truck type 1 (LTL) is used for point-to-point pickup, and truck type 2 (TL) is used for a maximum of 3 pickups including the first pickup supplier $s^{'}$ and its associated suppliers.

$$\sum_{s \in S} \delta_{s,s',t} \le del1_{s',t} + 3 \times \left(del2_{s',t}\right)$$

$$\forall s^{'} \in S; t \in \left\{t^{'}+1, t^{'}+2, \ldots, t^{'}+14\right\} \tag{10.24}$$

Constraints (10.25) and (10.26) ensure that weight is calculated corresponding to supplier s and $s^{'}$ only when supplier s is allocated to that first pickup supplier $s^{'}$ at a given time t. Constraint (10.27) ensures that if s is not allocated to its first pickup supplier $s^{'}$ at a given time t, then total weight corresponding to supplier s is 0.

$$Q^{''}_{s,s',t} \le Q^{'}_{s,t} + M\left(1 - \delta_{s,s',t}\right) \tag{10.25}$$

$$Q^{''}_{s,s',t} \ge Q^{'}_{s,t} - M\left(1 - \delta_{s,s',t}\right) \tag{10.26}$$

$$Q^{''}_{s,s',t} \le M\delta_{s,s',t} \tag{10.27}$$

$$\forall s^{'} \in S; t \in \left\{t^{'}+1, t^{'}+2, \ldots, t^{'}+14\right\}$$

Constraint (10.28) computes total weight of products carried from all suppliers including the first pickup supplier s' in time period t. In the constraint below, $Q^{''}_{s',t}$ represents the total weight corresponding to the first pickup supplier and its associated suppliers, which is equal to the sum of all the weights from the suppliers (including the first pickup supplier) who are associated with that first pickup supplier in time period t.

$$Q^{'''}_{s',t} = \sum_{s \in S} Q^{''}_{s,s',t}$$

$$\forall s^{'} \in S; t \in \left\{t^{'}+1, t^{'}+2, \ldots, t^{'}+14\right\} \tag{10.28}$$

Constraint (10.29) ensures that total weight of order quantity from the first pickup supplier should be less than or equal to the truck capacity of the particular truck type that is assigned to the first pickup (medoid) supplier.

$$Q^{'''}_{s',t} \leq \left(del1_{s',t} \times C1\right) + \left(del2_{s',t} \times C2\right)$$

$$\forall s^{'} \in S; t \in \left\{t^{'}+1, t^{'}+2, \ldots, t^{'}+14\right\} \tag{10.29}$$

The below constraints are used to find the transportation cost of the truck allotted to first pickup supplier and its associated suppliers during time period t if supplier $s^{'}$ acts as a first pickup supplier. Constraint (10.30) ensures that transportation cost of truck type 1 (LTL) which is allotted to first pickup supplier and its associated suppliers during time period t is calculated with respect to the total weight $Q^{'''}_{s',t}$ shipped by the first pickup supplier including its associated suppliers. The cost is computed only when truck type 1 is used for transportation i.e.,$del1_{s',t} = 1$ else cost computation is 0 (constraint 10.30 becomes redundant). Constraint (10.31) ensures that transportation cost of truck type 2 (TL) which is allotted to first pickup supplier and its associated suppliers during time period t is calculated with respect to the total distance traveled shipped by the first pickup supplier including its associated suppliers. The cost is computed only when truck type 2 is used for transportation i.e., $del2_{s',t} = 1$, else cost computation is 0 (constraint 10.31 becomes redundant). Constraint (10.32) ensures that transportation cost is 0, if $s^{'}$ does not act as a first pickup supplier.

$$Cost_{s',t} \geq \left(Q^{'''}_{s',t} \times c1\right) - M\left(1 - del1_{s',t}\right) \tag{10.30}$$

$$Cost_{s',t} \geq \left(D_{s',t} \times c2\right) - M\left(1 - del2_{s',t}\right) \tag{10.31}$$

$$Cost_{s',t} \leq M \times \delta_{s',s',t} \tag{10.32}$$

$$\forall s^{'} \in S; t \in \left\{t^{'}+1, t^{'}+2, \ldots, t^{'}+14\right\}$$

Constraint (10.33) is used to compute cost if truck type 1(LTL) is used to pick up order quantity from first pickup supplier $s^{'}$, i.e., $del1_{s',t}$=1, which is multiplied by the minimum charge of using that truck. Constraint (10.34) is used to compute cost if truck type 2 (TL) is used to pick up order quantity from first pickup supplier $s^{'}$, i.e., $del2_{s',t}$=1, which is multiplied by the minimum charge of using that truck.

$$Cost1_{s',t} \geq OH1 \times del1_{s',t} \tag{10.33}$$

$$Cost2_{s',t} \geq OH2 \times del2_{s',t} \tag{10.34}$$

$$\forall s^{'} \in S; t \in \left\{t^{'}+1, t^{'}+2, \ldots, t^{'}+14\right\}$$

Constraint (10.35) computes the volume occupied by the end inventory in the warehouse. This constraint is used to optimize the warehouse storage cost.

$$V_{p,s,t} \geq Inv_{p,s,t} \times \frac{vol}{unit_{\,p,s}}$$

$$\forall s \in S;\ p \in P(s);\ t \in \left\{t^{'}+5, t^{'}+6, \ldots, t^{'}+18\right\} \tag{10.35}$$

Constraint (10.36) ensures that at the end of given $t^{'}$, beginning inventory for next cycle gets updated from the previous cycle.

$$BI_{\,p,s,t'+15} = BI_{\,p,s,t'+1} + \sum_{\tau=t^{'}+1}^{t^{'}+14} Q_{p,s,\tau} - \sum_{\tau=t^{'}+5}^{t^{'}+18} D_{p,s,\tau}$$

$$\forall s \in S;\ p \in P(s);\ t^{'} \in \{0, 14\}. \tag{10.36}$$

Phase 3: Actual routing cost calculation. The constraint in the case of a TL (milk-run) type of truck is to choose the suppliers in such a way that they are within a radius of 200 kms from the first pickup supplier. This constraint is provided in view of making dense clusters for TSP. In our MILP model, the actual travel distance by the trucks has been approximated by considering the first supplier visited acts as a medoid in the cluster. This umbrella approximation of distance computation is shown in Fig. 10.1. Constraint (10.23) in the MILP model (Phase 2) is used to compute the approximate distance traveled by the truck. The total approximate travel distance consists of to and fro distance from the warehouse to the first pickup (notional medoid) supplier, and distance between the first pickup (notional medoid) and its associated suppliers.

The actual routing is shown in Fig. 10.2 considering $s^{'}$ as the first pickup supplier and s_1 and s_2 as its associated suppliers. The best alternative out of the two routings given is the one which is close to the warehouse, $s_1 \rightarrow wh$ or $s_2 \rightarrow wh$. The actual travel cost is then computed for the best routing.

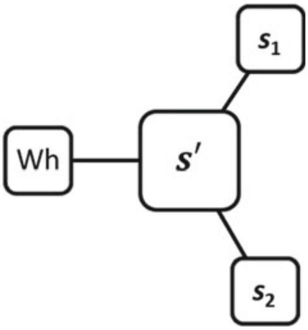


Fig. 10.1 Umbrella approximation for routing

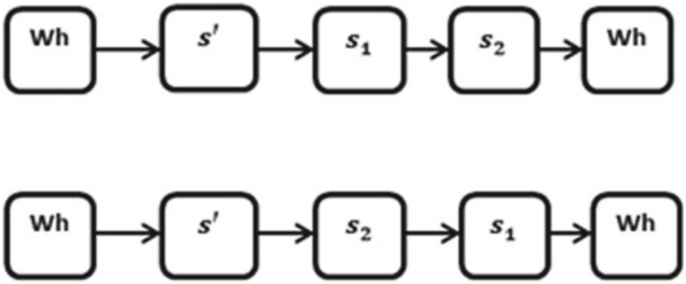


Fig. 10.2 Actual routing

10.4 Computational Results

The proposed MIP model for clustering is executed on the dataset and clusters are formed as in the sample Fig. 10.3. As the problem becomes computationally difficult to solve, when the number of parts and suppliers increase, clustering model in phase 1 is used to decompose the larger problem into small clusters. The proposed MILP model for order pickup and transport planning is then executed on one of the cluster consisting of 30 parts, 7 suppliers, and for a planning period of 28 days. The problem instances are solved using ILOG CPLEX 20.1.0.0 on a PC with INTEL® Xeon ® CPU E5-2620 v3@ 2.40 GHz 2 Processor, 64.0 GB RAM (10 GB used). After the clusters are obtained, a preprocessing check is also done to ensure that the demand on the first day doesn't exceed truck capacity. If the demand for the parts on the first day from any supplier exceeds the truck capacity, a separate TL truck is used to pick up the extra load from that supplier. The proposed MILP model minimizes the transportation, inventory, and warehouse storage costs, and yields the order pickup quantity and the transport plan. The model is run twice for every cluster, each with a time span of 14 days, and inventory is adjusted for the second period such that end inventory on the 14th day is equal to the beginning inventory for the second period, which starts from 15th day onwards.

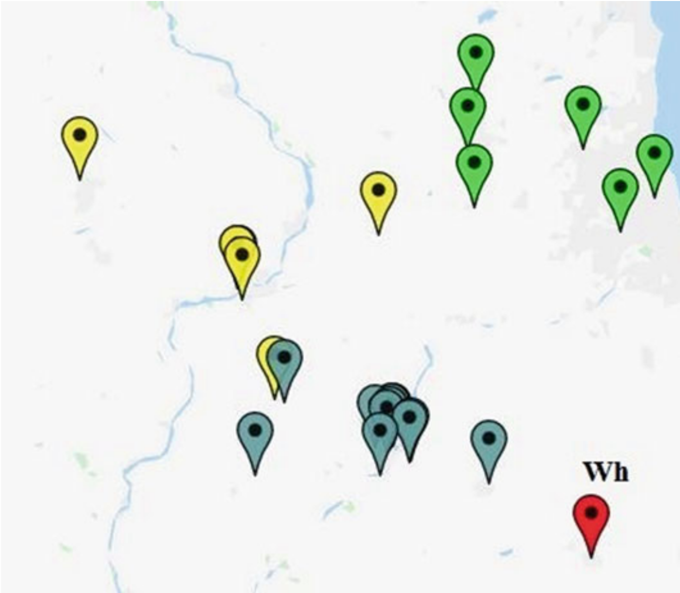


Fig. 10.3 A representative dataset to show the clusters formation after execution of the proposed MIP model for clustering

Ramkumar et al., (2012) proposed a MILP formulation for a many-to-many multi-product multi-depot IRP problem. They were able to run small instances consisting of two vehicles, two suppliers, three customers, and three periods. They were not able to solve the bigger problem instances to optimality in eight hours of CPU time. Coelho & Laporte, (2013b) proposed a MILP model for multi-product multi-vehicle IRP which could solve instances with up to five parts, five vehicles, three periods, and 30 customers.

The MILP model is solved for 2 time periods, i.e., in the first cycle the order placement is from 1 to 14 days for the demand for 1–18 days (considering lead time of 4 days), in the second cycle order placement is from 15 to 32 days (considering initial 4 days demand as 0) (Table 10.1).

Numerical Illustration.

The application of the proposed model is done on a sample data, which is given in the link below.

Table 10.1 Results after solving the proposed MILP model for one time period

CPU Time (s)	Objective value	Gap (%)
3600	6117.571	25.74%
7200	6117.488	25.74%
10,800	5692.245	20.19%

https://docs.google.com/document/d/1H3-E8PvbiJJogL9Efdd0ds_lmSBUXhiD/edit?usp=sharing&ouid=103465411975588874146&rtpof=true&sd=true

10.5 Conclusion and Future Work

One of the tactical decisions to be taken by managers is to design the schedule of pickup of raw materials from the suppliers, decide on the pickup quantity, and design routes. We have proposed an approach to solve the order pickup and transportation planning problem prevailing in many inbound logistics industries. The proposed model for clustering can be applied to a large-scale setting for breaking a large dataset of suppliers to smaller clusters. The proposed MILP model for order pickup is computationally expensive for the given dataset instance but can be implemented on many more instances for future work. The overall approach can also be used in a Decision Support System (Andersson et al., 2010). The approach can be useful in automobile industries and retail industries. Time window constraints can also be incorporated in the model and lead time variability can also be accounted in the future work.

References

1. Alegre J, Laguna M, Pacheco J (2007) Optimizing the periodic pick-up of raw materials for a manufacturer of auto parts. Eur J Oper Res 179(3):736–746. https://doi.org/10.1016/j.ejor.2005.03.063
2. Andersson H, Hoff A, Christiansen M, Hasle G, Løkketangen A (2010) Computers and operations research industrial aspects and literature survey : combined inventory management and routing. Comput Oper Res 37(9):1515–1536. https://doi.org/10.1016/j.cor.2009.11.009
3. Bertazzi L, Speranza MG, Ukovich W (1997) Minimization of logistic costs with given frequencies. Transp Res Part B: Methodol, 31(4 PART B), 327–340. https://doi.org/10.1016/s0191-2615(96)00029-x
4. Carter MW, Farvolden JM, Laporte G, Xu J (1996) Solving an integrated logistics problem arising in grocery distribution. Infor 34(4):290–305. https://doi.org/10.1080/03155986.1996.11732310
5. Coelho LC, Cordeau JF, Laporte G (2014) Thirty years of inventory routing. Transp Sci 48(1):1–19. https://doi.org/10.1287/trsc.2013.0472
6. Coelho LC, Laporte G (2013) A branch-and-cut algorithm for the multi-product multi-vehicle inventory-routing problem. Int J Prod Res 51(23–24):7156–7169. https://doi.org/10.1080/00207543.2012.757668
7. Coelho LC, Laporte G (2013) The exact solution of several classes of inventory-routing problems. Comput Oper Res 40(2):558–565. https://doi.org/10.1016/j.cor.2012.08.012
8. Dror M, Ball M, Golden B (1985) A computational comparison of algorithms for the inventory routing problem. Ann Oper Res 4:3–23. https://doi.org/10.1007/BF02022035
9. Gaur V, Fisher ML (2004) A periodic inventory routing problem at a supermarket chain. Oper Res 52(6):813–822. https://doi.org/10.1287/opre.1040.0150
10. Mjirda A, Jarboui B, Macedo R, Hanafi S (2012) A variable neighborhood search for the multi-product inventory routing problem. Electron Notes Discret Math 39:91–98. https://doi.org/10.1016/j.endm.2012.10.013

11. Moin NH, Salhi S, Aziz NAB (2011) An efficient hybrid genetic algorithm for the multi-product multi-period inventory routing problem. Int J Prod Econ 133(1):334–343. https://doi.org/10.1016/j.ijpe.2010.06.012
12. Mosca A, Vidyarthi N, Satir A (2019) Integrated transportation—inventory models: A review. Oper Res Perspect, *6*(February). https://doi.org/10.1016/j.orp.2019.100101
13. Preethi KR, Rajendran C, Viswanathan S (2018) Mathematical models and lower bounds for an inventory routing problem with backorders. In: Proceedings of International Conference on Computers and Industrial Engineering, CIE, 2018
14. Ramkumar N, Subramanian P, Narendran TT, Ganesh K (2012) Mixed integer linear programming model for multi-commodity multi-depot inventory routing problem. Opsearch 49(4):413–429. https://doi.org/10.1007/s12597-012-0087-0
15. Sofianopoulou S (2014) A heuristic approach for an inventory routing problem with backorder decisions. Oper Reserch 6:49–57
16. Speranza MG, Ukovich W (1994) Minimizing transportation and inventory costs for several products on a single link. Oper Res 42(5):879–894. https://doi.org/10.1287/opre.42.5.879

Chapter 11
Performance Evaluation of Trucking Industry Using BSC and DEA: A Truck Driver's Perspective

Krishna Kumar Dadsena, S. P. Sarmah, and V. N. A. Naikan

Abstract Roads transport lions-share of the freights as they can reach each and every corner of the country. In road freight transport, truck is more preferable due to its advantages over railways due to its ability to reach beyond geographical barriers, quantity barriers, and ease of availability of trucks. In order to achieve efficient and reliable operations of the trucking industry the truck drivers play a vital role. As one of the major issues the trucking industry facing is shortage of truck drivers. The objective of this study is to identify the performance measures and efficiency calculation. In order to this study adopted and hybrid approach by combining Balance Score Card (BSC) and Data Envelopment Analysis (DEA). The present study considers a case of Indian trucking industry, and the efficiency of truck drivers has been calculated. This study helps the organization in reciprocal learning, by focusing on the specific factors or perspectives for efficiency improvement. The results of the study can also be used in strategic implementation to improve the performance of the trucking industry.

Keywords BSC · DEA · Truck drivers · Trucking industry

11.1 Introduction

As the trucking industry is dramatically growing in recent times, its efficient management has become a challenging task. To achieve growing customer expectations in the competitive market and poor environment condition, the management attention needs to be more focused toward performance improvement. With the evolution of the trucking industry, the managers started focusing on the performance, productivity,

K. K. Dadsena (✉)
Department of Operations Management, Indian Institute of Management Ranchi, Ranchi, India
e-mail: Krishnakumar.dadsena@iimranchi.ac.in

V. N. A. Naikan
Subir Chowdhury School of Quality and Reliability, IIT Kharagpur, Kharagpur, India

S. P. Sarmah
Department of Industrial and Systems Engineering, IIT Kharagpur, Kharagpur, India

M. K. Tiwari et al. (eds.), *Applications of Emerging Technologies and AI/ML Algorithms*, Asset Analytics, https://doi.org/10.1007/978-981-99-1019-9_11

and efficiency of individual entity. At present, trucking sector is facing a few crucial issues. For example, massive shortage of truck drivers, due to inappropriate educational qualification, permissible driver's age, experience, and gender are the crucial issues. Analyzing the impact of human resource management on performance has become one of the main challenges for the industry. This encourages the management to take a strategic step toward the efficient and effective driver management system. Such system connects human resource management to their organizational performance. These factors open a new dimension of research to the scholars to explore the key performance indicators on the perspective of truck drivers. Performance is normally assessed based on an evaluation of the responses and estimation of the important factors. The traditional technique of performance evaluation is determination of individual job-related actions and their outcomes considering a particular environment. However, in the trucking company, the performance measurement does not only depend on output of transport service in terms of ton-km. The job satisfaction, organizational commitment, motivation of the employees (truck drivers) is also needed to be taken into account. Thus, the operational performance of trucking industry depends more on the close interrelations with driver's perspective and their job satisfaction criteria. In order to better understand such key performance indicators, we referred a field survey and key literature [15–17]. Some of the typical measurements has been considered for this study. The aim of this study is to evaluate past performance so that it could be utilized in the planning phase.

Considering the truck driver's point of view, the trucking company generally concerns the pursuit of a clearly defined mission and vision which is primarily focused on the economic concern. However, the challenges faced by the management are to translate their mission and vision into targets. This target should not only consistently focus on the economical effectiveness but also on the social efficiency. Therefore, the trucking industry must consider the truck driver's perspective, which may help to overcome the problem of truck driver shortage. Several methods are available which aim to assess the activities carried out by business organizations in their performance evaluation. Some of the methods are Balanced Scorecard (BSC), Data Envelopment Analysis (DEA), Ratio Analysis, Multi-Criteria Decision-Making (MCDM), etc. Among these techniques, BSC is extensively used to align business activities to the organizational strategies by strategic planning and management [1]. The Balanced Scorecard (BSC) system is a popular performance management technique which integrates key measurement indices considering the organizational goals and categories into four perspectives: Financial, Customer, Internal Business Process, and Learning and Growth [2]. On the other hand, DEA is a mathematical approach for identifying and analyzing the best practice of peer decision-making units (DMUs), in the presence of multiple inputs and multiple outputs [3]. In this paper, we propose an integrated approach that combines the BSC and DEA. The motive is to evaluate the performance of the trucking industry.

The remaining parts of the paper are organized as follows. Section 11.2 presents a brief review of the literature on performance evaluation, BSC, and DEA. Section 11.3 presents the methodology of the proposed approach. Section 11.4 produces the results and discussion. Finally, Sect. 11.5 summarizes the contributions of the research and future research scopes in this direction.

11.2 Literature Review

The growing importance of performance measurement approach is changing the way managers run their businesses [4]. The BSC is a formal management tool which provides a realistic framework which integrates the key stakeholders including owners, customers, and employees. "Balance" reflects the attempt to capture financial and non-financial measurements, short-term actions as well as long-term strategic objectives, internal and external performance perspectives, as well as qualitative-subjective and quantitative-objective measures [5]. Even though the popularity of BSC among practitioners and academia [3, 6, 7], authors have identified several limitations.

1. The trade-offs between the different criteria are not specified by BSC.
2. Objective weighting scheme for the performance measures is not specified and it causes more complexity because of lack of common measurement scale.
3. BSC may also fail to identify inefficiency in the use of resources.

These limitations mentioned previously indicate the need of DEA for BSC as complementary tool as it can be a helpful tool to deal with these types of issues. DEA is a linear programming based on a non-parametric method introduced [3]. Multiple measures of inputs and outputs are used in relative efficiency calculation for a group of DMUs. In addition to this, it also provides information related to the inefficient DMUs. DEA has already been satisfactorily employed in the logistics industry [8] and in other segments such as airline industry [9], airports [10, 11], road passenger transport [12], and freight transport [13, 14]. However, prior studies discussed that the DEA application in trucking industry needs more attention [15–17].

Advantages of DEA:

1. DEA permits to analyze multiple inputs and output factors simultaneously.
2. DEA solves an optimization problem and gains its weights result.
3. DEA is a non-parametric approach.
4. It reduces the bias in parameter estimation.

An integrated BSC and DEA approach helps to overcome important limitations of the BSC by adding the basis for enhanced performance assessment [18, 19]. The joint BSC-DEA model helps the managers in decision-making process by providing a conceptual framework for the performance as well as the efficiency evaluation of their operational process [20]. The importance of performance evaluation of a service has become very difficult due to increasing complexity and linkage between different operations and strategies. The authors have suggested an integrated BSC-DEA approach for performance evaluation and validated through a case of automotive industry [21]. An integrated BSC-DEA approach provides strength in performance measurement internally in relation with methodology and external by adding more operational insights to improve the organization performance [22]. However, to the best of our knowledge this is the first study which has developed and integrated approach using BSC and DEA to evaluate the performance of the trucking industry from truck driver's perspective.

11.3 Methodology

BSC is used to identify and specify the casual relationship between the success factors based on the vision, mission, and future perspectives of the organization. Figure 11.1 demonstrates the details of the relationship among the factors based on the four perspectives of financial, customer, internal process, and learning and growth.

The elements and indicators presented in Fig. 11.1 of the BSC are designed and further validated through a case of Indian trucking industry. The efficiency assessment has been performed by comparative study between the inputs and outputs measures which is presented in Table 11.1 obtained from different perspectives.

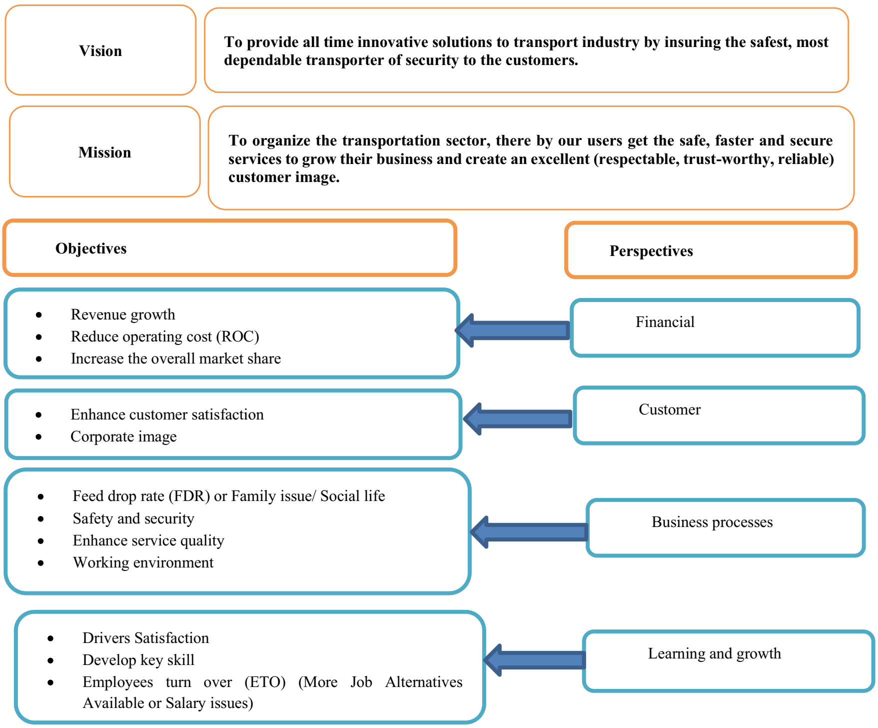


Fig. 11.1 The vision and mission strategy

Table 11.1 Input and output measures for DEA models

Output	Measures	Objective	Input	Measure	Objective
Develop key skill	This measures the capability to organize focused training and development skill for the truck drivers	Driving and training institute will motivate this profession	More Job Alternatives Available or Salary issues	This indicator indicates that number of truck drivers resigned (due to more job opportunity or salary issue)	To help the effective salary/ incentive for the truck drivers received
Employee satisfaction	It measures the number of the complaints by drivers	Happy employees work harder, more productively, and remain on staff longer	Working environment or Driver Turnover Rate (DTR)	Interview of drivers to examine your workplace culture, compensation packages, and benefits	To create a positive and pleasure working environment from their job
Safety and security	This measures the present image with the desired image for the organization in terms of safety and security of the drivers	Enhance job satisfaction and image of the organization	Family issue/ Social life or Feed drop rate (FDR)	This measures the average rate of feeds dropped in a month because of issues related to social life	This can help to increase service quality

Table 11.2 Notations of the sets, parameters

n	Number of decision making units (j = 1, 2, … n)
s	Outputs (r = 1, 2, …, s)
m	Inputs (i = 1, 2, …, m)
v_i	Weights assigned to *i*th input
u_r	Weights assigned to *r*th output
y_{rj}	Amount of output r produced by DMU j;
x_{ij}	Amount of input i used by DMU j;

The DEA method is basically data-oriented approach, for efficiency evaluation. In which the for the number of DMUs multiple inputs and outputs take place in calculation of efficiency. The efficiency of Decision-Making Units (DMUs), which converts multiple inputs into multiple outputs (Table 11.2).

$$\text{Maximize efficiency score } (\emptyset) = \frac{\sum_{r=1}^{s} u_r y_r j}{\sum_{i=1}^{m} v_i x_{ij}} \tag{11.1}$$

Subject to

$$\sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \leq 0 \tag{11.2}$$

$$\sum_{i=1}^{m} v_i x_{ij} = 1 \tag{11.3}$$

11.4 Results and Discussions

The integrated approach is validated through the Indian trucking industry. At initial stage, preliminary BCS approach is applied that provides a balanced portfolio for performance measurements of this study. The BSC helps to select right measures having the strongest causal linkages between the learning and growth with the internal business process. In the second stage, the measures from the BSC has been identified as the inputs and output variables for the DEA model. We selected a set of ten truck drivers and conducted personal interviews with these drivers. The data collection has been performed using seven Likert scale (Strongly disagree, Disagree, slightly disagree, can't say, slightly agree, Agree, strongly agree) for ten truck drivers (Table 11.3).

In this approach, the integration of DEA model leads to select the right indicators utilized efficiently. This allows the management to focus on strategic and operational goals considering the truck driver's perspective. The integrated approach provides a systematic way of identifying and measuring the performance with higher accuracy. Thus, adding more rationale evidence helps in adoption of strategic steps to understand the issues related to the efficiency measurement. The efficiency score and the ranking of the truck drivers is shown in Table 11.4.

From the results, it is observed that particularly some delegations (e.g., DMU1, DMU4, and DMU6) require attention on learning and growth and internal processes perspective. This study helps the management to conduct brainstorming sessions with different entities of the organization. This also helps to understand the relevant issues from the truck driver's perspective. This study attempts to develop a strategic map by considering the issues related to truck driver's performance. The development of BSC from the trucking industry viewpoint is a significant contribution. It focuses on identification of some of the critical issues and performance measurement of the industry. Moreover, the integration of the BSC framework with DEA is helpful to consolidate the performance results to identify and rank the decision-making units. It also suggests appropriate learning strategies. This study is useful to the managers to identify the main reasons behind the poor efficiency and performance of the trucking industry from the truck driver's perspective.

Table 11.3 Data from the truck drivers for the evaluation criteria

DMU	Age	Qualification	Experience (Years)	Develop key skill	Employee satisfaction	Safety and security	Salary issues	Working environment	Family issue/ social life
DMU1	25–30	UG	5–10	6	5	7	5	7	6
DMU2	24–24	12th	<5	6	5	5	6	5	7
DMU3	31–40	10th	15–20	5	5	5	6	7	6
DMU4	25–30	UG	5–10	5	4	7	6	7	6
DMU5	20–24	12th	<5	6	4	6	6	6	5
DMU6	<20	12th	<5	5	6	6	7	6	6
DMU7	31–40	10th	15–20	5	5	6	6	6	7
DMU8	25–30	UG	<5	6	5	5	6	7	6
DMU9	31–40	12th	5–10	5	5	7	7	6	6
DMU10	31–40	UG	5–10	6	6	6	6	5	7

Table 11.4 The efficiency score and ranking

DMU	Efficiency score	Rank
DMU1	0.8370	8
DMU2	0.9074	6
DMU3	1.0000	2
DMU4	0.7775	9
DMU5	1.0000	1
DMU6	0.6705	10
DMU7	1.0000	4
DMU8	0.8978	7
DMU9	1.0000	3
DMU10	0.9517	5

11.5 Conclusion

This paper attempted to demonstrate the performance of trucking industry considering the truck driver's perspective. Management should not only aim to achieve effective operation management but also efficient human resource management must be taken into consideration. For this, we have reviewed the literature that dealt directly or indirectly with the issues related to the truck drivers' point of view. This approach focuses mainly on learning and growth, business process perspective and provides good and acceptable results in decision-making. Furthermore, this study reveals how the factors like educational qualification and the age of the drivers reflect their mindset and efficiency toward their profession.

This study helps managers in appropriate decision-making and the development of strategic steps toward controlling the shortage of the truck drivers. This study reveals the conviction and expectation of the truck drivers toward their profession. From the result, it can also be observed the way educational qualification and driver's age factor affects their efficiency. This strategic map guides the management in improving the performance of the industry by helping in selection of right measurements indicator. In the future, this can be extended to multi-objective approach considering different perspectives.

References

1. Varmazyar M, Dehghanbaghi M, Afkhami M (2016) A novel hybrid MCDM model for performance evaluation of research and technology organizations based on BSC approach. Eval Program Plann 58:125–140
2. Kaplan RS, Norton DP (2001) Transforming the balanced scorecard from performance measurement to strategic management: part I. Account Horiz 15(1):87–104
3. Charnes A, Cooper WW, Rhodes E (1978) Measuring the efficiency of decision making units. Eur J Oper Res 2(6):429–444

4. Chiang C-Y, Lin B (2009) An integration of balanced scorecards and data envelopment analysis for firm's benchmarking management. Total Qual Manag 20(11):1153–1172
5. Phusavat K, Kanchana R (2008) Competitive priorities for service providers: perspectives from Thailand. Ind Manag Data Syst 108(1):5–21
6. Maestrini V, Luzzini D, Maccarrone P, Caniato F (2017) Supply chain performance measurement systems: a systematic review and research agenda. Int J Prod Econ 183:299–315
7. Staš D, Lenort R, Wicher P, Holman D (2015) Green Transport balanced scorecard model with analytic network process support. Sustainability 7(11):15243–15261
8. Ferrari C, Migliardi A, Tei A (2018) A bootstrap analysis to investigate the economic efficiency of the logistics industry in Italy. Int J Log Res Appl 21(1):20–34
9. Schefczyk M (1993) Operational performance of airlines: an extension of traditional measurement paradigms. Strateg Manag J 14(4):301–317
10. Kottas AT, Madas MA (2018) Comparative efficiency analysis of major international airlines using Data Envelopment Analysis: exploring effects of alliance membership and other operational efficiency determinants. J Air Transp Manag 70:1–17
11. Merkert R, Hensher DA (2011) The impact of strategic management and fleet planning on airline efficiency—a random effects Tobit model based on DEA efficiency scores. Transp Res Part A: Policy Pract 45(7):686–695
12. Venkatesh A, Kushwaha S (2018) Short and long-run cost efficiency in Indian public bus companies using Data Envelopment Analysis. Socioecon Plann Sci 61:29–36
13. Wolff MGdC, Caldas MAF (2018) A model for the evaluation of Brazilian Road transport: a sustainable perspective. J Adv Transp 2018
14. Melo IC, Junior PNA, Perico AE, Guzman MGS, Rebelatto DADN (2018) Benchmarking freight transportation corridors and routes with data envelopment analysis (DEA). Benchmark: Int J 25(2):713–742
15. Mejza MM, Corsi TM (1999) Assessing motor carrier potential for improving safety processes. Transp J 38(4):36–50
16. Odeck J, Hjalmarsson L (1996) The performance of trucks—an evaluation using data envelopment analysis. Transp Plan Technol 20(1):49–66
17. Poli PM, Scheraga CA (2001) A quality assessment of motor carrier maintenance strategies: an application of data envelopment analysis. Quart J Bus Econ 25–43
18. Amado CA, Santos SP, Marques PM (2012) Integrating the data envelopment analysis and the balanced scorecard approaches for enhanced performance assessment. Omega 40(3):390–403
19. Kazemkhanlou H, Ahadi H (2015) An integrated approach using data envelopment analysis and balanced score card to supply chain performance evaluation. In: 2015 International Conference on Industrial engineering and operations management (IEOM). IEEE
20. Kádárová J, Durkáčová M, Teplická K, Kádár G (2015) The proposal of an innovative integrated BSC–DEA model. Procedia Econ Finance 23:1503–1508
21. Tan Y, Zhang Y, Khodaverdi R (2017) Service performance evaluation using data envelopment analysis and balance scorecard approach: an application to automotive industry. Ann Oper Res 248(1–2):449–470
22. Basso A, Casarin F, Funari S (2017) How well is the museum performing? a joint use of DEA and BSC to measure the performance of museums. Omega

Chapter 12
E-procurement: An Emerging Tool for Pharmaceutical Supply Chain Management

Esha Saha, Pradeep Rathore, and Bhargav Anne

Abstract The movement of commodities, services and information in the pharmaceutical industry should be planned to effectively turn raw materials into completed dosage forms of medications. Pharmaceutical manufacturing companies frequently purchase huge amounts of active pharmaceutical ingredients which require efficient procurement systems to maintain everyday supply and activities. One of the major goals of pharmaceutical supply chain management is to efficiently apply information technology to its procurement systems. This research examines the Indian pharmaceutical companies' managers' perspectives about e-procurement technologies through a qualitative study. Based on extensive literature review and the obtained results, a research model for examining the pharmaceutical companies' behaviour to use e-procurement technologies is proposed. The proposed model is an extension of the Technology acceptance model. The identified factors for the model are data integration, performance expectancy, cost and user benefit, innovation, infrastructure, operations, individual attributes, perceived usefulness and perceived ease that will influence pharma companies' attitude-intention-behaviour to use e-procurement technologies.

Keywords E-procurement · Pharmaceutical industry · Supply chain management · Qualitative study · Extended technology acceptance model

E. Saha (✉) · P. Rathore
Paari School of Business, SRM University, Amaravati, Andhra Pradesh, India
e-mail: eshasaha83@gmail.com

P. Rathore
e-mail: pradeeprathore076@gmail.com

B. Anne
Institute of Management Technology, Hyderabad, India

M. K. Tiwari et al. (eds.), *Applications of Emerging Technologies and AI/ML Algorithms*, Asset Analytics, https://doi.org/10.1007/978-981-99-1019-9_12

12.1 Introduction

The procurement function in pharmaceutical industries includes purchasing and supply management of raw materials like active pharmaceutical ingredients along with other medical supplies, equipment, and instruments [1]. A traditional procurement process is characterized by paper-based systems and manual tasks including managing suppliers, requesting, and issuing orders, preparing contracts, negotiating with suppliers, generating invoices and bills. This traditional procurement process is time-consuming, prone to human errors and expensive in resource terms [2]. Electronic procurement, often known as e-procurement, is an online-based facility for procurement. It is an application of information and communication technology to procurement operations [3]. E-procurement is currently a powerful tool for improving the effectiveness, efficiency and service quality of its users. As a result, it has recently received a lot of attention in the corporate world. Many industries have now identified e-procurement as a priority and have implemented and adopted e-procurement in their varied procurement procedures. It has the potential to streamline all areas of the procurement process while decreasing transaction time and cost, simplifying processes and enhancing standardization [4]. Among many sectors that have implemented e-procurement platforms, pharmaceutical companies have emphasized the adoption of e-procurement [2, 3, 5, 6].

E-procurement ensures service improvements and cost effectiveness in pharma supply chains by enabling end-to-end integrated material purchases from procurement need identification to final payment through the internet. Basic technologies such as e-ordering and e-tendering, as well as advanced systems such as computerized procure to pay process, artificial system-based supply management, blockchain-based product information record and cloud-based supplier risk management may be used in e-procurement [4, 7]. Therefore, it is crucial to learn how pharmaceutical company managers who are involved in procurement operations view the use of e-procurement systems. This has motivated us for this study, and it adds the following novelty to the literature. First, best to our knowledge, this is the first study that examines the pharmaceutical companies' managers' opinions about e-procurement systems with specific focus on Indian pharma companies. Second, this study uses the TAM theory [8] and extends it to the field of e-procurement in the pharmaceutical supply chain. Third, this research work provides supporting evidence regarding the perspectives of pharma managers towards implementation of e-procurement system for pharmaceutical supply chain management. Fourth, it also identifies factors influencing the attitude-intention-behaviour of Indian pharmaceutical companies towards the use of e-procurement systems from the theoretical standpoint of the TAM. In addition, this study provides answers to the following research questions:

RQ1. What are the perspectives of pharma managers towards implementation of e-procurement system for pharmaceutical supply chain management?

RQ2. What are the factors influencing the attitude-intention-behaviour of Indian pharmaceutical companies towards the use of e-procurement systems from the theoretical standpoint of the Technology acceptance model (TAM)?

The rest of the chapter has been structured in the following way: Literature review is presented in Sect. 12.2. Section 12.3 discusses the research methodology. The results are shown in Sect. 12.4. Finally, limitations with future research directions and conclusion are highlighted in Sect. 12.5.

12.2 Literature Review

The pharmaceutical procurement process includes identifying and assessing suppliers, selecting the best vendor for purchasing goods including raw materials, machinery, and equipment to produce medications, and monitoring their performance [1]. Identifying suppliers who can provide cutting-edge technologies, cooperative relationships, long duration agreements, and smart contracts to address challenges in uncertain situations and disruptions, cloud-based ordering systems, electronic data interchange, online ordering, and automated processes are just a few of the advanced services that are required for modern sourcing [5, 9, 10]. These services may be provided with the assistance of e-procurement platforms. It can aid manufacturers in making purchasing decisions and increase market knowledge of suppliers and their products. In addition, cutting-edge technologies like online supplier performance tracking, AI-based supplier relationship management, fully interactive sourcing support, intelligent servicing, blockchain-based security, and Internet of Things (IoT)-based risk management systems can help in advancing and improving the pharmaceutical sourcing process [6]. E-procurement and ambitions to improve procurement operations were highlighted by Bag et al. [4]. To pinpoint the procurement challenges causing drug shortages, Chebolu-Subramanian and Sundarraj [2] worked with decision-makers and the pharmaceutical industry. AlNuaimi et al. [11] conducted research on the rise of big data technology to modernize and create a green e-procurement system from the conventional one. A review study related to supply and inventory model in the supply chain, which controls the raw material inventory in procurement operations, works in progress, and finished goods, was most recently given by Utama et al. [12].

From the review of literature, it has been noted that there is a significant gap between strategizing for new technology adoption and implementing such technology to obtain the intended benefits. To introduce new technology (such as e-procurement), managers' readiness must be assessed in order to give guidelines capable of handling any issues after adoption. Therefore, this research examined pharmaceutical companies' managers' opinions about e-procurement systems with specific focus on Indian pharma companies.

12.3 Methodology

According to the previous research, India is the world's largest supplier of generic medications, and third in pharmaceutical manufacturing, and fourteenth in terms of value [13]. Furthermore, the Indian pharmaceutical business is estimated to be worth approximately $120–130 billion by 2030 [13]. With the entrance of new digital technologies, the pharmaceutical business in India is undergoing a significant transformation [14]. It is anticipated that digital technologies would significantly alter pharmaceutical purchasing and supply management. Therefore, in this study data is collected from experts of Indian pharmaceutical industry. The data collection method is explained in detail in Sect. 12.3.1.

12.3.1 Data Collection and Analysis

With the use of an online survey platform, data was gathered. Twelve open-ended questions from the literature [1, 4, 15, 16] were included in the questionnaire. The inquiries are shown in Table 12.1. The survey is conducted online, and questionnaire was distributed to the pharmaceutical industry experts using email and many social networking sites. The details about pharmaceutical managers were gathered from websites and online directories. Numerous pharmaceutical managers in operations and supply chain received the questionnaire. While sharing the Google Form link, the study's goals and purpose were clearly stated. Since the focus of our research is related to pharmaceutical sector in India, hence the selected samples were restricted to businesses with operations in India. After data collection, the analyses of qualitative data are carried out.

12.4 Results and Discussions

The study examined Indian pharmaceutical companies' managers' opinions about e-procurement technologies through a qualitative study. Based on the results, when asked whether pharma managers have implemented or planning to implement e-procurement systems, most of the pharma managers responded that they are planning to use e-procurement systems. They believed 'e-procurement is an advancement over manual procurement activities'. Using the e-procurement systems can have substantial cost savings, less manual error, real-time information sharing and more transparency. Besides, for implementation of e-procurement systems, pharma companies do not prefer outsourcing of the activity, rather they will select an e-provider or develop their own e-procurement system.

Table 12.1 Questionnaire for the study

	Open-ended questions
1	Does your company presently use or planning to implement e-procurement systems?
2	Will you select e-provider, building your own system or outsourcing?
3	What kind of goods and services will your company procure through a web-based system?
4	What type of transaction your company will involve in the e-procurement?
5	What type of decision-making process will your company involve in the e-procurement?
6	Which type of costs will your company incur while implementing e-procurement?
7	Will you look for return on investment in e-procurement technology?
8	Do you think incorporating e-procurement will be challenging?
9	What is your viewpoint about the future of e-procurement?
10	How advanced e-procurement technologies may enhance your company's supply chain?
11	List some of the specific issues in the implementation of e-procurement in pharma
12	Any other thoughts you would like to share about e-procurement?

Source: [1, 4, 15, 16]

Further, all the pharma managers aim for 'strategic sourcing' using the e-procurement procedure. Most of the pharma companies pursue return on investment in e-procurement technology as a strategic/long-term goal. Among the various costs to be incurred by the company for e-procurement system, pharma managers mentioned technical equipment costs and training-related costs to have major contribution. Some of the specific issues mentioned by the managers in the implementation of e-procurement are 'understanding e-procurement system, handling software-related issues, frequently changing environment and unwillingness of employees to accept the change'. Moreover, the perspective of pharma managers is that advanced e-procurement technologies can enhance the pharma supply chain. Overall, the pharma managers' opinion is that e-procurement technologies will help the company to reduce the manual work involved in the organization. It will also ensure better decision-making, which in turn results in efficient operations and a reduction in overall supply chain cost. Based on the perspective of the pharma managers and extensive review of existing literature, we have proposed a conceptual model (Fig. 12.1) extending the Technology acceptance model (TAM) [8] and Theory of planned behavior (TPB) [17] to examine the pharmaceutical companies' behaviour to use e-procurement technologies. The extended TAM-TPB model is composed of major 13 constructs.

According to the model, data integration, performance expectancy, cost benefit and user benefit influence perceived usefulness of e-procurement technology in the pharma sector [18]. Innovation capabilities, infrastructure, operations and processes, and individual attributes influence perceived ease-of-use of the e-procurement technology in the pharma sector [19, 20]. Finally, perceived usefulness and perceived ease-of-use of the e-procurement technology influence pharma companies' attitude-intention-behaviour to use e-procurement services.

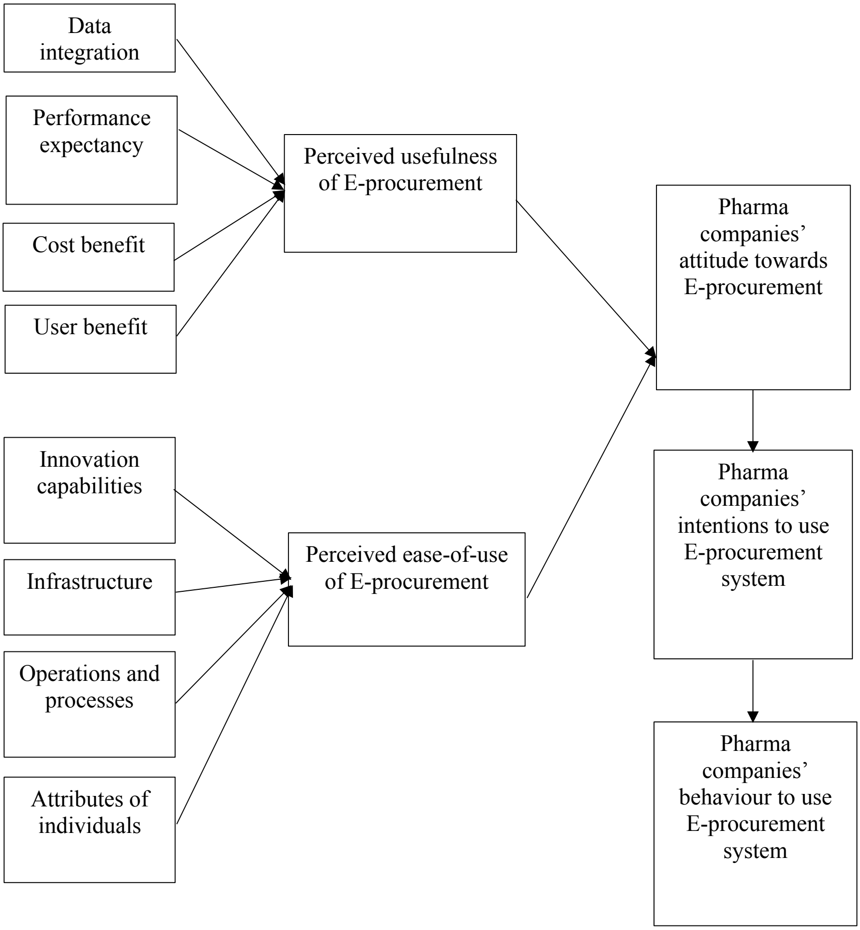


Fig. 12.1 A proposed research framework

12.5 Limitations, Future Scope and Conclusion

This study has conducted a qualitative study to understand the perspectives of pharma managers regarding e-procurement and to identify the factors that will influence pharma companies' behaviour to use e-procurement technologies. Thus, along with the qualitative study it is recommended to conduct a quantitative study to identify the significant factors influencing pharma companies' behaviour to use e-procurement technologies. However, it has been noted that there is a significant gap between strategizing for new technology adoption and implementing such technology to obtain the intended benefits. To introduce a new technology (such as e-procurement), managers' readiness must be assessed to give guidelines capable of handling any issues after adoption. Therefore, this research examined pharmaceutical companies' managers' opinions about e-procurement systems with specific focus on Indian pharma companies. A future research model for examining the pharmaceutical companies' behaviour to use e-procurement technologies is also proposed.

References

1. Ding B (2008) Pharma Industry 4.0: literature review and research opportunities in sustainable pharmaceutical supply chains. Process Saf Environ Prot 119:115–130
2. Chebolu-Subramanian V, Sundarraj RP (2021) Essential medicine shortages, procurement process and supplier response: a normative study across Indian states. Soc Sci Med 278:113926
3. Dassanayake KMDR, Koggalage PD, Kulasuriya PKSS (2020) Implementation of e-procurement for pharmaceuticals: perspectives of the staff of the state pharmaceuticals corporation (SPC) of Sri Lanka. Int J Procur Manag 1(1):1
4. Bag S, Wood LC, Mangla SK, Luthra S (2020) Procurement 4.0 and its implications on business process performance in a circular economy. Resour Conserv Recycl 152:104502
5. Caridi M, Cavalieri S, Pirovano C, Diazzi G (2004) Assessing the impact of e-procurement strategies through the use of business process modelling and simulation techniques. Product Plan Control 15(7):647–661
6. Hertrampf A, Müller H, Menezes JC, Herdling T (2015) Advanced qualification of pharmaceutical excipient suppliers by multiple analytics and multivariate analysis combined. Int J Pharm 495(1):447–458
7. Nicoletti B (2018) The future: procurement 4.0. In: Nicoletti B (ed) Agile procurement. Palgrave Macmillan, Cham, pp 189–230
8. Davis FD (1989) Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q Manag Inf Syst 13(3):319–339
9. Monczka RM, Handfield RB, Giunipero LC, Patterson JL (2009) Purchasing and supply chain management. Cengage Learning, South-Western
10. Scuotto V, Caputo F, Villasalero M, Del Giudice M (2017) A multiple buyer—supplier relationship in the context of SMEs' digital supply chain management. Product Plan Control 28(16):1378–1388
11. AlNuaimi BK, Khan M, Ajmal MM (2021) The role of big data analytics capabilities in greening e-procurement: a higher order PLS-SEM analysis. Technol Forecast Soc Chang 169:120808
12. Utama DM, Santoso I, Hendrawan Y, Dania WAP (2022) Integrated procurement-production inventory model in supply chain: a systematic review. Oper Res Perspect 9:100221
13. IBEF. https://www.ibef.org/industry/pharmaceutical-india.aspx. Last accessed on 2022/05/23
14. Singh K (2020) What ails India's pharmaceutical sector? The Economic Times
15. Kothari T, Hu C, Roehl WS (2005) e-Procurement: an emerging tool for the hotel supply chain management. Int J Hosp Manag 24(3):369–389
16. Caridi M, Cavalieri S, Diazzi G, Pirovano C (2004) Assessing the impact of e-procurement strategies through the use of business process modelling and simulation techniques. Product Plan Control 15(7):647–661
17. Ajzen I (1991) The theory of planned behavior. Organizational behavior and human decision processes 50(2):179–211
18. Zayyad M, Toycan M (2018) Factors affecting sustainable adoption of e-health technology in developing countries: an exploratory survey of Nigerian hospitals from the perspective of healthcare professionals. PeerJ 6:e4436
19. Ross J, Stevenson F, Lau R, Murray E (2016) Factors that influence the implementation of e-health: a systematic review of systematic reviews (an update). Implementation Sci 11(1):1–12
20. Saha E, Rathore P, Parida R, Rana NP (2022) The interplay of emerging technologies in pharmaceutical supply chain performance: an empirical investigation for the rise of Pharma 4.0. Technol Forecast Soc Change 181:121768

Chapter 13
Risk and Feasibility of Sustainable Techno-Eco-Env Green Supply Chain

Kapil Manohar Gumte and K. Nageswara Reddy

Abstract To evaluate the economic risk, a novel supply chain network is designed, which is a multi-period, multi-echelon in nature. The dynamic profit is evaluated in terms of net present worth (NPW), involving fluctuating money's value with time and depreciation to manage demand–supply ratios. The designed mixed integer nonlinear programming (MINLP) supply chain model incorporates technical aspects of potential site locations for manufacturing, import, inventory, and retailer, along with feasible and flexible connectivity for distribution. Economically, detailed capital and operating cost expenses are evaluated to find the cost distribution. Environmentally, greenhouse gas emissions are calculated throughout the different phases of the model's life cycle via life cycle assessment (LCA) following the government's carbon emission injunctions for the penalty of $35 per ton carbon equivalent. The SC feasible and infeasible operating conditions are identified via sensitive parameters to assist the overall strategic decision. Here, the sensitive parameter, specifically the transport cost below ₹ 3.25 per km per kg, is found to be a crucial parameter. When unit transport cost increases by 116.67%, the entire project becomes infeasible and risky, making NPW negative. Overall, the model incorporates the technical, economical, and environmental aspects as a step toward a sustainable supply chain by taking the cash crop cotton as the case study product for Pune city.

Keywords Supply chain network design · MINLP · Greenhouse gas emissions · Economic risk · Sustainable planning

---

K. M. Gumte · K. N. Reddy (✉)
Operations and Supply Chain, Indian Institute of Management Jammu, Jammu, India
e-mail: knsr481@gmail.com

K. M. Gumte
e-mail: kapil@iimj.ac.in

M. K. Tiwari et al. (eds.), *Applications of Emerging Technologies and AI/ML Algorithms*, Asset Analytics, https://doi.org/10.1007/978-981-99-1019-9_13

13.1 Introduction

The cost of cotton is increasing at timely intervals (see Fig. 13.1) due to supply shortage and higher demand, which categorically implies a demand–supply imbalance [1]. Here, the supply chain (SC) plays a vital role in meeting these demands by managing scarce resources as one of the aspects of cost reduction [2] and further can help in evaluating financial risk. After focusing firstly on economics, secondly, the SC has to take care of environmental regulations as per government injunction.

The earlier researchers have done theoretical and conceptual work in the domain of the cotton industry [3] and proposed significant suggestions related to procurement, purchase, planning, utilities, manufacturing, inventory, quality control, distribution, sales, etc. However, they missed implementing these ideas using SC quantitative modeling for real-time business. Qualitative SC models were available, e.g., Nazeer and Fuggate [1] proposed a qualitative supply chain framework based on 18 critical success factors (CSF) using a survey from domain technical experts. Bevilacqua et al. [4] used the cotton supply chain network to highlight the environmental analysis of cotton yarn and mentioned the specific upstream SC points to bring down the carbon emissions. Apart from being qualitative, these models missed the other factors related to SC's technical and economic aspects.

There did exist other types of models for cotton demand and supply, e.g., the model by Thakare et al. [6] used the linked Cobs-Douglas profit function to relate output supply and input demand, but this model has a limited scope which misses the holistic approach of layer by layer evaluation, as done by SC models. Moreover, profits obtained did not consider the depreciation and fluctuating time value of money over the time horizon.

The current paper tries to overcome the knowledge gaps mentioned by providing a quantitative SC model with the following contributions:

1. The technical approach takes care of

    - Network structure, Facility locations, and connectivity
    - Uninterrupted flow across the supply chain
    - External import in case of unmet demand to avoid stockouts
    - Warehouse distribution with inventory

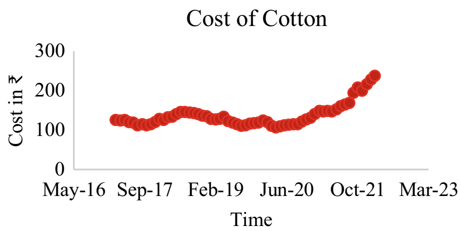


Fig. 13.1 Scatter plot showing the increasing cost of Cotton in India (Source: IndexMundi) [5]

2. Environmentally, greenhouse gas emissions based on life cycle assessment (LCA) are included to keep a check on the pollution level.
3. Economically, the fluctuating time value of money and depreciation is incorporated using the net present worth (NPW) approach. Categorically, costing is done in terms of capital expenditure (CAPEX) as a fixed cost and operating expenditure (OPEX) as transport, inventory, maintenance, and imports as a variable cost.

13.2 Model

Three-layered supply chain model is developed using multi-time period for cotton as the final product (see Fig. 13.2). The raw cotton is processed at manufacturing units, after which these are sent to the warehouses for inventory and distribution. Finally, the product is sent to demand retailers at the last layer. To avoid stockouts, the SC incorporated via external imports at distributors to avoid revenue losses and service level losses. The material flow from left to right, and money flows from right to left across the SC.

It is recommended that the reader goes through Appendix A nomenclature to understand the Equation subscript and acronyms along with Equations explanation in upcoming passages. Further, the model is developed to maximize the profit, which is calculated in terms of net present worth (NPW) (Eqs. 13.1 and 13.2), where the time value of money, earning $Earning$ (Eq. 13.3), depreciation $Depr$ (Eq. 13.4) capital $Capex$ (Eq. 13.7), and operating $Opex$ (Eq. 13.6) expenditures are involved.

$$\text{Max profit } NPW \tag{13.1}$$

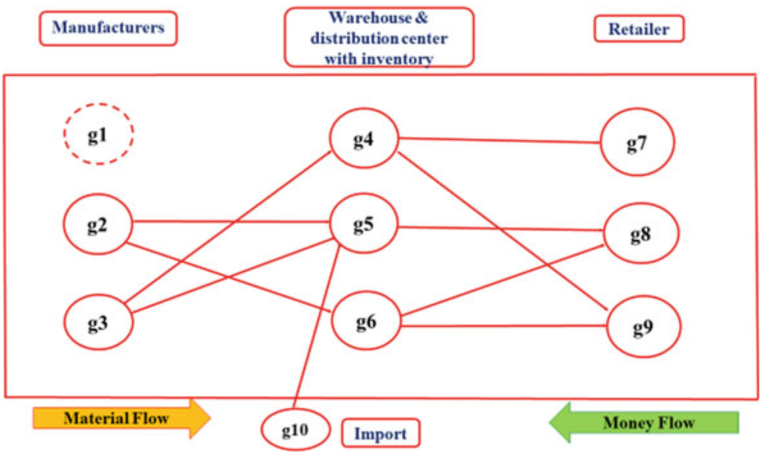


Fig. 13.2 Schematic view of designed supply chain network for time period t_3. The dotted circle for site g_1 shows the site not selected (Source: Authors)

$$NPW = \sum_{t} \frac{1}{(1+\alpha)^{N-1}}((Erning_t - Csting_t - Dpr_t) \times (1 - \Phi) + Dpr_t)\forall t \tag{13.2}$$

$$Erning_t = \sum_{p,ru} Dmd_{p,r,t} SP_{p,t}\ \forall t \tag{13.3}$$

$$Dpr_t = \frac{i}{((1+i)^N - 1)}(0.5 \times InfC_t)\ \forall t \tag{13.4}$$

$$Csting_t = Opex_t + Capex_t + GHGeC_t\ \forall t \tag{13.5}$$

$$Opex_t = InfMC_t + TranC_t + Invtd_t + ImptC_t + PrdC_t\ \forall t \tag{13.6}$$

$$Capex_t = InfC_t\ \forall t \tag{13.7}$$

The infrastructure cost $InfC$ involves a binary variable Y to be multiplied by site establishment costs across the supply chain (Eq. 13.8). The infrastructure maintenance cost $InfMC$ (Eq. 13.9) comes under operating cost $Opex$ and is calculated when the plant is not selected such that plant can be reused in the upcoming time period if selected again by the algorithm. The operating cost component further involves transport cost $TranC$ (Eq. 13.10), inventory cost $InvtCd$ (Eq. 13.11), and import cost $ImptC$ (Eq. 13.12), production or manufacturing or processing cost $PrdC$ (Eq. 13.13). The carbon greenhouse gas emission cost $GHGeC$ cost (Eq. 13.14) is also considered in the model. Equations 13.8–13.14 are based on the unitary method.

$$InfC_t = \sum_{m} Ym_{m,t}Cm_{m,t} + \sum_{d} Yd_{d,t}Cd_{d,t} + \sum_{im} Yim_{im,t}Cim_{im,t},\ \forall t \tag{13.8}$$

$$\begin{aligned} InfMC_t = &\sum_{m,t}(1 - Ym_{m,t})CMm_{m,t} \\ &+ \sum_{d,t}(1 - Yd_{d,t})CMd_{d,t} + \sum_{im,t}(1 - Yim_{im,t})CMim_{im,t},\ \forall t \end{aligned} \tag{13.9}$$

$$\begin{aligned} TranC_t = \sum_{l} UTC_l \Bigg(&\sum_{p,m,d} dmd_{m,d}Xmd_{p,m,d,l,t}Qmd_{p,m,d,l,t} \\ &+ \sum_{p,im,d} dimd_{im,d}Ximd_{p,im,d,l,t}Qimd_{p,im,d,l,t} \\ &+ \sum_{p,d,r} ddr_{d,r}Xdr_{p,d,r,l,t}Qdr_{p,d,r,l,t} \Bigg),\ \forall t \end{aligned} \tag{13.10}$$

$$InvtdC_t = \sum_{p,d} UICd_{p,du,t} invtd_{p,d,t}, \forall t \tag{13.11}$$

$$ImptC_t = \sum_{p,im} UImC_{p,t} impt_{p,im,t}, \forall t \tag{13.12}$$

$$PrdC_t = \sum_{p,d} UprdC_{p,t} Pm_{p,d,t}, \forall t \tag{13.13}$$

$$GHGeC_t = \Omega Emis_t, \forall t \tag{13.14}$$

Binary variables {0, 1} i.e., X, Y and constraints Eqs. (13.15–13.19) are used to make sure that if site Y is selected, then connectivity X is followed to that site location.

$$\sum_{p,d,l} Xmd_{p,m,d,l,t} \leq Ym_{m,t}, \forall t, \forall m \tag{13.15}$$

$$\sum_{p,m,l} Xmd_{p,m,d,l,t} \leq Yd_{d,t}, \forall t, \forall d \tag{13.16}$$

$$\sum_{p,r,l} Xdr_{p,d,r,l,t} \leq Yd_{d,t}, \forall t, \forall d \tag{13.17}$$

$$\sum_{p,d,l} Ximd_{p,im,d,l,t} \leq Yim_{im,t}, \forall t, \forall im \tag{13.18}$$

$$\sum_{p,im,l} Ximd_{p,im,d,l,t} \leq Yd_{d,t} \forall t, \forall d \tag{13.19}$$

Equation 13.20 shows the greenhouse gas emissions $GHGe$ calculation across the life cycle assessment (LCA) involving transport, production, and inventory, respectively, where the value of Ω obtained is \$ 35 per tCO_2e [7].

$$GHGeC_t = \Omega(Emis_t) = \Omega\left(EmisInvt_t + EmisManf_t + EmisTrans_t\right)\forall t \tag{13.20}$$

$$EmisInvt_t = \sum_{p,d} fd_p invtd_{p,d,t}, \forall t \tag{13.21}$$

$$EmisManf_t = \sum_{p,m} fm_p Pm_{p,m,t}, \forall t \tag{13.22}$$

$$EmisTrans_t = \sum_{l} f_l \left(\sum_{p,d,r} ddr_{d,r} Xdr_{p,d,r,l,t} Qdr_{p,d,r,l,t} \right.$$

$$+ \sum_{p,im,d} dimd_{im,d} Ximd_{p,im,d,l,t} Qimd_{p,im,d,l,t}$$

$$+ \sum_{p,m,d} dmd_{m,d} Xmd_{p,m,d,l,t} Qmd_{p,m,d,l,t} \Bigg), \forall t \quad (13.23)$$

Next, Eqs. 13.24–13.33 represent the material flow across the supply chain. Equation 13.24 shows the production at the manufacturing site Pmu following plant capacity $PmMax$ and distribution Qmd (Eq. 13.25) to the following distribution layer with transport limitations $QmdMax$ (Eq. 13.26).

$$Pm_{p,m,t} \leq PmMax_{p,m,t} Ym_{m,t}, \forall t, \forall p, \forall m \quad (13.24)$$

$$Pm_{p,m,t} = \sum_{d,l} Qmd_{p,m,d,l,t}, \forall t, \forall p, \forall m \quad (13.25)$$

$$Qmd_{p,m,d,l,t} \leq QmdMax_{p,m,d,l,t} Xmd_{p,m,d,l,t}, \forall t, \forall p, \forall m, \forall d, \forall l \quad (13.26)$$

The distribution layer has inventory $invtd$ (Eq. 13.27) and import connections $import$ (Eq. 13.28) with capacity constraints $invtdMax$ and $imptMax$ where inventory comes in from the previous time period, manufacturers, and import sites and goes out to the following retailers (Eq. 13.29).

$$\begin{aligned} invtd_{p,d,t} =& invtd_{p,d,t-1} + \sum_{m,l} Qmd_{p,m,d,l,t} \\ &+ \sum_{im,l} Qimd_{p,im,d,l,t} - \sum_{r,l} Qdr_{p,d,r,l,t}, \forall t, \forall p, \forall d \end{aligned} \quad (13.27)$$

$$invtd_{p,d,t} \leq invtdMax_{p,d,t} Yd_{d,t}, \forall t, \forall p, \forall d \quad (13.28)$$

$$impt_{p,im,t} \leq imptMax_{p,im,t} Yim_{im,t} \forall t, \forall p, \forall im \quad (13.29)$$

The import distribution to the distributor layer $Qimd$ (Eq. 13.30) has transport limitations $QimdlMax$ (Eq. 13.31), followed by a similar constraint $Qdrl$ to the following retail layer $QdrlMax$ (Eq. 13.32) till all the demand Dmd is met (Eq. 13.33). The model makes sure that demand is met under all circumstances.

$$impt_{p,im,t} = \sum_{d,l} Qimd_{p,im,d,l,t}, \forall t, \forall p, \forall im \quad (13.30)$$

$$Qimdl_{p,im,d,l,t} \leq QimdlMax_{p,im,d,l,t} Ximd_{im,d,t}, \forall t, \forall p, \forall im, \forall d, \forall l \quad (13.31)$$

$$Qdrl_{p,d,r,l,t} \leq QdrlMax_{p,d,r,l,t} Xdr_{d,r,t}, \forall t, \forall p, \forall d, \forall r, \forall l \quad (13.32)$$

$$\sum_{d} Qdrl_{p,d,r,l,t} = Dmd_{p,r,t}, \forall t, \forall p, \forall r, \forall p, \forall l \tag{13.33}$$

13.3 Methodology

The methodology adopted here is NP-hard mixed integer nonlinear programming (MINLP). The optimization model is solved in a 32 GB i7 processor system with a GAMS_38.3 programming version where a Dicopt solver is used to solve 93 discrete integer variables, 435 single equations, and 252 single variables of a computationally challenging model.

13.4 Data Collection

The data is collected from authenticated sources and is not shown here due to space constraints and can be made available upon asking.

13.5 Results and Discussions

13.5.1 Site Location and Connectivity

From Table 13.1, one can observe that number of sites needed for the last time period t_3 is one less than earlier time periods t_1 and t_2, as this can be seen for one reduced site selection for manufacturing unit. This happened because demand for the third time period is fulfilled via leftover inventories at the distributor site from previous time periods (Eq. 13.28) to save the manufacturing cost (Eq. 13.14) and inventory cost (Eq. 13.12) as per SC objective to maximize profit (Eq. 13.1). Now, the way these sites are connected is shown in Fig. 13.2 for time period t_3.

Table 13.1 Site selection for entire time horizon (Source: Authors)

Site selection	t_1	t_2	t_3
Manufacturing (m)	3	3	2
Distributor (d)	1	1	1
Importer (im)	3	3	3
Retailer (r)	3	3	3
Total	10	10	9

13.5.2 Cost Analysis and GHGe

The distribution of expenses across the time period is shown in Table 13.2, where costing decision variables values are shown. The capital expenditure contributes 6.18%, whereas operating expenditure does 93.82%. Out of the total cost (see Fig. 13.3), transport bears 49.88% (maximum), followed by production 32.91%, import 10.98%, infrastructure 6.18% cost, and other important sections. Transport bears more cost due to logistics and distribution, but imports clearly indicate insufficient indigenous production and a supply shortage. Hence external import plays a vital role in fulfilling the demand at the expense of a higher per-unit cost of cotton to avoid customer's credibility, which is difficult to quantify.

In Table 13.2 and Fig. 13.2, the GHGe cost is 0.01%. The value of 0.01% appears to be small in percentage as it only involves a single city; instead, the contribution will

Table 13.2 Cost analysis of SC (Source: Authors)

	t_1	t_2	t_3	Total	
NPW (₹)	3,148,261	5,079,417	4,135,368	12,363,046.4	
Total cost (₹)	9,847,280	11,819,597	9,402,719	31,069,595.1	100%
Capital cost (₹)	690,000	690,000	540,000	1,920,000	6.18%
Infrastructure cost (₹)	690,000	690,000	540,000	1,920,000	
Depreciation (₹)	104,229.6	104,229.6	81,571	290,030.211	
Operating cost (₹)	9,157,280	11,129,597	8,862,719	29,149,595.1	93.82%
Production cost (₹)	3,255,000	4,007,500	2,962,500	10,225,000	
Transport cost (₹)	5,111,550	5,806,260	4,579,590	15,497,400	
Inventory cost (₹)	2500	2500	0	5000	
Infrastructure maintenance cost (₹)	0	0	7500	7500	
Import cost (₹)	787,500	1,312,500	1,312,500	3,412,500	
Greenhouse gas emissions (tCO_2e)	0.28	0.32	0.24	0.84	
Emission cost (₹)	729.76	836.56	628.79	2195.10	0.01

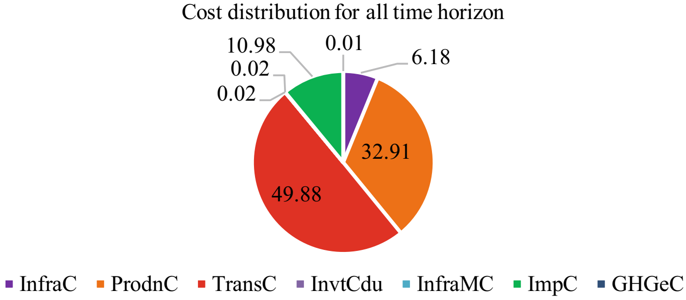


Fig. 13.3 Overall cost distribution across SC (Source: Authors)

Table 13.3 Effect of transport cost on profit NPV (Source: Authors)

Unit transport cost	% Transport cost increase	*NPW*
1.5	0	1.24E+07
2.5	66.66667	2.63E+06
3	100	8.47E+05
3.25	116.6667	−1.89E+06

be proportionally large as geographical terrain increases. Hence GHGe cost cannot be neglected.

13.5.3 Threshold Value of NPV

For a project to be economically feasible, the NPW must be zero or positive. When sensitivity analysis is done, the transport cost is found to have a major effect on NPV in relation to other parameters such as production cost and import cost, etc., and is not shown here due to space limitation. When the unit transport cost is made at 116.67% of the initial value, the profit NPW obtained is negative, making the project economic infeasibility, and this is where NPW reaches threshold value mainly due to transportation. This also means SC should operate with unit transport costs below ₹ 3.25 per km per (Table 13.3).

13.6 Conclusion

We presented a novel green supply chain network design model in the current work, incorporating sustainable techno-economical-environmental multi-period, multi-echelon MINLP-based NP-hard problem. The model not only finds the optimal location and connectivity to have flexible, dynamic SC but also does the greenhouse gas emission calculation and its penalty across the various LCA. Here OPEX incorporating the time value of money for profit NPV is found to dominate the expenses (93.82%) in which transport cost is the major (49.88%) and most sensitive one causing to project risk. The project's pivotal point is that the threshold value of unit transport cost of ₹ 3.25 per km per kg corresponds to 116.67% raise.

Appendix A—Nomenclature

It is highly advised to go through this nomenclature section with an explanation of subscripts and relate with model Eqs. 13.1–13.30 and its explanation in Sect. 13.2. Equations use values of scalars involving N as total time period, α as a discount factor, Ω as carbon penalty, i as annual interest rate, Φ as Goods and Service Tax (GST). All location sets are encompassed by g with subsets of manufacturer m, importer im, distributor d, and customer c sites across the echelons. Sets of products p (single) and a Transport mode set l (single) are used along with each time t period. SP stands for selling price, C represents infrastructure establishment costing, and CM represents infrastructure maintenance cost followed by the respective supply chain layer name, e.g., Cm and CMm represents manufacturing plant establishment cost and maintenance cost per plant respectively. U shows unit costing parameter, i.e., UTC, $UICd$, $UImC$, $UprdC$ represents unit transport cost, unit import cost, and unit production cost, respectively followed by their respective supply chain layer subset f_l, fm, fd show the emission factor for transport, manufacturing unit, and inventory unit. d show the distance followed by the site location acronym indicating start to end destination in the supply chain. Q shows the transfer quantity between two layers followed by connecting layer sites acronym and explicitly indicating maximum quantity Max. Y and X indicate binary integers for site selection and connectivity between sites, followed by the site location acronym.

References

1. Nazeer S, Fuggate P (2019) Sustainability framework for farm level cotton supply chain management. In: Proceedings of international conference on industrial engineering operations management, vol 2019, no MAR, pp 3751–3757
2. Simchi-Levi D, Kaminsky P, Simchi-Levi E (2004) Managing the supply chain: definitive guide. Tata McGraw-Hill Education
3. Rieple A, Singh R (2010) A value chain analysis of the organic cotton industry: the case of UK retailers and Indian suppliers. Ecol Econ 69(11):2292–2302. https://doi.org/10.1016/j.ecolecon.2010.06.025
4. Bevilacqua M, Ciarapica FE, Mazzuto G, Paciarotti C (2014) Environmental analysis of a cotton yarn supply Chain. J Clean Prod 82:154–165. https://doi.org/10.1016/j.jclepro.2014.06.082
5. Indexmundi Cotton monthly price—Indian rupee per kilogram. https://www.indexmundi.com/commodities/?commodity=cotton&months=60&currency=inr. Accessed 29 April 2022
6. Thakare SS, Shende NV, Shinde KJ (2012) Mathematical modeling for demand and supply estimation for cotton in Maharashtra. Int J Sci Res Publ 2(3):1–5
7. Capital CC (2021) Value of carbon market update 2021. https://carboncreditcapital.com/value-of-carbon-market-update-2021-2/. Accessed 07 May 2022

Chapter 14
Steel Price Forecasting for Better Procurement Decisions: Comparing Tree-Based Decision Learning Methods

Ravi Ram Reddy Palvai and Arshinder Kaur

Abstract Most of the manufacturing firms are exposed to the risk of Commodity price volatility which can have a substantial impact on their operational and financial decisions. One of the most used metal commodities in manufacturing industries is Steel and the known fact is that steel prices vary over the time making it difficult for procurement decisions. With the current advancements in the field of Artificial intelligence and Machine Learning, there is a growing emphasis on development of accurate forecasting methods for commodity prices that plays an important role in procurement decisions for manufacturing firms. The current paper aims to develop ML-based forecasting models by employing Tree-based algorithms namely Regression trees and Random Forests to forecast steel prices. Past 10 years monthly historical data of several variables impacting the prices of Steel and the prices of steel are considered for forecasting Steel Prices. The results reveal that the proposed methods present a promising forecast with high accuracy. The major performance metric used in this study to measure forecasting accuracy is Mean Absolute Percentage Error (MAPE). The tree-based models used in this paper gave MAPE of less than 5%, indicating that they outperform other traditional forecasting approaches used in the literature, making purchase decisions easier.

Keywords Steel price forecasting · Procurement · Machine learning (ML) · Regression trees · Random forest

R. R. R. Palvai (✉) · A. Kaur
Department of Management Studies, Indian Institute of Technology, Madras, India
e-mail: palvai.raviram@gmail.com

A. Kaur
e-mail: arshinder@iitm.ac.in

M. K. Tiwari et al. (eds.), *Applications of Emerging Technologies and AI/ML Algorithms*, Asset Analytics, https://doi.org/10.1007/978-981-99-1019-9_14

14.1 Introduction and Background

Steel is an essential metal commodity and is used as a primary raw material in many manufacturing industries. It is the most essential raw material for the construction sector and auto manufacturers, while the auto sector is the second largest sector in steel consumption only next to the Construction sector making it clear that these industries' operational costs depend on Steel. So, any fluctuation in the steel prices has a direct impact on the firm's profitability. The steel prices being highly volatile, there is a heavy need in forecasting these prices in order to make profitable procurement decisions. If the procurement plan is not appropriate, controlling the cost of steel purchasing is rather risky since the steel prices are highly volatile. Hence, a proper forecasting model with better accuracy is required in order to make purchase decisions efficient.

The market for mineral commodities is highly volatile and constantly changing. Traditional forecasting methods such as Exponential Smoothing, Autoregressive Integrated Moving Average (ARIMA), Autoregressive Conditional Heteroskedasticity (ARCH), and Generalized ARCH (GARCH), as well as econometric models, may not accurately capture the dynamic and time-related nature of metal commodity markets when trying to predict their prices. While time series models like ARIMA may be able to represent past trends with some degree of accuracy, they may not be able to predict sudden changes effectively because of the use of linear models [1].

Many works have been studied to forecast the prices of steel by using traditional forecasting methods. Adli [2] used ARIMAX model with exploratory variables to forecast steel prices in Turkey. Malanichev and Vorobyev [3] used Multiple Regression model to forecast global annual steel prices. But all these models used in the literature to forecast steel prices didn't give promising results. Therefore, advanced forecasting techniques using Machine Learning (ML) can be used to achieve forecasting results with higher accuracy when compared to traditional forecasting methods. ML techniques can deal with huge amounts of data and they are able to find hidden patterns and can learn from them.

This paper considers such ML techniques known as Tree-based Decision learning models that can be used to predict the Steel prices. Decision tree algorithm forms the basis of these models. A Decision tree works by splitting a data set in order to train a model through a recursive partitioning process, and then the model is used to predict the value of a target variable based on the independent variables [4]. Liu et al. [5] showed that Decision Trees method is capable of accurately and reliably predicting copper prices in both short term and long term. Chen and Xin [6] proved that Decision tree models performed better than the conventional models like Multiple Linear Regression and ARIMA in predicting Crude Oil prices.

We can measure the performance of this tree-based models on the basis of various metrics such as Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE), and Root mean square error (RMSE). MAPE is used as the performance metric in this study.

14.2 Methodology

In this section, we discuss various decision learning models considered for the purpose of forecasting steel prices. These models are also known in the machine learning literature as tree-based methods [7]. The tree-based methods applied in our methodology include Regression trees and an Ensemble method called Random Forest.

14.2.1 Regression Trees

Decision tree learning is a prediction tool commonly used in data mining [4]. Decision trees can be applied to both regression and classification problems. Decision trees applied to Regression problems where the output variable is numerical are called Regression trees [4, 7].

Decision trees are tree-shaped diagrams in which a data set is divided into several subsets using a sequence of dichotomous classifications. The proportion of unexplained variance decreases with each sub-split, but at the cost of a more complicated model. The end product is a tree with so-called decision nodes and leaf nodes. The nodes in the final layer of the tree are called leaf nodes. Specifically, following a path from the root node (top) to a leaf node (bottom) yields a set of decision rules that predict the desired output variable [8].

The tree structure is created by partitioning the data recursively using impurity splitting criteria until a leaf node is obtained. The process follows a top-down approach, starting at the top of the tree where all the observations are initially grouped together. The predictor space is divided into smaller regions with each split. The tree-building process is considered greedy, which means that the best split is selected at each step based on the available data at that point, rather than looking ahead and choosing a split that may lead to a better tree in the future step.

To perform recursive binary splitting, we start by selecting a predictor Xj and a cutpoint "k" that splits the predictor space into two regions: $\{X|Xj < k\}$ and $\{X|Xj \geq k\}$. We select the predictor and cutpoint combination that results in the maximum reduction of the Residual Sum of Squares (RSS). This involves considering all predictors X1, …, Xp, and all possible values of the cutpoint "k" for each predictor. The objective is to obtain the lowest possible RSS for the resulting tree. We then repeat the process by identifying the best predictor and cutpoint to further split the data into one of the two previously identified regions. This results in three regions, and we again look for the best predictor and cutpoint to minimize the RSS within each of these regions. This process continues until a stopping criterion is met (with a maximum depth or minimum node size as the stopping criteria). It is essential to note that the regions formed are based on the value of a single predictor Xj and cutpoint. This process is repeated for each region created, and the resulting tree is built recursively.

The goal is to find the optimal predictor and cutpoint combination that leads to the lowest possible RSS, which helps in identifying the best splits for the tree [7].

14.2.2 Random Forest

Random forest is an ensemble technique that consists of multiple Decision trees and is used for both classification and regression problems [9]. For classification tasks, the output of the random forest is the class selected by most trees. For regression tasks, the mean or average prediction of the individual trees is returned [10]. Random Forests for regression extend the single regression tree by building multiple regression trees based on random subsets of training data and predictors. Forecasts based on random forests take the average of the multiple fitted regression trees' predictions.

Regression trees tend to overfit with high variance and low Bias. Using a statistical technique known as Bagging (Bootstrap Aggregating), Random Forest can reduce the variance while leaving the low bias roughly unchanged [11]. Bagging is a technique used to increase the accuracy of any statistical methods (regression trees in our case). It involves creating multiple prediction models by fitting separate regression trees on random subsamples taken from the training observations. The final prediction is then calculated as an average of all the resulting predictions, which reduces variance. This approach improves the accuracy of the model by reducing overfitting and increasing stability. By generating multiple subsamples and fitting models on each, bagging helps to create more robust models that can generalize well to unseen data [12].

In random forest method, after the first random procedure which is bagging which generates a number of new training sets by randomly sampling with replacement from original training set and combines the results of trees fitted by new training sets obtained above, there is also a second random procedure which is random subspace selection, which selects a random subset of features at each node of the tree. The main reason for doing this is to reduce the correlation of individual trees (i.e., the predictive ability being similar across the regression trees) in random forests [13].

14.3 Data Description and Empirical Results

14.3.1 Variable Selection and Data Collection

Several variables influence steel prices. Some of the key explanatory variables taken into account while forecasting Steel prices include

(i) **Supply and Demand of Steel** Like any other commodity in the world, Steel price is mainly affected by its Supply and Demand. For our study, Supply and Demand of Steel are considered in terms of total Production Quantity and

Consumption Quantity of Steel in India, Production and Consumption Quantity of Steel in China and Global Production Consumption Quantity of Steel. Production and Consumption Quantity of Steel in China is also considered because China is the biggest producer and consumer of Steel hence a deciding factor.

(ii) **Steel prices in China** China being the biggest producer and consumer of steel in the world, Steel price in China can be a deciding factor for the global steel prices and also Indian Steel prices.

(iii) **Raw material Prices** The price and availability of raw materials in the production of steel also plays a major role in deciding Steel prices. Major raw materials involved in the production of Steel include Iron Ore and Coking Coal. Hence, Iron Ore and Coking Coal prices are considered as independent variables in forecasting Steel prices.

(iv) **Crude Oil Prices** Energy prices affect the price of almost every commodity. Hence, Crude Oil price is considered as one of the affecting variables for the steel prices.

(v) **Macroeconomic Indicators** Most of the commodity prices are related to the economic activity in that country. Hence, Macroeconomic indicators like YoY GDP growth, CPI Inflation rate, Interest rates (repo rate in India), Exchange rates (USD-INR), Index of Industrial Production, Purchasing Managers Index are considered as independent variables in forecasting Steel prices.

Based on the explanation above and subject to availability, Monthly data for all of the following factors has been collected over the last 10 years, i.e., from Jan'2012 to Dec'2021. The variable list is shown in Table 14.1.

14.3.2 Correlation Between Steel Price and Selected Variables

We first have to confirm that the steel prices are indeed correlated with all the selected variables. To quantify the observed qualitative relationships between Steel and other selected factors, we computed Pearson cross-correlation coefficients. The Pearson cross-correlation coefficient, which ranges from +1 to 1, measures the linear correlation, and thus the degree of linear dependency, between two variables. A coefficient of +1 is total positive correlation, 0 is no correlation, and −1 is total negative correlation. In more detail, Pearson's correlation coefficient is the covariance of two variables divided by the product of their standard deviations [5, 14]. For example, for the Steel price, P_s, and Iron Ore price, P_i, the correlation coefficient is defined as

$$\rho_{s,i} = \frac{\mathrm{cov}(P_s, P_i)}{\sigma(s) * \sigma(i)} \tag{14.1}$$

Table 14.1 Variables affecting steel prices

S. no.	Variables affecting Indian steel prices	Units	Source
1	Production quantity of steel in India	MnT	worldsteel.org
2	Consumption quantity of steel in India	MnT	worldsteel.org
3	Production quantity of steel in China	MnT	worldsteel.org
4	Consumption quantity of steel in China	MnT	worldsteel.org
5	Global production quantity of steel	MnT	worldsteel.org
6	Global consumption quantity of steel	MnT	worldsteel.org
7	Steel prices in China	USD/Ton	Steelmint
8	Price of iron ore in India (fines, 64%)	INR/Ton	Steelmint
9	Price of iron ore in China (fines, 62%)	USD/Ton	Steelmint
10	Price of coking coal premium CNF (Aus)	USD/Ton	Steelmint
11	Crude oil import basket price	USD/Barrel	ppac.gov.in
12	GDP growth of India	%	stats.oecd.org
13	CPI inflation rate	%	investing.com
14	Interest rates (repo rates)	%	investing.com
15	Exchange rates (USD-INR)	INR	investing.com
16	Change in industrial production (IPI)	%	investing.com
17	Purchasing managers index (PMI)	–	Markit

Table 14.2 Forecast performance of tree-based models for steel prices

	MAPE (%)		
	1 Month	2 Months	3 Months
Regression tree	5.18	4.47	4.18
Random forest	2.95	4.24	3.72

Except for crude oil prices, GDP growth, CPI inflation rate, change in industrial production, and the PMI index, all other explanatory variables are highly correlated (higher than 0.5). More specifically, the steel price is highly correlated with Iron Ore price and Coking Coal price which are direct raw materials used in the production of steel. This high correlation can be interpreted easily from the dependence of Steel prices on its raw material prices. Also, all the other variables which showed correlation with steel prices are considered as independent variables for forecasting steel prices. The reason for the lack of association between the above-mentioned variables, namely, crude oil price and other macroeconomic indicators, is beyond the scope of this study.

14.3.3 Framework and Results

The variables which are highly correlated with the steel prices are considered as independent variables and the target variable is Steel Prices in India (HRC, 2 mm).

The dataset with all the independent variables and target variable is first split into random training and test samples with test size of 25% and the remaining 75% as training data. The training sample is used to fit the model, while the test sample is used to evaluate the model's forecasting accuracy. We repeat this approach using the same random training and test sample for both Regression Tree and Random Forest model.

In these two models, we consider lagged independent variables to train the model i.e., if we are doing a 1-month price forecast, we take 1-month lagged exploratory variables corresponding to the present month's target variable. For example, we train the model using April month's explanatory variable data to forecast May month's steel price. Similarly, March month's explanatory variable data to forecast April month's steel price. In the same way, if we are doing a 2-month price forecast, we take 2-month lagged variables corresponding to the present month's target variable i.e., to forecast March month's Steel price, we use January month's explanatory variables.

To evaluate the model's forecasting performance, we use Mean Absolute Percentage Error (MAPE) as the measure of forecasting error. Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) can also be used as performance metrics but MAPE is more standard when measuring forecasting error [15]. Lower the MAPE value, higher will be the accuracy or performance. Formally, the MAPE is given by the following expression

$$\text{MAPE} = \frac{100\%}{n}\sum_{t=1}^{n}\left|\frac{A_t - F_t}{A_t}\right| \tag{14.2}$$

where A_t is the Actual value and F_t is the Forecast value.

After evaluating the MAPE of the test data to evaluate the model's performance, it was discovered that the MAPE values are substantially lower (less than 5%), indicating that the models' accuracy is quite good. The table below shows the MAPE values for both Regression Tree and Random Forest models for 1, 2, and 3 months forecasts.

From the Table 14.2, it can be noted that the tree-based models showed higher accuracy in forecasting Steel Prices. Random Forest model with lower MAPE outperformed Regression tree model in terms of accuracy.

14.4 Conclusions

The current work suggests Tree-based Machine Learning algorithms to forecast steel prices instead of traditional forecasting methods used in the literature to increase the performance of the forecasting model. We have used several explanatory variables including some macroeconomic variables that showed correlation with the steel prices. Using monthly data for the past 10 years, our empirical analysis shows that Random Forest model outperformed Regression tree model in terms of accuracy. Although steel price has been used as the target variable in our study, these tree-based methods are easily applicable to forecast other commodity prices as well, except that the explanatory variables might differ. In this study though the models are used for short-term forecast only (because of limited past data), it can also be extended for long term if we have enough data (or weekly data for 10 years). There are also several advantages in using these tree-based methods compared to the traditional regression methods including the fact that these models don't need to make any assumptions about how the factors are related or distributed. More advanced tree-based ensemble models like Adaptive Boosting (AdaBoost), Gradient Boosting Decision trees (GBDT), Extreme Gradient Boosting trees (XGBT), etc., can be used to increase the performance of the forecast when dealt with huge data.

This accurate forecasting of Steel can hugely benefit procurement decisions for construction and automobile industries since steel is their major raw material. One can develop an optimal order plan based on the predicted steel price, which can have a significant impact on the firm's profitability.

References

1. Cortez CT, Saydam S, Coulton J, Sammut C (2018) Alternative techniques for forecasting mineral commodity prices. Int J Min Sci Technol 28(2):309–322
2. Adli KA (2020) Forecasting steel prices using ARIMAX model: a case study of Turkey. Int J Bus Manag Technol
3. Malanichev AG, Vorobyev PV (2011) Forecast of global steel prices. Stud Russ Econ Dev 22(3):304–311
4. Breiman L, Friedman JH, Olshen RA, Stone CJ (2017) Classification and regression trees. Routledge
5. Liu C, Hu Z, Li Y, Liu S (2017) Forecasting copper prices by decision tree learning. Resour Policy 52:427–434
6. Chen E, He XJ (2019) Crude oil price prediction with decision tree-based regression approach. J Int Technol Inf Manag 27(4):2–16
7. James G, Witten D, Hastie T, Tibshirani R (2013) An introduction to statistical learning, vol 112. Springer, New York, p 18
8. Namazkhan M, Albers C, Steg L (2020) A decision tree method for explaining household gas consumption: the role of building characteristics, socio-demographic variables, psychological factors and household behaviour. Renew Sustain Energy Rev 119:109542
9. Breiman L (2001) Random forests. Mach Learn 45(1):5–32
10. Kam HT (1995) Random decision forest. In: Proceedings of the 3rd international conference on document analysis and recognition, vol 1416. Montreal, Canada, August, p 278282

11. Breiman L (1996) Bagging predictors. Mach Learn 24(2):123–140
12. Díaz JD, Hansen E, Cabrera G (2020) A random walk through the trees: forecasting copper prices using decision learning methods. Resour Policy 69:101859
13. Zheng J, Liu Y, Ge Z (2022) Dynamic ensemble selection based improved random forests for fault classification in industrial processes. IFAC J Syst Control 20:100189
14. Pearson K (1895) Notes on regression and inheritance in the case of two parents. Proc R Soc Lond 58:240–242
15. Makridakis S (1993) Accuracy measures: theoretical and practical concerns. Int J Forecast 9(4):527–529

Chapter 15
An Interactive Game Theory Analytics to Model the Panic Buying When the Downstream Supply Chain Channel Partners Undergo Horizontal Coopetition

Sarin Raju, T. M. Rofin, and S. Pavan Kumar

Abstract During COVID-19, many large economies like European Union, the U.K., Japan and Mexico officially encouraged companies to undergo coopetition to counter commotions, hoardings and stock-piling. Observing the benefits, many companies and economies are formally coming up with official norms for coopetition. This study suggests a novel interactive strategic algorithm using Game Theory to model the pricing decisions and quantify the optimal order quantity and profit during coopetition. For the analysis, we assumed a dual-channel supply chain consisting of the manufacturer, retailer and e-tailer. The Stackelberg game was used to model the interaction between the upstream and downstream partners. We assumed coopetition between the downstream retailer and e-tailer and modelled the interaction. Later we analysed the performance of the channel partners under normal buying conditions and panic buying and compared them to derive the propositions. During the coopetition model, the retailer's optimal price, order quantity, and profit increased during panic buying, whereas only the optimal price increased for the e-tailer. Manufacturer could reap maximum benefits during the panic buying period. The interactive analytics developed can aid the management practitioners in developing decision support systems and multi-agent systems during coopetition.

Keywords Game theory · Multi-agent systems · Coopetition · Crisis cartel

S. Raju (✉) · S. Pavan Kumar
National Institute of Technology Karnataka (NIT-K), Surathkal, Mangaluru, India
e-mail: sarinraju.207sm004@nitk.edu.in; mail2rajusarin@gmail.com

T. M. Rofin
National Institute of Industrial Engineering (NITIE), Mumbai, Maharashtra, India

M. K. Tiwari et al. (eds.), *Applications of Emerging Technologies and AI/ML Algorithms*, Asset Analytics, https://doi.org/10.1007/978-981-99-1019-9_15

15.1 Introduction

Competing businesses may cooperate for a variety of reasons, and there are numerous ways that the customers and the companies can get benefit from these collaborations. This idea leads to coopetition [1–3], a business condition where the organisation competes and cooperates simultaneously. For example, many competing authorities of large economies like the European Union, U.K., Japan, Mexico have issued official guidelines for companies to pool their experience and identify best practices to collaborate, cooperate and compete during COVID-19 [4–7]. They suggest that this will ensure supply, avoid supply-chain commotions, hoardings, stock-piling, etc. Observing the benefits, many companies and economies are formally coming up with official norms for coopetition. For instance, Pfizer and BioNTech reached a deal in March 2020 to work together on COVID-19 vaccine. In accordance with this business environment, we develop an interactive analytics to model the interaction of the supply chain partners when the downstream partners undergo coopetition. Later we analyse the impact of panic buying [8–10] on the same supply chain. As far as we are aware, this is the first game-theoretic study that examines the effects of panic buying or supply chain disruptions when the downstream partners engage in coopetition.

Panic buying occurs when people engage in surplus buying and storing goods extensively in their homes, expecting a shortage of supplies in the near future. This usually affects supply chains and creates drastic and severe disruptions. Panic buying usually extends from the period between the declaration of the lockdown and the stringent implementation of lockdown. In most countries, there will be a minuscule time between the declaration of lockdown by the authorities and implementation of the lockdown—this span for about 48 h in most of the countries [11, 12]. During panic buying, people prefer to have immediate gratification [13]. As a result, consumers depended more on retailers. Ambiguity about the future also forced people to avoid the e-tailer channel during the initial period. Based on the above-mentioned business condition, we formulate the following research objective.

RO:1 Develop an interactive analytics to analyse the supply chain partners' performance when the downstream partners undergo coopetition.

RO:2 Analyse the impact of panic buying on downstream channel partners when the downstream channel partners undergo coopetition.

The interactive analytics developed was found to be successful in modelling the pricing decisions of the normal period and panic buying period. The optimal price of all the downstream channel partners increased, whereas the optimal order quantity increased only for the retailer. Consequently, the optimal profit increased only for the retailer. The manufacturer was the channel partner who could reap maximum benefit from this condition as a tremendous increase was observed in his profit.

In the following section, we report an account of the existing studies in the field.

15.2 Literature Review

The internet has reached every corner of the world. Thanks to the development of higher generation telecom facilities. Along with that, this period witnessed the growth of many prominent third-party logistics companies like FedEx, UPS and Dehlivery. As a result, many companies like IBM, Apple, Dell, HP, Nike and Walmart started opening their own online channel to reach out to the customer. This gives rise to a concept called partially integrated dual-channel supply chain, in simple Dual-Channel Supply Chain (DCSC) [14]. After that, there were many studies conducted to study the pricing [15, 16], inventory policies [17, 18] and channel coordination [19, 20] in integrated DCSC. A change in the DCSC studies was brought by Rofin and Mahanty [21] by bringing e-tailers into the studies and thereby adding decentralised DCSC into the literature. Our study mainly addresses the disruptions in decentralised DCSC. Since there are no studies in the field, we are reviewing the disruption studies in integrated DCSC.

Due to the intrinsically complex behaviour of supply chains and the dynamic nature of the market, disruptions [22–25] are usual. The disruptions in a partially integrated DCSC were first studied by Huang et al. [24], using a linear demand equation and Stackelberg game. Using the same equation and game, Huang et al. [26] studied the disruption in production cost and found that the production costs need to be revised during disruption. He also found that the actual production cost has some robustness during disruption. Cao [27] studied the disruptions in the demand using the Stackelberg game and derived the optimal profit during normal and demand disruptions. He then compared them to quantify the information value and developed a revenue sharing method for coordination. Zhang et al. [28] found that coordinating a dual-channel supply chain could achieve the integrated profit of the supply chain and alleviate the channel conflict.

Soleimani et al. [22] analysed optimal decisions in a dual-channel supply chain under simultaneous demand and production cost disruptions. They discovered a strong correlation between production costs and demand interruptions, which affects price and production choices. Yan et al. [29] show that the amount of change in decision-making is a linear function of the amount of demand disruption, and it is independent of the risk-averse coefficient. They learned that the two models' optimal system production exhibits stability. Tang et al. [30] analysed channel competition and coordination of a dual-channel supply chain during the simultaneous disruption of demand and cost.

Huang et al. [31] made a change in the conventional DCSC disruption studies by introducing the concepts of different supply chain power structures: the manufacturer Stackelberg game, the retailer Stackelberg game and the vertical Nash game. They improved the performance of the supply chain by using the virtual production cost function. Service and pricing strategies with competition and cooperation in a DCSC with demand disruption were studied by Pi et al. [32]. They derive the Stackelberg-Nash equilibrium in the retailer cooperation and independence models using game theory and the two-stage optimisation method. Rahmani and Yavari [25] analysed the pricing policies of dual-channel green supply chain during demand disruptions and found that when a demand disruption occurs, the original production quantity determined based on the predicted demand will have some robustness in both the centralised and decentralised dual-channel green supply chain.

By analysing the existing academic studies, we found scant literature in the following areas.

(a) To the best of our knowledge, there is no game-theoretic analytics for the coopetition between downstream channel partners in a DCSC.
(b) Very scant literature analyses panic buying disruption in a decentralised DCSC.

We use the following analytics to address the research gaps.

15.3 Interactive Analytics and Models

We are assuming a DCSC [29, 33] consisting of a manufacturer and two downstream channel partners [34]. We employ Stackelberg game [25, 31] to study the interaction between the upstream channel partner and downstream partners. We also assume that the retailer and e-tailer enter into coopetition [1–3] to reduce the disruption. We propose a basic linear demand function, $D = a - \lambda p$ [24, 25] to relate the relationship between the retailer and e-tailer. Here, the term a stands for 'base market potential', which is defined as the level of demand for the good at zero cost. We also assume that the demand of the product is sensitive to the price [31, 35, 36] and is denoted by own price elasticity, λ. It can be defined as the change in demand due to a unit change in price. These presumptions led to the derivation of the following equations.

$$\text{Demand for the retailer, } D_r = a_r - \lambda p_r + \gamma p_e \tag{15.1}$$

$$\text{Demand for the e-tailer, } D_e = a_e - \lambda p_e + \gamma p_r \tag{15.2}$$

Here γ signifies cross-price elasticity. Additionally, we suppose that $\lambda, \gamma > 0$ and $\lambda > \gamma$.

We presume that the manufacturer and downstream channels engage in a Stackelberg game. Being the Stackelberg leader, the manufacturer sets the wholesale price.

$$\begin{aligned}\text{Profit of the Retailer} = \pi_r &= (P_r - w)D_r \\ &= (P_r - \mathrm{w})(a + \gamma P_e - \lambda P_r)\end{aligned} \tag{15.3}$$

$$\begin{aligned}\text{Profit of the e-tailer} = \pi_e &= (P_e - \mathrm{w})D_e \\ &= (P_e - \mathrm{w})(a - \lambda P_e + \gamma P_r)\end{aligned} \tag{15.4}$$

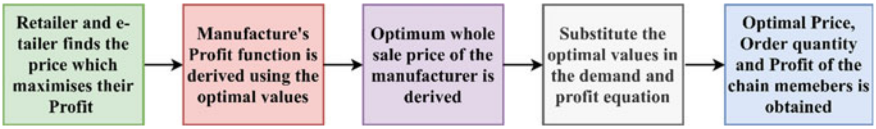


Fig. 15.1 Backward induction analytics

$$\text{Profit of the Manufacturer} = \pi_m = (\mathrm{w} - s)(Q_r + Q_e) \tag{15.5}$$

Here, Q_r, Q_e represent the relative order quantities of retailers and online retailers, and s stands for the unit production cost. Backward induction can be used to determine the wholesale price, w as shown in Fig. 15.1. We have assumed disruption in base market potential [25, 27] and $\Delta \mathrm{a}_r$ represent the disruptions in the base market potential of retailer.

Based on the above-mentioned basic equations and backward induction analytics shown in Fig. 15.1, we have formulated the following pseudocode for the study.

Pseudocode of the novel interactive analytics when downstream channels undergo coopetition

Interactive Analytics (Stackelberg game + Coopetition) {
return optimum values
}
If strictly concave ($\pi_t = \pi_r + \pi_e$) {
  return P_e^* & P_r^* ϵ arg max $\pi_t(P_r, P_e)$
}
    For $P_r = P_r^*$ and $P_e = P_e^*$ {
      return Q_r^* and Q_e^*
    }
      for $P_r = P_r^*$, $P_e = P_e^*$, $D_r = Q_r^*$ and $D_e = Q_e^*$ {
        return π_m
      }
        If strictly concave (π_m) {
          $w^* \epsilon$ arg max $\pi_m(w)$
            For $w = w^*$
              return $\pi_r^*, \pi_e^*, \pi_m^*$
        }

Based on the interactive analytics developed, we have derived the optimal conditions for normal and panic buying periods (Please see Table 15.1).

In the next section, we report the propositions by comparing the two periods.

15.4 Propositions

Proposition 15.1 *The optimal price of the retailer increased during the panic buying when the downstream channel partners undergo coopetition by* $\frac{\Delta \mathrm{a}_r \lambda(10\lambda - 7\gamma)}{4(\gamma - \lambda)(\gamma^2 - 4\lambda^2)}$.

Proposition 15.2 *The optimal price of the e-tailer increased during the panic buying when the downstream channel partners undergo coopetition by* $\frac{\Delta \mathrm{a}_r (5\gamma\lambda + 2\lambda^2 - 4\gamma^2)}{4(\gamma - \lambda)(\gamma^2 - 4\lambda^2)}$.

Table 15.1 Equilibrium decisions when the retailer and e-tailer undergo coopetition in a DCSC

	Normal buying period	Panic buying period
P_r	$\frac{2s(\lambda^2-\gamma^2)+(5\gamma+\lambda)a_e+(\gamma+5\lambda)a_r}{8(\lambda^2-\gamma^2)}$	$\frac{\Delta \mathrm{a}_r\gamma-2s\gamma^2+5\Delta \mathrm{a}_r\lambda+2s\lambda^2+(5\gamma+\lambda)a_e+(\gamma+5\lambda)a_r}{8(\lambda^2-\gamma^2)}$
P_e	$\frac{2s(\lambda^2-\gamma^2)+(\gamma+5\lambda)a_e+(5\gamma+\lambda)a_r}{8(\lambda^2-\gamma^2)}$	$\frac{5\Delta \mathrm{a}_r\gamma-2s\gamma^2+\Delta \mathrm{a}_r\lambda+2s\lambda^2+(\gamma+5\lambda)a_e+(5\gamma+\lambda)a_r}{8(\lambda^2-\gamma^2)}$
Q_r^*	$\frac{1}{8}(2s(\gamma-\lambda)-a_e+3a_r)$	$\frac{1}{8}(3\Delta \mathrm{a}_r+2s\gamma-2s\lambda-a_e+3a_r)$
Q_e^*	$\frac{1}{8}(2s(\gamma-\lambda)+3a_e-a_r)$	$\frac{1}{8}(-\Delta \mathrm{a}_r+2s\gamma-2s\lambda+3a_e-a_r)$
w	$\frac{2s(\lambda-\gamma)+a_e+a_r}{4\lambda-4\gamma}$	$\frac{\Delta \mathrm{a}_r-2s\gamma+2s\lambda+a_e+a_r}{4\lambda-4\gamma}$
π_r	$\frac{(2s(\gamma-\lambda)-a_e+3a_r)(2s(\gamma^2-\lambda^2)+(3\gamma-\lambda)a_e-(\gamma-3\lambda)a_r)}{64(\lambda^2-\gamma^2)}$	$\frac{(3\Delta \mathrm{a}_r+2s\gamma-2s\lambda-a_e+3a_r)(\Delta \mathrm{a}_r\gamma-2s\gamma^2-3\Delta \mathrm{a}_r\lambda+2s\lambda^2+(-3\gamma+\lambda)a_e+(\gamma-3\lambda)a_r)}{64(\lambda^2-\gamma^2)}$
π_e	$\frac{(2s(\gamma-\lambda)+3a_e-a_r)(2s(\gamma^2-\lambda^2)-(\gamma-3\lambda)a_e+(3\gamma-\lambda)a_r)}{64(\lambda^2-\gamma^2)}$	$\frac{(\Delta \mathrm{a}_r-2s\gamma+2s\lambda-3a_e+a_r)(3\Delta \mathrm{a}_r\gamma+2s\gamma^2-\Delta \mathrm{a}_r\lambda-2s\lambda^2-(\gamma-3\lambda)a_e+(3\gamma-\lambda)a_r)}{64(\gamma^2-\lambda^2)}$
π_m	$\frac{(2s(\gamma-\lambda)+a_e+a_r)^2}{16(\lambda-\gamma)}$	$\frac{(\Delta \mathrm{a}_r+2s\gamma-2s\lambda+a_e+a_r)^2}{16(\lambda-\gamma)}$

Proposition 15.3 *The optimal order quantity of the retailer increased during the panic buying when the downstream channel partners undergo coopetition by* $\frac{\Delta \mathrm{a}_r(\gamma-6\lambda)\lambda}{4(\gamma^2-4\lambda^2)}$.

Proposition 15.4 *The optimal order quantity of the e-tailer decreased during the panic buying when the downstream channel partners undergo coopetition by* $\frac{\Delta \mathrm{a}_r\lambda(2\lambda-3\gamma)}{4(\gamma^2-4\lambda^2)}$.

Proposition 15.5 *The optimal profit of the retailer increased during the panic buying when the downstream channel partners undergo coopetition by*

$$\frac{1}{256(\gamma^2-4\lambda^2)^4}\lambda(\lambda(\Delta \mathrm{a}_r\gamma-2s\gamma^2-6\Delta \mathrm{a}_r\lambda-2s\gamma\lambda+4s\lambda^2$$
$$+(-3\gamma+2\lambda)a_e+(\gamma-6\lambda)a_r)^4-16(\gamma^2-4\lambda^2)^2(-2s(\gamma^2+\gamma\lambda-2\lambda^2)$$
$$+(-3\gamma+2\lambda)a_e+(\gamma-6\lambda)a_r)^2)$$

Proposition 15.6 *The optimal profit of the e-tailer decreased during the panic buying when the downstream channel partners undergo coopetition by* $\frac{1}{16(\gamma^2-4\lambda^2)^2}\Delta \mathrm{a}_r(3\gamma-2\lambda)\lambda(3\Delta \mathrm{a}_r\gamma+4s\gamma^2-2\Delta \mathrm{a}_r\lambda+4s\gamma\lambda-8s\lambda^2-2(\gamma-6\lambda)a_e+(6\gamma-4\lambda)a_r)$.

Proposition 15.7 *The optimal profit of the manufacturer increased during the panic buying when the downstream channel partners undergo coopetition by* $\frac{\Delta \mathrm{a}_r\lambda(x+4s\gamma-4s\lambda+2a_e+2a_r)}{8(\gamma-2\lambda)(\gamma-\lambda)}$.

15.5 Discussions

The main contribution of the study is the novel analytics to model coopetition and mathematical modelling of panic buying when the retailer and e-tailer undergo coopetition. The models for both pre-pandemic and panic buying were developed and analysed to find the impact of panic buying. It was observed that the decision variables, the price of the downstream partners, increased during the panic buying period. But the optimal order quantity only increased for the retailer. This period also witnessed a high increase in the profit of the manufacturer. A similar trend was observed for the retailer. But e-tailer couldn't reap benefits from this period, and profit significantly decreased. This study adds to the body of academic knowledge since it is the first to focus on the panic buying that occurs in DCSC when the downstream channel partners undergo coopetition. Also, this study is the first one that focuses on asymmetrical disruptions.

The study also aids managers in developing decision support systems when downstream partners undergo coopetition and to make optimal decisions during pandemic disruptions. It was found that, during the initial period of pandemics, the retailer should increase the stock of goods. Else he will face stockout which will affect his profit. On the other hand, e-tailer needs to curb the price of the product. The increase in the product price during the panic buying period decreased the demand and, thereby, the optimum order quantity of the e-tailer. This significantly created a sharp decrease in his profit. The initial days of the pandemics are honoured to be the period where the manufacturer can reap maximum benefit from the market. The manufacturer must be able to escalate the entire process to increase production. Else it will result in market punishment and, thereby decrease in profit. The interactive analytics developed can be used to make decision support systems for collaborative models and as an input to multi-agent systems.

15.6 Conclusions

The study's primary purpose is twofold: (1) to present interactive analytics to model the coopetition between downstream channel partners in a supply chain and (2) to analyse the impact of the panic buying when the downstream channel partners undergo coopetition. We have developed novel analytics using the Stackelberg and cartel model to study the coopetition among downstream channel partners in a Dual Chanel Supply Chain. With the aid of the Stackelberg and cartel games, we modelled the interaction and compared the normal period with panic buying to quantify the increase or decrease of the optimal price, optimal order quantity and optimal profit. The study can be extended in different dimensions. We have analysed only the panic buying period. Analysing the lockdown period and associated disruption will tell a different story.

References

1. Massari GF, Giannoccaro I (2021) Investigating the effect of horizontal coopetition on supply chain resilience in complex and turbulent environments. Int J Prod Econ 237:108150. https://doi.org/10.1016/J.IJPE.2021.108150
2. Bengtsson M, Kock S (2000) "Coopetition" in business networks—to cooperate and compete simultaneously. Ind Mark Manag 29:411–426
3. Dorn S, Schweiger B, Albers S (2016) Levels, phases and themes of coopetition: a systematic literature review and research agenda. Eur Manag J 34:484–500. https://doi.org/10.1016/j.emj.2016.02.009
4. COFECE. COFECE's position in terms of the application of the Federal Economic Competition Law in the face of the health emergency. Date of publication: 03-27-2020. Communication number: COFECE-012-2020—Federal Economic Competition Commission. https://www.cofece.mx/postura-cofece-ante-emergencia-sanitaria/. Last accessed 20 Sept 2021
5. Competition and Markets Authority (2020) CMA approach to business cooperation in response to COVID-19
6. European competition network: competition policy. https://ec.europa.eu/competition-policy/european-competition-network_en. Last accessed 20 Sept 2021
7. OECD (2020) Co-operation between competitors in the time of COVID-19

8. Yuen KF, Wang X, Ma F, Li KX (2020) The psychological causes of panic buying following a health crisis. https://doi.org/10.3390/ijerph17103513. www.mdpi.com/journal/ijerph
9. Prentice C, Chen J, Stantic B (2020) Timed intervention in COVID-19 and panic buying. J Retail Consum Serv 57:102203. https://doi.org/10.1016/j.jretconser.2020.102203
10. Ahmad AR, Murad HR (2020) The impact of social media on panic during the COVID-19 pandemic in Iraqi Kurdistan: online questionnaire study. J Med Internet Res 22:1–11. https://doi.org/10.2196/19556
11. Pantano E, Pizzi G, Scarpi D, Dennis C (2020) Competing during a pandemic? Retailers' ups and downs during the COVID-19 outbreak. J Bus Res 116:209–213. https://doi.org/10.1016/j.jbusres.2020.05.036
12. Hoffman C (2020) Amazon & the covid-19 crisis
13. Kulkarni SR. Effect of covid-19 on the shift in consumer preferences with respect to shopping modes (offline/online) for groceries: an exploratory study. http://www.iaeme.com/IJM/index.asp581http://www.iaeme.com/IJM/issues.asp?JType=IJM&VType=11&IType=10http://www.iaeme.com/IJM/issues.asp?JType=IJM&VType=11&IType=10. Last accessed 27 May 2021
14. Chiang WYK, Chhajed D, Hess JD (2003) Direct marketing, indirect profits: a strategic analysis of dual-channel supply-chain design. Manag Sci. https://doi.org/10.1287/mnsc.49.1.1.12749
15. Kurata H, Yao DQ, Liu JJ (2007) Pricing policies under direct vs. indirect channel competition and national vs. store brand competition. Eur J Oper Res 180:262–281. https://doi.org/10.1016/j.ejor.2006.04.002
16. Hua G, Cheng TCE, Wang S (2011) Electronic books: to "e" or not to "e"? A strategic analysis of distribution channel choices of publishers. Int J Prod Econ. https://doi.org/10.1016/j.ijpe.2010.11.011
17. Yao DQ, Yue X, Mukhopadhyay SK, Wang Z (2009) Strategic inventory deployment for retail and e-tail stores. Omega 37:646–658. https://doi.org/10.1016/j.omega.2008.04.001
18. Takahashi K, Aoi T, Hirotani D, Morikawa K (2011) Inventory control in a two-echelon dual-channel supply chain with setup of production and delivery. Int J Prod Econ (Elsevier) 403–415. https://doi.org/10.1016/j.ijpe.2010.04.019.
19. Chen J, Zhang H, Sun Y (2012) Implementing coordination contracts in a manufacturer Stackelberg dual-channel supply chain. Omega. https://doi.org/10.1016/j.omega.2011.11.005
20. Saha S (2016) Channel characteristics and coordination in three-echelon dual-channel supply chain. Int J Syst Sci. https://doi.org/10.1080/00207721.2014.904453
21. Rofin TM, Mahanty B (2018) Optimal dual-channel supply chain configuration for product categories with different customer preference of online channel. Electron Commer Res 18:507–536. https://doi.org/10.1007/s10660-017-9269-4
22. Soleimani F, Arshadi Khamseh A, Naderi B (2016) Optimal decisions in a dual-channel supply chain under simultaneous demand and production cost disruptions. Ann Oper Res 243:301–321. https://doi.org/10.1007/s10479-014-1675-6
23. Nagurney A (2021) Supply chain game theory network modeling under labor constraints: applications to the Covid-19 pandemic. Eur J Oper Res 293:880–891. https://doi.org/10.1016/j.ejor.2020.12.054
24. Huang S, Yang C, Zhang X (2012) Pricing and production decisions in dual-channel supply chains with demand disruptions. Comput Ind Eng 62:70–83. https://doi.org/10.1016/j.cie.2011.08.017
25. Rahmani K, Yavari M (2019) Pricing policies for a dual-channel green supply chain under demand disruptions. Comput Ind Eng 127:493–510. https://doi.org/10.1016/j.cie.2018.10.039
26. Huang S, Yang C, Liu H (2013) Pricing and production decisions in a dual-channel supply chain when production costs are disrupted. Econ Model 30:521–538. https://doi.org/10.1016/j.econmod.2012.10.009
27. Cao E (2014) Coordination of dual-channel supply chains under demand disruptions management decisions. 37–41. https://doi.org/10.1080/00207543.2014.938835
28. Zhang P, Xiong Y, Xiong Z (2015) Coordination of a dual-channel supply chain after demand or production cost disruptions. Int J Prod Res 53:3141–3160. https://doi.org/10.1080/00207543.2014.975853

29. Yan B, Jin Z, Liu Y, Yang J (2018) Decision on risk-averse dual-channel supply chain under demand disruption. Commun Nonlinear Sci Numer Simul 55:206–224. https://doi.org/10.1016/j.cnsns.2017.07.003
30. Tang C, Yang H, Cao E, Lai KK (2018) Channel competition and coordination of a dual-channel supply chain with demand and cost disruptions. Appl Econ 50:4999–5016. https://doi.org/10.1080/00036846.2018.1466989
31. Huang S, Chen S, Li H (2018) Optimal decisions of a retailer-owned dual-channel supply chain with demand disruptions under different power structures. Int J Wirel Mob Comput 14:277–287. https://doi.org/10.1504/IJWMC.2018.092370
32. Pi Z, Fang W, Zhang B (2019) Service and pricing strategies with competition and cooperation in a dual-channel supply chain with demand disruption. Comput Ind Eng 138:106130. https://doi.org/10.1016/j.cie.2019.106130
33. Mohsenzadeh Ledari A, Arshadi Khamseh A (2018) Optimal decisions in a dual-channel supply chain for the substitute products with special orders under disruption risk and brand consideration. Int J Eng Trans B Appl 31:759–769. https://doi.org/10.5829/ije.2018.31.05b.11
34. Rofin MB (2020) Impact of wholesale price discrimination by the manufacturer on the profit of supply chain members. Manag Decis. https://doi.org/10.1108/MD-11-2019-1644
35. Raju S, Rofin TM, Pavan Kumar S (2021) Delivery lead time competition between e-tailers: a game-theoretic study. In: 2021 International conference on data analytics for business and industry, ICDABI 2021, pp 542–546. https://doi.org/10.1109/ICDABI53623.2021.9655934
36. Rofin TM, Mahapatra MS, Mahanty B (2021) Impact of green retail operations on the profit of the manufacturer and the retailer under different pricing strategies. Opsearch 58:125–143. https://doi.org/10.1007/s12597-020-00478-1

Chapter 16
Supply Chain Data Analytics for Predicting Delivery Risks Using Machine Learning

Arun Thomas and Vinay V. Panicker

Abstract The most difficult predictive challenge in supply chain disruption management is order delivery delay. Identifying the risk in delivering an order in the scheduled time will help the company to focus on the prioritized orders to mitigate the disruption before its occurrence. This research presents a machine learning-based predictive model for delivery risk prediction of different product orders. The proposed approach deals with an imbalanced class problem, where the frequency of orders which have the delivery risk is rare when compared to the orders that do not. The Area Under the Curve (AUC) score is the selected performance metric for the proposed risk prediction problem. With a comparative analysis, it is found that the Random Forest model in Synthetic Minority Over-sampling Technique (SMOTE) with the Tomek link gives a better performance with an AUC score of 0.80. It is also found that the Random Forest model performs better in SMOTE and SMOTE Tomek oversampling methods, whereas K-Nearest Neighbour (KNN) performs well in the random oversampling technique.

Keywords Delivery risk · Machine learning · Supply chain disruptions · Prediction

16.1 Introduction

A disruption in the supply chain is an unanticipated or unplanned occurrence that disrupts the usual flow of goods, materials, or services throughout an entire supply chain. Modern supply chains are becoming increasingly vulnerable to disturbances, and a disruption in one region of the world may affect the performance of companies all over the world [1]. A study conducted by Business Continuity Institute in 2019 reveals that every year, more than 56% of organizations in the world experience a supply chain disruption, and the firms have started to analyze the disruptions more

A. Thomas · V. V. Panicker (✉)
National Institute of Technology Calicut, Kozhikode, India
e-mail: vinay@nitc.ac.in

M. K. Tiwari et al. (eds.), *Applications of Emerging Technologies and AI/ML Algorithms*, Asset Analytics, https://doi.org/10.1007/978-981-99-1019-9_16

seriously [2]. Delivery delay of products is an import type of disruption occurred on the downstream side of a supply chain. Recently, the dynamics of companies' competing efforts have had an impact on efforts to handle orders so that goods can reach consumers at the appropriate time and in the right place [3].

This work presents the application of supervised learning algorithms in delivery risk prediction of orders in supply chain disruption management with the combination of ensemble classifiers and sampling methods.

16.2 Literature Review

Data analytics in the supply chain are categorized into three main categories such as prescriptive, descriptive, and predictive analytics [4, 5]. Prescriptive analytics entails using predictions to optimize the existing situation and taking action to progress toward a more desirable state. Several articles discussed the application of big data and machine learning to predict disruptions in the supply chain. Dani [6] proposed a proactive approach to identify the potential risk source by using data mining. De Santis et al. [7] presented a predictive model for the backorder imbalanced class problem, where the frequency of items that are back-ordered is very less compared to those that are not back-ordered. Shajalal et al. [8] also discussed the same backorder prediction problem considering the same data set and developed a deep neural network model for the prediction of probable backorders. Brintrup et al. [9] discussed how an original equipment manufacturer may use data analytics to identify first-tier supply disruptions using historical data and compared the performance of different machine learning algorithms for the selected prediction problem. Singh et al. [10] presented a public distribution system model with three different cases to analyze food supply chain disruptions.

16.3 Methodology

16.3.1 Dataset

The data set was downloaded from a public repository (https://data.usaid.gov/d/a3rc-nmf6) which includes shipment and pricing data of the health commodity supply chain from 2006 to 2015. Table 16.1 lists the most relevant variables in the dataset and their descriptions. The steps followed in the development of the risk prediction models are shown in Fig. 16.1.

Table 16.1 Summary of the dataset

Dataset	Number of features	Positive class	Negative class	Total	Imbalance ratio
Supply chain Shipment data	33	1186	9138	10,324	12:88

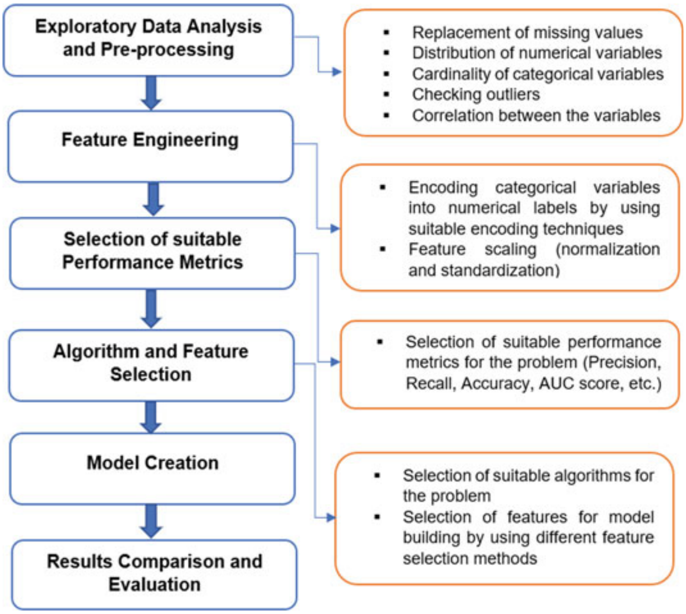


Fig. 16.1 Steps followed in the development of the risk prediction model

16.3.2 Exploratory Data Analysis and Pre-processing

After removing the duplicate details, processing the missing values, and initial feature selection, it was found that the dataset contains 10,324 orders of five product groups from 73 vendors of which 12% of orders are delivered later than the scheduled date. The positive class in the analysis means the orders which are late, and the negative class means the orders which are delivered on or before the scheduled date. However, upon closer inspection, it is found that the occurrence of the negative class, i.e., the orders delivered on time is substantially greater than the positive class, i.e., the orders delivered late. So, the dataset is highly imbalanced, and this imbalance may cause false-positive classification errors in the results.

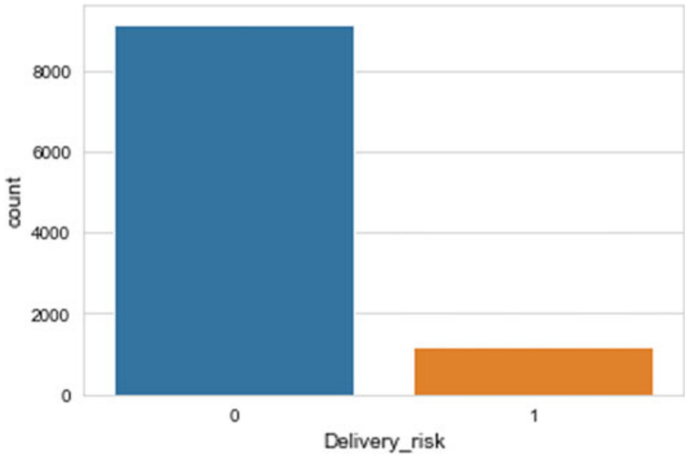


Fig. 16.2 Class imbalance statistics of the dataset

16.3.2.1 Handling the Class Imbalance

A dataset is unbalanced in supervised learning when the number of instances of a specific class of interest is infrequent when compared to the other classes. There are numerous classification issues of this sort in the real world such as pollution detection, and risk management, hence this is an interesting subject to explore [11] and [12]. Figure 16.2 shows the class imbalance statistics of the problem.

16.3.2.2 Synthetic Minority Over-Sampling Technique

In this method, the balancing of the two classes is done by creating artificial data points based on similarities in feature space between the existing minority class by considering the K-nearest neighbors in the Euclidean space.

The dataset contains five numerical features. Figure 16.3 shows the correlation heatmap of all numerical features. From the figure, it is found that the features of line-item quantity and line-item value are highly correlated with a correlation value of 0.84, and one of these features is removed from further analysis.

16.3.3 Feature Engineering

The process of transforming raw data into features that may be used to develop a machine learning prediction model is known as feature engineering. It is found that out of the initially selected 19 input features, 14 features are categorical. To build a predictive machine learning model with these features, the categorical features

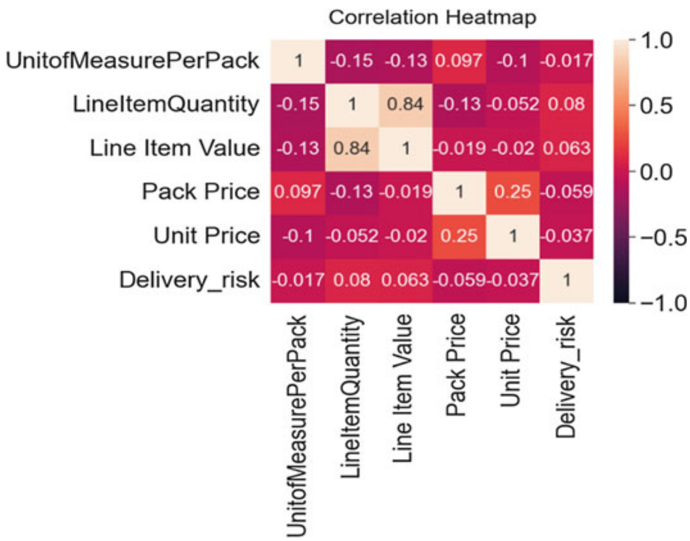


Fig. 16.3 Correlation heatmap of all numerical features

must be converted into numerical values. Table 16.2 shows the number of unique values in each categorical feature. The categorical features are encoded by using two feature encoding techniques. The features with several unique values of less than ten are encoded by using the One Hot encoding method and the remaining features are encoded by the frequency encoding method.

Table 16.2 Number of unique values in categorical features

Sl. No.	Categorical feature name	Unique values
1	Project code	142
2	Country	40
3	Managed by	4
4	Fulfill via	2
5	Vendor INCO term	8
6	Product group	5
7	Sub classification	6
8	Vendor	73
9	Molecule/test type	86
10	Brand	48
11	Dosage	54
12	Dosage form	17
13	Manufacturing site	88
14	First line designation	2

Table 16.3 Confusion matrix

	Predicted negative	Predicted positive
Actual negative	True negative (TN)	False positive (FP)
Actual positive	False negative (FN)	True positive (TP)

16.3.4 Selection of Suitable Performance Metrics

In a binary classification problem, accuracy, precision, and recall are the commonly used metrics to evaluate a classifier's performance. The confusion matrix shown in Table 16.3 is used to compute these standard assessment criteria.

Accuracy (Acc) is the most used performance evaluation metric of any classification algorithm, which can be calculated by "Eq. 16.1".

$$\mathrm{Acc} = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN}} \tag{16.1}$$

If the dataset is balanced, accuracy can be selected as a standard performance metric for classification problems. The accuracy is not a good performance indicator of the classifier when the output variable is imbalanced. In such an imbalanced class context, selecting a suitable assessment metric that achieves the necessary performance is critical. In these cases, the class-independent metrics, and the area under Receiver Operating Characteristic (ROC) curve are the preferable alternative assessment metrics [13]. Recall and precision are two more metrics that are frequently used as suitable measures for estimating the performance of the balance classification problem.

The AUC score is the commonly used performance assessment metric on class imbalanced problems. In this work also AUC score is used to validate the performance of the proposed machine learning classification algorithms.

16.3.5 Algorithm and Feature Selection

In every machine learning project, feature selection is critical. To develop an effective machine learning algorithm, the features must be independent, informative, and discriminating. From the literature review, it is found that most of the risk predictive models are developed by using K-Nearest Neighbour, Random Forest, Logistic Regression, and Support Vector Machine (SVM).

Table 16.4 Parameters used for different algorithms

Logistic regression	K-nearest neighbour	Support vector machine	Random forest
Tol = 0.0001, C = 1 solver = 'lbfgs' Max iterations: = 25,000	neighbours = 3 metric = 'minkowski' p = 2	probability = True, C = 0.01	criterion = 'gini', max_features = "log2", min_samples_leaf = 10, max_depth = 30, n_estimators = 100

16.4 Results and Discussions

16.4.1 Results Comparison and Evaluation

The delivery risk prediction problem is solved by using four different machine learning algorithms including K-Nearest Neighbour, Random Forest, Logistic Regression, and Support Vector Machine. The parameters used for different algorithms are given in Table 16.4.

The proposed models are solved by using three different oversampling methods including random oversampling, SMOTE, and SMOTE Tomek methods. Figure 16.4 shows the ROC curve of different algorithms used in model creation.

16.4.2 K-fold cross-validation

It is frequently used in machine learning modeling approaches to compare and select a model for a particular predictive modeling challenge. In order to minimize overfitting in training, a stratified fivefold cross-validation technique is used. Tables 16.5, 16.6, and 16.7 show the performance metrics values of different algorithms after cross-validation by considering three different over-sampling methods.

AUC score is the selected performance metric for the proposed risk prediction problem, and from the analysis, it is found that the Random Forest model with SMOTE Tomek oversampling method gives a better performance with an AUC score of 0.80. It is also found that the Random Forest model performs better in SMOTE and SMOTE Tomek oversampling methods, whereas KNN performs well in the Random oversampling method.

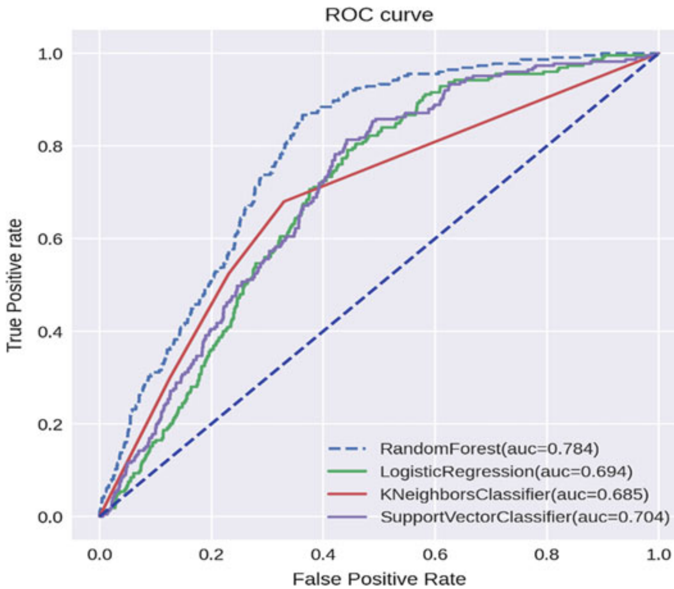


Fig. 16.4 Receiver operator characteristic curve of different algorithms

Table 16.5 Performance metrics values with SMOTE oversampling

Algorithm/Performance metrics	Precision	Recall	F1_score	Accuracy	AUC_score
Logistic regression	0.62	0.78	0.69	0.65	0.69
KNN	0.75	0.77	0.74	0.73	0.69
SVM	0.62	0.78	0.71	0.66	0.70
Random forest	0.78	0.85	0.84	0.82	0.79

Table 16.6 Performance metrics values with random oversampling

Algorithm/Performance metrics	Precision	Recall	F1_score	Accuracy	AUC_score
Logistic regression	0.61	0.77	0.68	0.64	0.69
KNN	0.76	0.74	0.69	0.72	0.71
SVM	0.63	0.77	0.71	0.67	0.69
Random forest	0.78	0.81	0.85	0.84	0.78

Table 16.7 Performance metrics values with SMOTE Tomek oversampling

Algorithm/Performance metrics	Precision	Recall	F1_score	Accuracy	AUC_score
Logistic regression	0.63	0.77	0.69	0.66	0.70
KNN	0.81	0.79	0.80	0.79	0.70
SVM	0.63	0.79	0.72	0.68	0.71
Random forest	0.80	0.82	0.79	0.81	0.80

16.5 Conclusion

16.5.1 Conclusion

Product delivery risk prediction is one of the major predictive challenges in supply chain disruption management that many companies are confronted with today. This work presented a machine learning-based predictive system for product delivery risk using the available historical data. The prior identification of the risk in delivering the product in the scheduled time will help the company to focus on the prioritized orders to mitigate the disruption before its occurrence. The overall service performance of a company can be improved by adopting these predictive systems.

The proposed predictive approach deals with an imbalanced class problem and which is efficiently handled by different oversampling techniques including Random oversampling, SMOTE, and SMOTE Tomek oversampling methods. Since delivery-risky items (positive class) are uncommon compared to non-risky products (negative class), several specific methodologies and metrics are used in the design, development, and assessment of models in the unbalanced class problem.

Random Forest model with SMOTE Tomek oversampling method gives a better performance with an AUC score of 0.80. It is also found that the Random Forest model performs better in SMOTE and SMOTE Tomek oversampling methods whereas KNN performs well in the Random oversampling method.

16.5.2 Future Research Avenues

Further research can be done to improve the prediction performance with new algorithms and optimize the learning process. Potential features can be derived from the existing features to improve the performance of the proposed model by different feature engineering techniques.

References

1. MacKenzie CA, Barker K, Santos JR (2014) Modelling a severe supply chain disruption and post-disaster decision making with application to the Japanese earthquake and Tsunami. IIE Trans 46:1243–1260. https://doi.org/10.1080/0740817x.2013.876241
2. Katsaliaki K, Galetsi P, Kumar S (2021) Supply chain disruptions and resilience: a major review and future research agenda. Ann Oper Res. https://doi.org/10.1007/s10479-020-03912-1
3. Kusrini E et al (2020) Risk mitigation on product distribution and delay delivery: a case study in an Indonesian manufacturing company. In: IOP conference series: materials science and engineering, vol 722, pp 012015. https://doi.org/10.1088/1757-899x/722/1/012015
4. Wang G, Gunasekaran A, Ngai EWT et al (2016) Big data analytics in logistics and supply chain management: certain investigations for research and applications. Int J Prod Econ 176:98–110. https://doi.org/10.1016/j.ijpe.2016.03.014

5. Baryannis G, Validi S, Dani S et al (2018) Supply chain risk management and artificial intelligence: state of the art and future research directions. Int J Prod Res 57:2179–2202. https://doi.org/10.1080/00207543.2018.1530476
6. Dani S (2009) Predicting and managing supply chain risks. In: Supply chain risk. Springer, Boston, pp 53–66
7. De Santis RB et al (2017) Predicting material backorders in inventory management using machine learning. In: IEEE Latin American conference on computational intelligence (LA-CCI). https://doi.org/10.1109/la-cci.2017.8285684
8. Shajalal M, Hajek P, Abedin MZ (2021) Product backorder prediction using deep neural network on imbalanced data. Int J Prod Res 1–18. https://doi.org/10.1080/00207543.2021.1901153
9. Brintrup A, Pak J, Ratiney D et al (2019) Supply chain data analytics for predicting supplier disruptions: a case study in complex asset manufacturing. Int J Prod Res 58:3330–3341. https://doi.org/10.1080/00207543.2019.1685705
10. Singh S, Kumar R, Panchal R et al (2020) Impact of covid-19 on logistics systems and disruptions in food supply chain. Int J Prod Res 59:1993–2008. https://doi.org/10.1080/00207543.2020.1792000
11. López V, Fernández A, García S et al (2013) An insight into classification with imbalanced data: empirical results and current trends on using data intrinsic characteristics. J Inf Sci 250:113–141. https://doi.org/10.1016/j.ins.2013.07.007
12. Galar M, Fernandez A, Barrenechea E et al (2012) A review on ensembles for the class imbalance problem: bagging-, boosting-, and hybrid-based approaches. IEEE Trans Syst Man Cybern 42:463–484. https://doi.org/10.1109/tsmcc.2011.2161285
13. Chawla NV, Bowyer KW, Hall LO et al (2002) SMOTE: synthetic minority over-sampling technique. J Artif Intell Res 16:321–357. https://doi.org/10.1613/jair.953

Chapter 17
Importance of Equitable Public Procurement of Food Grains in India for Sustainability

Maheswar Singha Mahapatra and Biswajit Mahanty

Abstract This paper has analyzed the effects of massive public procurement of food grains and their indirect impact on the environmental front. For decades, the practice of food grains procurement has created a barrier to a more equitable and sustainable procurement. We have applied the concept of system dynamics to analyze the critical loops that played a significant role in the current scenario. We have proposed equity as an alternative approach to counter the skewed procurement and sustainability issue.

Keywords Public procurement · Equity · System thinking

17.1 Introduction

The government of India has been procuring a significant amount of rice and wheat at the minimum support price (MSP) for the past few decades. Such public procurement has helped India become self-sufficient in food grains production, provide a safety net for farmers, and run the massively subsidized food grains distribution program for more than 80 crore targeted people [4]. However, farmers from only a few states have benefited largely instead of an equitable share of public procurement for farmers across all states. Due to such skewed procurement practices over the years, other side effects have emerged, such as rising cost of operation due to increased statutory charges, which is more than 10% of food subsidy [3]. Another side effect is the reduction in private procurement in those states, which skewed the public procurement of food grains further. Most importantly, the environmental aspects still lack urgent and critical attention from academia and policymakers.

M. S. Mahapatra (✉)
National Institute of Industrial Engineering (NITIE), Mumbai 400087, India
e-mail: msmahapatra@gmail.com

B. Mahanty
Department of Industrial and Systems Engineering, Indian Institute of Technology Kharagpur, West Bengal, Kharagpur 721302, India

M. K. Tiwari et al. (eds.), *Applications of Emerging Technologies and AI/ML Algorithms*, Asset Analytics, https://doi.org/10.1007/978-981-99-1019-9_17

Availability of water is a critical requirement for agriculture production. However, India is already categorized as a water-stressed country, and 78% of freshwater is used in the agriculture sector. Out of that, Paddy, wheat, and sugarcane consume almost 80% of irrigation water. Based on the water productivity report by [6], Punjab and Haryana are highly stressed, and Telangana is stressed out of the top few states in terms of rice production and procurement. As per the recent assessment of groundwater at the block level in India (2020), 78% of blocks in Punjab and 60% of blocks in Haryana are over-exploited in groundwater use, which is alarming. The existing public procurement policy with an assured minimum support price is probably one of the few crucial factors that led to the situation.

In this study, we have analyzed the effects of public procurement of food grains and how they become a barrier to sustainability. We have also presented the importance of equity to address skewed public procurement. It will also produce better results in terms of sustainability.

17.2 Methodology

We have used the concept of system dynamics to understand the issues mentioned above in the public procurement of food grains. A feedback loop is an essential concept of system dynamics that can be of two types, a reinforcing loop (positive loop) and a balancing loop (negative loop). Based on the relative strength of these loops, system behavior can be predicted. We demonstrate both types of loops below (see Fig. 17.1) for the population in a very simple way.

Similar behavior is observed in a few states with respect to public procurement and production of food grains, i.e., the reinforcing loop. More public procurement leads to more production, which increases public procurement further. One policy that facilitates an increase in public procurement is the open procurement policy. However, [5] proved that the total procurement in the past decade remains around one-third of total production. They failed to reject the hypothesis at a 5% significance

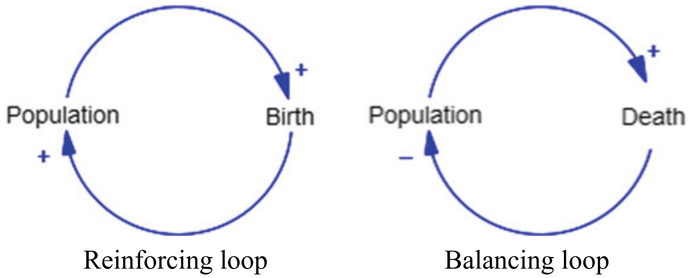


Fig. 17.1 Reinforcing loop and balancing loop

level. Although, for the past two years (2020 and 2021), the amount of food grains procurement was significantly higher than in previous years.

Therefore, we have investigated various critical loops that impact the overall system behavior and have also tried to match the trend of public procurement. Subsequently, if we can show significant similarities with a few well-known system archetypes, it will be easy to predict future outcomes.

17.3 Results

Analyzing the past data, we have observed that the share of procurement for a few states is almost double compared to their production contribution. For example, the production contribution of wheat for three states: Punjab, Haryana, and Madhya Pradesh, is around 45%, but more than 85% of wheat procurement is done there. Similarly, the production contribution of rice for six states: Andhra Pradesh, Chhattisgarh, Haryana, Odisha, Punjab, and Telangana, is around 40%, but more than 75% of rice is procured from there. These statistics establish the skewed nature of public procurement. The amount of procurement for rice and wheat is also increased steadily and can be found in Tables 17.1 and 17.2.

Now, it can be said that successful states become more successful in procurement which closely indicates the 'Success to the successful' system archetype as presented below in Fig. 17.2.

Table 17.1 Rice procurement in Lakh Metric Ton (LMT)

States	2013–14	2014–15	2015–16	2016–17	2017–18	2018–19	2019–20	2020–21
Andhra Pradesh	37.37	35.96	43.26	35.81	40	48.06	55.33	56.66
Telangana	43.53	35.04	15.81	33.76	36.18	51.86	74.54	94.53
Chhattisgarh	42.9	33.55	34.42	46.62	32.55	39.71	50.53	47.74
Haryana	24.06	20.15	28.61	35.83	39.92	39.41	43.07	37.89
Odisha	28.19	34.87	33.69	30.24	32.87	43.83	47.98	52.58
Punjab	81.06	77.86	93.5	110.52	118.33	113.34	108.76	135.89

Table 17.2 Wheat procurement in Lakh Metric Ton (LMT)

States	2013–14	2014–15	2015–16	2016–17	2017–18	2018–19	2019–20	2020–21
Punjab	108.97	116.41	103.44	106.45	117.06	126.91	129.12	127.14
Haryana	58.73	64.95	67.78	67.22	74.32	87.39	93.2	74
Madhya Pradesh	63.55	70.94	73.09	39.9	67.25	72.87	67.25	129.42

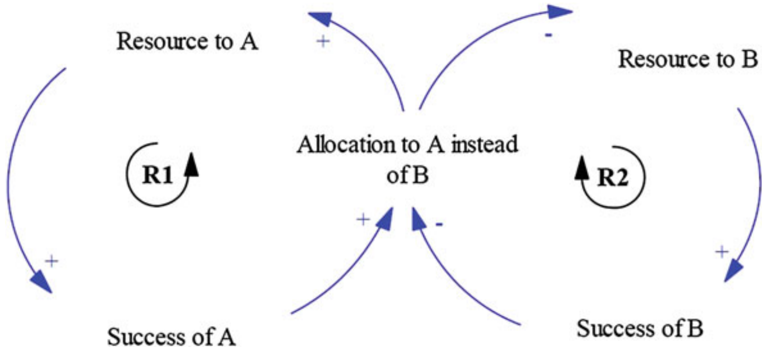


Fig. 17.2 Success to the successful (*Source* [1])

This skewed procurement has also created a scope for these successful states to increase statutory charges, making private procurement at mandi costlier. In this regard, a few critical loops are presented in Fig. 17.3.

If there is no check (as can be observed for open procurement policy), the share of private procurement is going to reduce as time passes as we can see a reinforcing loop as follows:

Public Procurement—Private Procurement—Unsold Production—Procurement Pressure for Public Agencies—Public Procurement.

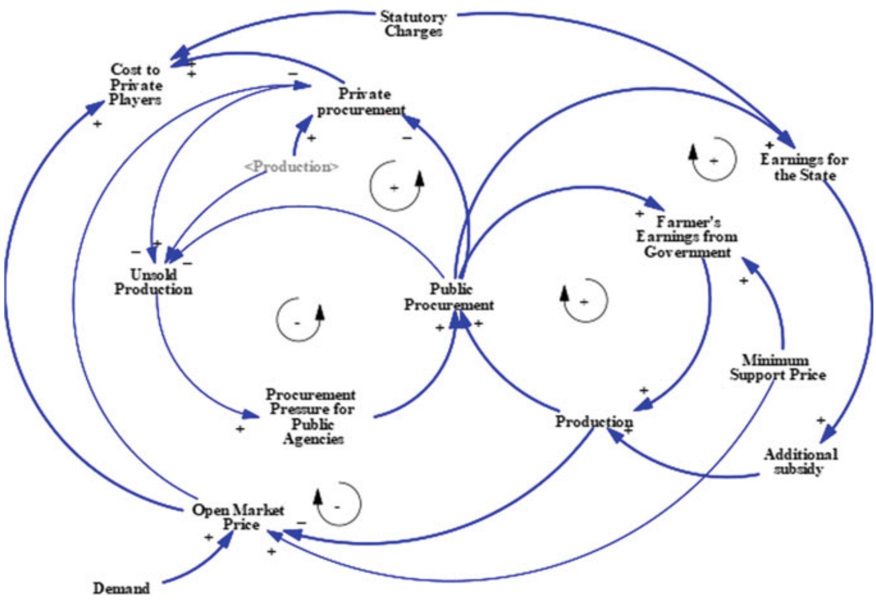


Fig. 17.3 Public procurement loops

This scenario is balanced by more public procurement, as can be seen through the balancing loop, i.e., Public Procurement—Unsold Production—Procurement Pressure for Public Agencies—Public Procurement.

However, this pattern is not financially viable in the long run. Therefore, the share of private procurement has to be increased. It will happen only when the share of public procurement will not encroach upon the share of private procurement. Therefore, we need to fix an equitable procurement strategy where each state would have a fair share based on well-defined criteria, such as demand, production, etc. It will put a threshold or a cap for procurement from any state. Over time, skewed procurement would turn into a more equitable public procurement. It will force the existing high food grains procuring states to go for crop diversification. As crop diversification is adopted, water productivity would increase and improve sustainability.

17.4 Conclusion

In this paper, we have analyzed public procurement of food grains which has its merits and demerits considering the length of operation. Apart from being self-sufficient in food grains production, running a subsidized food program for more than 80 crores people is a great success. However, some unintended side effects are also observed, such as the rapid increase in food subsidy, reduction in private procurement in many states, and over-exploitation of water resources for rice and wheat, which is alarming for a few states. Therefore, moving towards an equitable procurement will counter the skewed procurement practices at present and lead to better sustainability.

References

1. Barun W (2002) The system archetypes. https://www.albany.edu/faculty/gpr/PAD724/724WebArticles/sys_archetypes.pdf
2. Government of India, Central Ground Water Board (CGWB). http://cgwb.gov.in/
3. Gulati A, Gujral J, Nandakumar T, Jain S, Anand S, Rath S, Joshi P (2012) National food security bill challenges and options (Issue 2)
4. Mahapatra MS, Mahanty B (2020) Policies for managing peak stock of food grains for effective distribution: a case of the Indian food program. Socioecon Plann Sci 71:100773. https://doi.org/10.1016/j.seps.2019.100773
5. Mahapatra MS, Mahanty B (2021) Effective public procurement of food grains: a case of India under differential charges. Comput Ind Eng 154(January):107165. https://doi.org/10.1016/j.cie.2021.107165
6. Sharma BR, Gulati A, Mohan G, Manchanda S, Ray I, Amarasinghe U (2018) Water productivity mapping of major Indian crops. https://www.nabard.org/auth/writereaddata/tender/1806181128Water%20Productivity%20Mapping%20of%20Major%20Indian%20Crops,%20Web%20Version%20(Low%20Resolution%20PDF).pdf

Chapter 18
A Framework for 5G Enabled Vaccine Supply Chain Digital Twin

Mohd Juned, Purnima S. Sangle, and Manoj Kumar Tiwari

Abstract Real-time controlling the vaccine shortage and wastage within the supply chain network (SCN) is a critical challenge for healthcare professionals. However, in the literature, only limited studies are available on theoretical support and application of 5G to control the shortage and wastage of vaccines and no effective framework has been developed. Moreover, we are proposing a framework for a 5G enabled vaccine supply chain (VSC) digital twin to control the shortage and wastage of vaccines across the SCN. The 5G technologies enhance real-time visibility and connectivity of the supply chain (SC) at a highly granular level. This paper aims to propose an application of 5G technologies to provide the digital twin framework for improving visibility and connectivity that will help to make the real-time decision for controlling vaccine shortage and wastage within the SCN. As a result, this proposed digital twin framework for the vaccine supply chain will help improve delivery performance during pandemic and emergency times through highly granular visibility and connectivity across the network.

Keywords 5G · Vaccine supply chain · Framework · Digital twin

18.1 Introduction

During the pandemic, providing timely vaccination to everyone was a global challenge. Also, shortage and wastage of vaccines would be the major cause of death during a pandemic because it makes delays to provide timely vaccination. How can we

M. Juned (✉) · P. S. Sangle
Department of Analytics and Data Science, National Institute of Industrial Engineering (NITIE), Mumbai, India
e-mail: mohd.juned011.2017@nitie.ac.in

P. S. Sangle
e-mail: purnima@nitie.ac.in

M. K. Tiwari
National Institute of Industrial Engineering (NITIE), Mumbai, India

M. K. Tiwari et al. (eds.), *Applications of Emerging Technologies and AI/ML Algorithms*, Asset Analytics, https://doi.org/10.1007/978-981-99-1019-9_18

provide timely vaccination to everyone along with minimum shortage and wastage of vaccines? With this objective, we develop a 5G enabled vaccine supply chain framework to provide a road map for how we can minimize and control the shortage and wastage of vaccines within the supply chain using 5G and other disruptive technology applications. During the pandemic, vaccine distribution was the major challenge due to the involvement of various risk factors such as temperature, wastage, late delivery, and weak network connectivity. Moreover, disruptive digital technologies like 5G technology, IoT (internet of things), blockchain, cloud technology, RFID, and sensors have played a major role to improve visibility and connectivity across the supply chain network structure.

Dolgui and Ivanov [3] have discussed the major abilities of the digital supply chain (DSC) which would be heightened by 5G, i.e. visibility, intelligence, clarity, connectivity and networking. Most of the authors Wallace et al. [15], Guichard et al. [4], Lazarus et al. [8], and Duttagupta et al. [2] have discussed the vaccine wastage issues at each echelon of the supply network. The vaccine shortage problem under pandemic and emergency conditions has been addressed by McClellan et al. [10], Miranda-García et al. [11] Torjesen [14]. Torjesen [14] discussed the various cause of shortages like delayed shipment, and dependency on a single supplier.

The impact of various digital technologies on SC performance has been considered by Ivanov et al. [6, 7] with different prospective to minimize the disruption risk impact on the network structure. Dolgui and Ivanov [3] have discussed the importance of 5G technology to improve the connectivity between the supply chain stakeholders at strategic, operational, and cross-functional levels. The supply chain risk management issues were addressed by Ivanov et al. [5–7] who developed a digitalization framework after using various emerging digital technologies like IoT, cloud, and blockchain technology to control and minimize the various type of risk in the supply chain network.

Figure 18.1 shows the cellular evaluation from 1 to 5G (2.4–9.6 Kbps to >1 Gbps). 5G is a new network solution designed to address the problems of future communication demands, in which a vast number of intelligent devices will be able to connect at any time and in any location. These devices need exceptional communication network capabilities, such as near-zero latency and data speeds in the Gbps range.

Consequently, the 5th generation of cellular networks provides significant technical capabilities beyond those of 4G, such as—device connection, reduced energy consumption, unconventional resource virtualization, service-oriented resource allocation on demand, and automated management. A significant characteristic of 5G networks is their capacity to increase connection for many high-speed applications. Thus, 5G networks respond to the rising expectations for communication quality. By removing delay, 5G is anticipated to greatly enhance connection, capacity, and transfer rates [12].

Liu and Zheng [9] employed intelligent logistics technology to automatically gather vehicle information via radio frequency tags and transfer vehicle information to the back-end computer system so that vehicle information can be monitored in real-time. This time, we are aware of the amount of time required for the logistical distribution of these three firms using this technology. Cheng et al. [1] explained how

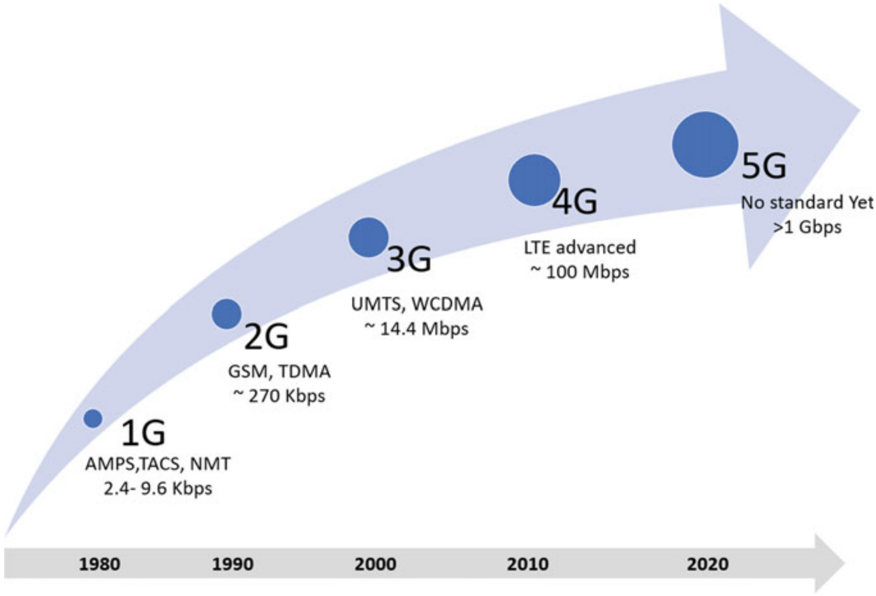


Fig. 18.1 Cellular evaluation (adopted from Surantha et al. [13])

5G's high bandwidth, low latency, and huge connections impact intelligent manufacturing. 5G has greater data throughput, shorter latency, lower power consumption, improved communication quality, and more node connections than 4G. High spectrum utilisation, powerful system performance, bidirectional signal transmission, low power consumption, and cost are its primary features. 5G works with blockchain, RFID, etc. It can track numerous transactions and database logs using blockchain, which promotes 5G applications.

Yan et al. [16] noted that 5G technology paired with new techniques and technologies such as AI, ML, and network intelligence technology may link the physical entities and data of industrial applications including digital supply chain to enhance industry growth.

18.2 5G Enabled Vaccine Supply Chain Framework: Digital Twin

The digital twin framework has been proposed by Ivanov et al. [6, 7] to minimize and control the disruption risk in the SC network. The authors have used the combined application of optimization, simulation, and data analytics to represent the data-driven framework. The proposed framework referred to the few disruptive technologies applications from the digital twin framework developed by Ivanov et al. [6, 7].

Figure 18.2 shows the data-driven framework of vaccine distribution. Broadly, it is a combination of the physical vaccine supply chain (i.e. manufacturer, national storage facility, regional hospital, health centre, and vaccination outreach) and digital supply chain (i.e. manufacturer data, national storage facility data, regional hospital data, health centre data, and vaccination outreach data). Moreover, the digital twin formed with the combination of optimization, simulation, artificial intelligence, and big data analytics. Also, use an application of various digital technologies to minimize the impact of risk and improve the performance of the SC in terms of visibility and connectivity to reduce wastage and shortage of vaccines during emergency conditions like a pandemic.

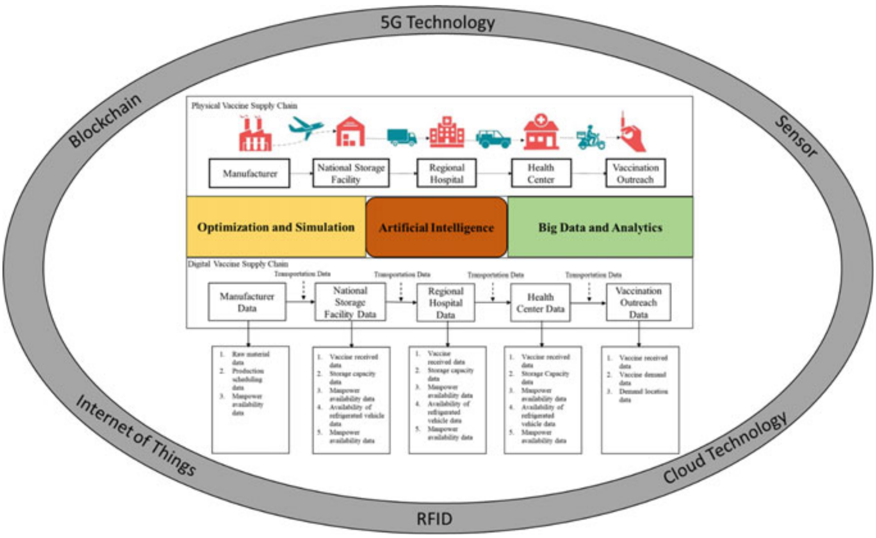


Fig. 18.2 5G enabled vaccine supply chain framework: digital twin

18.2.1 Timely Response

5G technology provides high speed (>1 Gbps) of data transmission across different stages of the vaccine supply chain network. Moreover, due high speed of data sharing every stakeholder in the SCN starts getting a fast or timely response. Also, get real-time feedback which helps to minimize the error rate.

18.2.2 Improved Visibility

The data and different digital technologies like 5G, IoT, sensors, and RFID help to enhance the real-time end-to-end VSC visibility. After performing three major operations like planning, monitoring, and control of SC with real-time data stemming and supported by the application of artificial intelligence and data analytics.

18.2.3 Improved Connectivity

The 5G technology enables a higher level of connectivity across the supply chain network structure. This leads to a high level of transparency between different nodes and traceability of operations of the supply network. Moreover, 5G technology also allows it to handle a large volume of information and it is processed close to the source as much as possible and transferred only information to the next level.

18.3 Expected Response of Proposed Framework

Simulation approach help to implement the proposed framework under different disruptive scenarios such as power breakdown, labour shortage and shortage of refrigerated vehicles etc. Virtually simulate the proposed framework of the vaccine supply chain based on captured data at different stages of the supply chain network. Figure 18.3 shows the impact of 5G along with the different disruptive technologies like blockchain, Internet of Things (IoT), Cloud, RFID, Sensors etc. to make the vaccine supply chain responsive and also help to improve the visibility, connectivity, timely response, and real-time feedback across the SCN.

The logistics and supply chain need to track the state changes from manufacturer to the national storage facility, national storage facility to the regional hospital, regional hospital to the health centre, and health centre to vaccination outreach. On the other hand, it needs to track the distribution of vaccines among stakeholders in the supply chain.

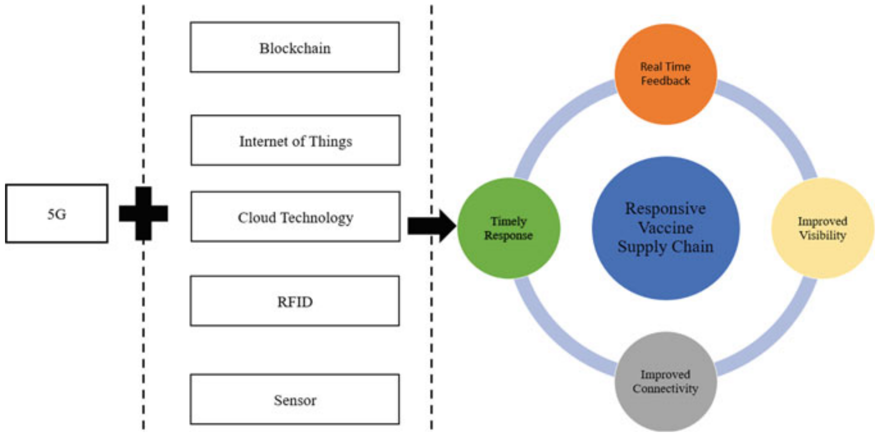


Fig. 18.3 Impact of 5G including other emerging technologies on vaccine supply chain performance

With the low delay and high-reliability characteristics of 5G technology, logistics tracking can be realized. The support of cloud computing technology is crucial to monitor and track all elements and the entire process.

18.4 Application of 5G Technologies on Vaccine Supply Chain

To check the applicability of emerging digital technologies in the SC it is broadly categorized into three levels strategic, operational, and cross-functional level.

18.4.1 Strategic Level

At the strategic level, we have considered new digital business models, digital supply chain ecosystems, mapping virtual and geographical SC footprints, sustainability, and resilience to minimize and control the risk impact on the supply chain.

18.4.2 Operational Level

At the operational level, we have considered demand forecasting, analytics, smart capacity planning, digital production planning, automated inventory, warehouse management, and intelligent fleet management to minimize and control the risk impact on the supply chain.

18.4.3 Cross-Functional Level

At the cross-functional level, we have considered autonomous and connected vehicles and SC logistics, smart cities, and smart homes as part of the supply chain ecosystem, E-Healthcare, and supply chain management to minimize and control the risk impact on the supply chain.

18.5 Discussion

The digitalization of the SC and its capability for practical implementation greatly depends on the data, its availability, and its accuracy. The group of technologies such as cloud, 5G, blockchain, sensor, and RFID help to enhance the visibility and connectivity of the data-driven digital supply chain (DSC) in real time.

Network-wise local digitalization enables the digital process and leads to a complete DSC. Moreover, some potential benefits of emerging digital technologies have been declared. Our paper aims to address the challenges in the VSC and the proposed framework to control the shortage and wastage of vaccines in the supply chain.

We established the discussion based on two major capabilities of the DSC which can be heightened by 5G technology, cloud technology, blockchain, sensor, IoT, and RFID i.e. visibility and connectivity within the SC network structure. Our analysis includes strategic and operational processes. Under operational process the transformation of manufacturing, warehouse, and distribution operations by end-to-end connectivity of devices. From the strategic perspective, the transformation of the SC network structure helps to improve the end-to-end visibility and connectivity of various stakeholders of the VSC.

18.6 Conclusion and Future Scope

The Internet 5G supply chain management system reduces the time for delivery of diverse components, controls manufacturing, storage, and distribution, and increases supply chain management efficiency. Pay attention to supply chain visualisation and information control at critical industrial positions. Monitoring and managing raw materials, assembly, quality control, sale, driving, and maintenance, and increasing production, storage, and distribution. Intelligent supply chain management must be visible and transparent, eliminate stock variations due to early or late delivery, improve automation and intelligence, and reduce human mistakes. It simplifies and streamlines the logistics process and gives decision-makers real-time monitoring and service.

The proposed VSC risk analytics framework could help supply chain professionals to minimize and control the impact of various types of disruptions like pandemics, war, tsunamis, etc. such disruptions usually lead the power breakdown, labour shortages and interruptions in connectivity. Also, the application of different technologies in VSC would make a positive impact and help to increase visibility and connectivity across the SC network. Finally, this will help to minimize and control the impact of risk within the VSC network structure. VSC risk control using emerging technologies also help to minimize the wastage and shortage of vaccine during emergency time. In the future, we can model this proposed framework to check its effectiveness and usefulness in the industry.

References

1. Cheng J, Yang Y, Zou X, Zuo Y (2022) 5G in manufacturing: a literature review and future research. Int J Adv Manuf Technol 1–23
2. Duttagupta C, Bhattacharyya D, Narayanan P, Pattanshetty SM (2017) Vaccine wastage at the level of service delivery: a cross-sectional study. Public Health 148:63–65
3. Dolgui A, Ivanov D (2022) 5G in digital supply chain and operations management: fostering flexibility, end-to-end connectivity, and real-time visibility through internet-of-everything. Int J Prod Res 60(2):442–451
4. Guichard S, Hymbaugh K, Burkholder B, Diorditsa S, Navarro C, Ahmed S, Rahman MM (2010) Vaccine wastage in Bangladesh. Vaccine 28(3):858–863
5. Ivanov D (2021) Digital supply chain management and technology to enhance resilience by building and using end-to-end visibility during the COVID-19 pandemic. IEEE Trans Eng Manag
6. Ivanov D, Dolgui A, Das A, Sokolov B (2019) Digital supply chain twins: managing the ripple effect, resilience, and disruption risks by data-driven optimization, simulation, and visibility. In: Handbook of ripple effects in the supply chain. Springer, Cham, pp 309–332
7. Ivanov D, Dolgui A, Sokolov B (2019) The impact of digital technology and Industry 4.0 on the ripple effect and supply chain risk analytics. Int J Prod Res 57(3):829–846
8. Lazarus JV, Karim SSA, van Selm L, Doran J, Batista C, Amor YB, Yadav P (2022) COVID-19 vaccine wastage in the midst of vaccine inequity: causes, types and practical steps. BMJ Glob Health 7(4):e009010
9. Liu Y, Zheng J (2022) Intelligent management of supply chain logistics based on 5G LoT. Clust Comput 1–10
10. McClellan M, Udayakumar K, Merson M, Edson G (2021) Reducing global COVID vaccine shortages: new research and recommendations for US leadership. Policy Brief. Duke Global Health Innovation Center
11. Miranda-García MA, Hoffelner M, Stoll H, Ruhaltinger D, Cichutek K, Siedler A, Bekeredjian-Ding I (2022) A 5-year look-back at the notification and management of vaccine supply shortages in Germany. Eurosurveillance 27(17):2100167
12. Rejeb A, Keogh JG (2021) 5G networks in the value chain. Wirel Pers Commun 117(2):1577–1599
13. Surantha N, Sutisna N, Nagao Y, Ochi H (2017) SoC design with HW/SW co-design methodology for wireless communication system. In: 2017 17th International symposium on communications and information technologies (ISCIT). IEEE, pp 1–6

14. Torjesen I (2021) Covid-19 vaccine shortages: what is the cause and what are the implications? BMJ: Br Med J (Online) 372
15. Wallace AS, Willis F, Nwaze E, Dieng B, Sipilanyambe N, Daniels D, Ryman TK (2017) Vaccine wastage in Nigeria: an assessment of wastage rates and related vaccinator knowledge, attitudes and practices. Vaccine 35(48):6751–6758
16. Yan K, Liu L, Xiang Y, Jin Q (2020) Guest editorial: AI and machine learning solution cyber intelligence technologies: new methodologies and applications. IEEE Trans Industr Inf 16(10):6626–6631

Chapter 19
Issues in Procurement and Distribution of Plantation Crops: Can AI-ML Technologies Offer Better Performance Outcomes?

Joshin John

Abstract In this paper, we explore the issues in supply chain procurement and distribution of plantation crops. Specifically, we explore challenges related to visibility and information sharing related to production data, problems in price disparity, supply–demand mismatch, and procurement auctioning. The context is analyzed under the lens of a case study on rubber crop plantation in the state of Kerala. We use simple comparison and review to explore the scope of applicability of data analytics and AI-ML for superior performance outcomes. The paper concludes with a discussion on implications to practitioners, limitations, and future scope of research.

Keywords Procurement · Distribution · Plantation crops · AI-ML technologies · Scope review · Case study

19.1 Introduction

While Artificial Intelligence (AI) and Machine Learning (ML) are often trumpeted as go-to-technologies to solve business problems of today, it may be remembered that Generative AI and AI-Augmented systems are currently near the peak of inflated expectations on Gartners' Hype Cycle [2]. In this paper, we critically examine the scope of application of data analytics, AI and ML technologies in competitive market auctioning for procurement of plantation crops as raw commodities. The rest of this paper is structured as follows: Sect. 19.2 describes the plantation crops business, and frames procurement related issues. A case illustration is presented in Sect. 2.2. In Sect. 19.3, we assess the current business practices and compare it with Data analytics and AI-ML enabled outcomes. The paper concludes with a discussion, detailing implication for practitioners, limitations of the study, and future scope of research.

J. John (✉)
Xavier Institute of Management and Entrepreneurship, Kochi, Kerala, India
e-mail: joshin@xime.org; joshin.john@iiml.ac.in

M. K. Tiwari et al. (eds.), *Applications of Emerging Technologies and AI/ML Algorithms*, Asset Analytics, https://doi.org/10.1007/978-981-99-1019-9_19

19.2 Plantation Crops, Procurement Process, and Related Issues

Plantation crops, although are generally considered under the broad category of agricultural crops, are different food crops. This is because unlike, agricultural food crops like paddy, wheat or maize, plantation crops generally fall under the category of cash crops, and are grown in agricultural estates. Major plantation crops include rubber, tea, coffee, cotton, fruit tree plantations, cocoa, oil seed, oil palm, spice plantations, etc.

While the produce from plantations also come under the category of commodities, there are important differences that distinguish them from typical food crops such as rice and wheat. One, the plantation crops typically have a longer life cycle. For instance, rubber has a typical lifespan of 30 years, tea and coffee on an average last for 40 years. This is very different from food crops that are seasonal, and have to be planted, cultivated and harvested during every season. Two, typically most of the plantation crops after basic processing have longer shelf life. After value addition, the time to perish under good storage conditions for these crops are even longer—sometimes even 2–3 years or more. While there are several issues in production and harvesting of plantation crops, we limit the scope of this paper to issues related to public procurement and distribution systems.

19.2.1 The Procurement and Distribution Process

The specifics of procurement and distribution process varies depending on the type of crop. While small farmers produce crops in small lots and sell it as raw commodity either to farmer cooperative societies, processing plants or in the commodity market, large plantation estates may have in-house facilities for value addition, and may be vertically integrated in the value chain. A generic representation of value chain is represented for a plantation crop (rubber) in Fig. 19.1. Similar, value chains exist for tea, coffee, etc. but we limit the scope of analysis to rubber in this paper for parsimony of analysis and brevity.

The procurement process is multi-staged, with some complexity in the mechanics of the supply chain network. At procurement point 1, the latex that is collected from the rubber trees is procured either by Rubber Processing Societies (RPS) or by private processing plants. The processing of rubber happens through a series of activities: emulsification wherein the latex is treated with acids and emulsifiers, then shaped into sheets in rolling machines, and further dried in smoke-houses, graded into different categories to be shipped. At procurement point 2, the dried rubber sheets are processed by different businesses such as tyre manufacturers, producers of rubber packaging materials, medical companies, etc. The price of procurement is the minimum of Minimum Support Price (MSP) set by the government (Rubber Board) or determined by market forces (above MSP). The procurement companies

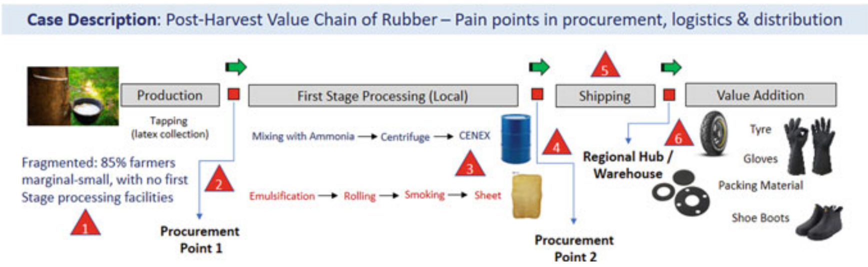


Fig. 19.1 Value chain for rubber with points of procurement indicated

also purchase synthetic rubber (from petroleum) as a substitute to natural rubber depending upon the prices.

19.2.2 Case Study: Procurement and Distribution Issues in Rubber Value Chain

In this subsection, we discuss the procurement issues in the value chain for rubber. We illustrate this using supply chain issues in the latex-rubber business in the state of Kerala. We select the rubber plantation business in the state as the unit of analysis for the following reasons. Kerala is known as the plantation state of India. The major plantation crops in Kerala are rubber, tea, coffee and cardamom. Kerala accounts for about seventy-five percent of 712,000 tonnes of rubber produced in India, and hence, the associated supply chain and public procurement practices has a significant impact on the market, and supply chain performance.

Of the 540,000 ha of land that are under rubber plantations in the state, 85% of the rubber plantation land are owned and operated by marginal and small farmers with less than two hectares of land. However, a majority of the marginal or small farmers do not own processing facilities such as rollers or smoke-houses. They sell the natural latex mix collected to processing plants. This procurement point 1 (refer Fig. 19.1) is often stacked against the interests of the farmer's with very less bargaining power as compared to the processing plant. The Rubber Processing Societies (established in 1986) was formed to protect the interests of small rubber growers and enable collective bargaining. While this helped to some extent to ensure fair price, the procurement process could also be manipulated against fair pricing at procurement point 2 (refer Fig. 19.1).

The raw materials sourced by value added processing units (tyre manufacturers, rubber-based packaging companies, medical companies, etc.) were purchased at procurement point 2. However, this point of procurement was vulnerable to price manipulation. This is due to the following reasons: (a) The tyre processing and other rubber manufacturing (value-addition) units in India, retained the option to import rubber from other countries such as Thailand and Indonesia, often flooding

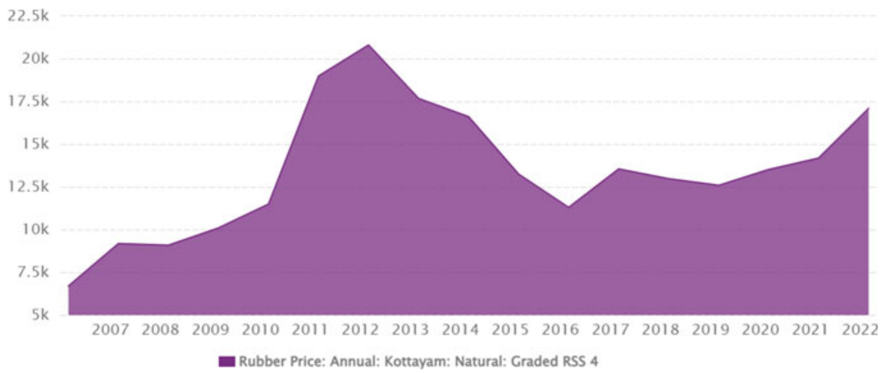


Fig. 19.2 Price of rubber in rupee (₹) per 100 kg from 2006 to 2022 (RSS4 Grade Kottayam)

the domestic market, leading to crash in rubber prices. The rubber price (RSS4 grade) over the last few years is shown in Fig. 19.2. (b) Free trade agreements with ASEAN countries in south east Asian region, that accounted to majority of world rubber production, often led to price manipulation in the domestic market which rendered the domestic rubber not competitive. This was despite supply shortage in the domestic markets.

19.2.2.1 Supply Control Options at Production Side for Domestic Planters, and Issues Faced by Stakeholders

While the local rubber producers (planters) in Kerala had options to control natural rubber supply by adjusting the frequency of latex tapping, production, and inventory levels, the use of these options was limited by the following: (a) While access to international and domestic commodity markets including pricing and volume details were available, there was little visibility on the volume of production at supply side on a real-time basis. Few large-scale producers had enterprise level MIS that could track inventory, but by and large, marginal and small rubber growers were fragmented with low level of information sharing, (b) production (latex tapping) was largely labor intensive. While labor force capacity could be planned for production, unavailability of local labor to work in plantations led to capacity constraints. Latex tapping in Kerala was unviable at rubber prices below ₹120 per kg. In recent years, much of the plantation labor force were migrant workers from other states seeking higher wages in the state of Kerala, which had an HDI of 0.79, the highest in India. (c) While the amount of latex produced by rubber growers in Kerala was almost three-quarters of national production, the processing and value addition in the value chain to a point near the final consumption was almost nil. This means significant erosion of potential margins, due to lack of consumption platforms that can be availed by the rubber growers for processing, and selling. (d) Uncertainty in weather patterns, and climate change also led to reduction in yield. The planters retained the option to switch to other crops, or practice multi-cropping, inter or mixed cropping practices in a limited scale allowed by government regulations. However, this option was hardly used due to lack of information, and prediction capabilities and insights for decision making, (e) inventory handling, and logistics costs are realized as hidden costs that can be optimized.

19.3 Comparing Current Business Practices with Data Analytics and AI-ML Enabled Outcomes

In this section we benchmark current business practices in the crop value chain with data-analytics and AI-ML enabled outcomes. While the context is grounded in the case described in the previous section, the insights may be generalizable to other plantation crops, unless there are significant differences in processes. The comparison is at a conceptual level given analytics capabilities that have been theoretically demonstrated. However, confirmation of the same is pending in an extension of this work involving quantitative treatment of the case discussed.

In Table 19.1, we list the business practices as-is, and the proposed treatment using data analytics or AI-ML technologies. While we observe that better real-time visibility through information sharing on production data using MIS starts at base level, diagnostic capabilities to analyze (why?) and handle demand supply mismatches, price fluctuations, procurement, inventory holding, supply chain segmentation, productivity, and auctioning related issues are promising.

Further, we propose to the scope of application in the given context, use of predictive analytics tools for production and inventory planning. OR tools for optimization may be best deployed for inventory and logistics optimization in the crop procurement and distribution supply chain. The proposed scope of application of data analytics tools, is summarized in Table 19.1.

19.4 Discussion

While we proposed several data analytics-based interventions to address the issues listed in procurement and distribution of plantation crops, it is important to discuss the implications of the same to practitioners. Firstly, the quality of data and information sharing capabilities are prime concerns [4]. The network-infrastructure, connectivity, integrity of data and the systemic capability are hygiene concerns. Once this is ready, the diagnostic capabilities can be set up, which is relatively easy, given the standard codes and algorithms available that can be suited to the context. However, the effectiveness of the supply chain with analytics will not only be depended on the physical infrastructure (distribution centers, storage facilities, transport vehicles

Table 19.1 Proposed treatment based on DA/ AI-ML interventions

PP #	Issue	Current practice	DA/AL-ML intervention	Algorithm/remark
1a	Production data	Manual periodic aggregation	Real-time aggregation: MIS/ ERP	–
2a	Fall in crop yield, crop infection	Nil (absorb losses)	Precision agriculture (e.g. Target Chem. Spray)	Image recognition (Diagnostic)-SVM
2b	Demand supply mismatch, price crash	Govt. procures excess crop, minimum support prices	Production control based on demand	Game theory: adversarial network
2c	Procurement price/volume disparities	Nil (procurement inefficiencies)	Block chain technology	Transparent ledgers for approvals
2d	Auctioning system	Commodity markets	Auctioning algorithms	Bidder Reco, GT-ANN
3	Planning/ forecast	Manual/macro level	Regression based	CNN, DNN
4	SC segmentation	Limited or no intervention	Clustering algorithms	K-Mean, affinity
5	Logistics optimization	Sub-optimal	Vehicle routing	NP-Hard
6	Inv. optimization	Inadequate	MILP	NP-Hard

etc.), but also on the development of analytic culture that is built around systematic data capture and data governance strategy [3].

Although we illustrated the public procurement context through means of a case study, it is limited to the rubber plantation crop in the state of Kerala in India. It is important to pursue and validate the procurement and distribution related issues listed in other geographies and for other crops. While this is a limitation of this study, it also offers exciting future research opportunities combining analytics in Industry 4.0 settings [1]. Confirmatory studies for the proposed treatments in improvement supply chain performance outcomes need to be carried out. Future research possibilities not only include application of diagnostic, predictive and prescriptive tools to improve supply chain practices, but also methodological improvements to develop superior algorithms suited to specific contexts.

References

1. Ivanov D, Tang CS, Dolgui A, Battini D, Das A (2021) Researchers' perspectives on Industry 4.0: multi-disciplinary analysis and opportunities for operations management. Int J Prod Res 59(7):2055–2078

2. Gartner's Hype cycle (2021) Available at: https://www.gartner.com/en/information-technology/glossary/hype-cycle. Last accessed 05 May 2022
3. Handfield R, Jeong S, Choi T (2019) Emerging procurement technology: data analytics and cognitive analytics. Int J Phys Distrib Logist Manag 49(10):972–1002
4. Soylu A, Corcho Ó, Elvesæter B, Badenes-Olmedo C, Yedro-Martínez F, Kovacic M, Posinkovic M, Medvešček M, Makgill I, Taggart C, Simperl E (2022) Data quality barriers for transparency in public procurement. Information 13(2):99

Chapter 20
Quantifying the Quality Grade of the Return Mobile Phone in the Context of a Retail Store

Satchidananda Tripathy and Akhilesh Kumar

Abstract Products including garments, clothing accessories, and consumer electronics are prone to suffer from short product life cycles as a result of fast technological advancements, manufacturers' marketing methods, and rapidly changing customer tastes. With the growing demand for the technology-driven product, the mobile phone's lifespan has fallen to three to six months. To compete with the global market, the mobile manufacturer cut down the price of mobile phones and provides upgraded specifications. For this reason, consumer wants to experience the new technology in every instance and exchange their mobile phone for a newer one. The performance of the mobile phone does not reduce, but to upgrade to modern technology, the customer intends to swap it with the fresh one. For this purpose, a circular economy comes into the picture. Remanufacturing is becoming a critical element of a circular economy where products are created, produced, used, and retrieved to avoid any kind of waste and decrease the extraction of raw materials. However, lots of studies are carried out on the quality grading of the returned core, but we did not find any of the quantified definite indexes for the quality of the yielding core. Thus, in these articles, we attempt to quantify the quality of returned core (mobile phone) by using MCDM and PCA / FA methods. Finally, the reduction equation is established to predict the quality of returned core, which reduces the remanufacturing cost by providing optimal shorting.

Keywords Remanufacturing · Optimal sorting · Quality of core returned · PCA · MCDM

S. Tripathy (✉)
Department of Industrial and Systems Engineering, Indian Institute of Technology, Kharagpur, West Bengal 721302, India
e-mail: satchi.iitkgp18@gmail.com

A. Kumar
Operations Management and Quantitative Methods, Indian Institute of Management Raipur, Chhattisgarh 493661, India
e-mail: akumar@iimraipur.ac.in

M. K. Tiwari et al. (eds.), *Applications of Emerging Technologies and AI/ML Algorithms*, Asset Analytics, https://doi.org/10.1007/978-981-99-1019-9_20

20.1 Introduction

Due to fast invention and development in science and technology, as well as customer behavior in chasing the newest innovation and style, technology-based products have a shorter life cycle. According to the Lebreton & Tuma [6] technology-based commodities such as mobile phones and computers have a quicker innovation cycle, which means the previous generation becomes obsolete more quickly, both functionally and psychologically. Similarly, Hsueh [5] noted that, as a consequence of technological advancements, the product life cycle in the electronics sector is shorter than previously, and as a result, an obsolete product might reach its end-of-use even if it is still in excellent condition. A shorter life-cycle has a negative contribution toward sustainability since there is an increase in product disposal. Customers prefer newer things and dump the old ones, and these choices would quickly deplete landfill space. Furthermore, due to unnecessarily increasing obsolescence, more natural resources and energy are spent to develop new items than is required. To make matters worse, electronic items are increasingly recognized as having a shorter and shorter life cycle, while the wastes are poisonous and unfriendly to the environment.

Several solutions, such as the life cycle approach, regulation, and society approach, have been established to reduce product disposal and waste. Dealing with products at their end-of-life is one part of the life cycle strategy. According to de Brito and Dekker [3] End-of-use returns are circumstances in which a consumer has the option to return a product at a given stage of its life cycle, such as leased cases and returnable containers. Hsueh [5] explored a distinct type of return, in which a product is returned because it has become obsolete and the consumer wants to purchase a new one. Remanufacturing is one option for managing items at their end of life, and it allows for regulatory compliance while preserving profitability. By the turn of the century, America had over 73,000 remanufacturing businesses. Some Chinese companies are also getting into the remanufacturing business. In Shanghai, for example, the first Chinese self-R&D toner cartridge recycling system was launched in 2006. Product return management, remanufacturing operations, and market development for remanufactured products are the three primary tasks in remanufacturing program.

However, many management issues in the remanufacturing industry have still to be researched and handled. The variability in the timing, amount, and quality of the returns is one of the most difficult aspects of remanufacturing. Paying end customers a collection fee to recover the discarded goods is one approach to alleviate this ambiguity. This method gives the company control over returned used products, and it is commonly employed in the remanufacturing business. Even so, not all of the collected cores can be remanufactured due to quality uncertainty. Therefore, the random yield in remanufacturing process is another important factor.

In the recent decade, there has been extensive research on remanufacturing and reverse logistics. While most of them are regarding the reverse logistics networks design and the inventory planning, the literature on used product acquisition, their exchange pricing, and the quality of returned core are limited. So, there is a need for a sorting technique according to the quality of returned core before going for disassembly.

In this research, we give a fresh sorting technique for the return smartphone based on factors such as potential component reusability, consumer identification information, consumer safety status, and external variables such as market demand behavior of second-hand smartphones. This document aims to enhance decision-making in remanufacturing operations by incorporating product life cycle information, especially product use stage data, into both ideal sorting strategies as well as end-of-life / end-of-use choices. To achieve this, two associated analyses are carried out: first, for each item unit, a reusability index is obtained based on the reusability of its parts, item characteristics, and item use stage information. Second, the reusability coefficient for each product category is used to find the composite index for obtaining the quality grade of the return core. For this purpose, we use the MCDM processes like TOPSIS and PCA/FA. The approach is based on the data gathered from the item's use point. The proposed construct is universal and can be applied to any data set that indicates the value or reusability of the item.

20.2 Research Methodology

To obtain detailed information on the lifespan of the outdated smartphone in Kharagpur and Medinipur, we engaged in conversation with a trader of the familiar second-hand mobile phone recycling market. Most of these traders indicate that local user was the leading source of second-hand mobile. Some stores are re-polishing and repairing their own, while others are sending them to intermediate restoring stores before reselling them to customers at rates below the original. We gather some data in this survey by preparing a questionnaire and studying the data to know the value of a mobile phone exchange in the retail shop. From the information gathered, we rank them by MCDM operations like TOPSIS and PCA / FA. In this study, we bring nine-factor with some level to forecast the old mobile phone's condition. We processed the raw data and handled the missing data by the average value of the corresponding variable before implementing the technique. We ranked the factor and developed a reduction equation. By using random factors in excel, the mobile phone situation is generated, and the quality of the exchange smartphone is predicted.

20.2.1 Determination of Closeness Ratio Using TOPSIS

We modeled our problem as $V_i = [v_i | i = 1, \ldots, m]$ several vendor options, $F = [f_k | k = 1, \ldots, n]$ factor affects the quality of return core(smartphone), and $W_j = \left[w_j | j = 1, \ldots, n\right]$, $w_j \geq 0$ weight of the factor, where $\sum_{j=1}^{n} w_j = 1$ the rank of the significance of the requirements, $X = \left[x_{ij} | i = 1, \ldots, m;\ j = 1, \ldots, n\right]$ the choice matrix, where the alternative's results score is based on the criterion. In this analysis equal weightage is given to each factor affecting the return core.

20.2.2 Determination of Composite Reduction Index (CRI) for Cell Phone Quality Evaluation

Since factors had different measurement scales; we normalized the entire factor by using this equation $X^{-} = \frac{x_{max}^{-} - x^{-}}{x_{max}^{-} - x_{min}^{-}}$ (As all the factors are a negative impact on the quality grade index) Then, the intermediate reduction indices IRI_j, related to each of the principal components j, must be calculated for weight calculation once the principal components or parameters have been acquired. This is accomplished by estimating an accumulation of weighted measures [4] : $IRI_j = \sum_{k=1}^{n} w_{kj} I_k$; Where IRI, an intermediate indicator of reusability; w, indicator weight; I, standardized indicator; j, component number; k, indicator number. The w_{kj} weights are obtained from the loading factor: $w_{kj} = \frac{(factorloading_{kj})^2}{eigenvalue_j}$. Composite reduction indexes were finally estimated as a weighted accumulation of intervening indexes of reusability: $= \sum_{j=1}^{m} \alpha_j IRI_j$; Where CI, composite reduction indices; j, the number of principal components; IRI, the intermediate reusability indicator; α_j the weight applied to the intermediate reusability indicator. This weight is calculated as follows: $\alpha_j = \frac{eigenvalue_j}{\sum_{j=1}^{m} eigenvalue_j}$.

20.2.3 Data Creation and Cleaning

Further for determining the quality grade of the product, we took help from the local vendors. The quality perceived by them is based on their visual attributes. Thus, to get detailed knowledge about their decision making we have devised a questionnaire (Appendix 1) which was distributed among local vendors. Our sample size consists of 18 data points. We surveyed the mobile phone retailers in Kharagpur and the Midnapore area and covered all the retailers who facilitate returns or exchange old phones for new ones. In Kharagpur, we obtained 12 sample data, collected from Tech Market, Gole Bazar, and Malancha, while we obtained 6 sample data in Medinapore majorly surveyed from School Bazar, Station Road, and Keranitola. The following factors were taken into the consideration for devising the questionnaire:

1. Effect of usage period on return price
2. Effect of the warranty period
3. Brand Value
4. The severity of Component Damage and Repair Cost
5. Various Operating Conditions and payoffs associated with the same

The primary survey dataset consisted of 30 variables and 18 data points. There were nuances in the dataset mainly due to the vendor's lack of experience in dealing with a particular brand, encountering particular operational/damage conditions of mobile. In this case, they were unable to provide us with the relevant data. Also, in case of mobile conditions that required minor repairing, the handset was repaired by a third party, and the reduction in exchange price due to repair was based on the price quoted by mobile repair personnel. Thus, vendors dealing with only exchanges were not able to provide us with the data related to repair costs. The following table describes the data imputation task for the variables (Table 20.1).

Table 20.1 Survey data cleaning

<table>
<tr><th>Factors</th><th>Associated variables</th><th>Data type</th><th>Data cleaning method</th></tr>
<tr><td>Reduction in case of operational condition</td><td>• Crack on screen
• Body/ back panel damage
• Crack on camera
• Dead Battery</td><td>Numerical (Range given in some cases)</td><td rowspan="2">– The range was replaced by the average
– Missing values were replaced by the mean of a dataset
– Replaced 'condition not purchased' case with dataset mean</td></tr>
<tr><td>Reduction due to component dysfunction (Repair cost)</td><td>• Fingerprint
• Volume/Power Button
• Visible Scratches
• Wifi
• Speakers
• GSM module</td><td>Numerical (Range given in some cases)</td></tr>
<tr><td>Additional benefits due to accessories</td><td>• Charger
• Earphone
• Original Bill</td><td>Numerical + Categorical</td><td>– The range was replaced by mean
– Removed 'Original Bill' column as many stated it was necessary</td></tr>
<tr><td>Reduction due to age factor</td><td>• 1–3 months old
• 3–6 months old
• 6–12 months old
• 12–24 months old
• More than 2 years</td><td>Numerical (Range given in some cases)</td><td>– The range was replaced by mean
– Removed the case of reduction due to warranty expiry as it was highly dependent on age</td></tr>
<tr><td>Brand Profitability</td><td>• Apple
• Vivo/Vivo
• Oneplus
• Xiaomi
• Micromax/ Intex
• Lenovo/ Huawei</td><td>Categorical(4categories)</td><td>– Imputed mode values for missing data
– Oneplus and Apple were highly correlated. For apple, missing data used values of Oneplus</td></tr>
</table>

20.3 Result and Discussion

Since the acquired core is indiscriminate and the quality of each core is different and unidentified before procurement (i.e., quality variability), the remanufacturer should estimate the quality of the acquired core based on historical pre-acquisition data. In this paper, we find the closeness value using TOPSIS and reduction index by PCA to predict the quality of return smartphones.

20.3.1 Survey Data Analysis

We analyzed the individual variables in our survey by plotting their graphs. The data given by all the vendors were highly correlated thus showing the obvious nature that reduction increases as the phone gets old (Fig. 20.1).

Next, we observe the effect of various operating conditions. The reduction in the case of conditions considered is quite significant. These reduction percentages are over the top of age reduction. During the survey, some of the vendors refused to take returns of such defaulted products and even some of them were taking the returns, they offered negligible prices for such pieces. The dead Battery case turned out to be quite severe as in such cases there is no guarantee that the phone might not even reboot. While high value of camera reduction is attributed to the trend of people looking for a good camera in excellent condition (Fig. 20.2).

As we observe (Fig. 20.3) the repairing cost, we can infer that the reduction is not high in such dysfunctional conditions. Sometimes the repairing task does not even include component replacement. While in some cases problem might be quite severe. Thus, a standard deviation of reduction is a bit high in such cases.

Survey data shows that the Indian Companies are way behind their counterpart in the return mobile market as 10 out of 16 vendors stated there is insignificant profit for Micromax/Intex models. Currently, leading brands in the mid- segment such as Oppo, Vivo, and Samsung are performing well while high-end phones of Apple and

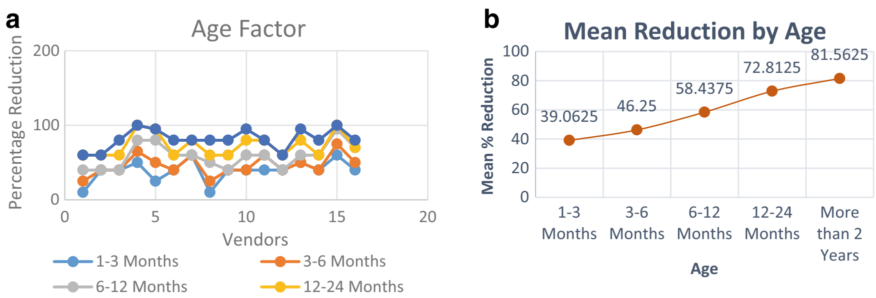


Fig. 20.1 **a** Survey Data for age. **b** Mean reduction due to age

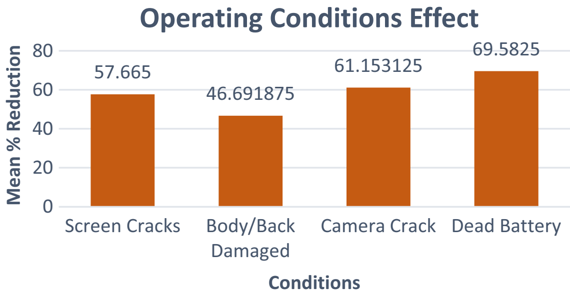


Fig. 20.2 Reduction in case of operating conditions

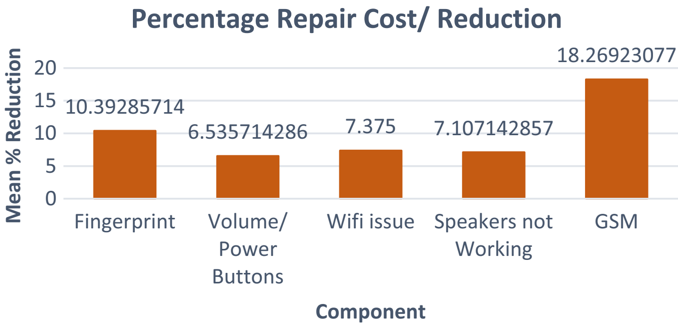


Fig. 20.3 Reduction in case of repairing involved

One Plus though having high profitability, suffer from a low customer on buyers as well as sellers' sides. Now, we applied the TOPSIS method to find out the importance of Factor with level for reducing the exchange price of smart-phone. To analyses the number of component influences, we use the Pearson correlation matrix and scree plot.

20.3.2 Ranking of Influencing Factors by TOPSIS and Their Correlation

From the closeness Value, we rank the influential factors and assign weightage to the factors. As the lower reduction factors give better quality so here we consider lower is the better case. From Table 20.2 we found that age is the main factor for the reduction in the quality of returned core. With increases in age, quality decreases.

In the TOPSIS method, we simply normalized the data, but the variable is correlated with each other, so we apply PCA to find out the composite reusability index that

Table 20.2 Closeness value

Factor	Closeness value	Ranking
1-3moth	0.248	1st
3-6 month	0.332	2nd
6-12 month	0.514	4th
12-24 month	0.702	8th
2 year	0.834	9th
Crack on screen	0.518	6th
Body	0.333	3rd
Camera	0.517	5th
Battery	0.603	7th

similarly predicts the quality grade. To analyse the number of component influences, we use the Pearson correlation matrix and scree plot.

The Pearson correlation matrix was ready for the retail information gathered and displayed in Table 20.3 within the parameters analysed.

Factors such as camera, battery, body, and crack on display indicate a significant connection with mobile phone age 1–3 months of 0.03, 0.23, 0.18, and -0.06 respectively. 3–6 months indicate a powerful connection with a Pearson correlation ratio of 0.79.6–12 months, 12–24 months, and above two years demonstrate a slight coefficient of correlation of 0.42, 0.45, and 0.41, respectively. When parameters have a powerful or moderate link, the transformation of linked parameters into uncorrelated parameters is crucial to predict the performance of a mobile phone effectively. An appropriate technique for transforming correlated parameters into uncorrelated parts is provided by the principal component assessment. Then it is possible to evaluate the number of PCs compelled in the research from the scree plot shown in Fig. 20.4.

It is noted that three principal components describing 75.12% of the complete variability are adequate in the scree matrix for the research, whose Eigen scores are greater than one. The PC1 accounts for 41.2%, PC2 accounts for 20.91%, and PC3 account for 12.91% of the complete variability as calculated by loads for a cumulative proportion by R.

20.3.3 Composite Reduction Index for Return Smartphone

Using three principal components we calculate the composite reduction index (CRI) by applying the above mention algorithm described in the method section (Table 20.4).

Interestingly, the composite reduction index shows that the age of mobile phone effect constant reduction after three months up to 12 months then again slightly more reduction due to the expiry of the warranty period and also the obsolesce of technology. After two years, an almost 46% reduction is there from the original price

Table 20.3 Pearson correlation matrix

	1–3 moth	3–6 month	6–12 month	12–24 month	2 year	Screen	body	camera	battery
1–3 moth	1								
3–6 month	0.79	1							
6–12 month	0.42	0.77	1						
12–24 month	0.45	0.82	0.87	1					
2 year	0.41	0.61	0.79	0.79	1				
screen	−0.06	−0.24	−0.32	−0.26	−0.26	1			
body	0.18	0.02	−0.12	−0.02	0.24	0.52	1		
camera	0.03	0.21	0.27	0.36	0.54	−0.18	0.27	1	
battery	0.23	0.01	−0.14	−0.09	0.01	0.54	0.51	−0.32	1

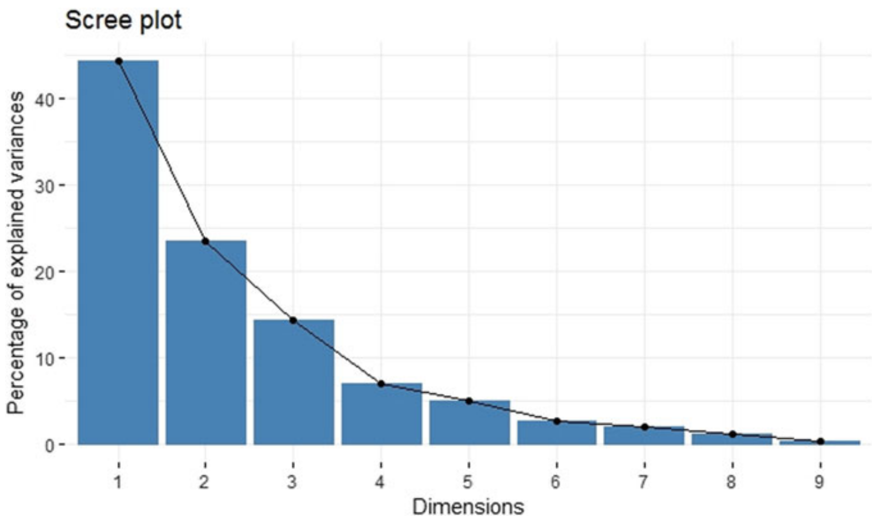


Fig. 20.4 Scree plot of nine mobile device performance factor parameters

Table 20.4 Composite reusability index for return smartphone

Factor	CRI
1–3 month	0.304
3-6 month	0.348
6-12 month	0.345
12-24 month	0.365
2 year	0.462
crack on screen	0.348
body	0.553
camera	0.438
battery	0.212

if the mobile is in good functional condition otherwise again reduce the price with dysfunctional the component. The battery of the smartphone can be replaced so it can affect less than another factor that is found by the Composite reduction index. From this analysis, we find that the composite reduction index gives better insight to grade the return smartphone. We see that with the age factor, the mobile phone's functional factor is also needed to predict the return mobile phone's quality. In our study, we consider four technical factors and found that if they do not work, then the exchange price will be reduced to 20–55%. Then, we try to establish an equation with dummy variables. For example, if present, the reduction will be applied, otherwise not.

Table 20.5 Predication of quality grade

Age	Screen	Camera	Battery	Body	Reduction with Factor	Quality
3	0	0	1	1	1.09	Recycles
2	0	1	0	0	0.44	High
1	1	0	1	1	0.65	Low
3	1	1	1	1	1.52	Recycles
1	1	1	0	1	0.60	Low
1	0	1	1	0	0.45	High
2	0	0	0	0	0.33	High
5	1	1	1	1	2.47	Recycles
5	1	1	1	0	1.97	Recycles
2	1	0	0	1	0.70	Low

$$Percentage\ reduction = 100\sum_{i=1}^{n} w_i \tau_i \beta_i \tag{20.1}$$

where $\tau_i =$ Age factor, β_i is the condition of returned mobile phone, n = number of influential factors, and w_i is the weight according to ranking found in the TOPSIS method

This Eq. 20.1 is used to predict the return mobile phone quality grade. Based on the reduction value, we categorized the return quality as high, moderate, low, and recycled as more than 80% as recycles, more than 60% as low, more than 50% as moderate, and less than 50% as high.

To illustrate this, we consider Age to be five levels, and the functional factor has two levels that work as 0 and dysfunctional as 1. Data were generated in Excel using the function = RANDBETWEEN (bottom, top). Some samples were shown in the Table 20.5.

We see that with the age factor, the mobile phone's functional factor also plays a vital role to predict the core quality. After two years the phone went to recycling and dismantled the part due to advanced technology available in the market.

20.4 Conclusion

A general index of reduction based on vital product features has been proposed for the measurement of return core quality level. To this end, we use the TOPSIS technique and the principal component analysis to calculate the reusability index for each mobile phone functional component. We compute the closeness value according to that ranking of the influential factors getting from the TOPSIS method. Composite reduction index (CRI) according to the weightage of the factors ranking from the TOPSIS method is used to predict the quality of returned core. We categorize them into three values. If the decrease is more than 80% as recycles, more than 60% as low, more than 50% as moderate, and less than 50% as high. Finally, to demonstrate this, we find Age to be five stages, and the vital variable has two stages that operate as 0 rather than as dysfunctional as 1. Using the function = RANDBETWEEN (bottom, top) data was produced in Excel.

In several respects, this job can be enhanced. The information we had was restricted to the information on the mobile phone only. Another Weight technique, like AHP, is also used to estimate the sustainability index. Clustering, supporting vector device, rainforest, and genetic algorithm ANFISS are then used to determine the performance grade using the secondary sustainability index.

References

1. Cai X, Lai M, Li X, Li Y, Wu X (2014) Optimal acquisition and production policy in a hybrid manufacturing/ remanufacturing system with core acquisition at different quality levels. Eur J Oper Res 233(2):374–382. https://doi.org/10.1016/j.ejor.2013.07.017
2. Deng QW, Liao HL, Xu BW, Liu XH (2017) The resource benefits evaluation model on remanufacturing processes of end-of-life construction machinery under the uncertainty in recycling price. Sustainability (Switzerland) 9(2). doi:https://doi.org/10.3390/su9020256
3. de Brito MP, Dekker R (2004) A framework for reverse logistics. Rev Log 3–27. https://doi.org/10.1007/978-3-540-24803-3_1
4. Gomez-Limon JA, Riesgo L (2008) Alternative approaches on constructing a composite indicator to measure agricultural sustainability
5. Hsueh CF (2011) An inventory control model with consideration of remanufacturing and product life cycle. Int J Prod Econ 133(2):645–652. https://doi.org/10.1016/J.IJPE.2011.05.007
6. Lebreton B, Tuma A (2006) A quantitative approach to assessing the profitability of car and truck tire remanufacturing. Int J Prod Econ 104(2):639–652. https://doi.org/10.1016/j.ijpe.2004.11.010
7. Mufazzal S, Muzakkir SM (2018) A new multi-criterion decision making (MCDM) method based on proximity indexed value for minimizing rank reversals. Comput Ind Eng 119:427–438. https://doi.org/10.1016/j.cie.2018.03.045
8. Yang C-H, Liu H-B, Ji P, Ma X (2016) Optimal acquisition and remanufacturing policies for multi-product remanufacturing systems. J Clean Prod 135:1571–1579. https://doi.org/10.1016/j.jclepro.2015.10.057
9. Yoon K (1987) A reconciliation among discrete compromise solutions. J Operat Res Soc 38(3):277–286. https://doi.org/10.1057/jors.1987.44
10. Zhou J, Deng Q, Li T (2018) Optimal acquisition and remanufacturing policies considering the effect of quality uncertainty on carbon emissions. J Clean Prod 186:180–190. https://doi.org/10.1016/j.jclepro.2018.03.150

Chapter 21
Integrated Blockchain Architecture for End-to-End Receivables Management of Indian MSMEs

Deepak Kumar, B. V. Phani, and Suman Saurabh

Abstract Small firms, also known as SMEs or MSMEs, face acute shortage of capital. These firms have huge pending receivables from large corporate buyers and government agencies which contributes significantly to this shortage. Despite a series of initiatives by the government, the issue remains critical to the survival and growth of MSMEs in India. This study explores the current issues and gaps in MSME receivables realization and government interventions in India, and proposes an integrated Blockchain-based architecture for end-to-end receivables management of Indian MSMEs using Ethereum. This will enable timely, cost-effective, transparent and competitive receivables realization for the small firms freeing up a significant amount of capital.

21.1 Introduction

Micro, Small and Medium enterprises (MSMEs) are recognized differently by different economies across the globe. Parameters like investment in plants and machinery, turnover, and the number of employees are used to classify firms into micro, small or medium. In India, these firms are classified into micro, small and medium based on investment in plant and machinery and turnover. Firms with investments in plants and machinery less than one crore and turnover less than five crores are termed as micro, with investments not exceeding ten crores and turnover below 50 crores as

D. Kumar (✉) · B. V. Phani · S. Saurabh
Indian Institute of Technology Kanpur, Kanpur, Uttar Pradesh, India
e-mail: kumard@iitk.ac.in

B. V. Phani
e-mail: bvphani@iitk.ac.in

S. Saurabh
e-mail: sumans@iitk.ac.in

M. K. Tiwari et al. (eds.), *Applications of Emerging Technologies and AI/ML Algorithms*, Asset Analytics, https://doi.org/10.1007/978-981-99-1019-9_21

small and investments less than 50 crores with turnover not exceeding 250 crores as medium enterprises. There are more than six crore such firms and constitute more than 99% of the firms in the country, contributing almost 30% to GDP, 40% to the total export and employing more than 100 million people. Despite their significant economic contributions, these firms face numerous challenges owing to their small size and limited resources.

These firms face perennial shortage of funds, which hinders their operation, growth and expansion. The opaque nature of such firms prevents access to credit. Besides, their sales on credit (receivables) tie up a considerable amount of funds eating up their working capital [1]. Significant delays in settlement of dues/payment of bills by large buyers, central and state departments, and public sector units (PSUs) have a negative impact on the recycling of money and business operations of these companies. Large purchasing corporations defer payments to their small suppliers to maintain the health of their cash conversion cycle [2]. This often makes MSMEs debt-ridden, eventually pushing them to the brink of insolvency.

The cash flow cycle of MSMEs gets stretched due to stuck and delayed payments, and most often, this is a deliberate strategy of big buyers. As per an estimate, an average of Rs. 15 lakh crore is owed to MSMEs at any point [3]. And this figure only represents the registered MSMEs which account for just 13% of all existing MSMEs in India. For micro enterprises, this problem is more severe, and more than 80% of them see payments delayed beyond the statutory timeline of 45 days, and as much as 45% see them delayed beyond 180 days. In this light, this paper attempts to

- Review the state of MSME dues/receivables through time and steps taken by the government to mitigate it.
- Find the current challenges and gaps in MSME receivables realization and the government's current interventions.
- Propose an integrated blockchain-based framework and architecture with Smart contract for end-to-end receivables management for MSMEs.

21.2 Literature Review

Working capital management (WCM) efficiency is vital for small firms as it represents over half of their total assets for production firms and an even greater proportion for trading and distribution firms, significantly affecting their liquidity and profitability [4]. Howorth and Westhead note that WCM has a greater impact on the profitability of MSMEs than large corporations [5], because of their higher percentage of current assets, insufficient liquidity and highly volatile cash flows [6]. A significant amount of academic literature points to a positive relationship between WCM efficiency and profitability and believes that firms can increase profitability through aggressive WCM to reduce the number of day's accounts receivable and inventories [3, 7–11].

The three components of working capital are receivables, payables and inventory. But, a significant part of the working capital requirement of an MSME arises due to long receivables realization cycles [2]. According to Jakpar et al. and Tanveer et al., the collection time of receivables and economic profitability have a positive association [12, 13]. For a sample of Japanese enterprises from 1990 to 2004, Nobanee et al. discovered a strong negative relationship between the length of the cash conversion cycle and profitability [14]. Poor debtor management has been the cause of 20% of bankrupt businesses in the UK. According to Peel and Wilson, if financial/working capital management methods in the small business sector could be dramatically improved, fewer businesses would fail and economic welfare would be significantly raised [15].

But most often, MSMEs are not in a position to do that due to their lopsided bargaining power in relation to large corporate buyers. The median collection duration for manufacturing MSMEs in 2018–19 was 65 days, whereas it's above 100 days for services [16]. A 2019 RBI survey found that 44% of manufacturing MSMEs and 27% of service MSMEs experience payment delays. According to a report by the Global Alliance for Mass Entrepreneurship and Krea University, 'Unlocking Credit for India's Job Creators', the average payment period for MSMEs ranges from three to six months, with the delay increasing with company size. 77% of micro, 69% of small and 65% of medium sized businesses experienced late payments.

The asymmetry in bargaining power lies at the core of this problem. Delaying payments to smaller suppliers allows large corporations to keep their working capital days under control. Also, the government and PSUs, with their bureaucratic structure, follow the traditional ways of purchasing, where there is no deadline for making payments, rendering small firms to wait perpetually for their receivables. The scale of the problem warrants an effective and immediate solution to ease the capital constraints of MSMEs.

21.3 Government Initiatives and Current Challenges

The government has taken various steps to help small businesses suffering from late payments for their goods and services. In February 1999, the Department of Company Affairs modified Schedule VI of the (then) Corporations Act, 1956 to require companies to declare outstanding dues to SSIs of 0.01 million or more that are more than 30 days old in their balance sheets. MSME Development Act, 2006, brought provision to penalize buyers who delay payment to MSMEs for more than 45 days by requiring them to pay interest to the supplier. India's Planning Commission formed a High Level Committee on Financial Sector Reforms in August 2007, which came up with a report in 2009, 'A Hundred Small Steps'. This report recommended securitizing trade credit, which can lower MSMEs working capital and financing needs [13].

The Reserve Bank of India, in March 2014, published a concept paper titled 'Micro, Small, and Medium Enterprises Factoring Trade Receivables Exchange'. Subsequently, it launched the Trade Receivables Discounting System (TReDS) in 2016 to address liquidity shortages in the MSME ecosystem. The purpose was to build a platform where MSME sellers, corporate buyers, and financers (banks and NBFCs) could submit, accept, discount, trade and pay MSMEs' invoices/bills. Three exchanges, namely, Receivables Exchange of India (RXIL), M1 Exchange and Invoice Mart Exchange, were launched in India between December 2016 and July 2017.

In October 2017, the Ministry of MSMEs introduced SAMADHAAN, a delayed payment tracking portal, where MSMEs can file applications for delayed payments from the buyers. Many Micro and Small Enterprises Facilitation Councils (MSEFC) are established in each state to settle the cases. The awareness remains low amongst MSMEs as less than 1% of all registered MSMEs have filed delayed payment petitions, corresponding to only 1.3% of their estimated unpaid payments. Also, due to their unequal bargaining position, MSMEs are reluctant to submit complaints against purchasers, despite huge receivables.

The efforts to handle the MSME receivables problem have made progress. However, there are still gaps and hurdles to overcome. For example, the MSMED Act of 2006 required payments to MSMEs to be paid within 45 days, but it is uncertain when the 45-day period begins. Similarly, TReDS and SAMADHAAN have the potential to be game changers, but some factors are hindering their effective utilization.

21.3.1 Problems with MSME SAMADHAAN

Though intended to provide MSMEs recourse to pending payments, MSME SAMADHAAN is too detailed and complex to meet all the requirements. The requisite paperwork is the first issue that occurs. All applicants on the SAMADHAAN portal must upload copies of various papers, including signed contracts, work orders, invoices and proof of product or service delivery to the buyer, to substantiate their claim. The MSMEs, in the majority of cases, lacked appropriate documentation to begin with. They had entered into contracts with their customers and offered goods or services to them without clearly defining the deliverables, and in many cases, without even signing a legal contract. Purchase orders and invoices, for example, were either missing or not legally admissible in other instances.

Second, the paperwork for each case is so detailed that going through it requires a significant amount of time and effort (thus the cost). A highly trained authority would need at least four hours to review and vet each document before making a judgment on the case [17]. Accordingly, to solve all pending case on the portal requires 400,000 man-hours, implying, an army of 10,000 devoted bureaucrats would have to work nonstop on SAMADHAAN cases for four months, even as new applications continue to pour in.

The third issue is the length of time of settlement. While most applications to the MSECF are unanswered, those that are picked up take weeks to process. Even starting the hearing on applications takes at least two weeks. Whenever an aggrieved MSME contacts the state council, the initial hearing takes at least two months, after which a notice is sent to both parties, giving them another two months to resolve.

The case is then referred to an arbitrator, who is given a three-month deadline to hold a hearing, but the buyers do not want to attend.

Fourth, there is a lack of enforcement of the litigation and compliance. The defaulter's failure to appear for hearings is typical. The buyer will ignore the hearings and continue to press the case for years because there is no penalty for doing so. Even if the arbitration results in an MSME being able to recover its dues, the company is compelled to go to a local civil court to do it. After that, it's a civil matter. Because the civil court is likewise exceedingly slow, it could take up to ten years to recoup the debts."

21.3.2 Problems with TReDS

TReDS was established to provide liquidity to MSMEs through discounting their bills or receivables. It is a platform that connects lenders, buyers, and micro, small, and medium businesses. Lenders compete to settle an MSME supplier's claims after the buyer acknowledges the invoice. After a predetermined period, the buyer repays the lender, which is usually a bank. The platform, however, is yet to be adopted by many stakeholders. Only 35% of corporates with a turnover of more than Rs. 500 crore which were mandated to onboard have registered themselves as of April 1, 2022. Many who signed up appear to have done so solely to comply with statutory guidelines. Very few of them have made a transaction yet. Despite this, in 2021–22, Rs. 36,000 crore invoices were discounted, and the growth looks promising. But some hurdles are impeding wider adoption of the platform.

To begin with, most MSMEs are unfamiliar with digital platforms and government processes, making them wary of going down this path. Also, despite having a registration, many MSMEs do not publish their invoices on the platform owing to undisclosed corporate pressure (buyers). Besides, large businesses are hesitant to submit invoices to the internet for two reasons. First, TReDS is a transparent system, so they'd have to pay the suppliers' invoices within 45 days of receiving the goods or services. Second, their competitors could identify their MSME suppliers by looking at their internet invoices.

The speed with which buyers approve invoices is the root of the problem, mainly the Central Public Sector Enterprises. Invoices are typically approved in 45–60 days by these companies. They usually settle the money by this time. However, the volume generated by PSUs is only about 7–8% of the total volume generated by the three platforms. The problem is that invoice approval occurs after the due date in many circumstances. This defeats the purpose of TReDS platforms in the first place. On the other hand, private corporations take between 5 and 15 days to acknowledge an invoice, after which the MSME supplier receives the approved payments.

The major problem for most online trade receivables exchanges has been getting the volumes and building a vibrant and liquid marketplace. In India, the exchanges are undergoing similar tests. The turnover ceiling should be cut from Rs. 500 crores to Rs. 250 crores to attract more businesses. Also, more awareness needs to be made among MSMEs to register on the platform and avail the liquidity option.

21.4 Blockchain-Based MSME Receivables Management Platform

The section above highlights the challenges inherent in the existing processes in MSMEs receivables realization. Extensive requirement of documents in proper formats, huge cost in form of time and effort to process them for arbitration, longer time period in settlement, deliberate incompliance, quality issues, fear and low bargaining power of small suppliers leave them with no option but to be at the mercy of big buyers. The other route, i.e. TReDS, though provides an alternative, suffers from similar issues like standardized documents requirement, technological aversion by MSMEs, corporate pressure, delay in invoice acknowledgment and approval, and thus liquidity.

A blockchain-based solution can begin with a few standardized templates (precisely defining the deliverables and the payment terms and conditions) inherent in smart contracts starting from the purchase order. This would prevent a significant proportion of disputes from forming in the first place. The standardization of contracts will aid both buyers and sellers in compiling documents, and MSME councils will be able to expedite the resolution of all cases. Officials would find it easier and faster to go over the essential elements in the standardized contracts, even if they had to use a manual information retrieval procedure because they would know exactly where to look. Modern technologies like computer vision and OCR, on the other hand, might be used to automate the extraction of information from contracts, reducing the time it takes for officials to make a decision from a few hours to a few minutes. The issue of invoice acknowledgment and approval delays will be eliminated, as real-time acknowledgment will allow for speedier discounting options.

21.4.1 Platform and Methodology

We used the most popular blockchain, Ethereum, which provides a well-tested ecosystem where many examples are available to use standard procedures and datatypes for developing decentralized applications. Ethereum smart contracts, implemented in Solidity, are well suited for applying BOSE and ABCDE methodology, and data recording and transaction management procedures are easily implemented. We follow the ABCDE method [18], an eight-step software development approach designed specifically for blockchain software systems. We will not describe the App System code phase or the system integration phase for brevity but will finish at the end of the design phases.

21.4.2 *System Design*

Following the ABCDE method [18], we provide the system design step by step below.

21.4.2.1 The Goal of the System

It aims to integrate the whole MSME receivables process, including the erstwhile discounting through TReDS and arbitration through MSME SAMADHAAN, and implement it over a blockchain architecture. This would reduce the process's duplicity and bring more transparency and formality to receivables management, beginning from the very point of purchase.

21.4.2.2 Actors

The system has the following actors:
MSME They are the suppliers of goods and services to big firms.
Buyer They are the entities to whom goods and services are supplied to by MSMEs. They comprise big corporates, central and state departments, and PSUs.
LSP (Logistic service provider) They are the transporters who help deliver goods from suppliers to the buyer.
Financer These are the institutions that finance the receivables and payables.
UAA (Udyog Aadhar Authority) This entity can be from the ministry of MSME and shall act as the system administrator and supervisory body.
MSEFC (Micro and Small Enterprises Facilitation Council) It shall act for dispute resolution and settlement.

21.4.2.3 User Stories

This section outlines the various activities actors on the platform undertake (see Fig. 21.1). UAA (Udyog Aadhar Authority) registers different actors and manages and controls the various actors' reading and writing accesses in the system.

The Buyer sends a purchase order until the supplier's orders remain unaccepted (MSME). After acceptance of the purchase order, it can raise a cancellation or modification request, which the supplier has the option to accept or reject. If the supplier accepts cancellation or modification, the smart contract (SC) will terminate, and a fresh SC will be needed to create a new purchase order. The buyer validates the billing invoice and delivery invoice after receiving the order (auto validate if not done within 24 h of LSP sending a delivery signal). It can raise an alert if the item is not delivered, and LSP flags the delivery signal. Then it can either pay the bill or opt to be factored (Reverse factoring).

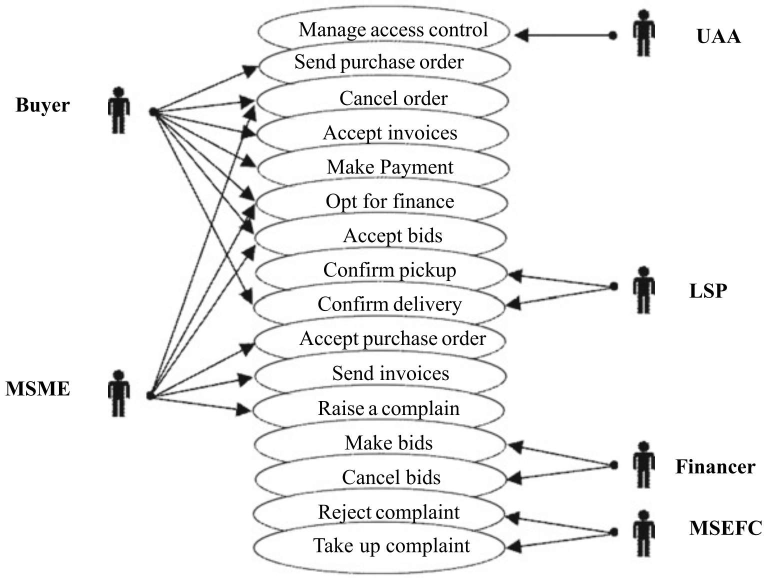


Fig. 21.1 The user stories of the system. Created by the Authors

The MSME (supplier) accepts the purchase order and sends the billing and delivery invoice to the buyer. It can raise the cancellation request after accepting, which the buyer has the right to accept or reject. It can accept the buyer's cancellation request, which will lead to the termination of the current contract. Then after delivery, it either receives the payment, opts for factoring, or raises a complaint to the MSEFC.

The LSP will have to make two entries: picked up and delivered. On the same day as the delivery date of the logistic partner, the buyer will have to approve the delivery, which will also entail the approval of bills and invoices.

The financers bid for financing the receivables (Factoring units). The seller or buyer gets the option to accept the bid depending upon who will bear the cost of financing. They pay the amount to the MSME (supplier) and receive payments later from the buyer. MSEFC receives the complaint, notifies the two parties involved and asks for mutual settlement, may reject a plea or accept the plea, and pronounces a verdict.

21.4.2.4 UML Sequence Diagram

UML sequence diagram represents the sequence of operations that the actors and the system perform. Figure 21.2 shows these interactions. It includes the purchase process between buyer and seller, the process of financing the receivables or payables, and the arbitration process.

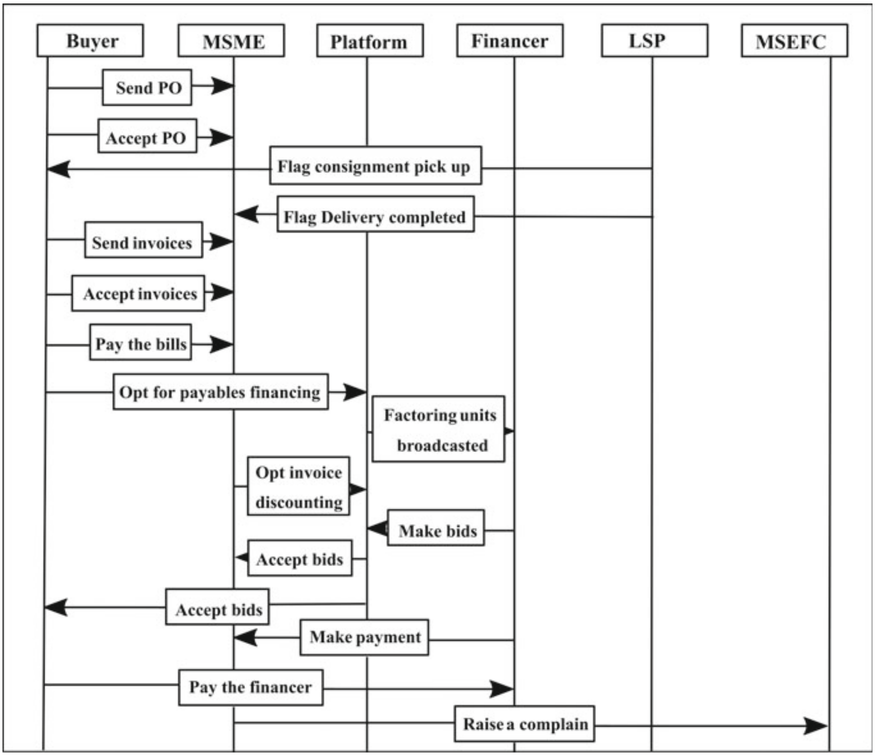


Fig. 21.2 UML sequence diagram. Created by the Authors

21.4.2.5 Smart Contracts UML Class Diagrams

The contracts UML diagrams in Fig. 21.3 show the structure of smart contracts and their integration. The managing authority, i.e. UAA assigns one of four typologies to every node that defines their roles. It also has a mechanism to suspend or blacklist nodes that do not act according to the network's rules or renege on their commitments or roles in the system. Every time a purchase order is sent, a receivables Management contract is created which in turn creates a purchase contract. The factoring and arbitration contracts are only created when either the buyer/seller opts for financing their payables/receivables or when the seller raises a complaint, respectively.

Each of the smart contracts consists of numerous variables and functions dealing with specific tasks. They also have events that are triggered when something significant occurs and can be detected by observer programs. Because SCs can't call external systems' operations directly, events offer a way for them to communicate with the outside world. Modifiers are Boolean functions that are called before a function is executed in all contracts. They have the ability to verify limitations and

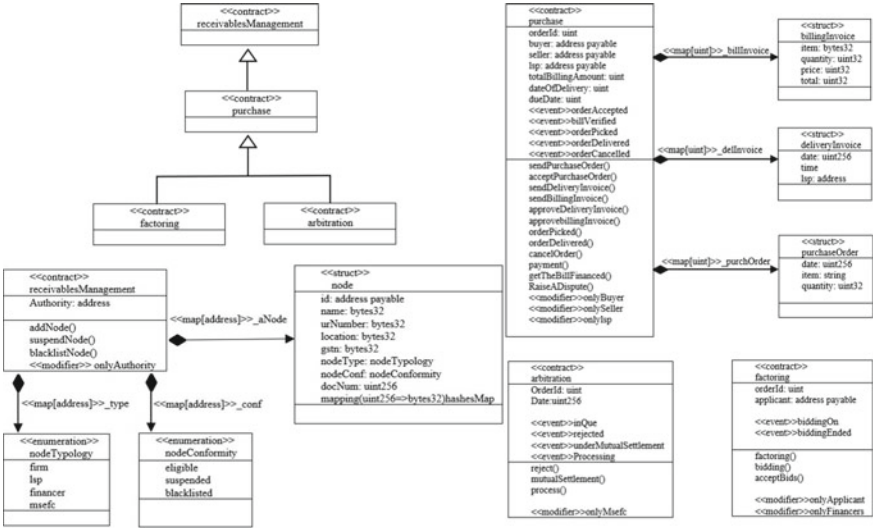


Fig. 21.3 Smart contracts inheritance, contracts 'purchase'; 'factoring'; 'arbitration' and 'recievablesManagement' (clockwise from top left). Created by the Authors

possibly halt the execution of a function. They also check whether the node accessing the particular contract has the authority to do so or not.

21.5 Conclusion

MSMEs have the potential to catalyze the economy's growth but often remain on the brink of survival owing to their inability to manage and raise timely capital. Due to the asymmetry of bargaining power, they remain at the mercy of big players who delay payments to these firms to keep their own working capital days in control. The government has made continuous efforts to address this by bringing mechanisms like MSMED Act, MSME SAMADHAAN, MSME facilitation councils and TReDS. Though these steps have realized some of the intended outcomes, gaps and challenges remain. Some of them include extensive documentation, fragmented processes, costly and illiquid discounting market, lengthy arbitration cycle, deliberate incompliance by large firms, quality issues and hesitant MSMEs owing to fear and low bargaining power.

The proposed blockchain platform integrates the whole receivable management process in a single system, removing the processes' duplicity and bringing more transparency to the system. A transparent purchase system and automated payment and settlement mechanism would reduce the asymmetry of power between the small and large firms and result in timely and fair-trade settlements. Also, the blockchain-based integrated financing of payables/receivables with more stakeholders and a transparent process can bring down the cost and illiquidity. The digitized and automated end-to-end process will also bring down the time and cost of litigation both for the parties and the government. The framework can be fine-tuned according to the need and can also be expanded to integrate more related processes and stakeholders.

References

1. Madishetti S, Kibona D (2013) Impact of accounts receivables management on the profitability of SMEs in Tanzania. Asian J Res Bank Finance 3(2):1–14
2. Chatnani NN (2018) Receivables management and supply chain finance for MSMEs: analysis of TReDS. Acad Strateg Manag J 17(3):1–8
3. Shin H-H, Soenen HL (1998) Efficiency of working capital and corporate profitability
4. Raheman A, Nasr M (2007) Working capital management and profitability-case of Pakistani firms. Int Rev Bus Res Papers 3(1):279–300
5. Howorth C, Westhead P (2003) The focus of working capital management in UK small firms. Manag Account Res 14(2):94–111
6. Peel MJ, Wilson N, Howorth C (2000) Late payment and credit management in the small firm sector: some empirical evidence. Int Small Bus J 18(2):17–37
7. Deloof M (2003) Does working capital management affect profitability of Belgian firms? J Bus Finance Account 30(3–4):573–588
8. Eljelly AM (2004) Liquidity-profitability tradeoff: an empirical investigation in an emerging market. Int J Commer Manag
9. Fernández-López S, Rodeiro-Pazos D, Rey-Ares L (2020) Effects of working capital management on firms' profitability: evidence from cheese-producing companies. Agribusiness 36(4):770–791
10. García-Teruel PJ, Martínez-Solano P (2007) Effects of working capital management on SME profitability. Int J Manag Finance
11. Lazaridis I, Tryfonidis D (2006) Relationship between working capital management and profitability of listed companies in the Athens stock exchange. J Financ Manag Anal 19(1)
12. Jakpar S, Tinggi M, Siang T, Johari A, Myint K, Sadique M (2017) Working capital management and profitability: evidence from manufacturing sector in Malaysia. J Bus Financ Aff 6(2):1–9
13. Tanveer B, Nazir MI, Khan MA, Khan MA, Razzaq S (2016) The impact of working capital management on firms financial performance: evidence from Pakistan. Int J Econ Financ Iss 6(3):1097–1105
14. Nobanee H, Abdullatif M, AlHajjar M (2011) Cash conversion cycle and firm's performance of Japanese firms. Asian Rev Account
15. Peel MJ, Wilson N (1996) Working capital and financial management practices in the small firm sector. Int Small Bus J 14(2):52–68
16. Prasad A, Srinivasan G, Pradhan H, Hamendra, Ghosh I, Sriram M (2020) MSME financing in India: key to the 5 trillion economy
17. Kumar A (2021) Why MSME SAMADHAAN scheme is not working, and how to fix it
18. Marchesi L, Marchesi M, Tonelli R (2020) ABCDE-agile block chain DApp engineering. Blockchain: Res Appl 1(1–2):100002

Chapter 22
Investigating the Key Enablers in Perishable Food Supply Chain Using DEMATEL and AHP—PROMETHEE

Malleswari Karanam, Krishnanand Lanka, Sai Nikhil Pattela, and Vijaya Kumar Manupati

Abstract Demand variations, disruptions, and environmental regulations stand as major setbacks in the perishable food supply chain (PFSC). Hence, in PFSC there is a need for identification of enablers and critical enablers that can help for better decisions to overcome the above mentioned issues by making the supply networks more resilient. In this chapter, first with the identified enablers the most influential ones have been recognized with the multi-criteria decision-making (MCDM) based Decision making trial and evaluation laboratory (DEMATEL) method. In addition to the ranking of enablers, we have determined the cause and effect relationship of the enablers for the perishable food supply chain using DEMATEL. Later, a hybrid MCDM i.e., Analytical Hierarchy Process (AHP) - Preference ranking organization method for enrichment evaluation (PROMETHEE II) has been used to rank the enablers and further identify the key enablers for the considered PFSC. The results obtained from the above models provide the critical enablers which further help the members in the supply chain be more resilient.

Keywords Perishable products · MCDM · DEMATEL · PROMETHEE II · AHP

M. Karanam (✉) · K. Lanka · S. N. Pattela
Department of Mechanical Engineering, National Institute of Technology Warangal, Warangal 506004, India
e-mail: malli.nittrichy@gmail.com

K. Lanka
e-mail: lanka@nitw.ac.in

V. K. Manupati
Operations and Supply Chain, National Institute of Industrial Engineering (NITIE), Mumbai 400087, India
e-mail: manupativijay@nitie.ac.in

M. K. Tiwari et al. (eds.), *Applications of Emerging Technologies and AI/ML Algorithms*, Asset Analytics, https://doi.org/10.1007/978-981-99-1019-9_22

22.1 Introduction

In recent years, digitalization and Industry 4.0 have become prominent in many industries for improving their performance and efficiencies. Integrating perishable food supply chain (PFSC) with digital technologies such as Radio frequency identification, Internet of things (IoT) etc., aids in the identification, traceability, and quality monitoring of perishable products throughout the supply chain (SC). Perishable products such as vegetables, dairy products, and fruits are constantly deteriorating both in quality and economic parameters. In order to maintain a hassle-free functioning and performance of food supply networks, these products need to reach from the supplier to the end customer within their limited product life [1].

Perishable food supply chains have to be resilient enough to tackle the demand uncertainties, disruptions, and environmental regulations. However, appropriate transportation selection also promotes the effort to preserve the quality, freshness, and safety of the perishable products, which helps reduce the overall losses [2, 35]. Moreover, economic, environmental, and societal risks may indeed be reduced with efficient PFSC management. The environmental considerations are to be amalgamated with the SCs, due to the rapid rise in the global temperatures. Developing sustainable SCs minimizes the total carbon emissions in the long run [3]. Perishable products have a significant influence on food security in many developing countries, and effectively managing them by identifying the key enablers facilitates the decision-making process.

Considering the necessity for PFSC and the aforementioned problems, this chapter has been focusing on the following goals:

1. To rank the enablers and identify the most key enablers for the PFSC using the Decision making trial and evaluation laboratory (DEMATEL) approach.
2. To determine the cause and effect relationship of the enablers for the PFSC using DEMATEL.
3. To rank the enablers and identify the most key enablers for the PFSC through integrated AHP—PROMETHEE II.
4. To identify the key enablers using DEMATEL and integrated AHP—PROMETHEE II

The remainder of this chapter has been carried out as follows: Sect. 2 describes the significance of enablers related to PFSC work and identifies the methodologies to rank these enablers. Section 3 explain the MCDM methodologies such as DEMATEL and integrated AHP—PROMETHEE II and identify the key PFSC enablers. Finally, Sects. 4 and 5 discuss the results and conclusions of the current study.

22.2 Related Work

A literature survey conducted to find the enablers of PFSC revealed fifteen enablers that impact the PFSC, and [4] in their study have explored the themes of PFSC and identified the hierarchy of each enabler using Total interpretive structural modeling—Cross impact matrix multiplication applied to classification (TISM—MICMAC) analysis. They explored the literature supporting the enablers' importance of PFSC. PFSC enablers were identified from literature to maintain food products significantly throughout the SC. The fifteen enablers are described in the following sub-section.

22.2.1 Enablers Related to PFSC

The fifteen enablers have been identified from the literature that are Radio frequency identification (RFID), IoT, Shelf Life (SL), Cold Storage (CS), Inventory Control (IC), Decision Making (DM), Third-party logistics (3PL), Vehicle Routing (VR), Unit Capacity (UC), Fuel Consumption (FC), Freshness keeping (FK), Cost–Benefit Analysis (CBA), Global Warming (GW), Carbon Emissions (CE), and Energy Utilization (EU).

Radiofrequency identification: RFID tags are used to search, identify, and track the products. RFID plays a vital role for perishable products upstream to downstream of the PFSC. Monitoring and tracking environmental conditions for delivering perishable products in cold chains have been made simple using RFID tags and other sensors [5].

Internet of Things: The IoT infrastructure includes a smart device that measures environmental parameters such as temperature, humidity, and gas concentrations, enables real-time computations and early reports [6].

Shelf Life: Perishable products are distinguished from non-perishable products due to their shelf life. The shelf life of perishables varies from one product to another product. The perishable product's shelf life and quality are assessed based on the deterioration [7, 8].

Cold Storage: The rate of deterioration of perishable products is substantially reduced due to cold storage. A temperature-controlled transit system is necessary to prevent perishable items from deteriorating from upstream to downstream of the PFSC [9].

Inventory Control: IC has a significant impact on perishable products. Due to the limited life period, there is a restriction on perishable products inventory, which may lead to not satisfying the consumer requirement [10].

Decision Making: Decision making has to be based on the deterioration of perishable products from supplier to the end consumer. For example, products with less shelf life have to be delivered first. Hence, DM is a significant influence on PFSC [11].

Third-party logistics: 3PL providers deliver the perishable products from the target to the destination within the time window to keep the freshness of the perishable products [12].

Vehicle Routing: VR plays a vital role in transporting the perishable products upstream to downstream of the PFSC. There are constraints in VR, such as the vehicle's capacity, different types of vehicles, and multi-compartment vehicles to differentiate the different perishable products [13].

Unit Capacity: Storage and Transportation unit capacity are considered during the PFSC. UC depends on the type of vehicle during the transportation of products [14].

Fuel Consumption: Perishable products are temperature-dependent, so more fuel consumption is required during the transportation of products to maintain the quality of perishable products [15]. The goal is to find the path that uses the least amount of fuel for both propulsion and refrigeration [16].

Freshness keeping: The selling cost of perishable products is decided based on freshness keeping. The perishable product's cost varies in all stages of the SC due to the freshness of the product (such as supplier, wholesaler, distributor, retailer, and vendor) [17].

Cost Benefit Analysis: It aims to minimize the total cost, including the transportation and operational costs of each stage of the SC [18].

Global Warming: Perishable products are required by the refrigeration system to maintain their quality. During the transportation of products, more greenhouse gas emissions occur due to a gas leakage [19].

Carbon Emissions: Carbon emissions trading policies are required to lead cold chain logistics firms in energy-saving and emission-reduction transformations of cold storage [20].

Energy Utilization: Refrigeration in transportation takes a lot of energy, resulting in many carbon emissions [6].

22.2.2 DEMATEL and AHP—PROMETHEE II Approaches

In this section, the authors illustrated various multi-criteria decision making approaches such as DEMATEL, AHP, PROMETHEE II, TOPSIS, VIKOR, ANP, and fuzzy concepts. Applying the MCDM approach helps identify the key criteria for a given problem. The related studies on MCDM are: Karanam et al. [4] identified the 15 enablers and their hierarchies for PFSC using TISM-MICMAC analysis. Wu et al. [21] studied the interrelationships and level of influence on food waste enablers using priority ranking and cause and effect diagrams using Hybrid AHP-DEMATEL. Kilic et al. [22] proposed the hybrid methodology N-DEMATEL-TOPSIS for investigating the environmental sustainability dimensions and performance comparison of various municipalities. Sharma et al. [23] applied DEMATEL to identify the most driving enabler on e-waste management in circular economies. Khanzode et al. [24] identified 14 enablers in manufacturing firms (sustainability and circular economy domains), later ranking and establishing cause and effect relations using the DEMATEL methodology. Li et al. [25] applied DEMATEL-PROMETHEE and identified 5 enablers and found that the energy efficiency parameter to be given the highest priority by governments for selecting the resources. Makan et al. [26] six alternatives were taken, and the results show that reactor technologies are more sustainable than closed technologies using PROMETHEE method. Gupta et al. [27] Prioritizing enablers such as nursing staff, doctors, and diagnostic technicians for service quality in the healthcare sector using MCDM methods.

22.2.3 Summary

Following a review of the related work, we observed that no one had employed DEMATEL and AHP—PROMETHEE II to determine the PFSC's most key enablers. DEMATEL was selected as the best alternative based on its merits, according to the literature. AHP integrated PROMETHEE-II methodology is employed to compare the results obtained in DEMATEL analysis. Therefore, we used the DEMATEL and AHP-PROMETHEE II for ranking the enablers of PFSC and identify the most key enablers.

22.3 Methodology

The current study goal is to build a unique research framework for identifying the key enablers of PFSC based on the prior studies and experts' ratings. In this chapter, MCDM methodologies are DEMATEL, integrated AHP –PROMETHEE II have been used to rank the enablers of PFSC and finally identify the key enablers. We have adopted the traditional DEMATEL method to determine rank and the causal relation between 15 selected enablers. Traditional DEMATEL over the other methods in MCDM because in complex systems, the impact of influences is better captured by the DEMATEL method [28]. Next, we have proposed one more MCDM method AHP—PROMETHEE II, to rank the enablers. The required weights for PROMETHEE II have been calculated through AHP. MCDM tool AHP is used to assign the priority weights to the enablers used to obtain the hierarchy levels [29]. Finally, we identified the key enablers through obtained rankings from the methods mentioned above. The proposed framework is described in Fig. 22.1.

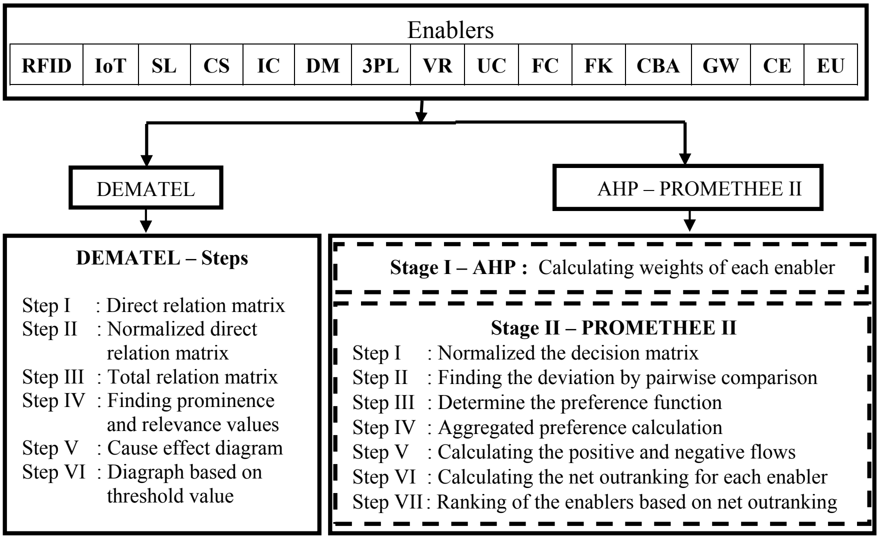


Fig. 22.1 Proposed framework

22.3.1 Dematel

In this section, DEMATEL approaches have been used to identify the key enablers for PFSC. The initial interrelationship among each enabler has been taken from the literature.

22.3.2 Dematel

In this section, DEMATEL approaches have been used to identify the key enablers for PFSC. The initial interrelationship among each enabler has been taken from the literature.

Step1: Generating the direct relation matrix (DRM)

A $n \times n$ matrix is generated from the experts' ratings on the enablers (A). a_{ij} denotes the influence of enabler i on enabler j. Experts have used 0 to 5 scale to represent the relationship among various enablers. The DRM is formulated based on the dominance rule from the experts' ratings as shown in Table 22.1.

Step 2: Normalized direct-relation matrix

We calculate the normalized values of the matrix DRM from the previous step. For each row, the sum of the rating $\sum a_j$ values are calculated. Then, of all the sum

Table 22.1 Direct relation matrix (DRM)

Enabler	RFID	IoT	SL	CS	IC	DM	3PL	VR	UC	FC	FK	CBA	GW	CE	EU
RFID	1.0	4.8	4.3	4.6	4.9	4.9	4.2	4.1	0.0	0.0	0.0	4.3	3.1	0.0	4.2
IoT	4.8	1.0	0.0	4.9	4.4	4.4	0.0	0.0	0.0	0.0	0.0	4.2	3.1	0.0	4.5
SL	4.9	4.5	1.0	4.5	4.5	4.5	4.0	4.0	0.0	4.2	4.0	4.9	0.0	3.2	4.2
CS	4.5	4.0	0.0	1.0	4.2	4.2	0.0	0.0	0.0	4.2	0.0	4.9	4.8	4.4	4.6
IC	4.5	4.0	4.0	4.9	1.0	4.0	4.9	4.0	0.0	3.2	4.1	4.5	4.9	3.4	4.9
DM	0.0	4.4	4.5	4.8	4.0	1.0	4.4	4.2	0.0	0.0	0.0	4.8	4.0	3.4	4.2
3PL	0.0	4.2	3.4	4.1	4.0	4.0	1.0	4.0	0.0	0.0	0.0	4.3	4.2	0.0	3.4
VR	3.4	3.3	4.9	4.2	4.0	4.0	4.9	1.0	0.0	0.0	4.2	4.2	3.4	0.0	3.4
UC	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0
FC	0.0	0.0	0.0	4.1	0.0	0.0	0.0	0.0	0.0	1.0	4.8	0.0	4.9	4.9	4.9
FK	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4.9	1.0	0.0	4.5	4.0	4.5
CBA	0.0	4.3	4.5	4.6	4.9	4.9	4.4	4.2	0.0	3.3	0.0	1.0	4.3	3.3	4.9
GW	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	4.9	4.3
CE	0.0	0.0	0.0	3.4	0.0	0.0	0.0	0.0	0.0	4.5	0.0	0.0	4.7	1.0	4.9
EU	0.0	4.2	0.0	4.0	0.0	0.0	0.0	0.0	0.0	4.0	4.5	4.3	4.7	4.9	1.0

values, the maximum value is selected as sum_{max}. Every rating in the initial decision matrix is divided by the sum_{max} to get the normalized values.

Step 3: The total relation matrix (R)

An identity matrix (I) similar to the size of matrix Y is considered. Then, the matrix (I-Y) is computed. The matrices Y and inverse of (I-Y) are multiplied using the matrix product method to get the total relationship matrix (R) (Table 22.2).

$$R = Y \times (I - Y)^{-1} \tag{22.1}$$

Step 4: Prominence and Relevance

The sum of the rows (K_j) and columns (L_i) in the matrix R is found

$$K_j = \sum_{i=1}^{n} r_i \tag{22.2}$$

$$L_i = \sum_{j=1}^{n} r_j \tag{22.3}$$

Table 22.2 Total relation matrix

Enabler	RFID	IoT	SL	CS	IC	DM	3PL	VR	UC	FC	FK	CBA	GW	CE	EU	K_j
RFID	0.176	0.199	0.260	0.173	0.223	0.208	0.199	0.240	0.000	0.156	0.149	0.184	0.064	0.094	0.158	2.483
IoT	0.210	0.152	0.217	0.144	0.178	0.158	0.152	0.133	0.000	0.055	0.053	0.176	0.040	0.132	0.057	1.857
SL	0.258	0.194	0.275	0.156	0.215	0.199	0.194	0.244	0.000	0.101	0.149	0.193	0.142	0.160	0.155	2.635
CS	0.163	0.149	0.195	0.062	0.165	0.106	0.149	0.186	0.000	0.055	0.053	0.087	0.049	0.126	0.057	1.602
IC	0.206	0.184	0.241	0.086	0.211	0.210	0.133	0.229	0.000	0.149	0.146	0.116	0.142	0.151	0.167	2.371
DM	0.211	0.119	0.229	0.085	0.203	0.188	0.169	0.218	0.000	0.147	0.139	0.168	0.060	0.069	0.148	2.153
3PL	0.195	0.159	0.195	0.070	0.181	0.172	0.159	0.189	0.000	0.122	0.129	0.097	0.052	0.062	0.084	1.866
VR	0.229	0.176	0.254	0.147	0.145	0.191	0.176	0.225	0.000	0.160	0.092	0.179	0.133	0.128	0.163	2.398
UC	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.018	0.000	0.000	0.000	0.000	0.000	0.000	0.018
FC	0.144	0.017	0.145	0.061	0.028	0.026	0.017	0.108	0.000	0.007	0.007	0.140	0.106	0.013	0.007	0.827
FK	0.128	0.006	0.132	0.108	0.025	0.023	0.006	0.092	0.000	0.003	0.003	0.063	0.039	0.005	0.003	0.635
CBA	0.206	0.187	0.222	0.087	0.199	0.176	0.187	0.211	0.000	0.149	0.142	0.115	0.069	0.072	0.151	2.174
GW	0.042	0.004	0.098	0.019	0.012	0.011	0.004	0.019	0.000	0.002	0.002	0.106	0.010	0.003	0.002	0.336
CE	0.120	0.014	0.121	0.113	0.015	0.014	0.014	0.029	0.000	0.006	0.006	0.111	0.022	0.012	0.006	0.604
EU	0.166	0.040	0.109	0.124	0.121	0.116	0.040	0.134	0.000	0.021	0.020	0.156	0.107	0.027	0.022	1.202
L_i	2.453	1.6	2.69	1.435	1.921	1.799	1.6	2.255	0.018	1.135	1.09	1.891	1.036	1.055	1.183	

Then prominence ($K_j + L_i$) and net effect (K_j - L_i) values are calculated for each enabler. The value of ($K_j + L_i$) is proportional to the impact of enabler i on the other enablers. If the net effect value is greater than zero, the enabler belongs to the cause family; otherwise, it goes into the effect category. The ranking of the enablers is done using (K_j - L_i) values and is shown in Table 22.3.

Step 5: Cause and effect diagram

The cause and effect diagram has been plotted based on the ($K_j + L_i$) and (K_j - L_i) obtained from the R. Here, x axis is considered as prominence ($K_j + L_i$) and y axis is considered as relevance (K_j - L_i) values. The causes are shown in green colour, and the effects are shown in orange colour in the cause and effect diagram. Figure 22.2 illustrates the cause and effect diagram.

Step 6: Diagraph based on the threshold value

To explain a structural relationship among the enablers, a threshold value helps in filtering out the negligible relationships in matrix R. The threshold value (0.103) is determined as the average value of R. To further investigate the interplay between the enablers, each enabler value from the R is compared with the threshold value, and the relation between various causes and effects are portrayed in the digraph (Fig. 22.3).

22.3.3 Integrated AHP-PROMETHEE II

This section uses an integrated AHP and PROMETHEE II to rank the PFSC enablers. The fifteen enablers have been identified from the literature, and the pairwise comparisons among the enablers are collected from expert ratings. AHP is used to calculate the weights of each enabler. These weights have been used in PROMETHEE II to rank the enablers.

22.3.3.1 Ahp

Step 1: Constructing a decision hierarchy

A pairwise comparison matrix is constructed such that the diagonal elements are equal to 1, as an enabler has equal importance with itself. The elements of the matrix correspond to the pairwise comparisons among the enablers using a 1 to 9 scale in Table 22.4. For judgment consistency, $a_{ij} = k$ implies that $a_{ij} = 1/k$ [30].

Step 2: Normalizing the comparison matrix

A normalized matrix is calculated by dividing each element with its column total ($a_{ij}/\sum a_j$) to sum up the column elements of the resulting normalized matrix to 1. Then, the average of the elements in each row is calculated to find the criteria weights (W_i).

Table 22.3 Prominence and relevance values

Enabler	SL	RFID	VR	IC	CBA	DM	3PL	IoT	CS	EU	FC	FK	CE	GW	UC
$(K_j + L_i)$	5.325	4.937	4.653	4.292	4.064	3.951	3.466	3.456	3.036	2.386	1.962	1.725	1.659	1.371	0.036
$(K_j - L_i)$	−0.06	0.030	0.142	0.450	0.283	0.354	0.266	0.257	0.167	0.019	−0.307	−0.455	−0.452	−0.700	0.000
Rank	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15

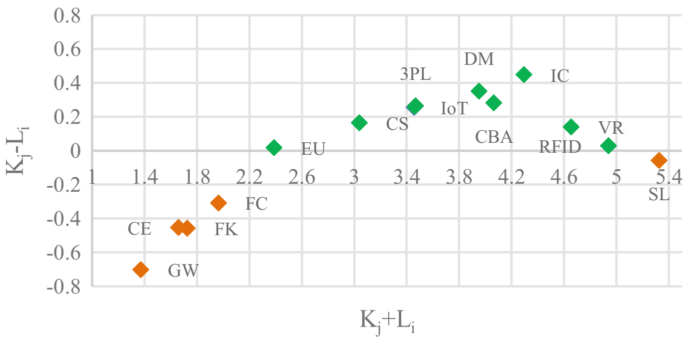


Fig. 22.2 DEMATEL—cause and effect diagram

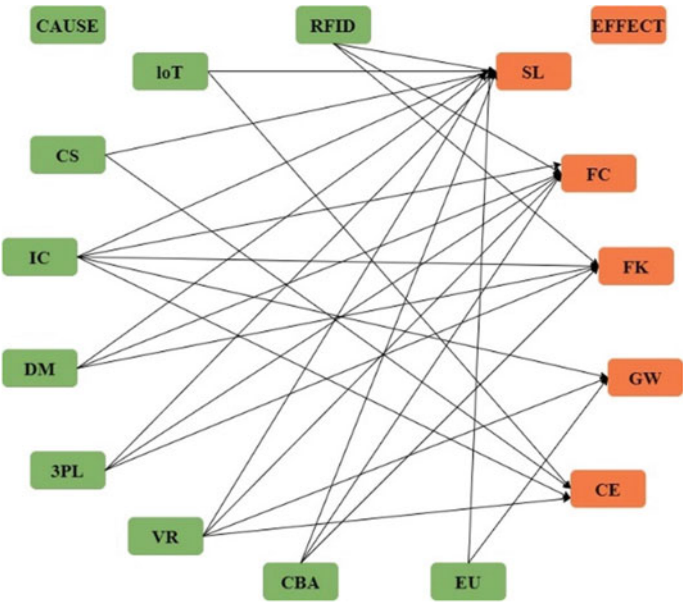


Fig. 22.3 DEMATEL—Diagraph based on the threshold value

Table 22.4 Value in AHP matrix and level of importance

Value	1	3	5	7	9	2,4,6,8
Level of importance	Equal	Weak	Strong	Very strong	Extreme	Intermediate values

Step 3: Consistency check

We find critical ratio (CR) and critical index (CI) values to check the consistency of the w_i found in the previous step. The eigenvalue λ is calculated for each row

Table 22.5 Number of dimensions and RI values

N	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
RI	0	0	0.58	0.9	1.12	1.24	1.32	1.41	1.45	1.49	1.51	1.48	1.56	1.57	1.59

by multiplying W_i and weighted sum value for each row (WS_i). The maximum eigenvalue λ_{max} is the mean of all the eigenvalues thus obtained. The critical ratio is calculated using the formula:

$$CR = \frac{\lambda_{\max} - n}{n - 1} \tag{22.4}$$

where "n" is the number of enablers considered. The obtained CR value is

CR = 2.1673 – 15 / 15 – 1 = > 0.1548 where n = 15.

Later, the critical index is found using the relation

$$CI = CR/RI \tag{22.5}$$

where RI is the Random Index (Table 22.5) which depends on the number of criteria/ dimensions/enablers used in the comparison matrix.

$CI = 0.1548 / 1.59 = > 0.0974 < 0.1$. For the comparison matrix to be consistent, the CI value should be less than 0.1, or inconsistency should not be higher than 10%. Table 22.6 shows the weights obtained in this step, which are used in the second stage to arrive at the final ranking of the enablers.

22.3.4 Promethee Ii

Step 1: Normalize the decision matrix

Every enabler is classified into two groups, viz. beneficial and non-beneficial. The criteria whose lower limit is necessary; it is known as the non-beneficial criteria. When the higher values are desired, they fall into beneficial criteria [31]. In this chapter, enablers falling in non-beneficial criteria are FC, CBA, GW, and CE, the rest are of beneficial criteria. The normalized values thus calculated are in Table 22.7.

$$N_{ij} = \frac{a_{ij} - a_{j,\min}}{a_{j,\max} - a_{j,\min}} \tag{22.6}$$

Table 22.6 Enabler and weight calculated

Enabler	RFID	IoT	SL	CS	IC	DM	3PL	VR	UC	FC	FK	CBA	GW	CE	EU
Weight	0.133	0.141	0.109	0.104	0.082	0.07	0.053	0.048	0.049	0.051	0.042	0.031	0.034	0.028	0.026

Table 22.7 Normalized decision matrix

Enabler	RFID	IoT	SL	CS	IC	DM	3PL	VR	UC	FC	FK	CBA	GW	CE	EU
RFID	0.204	1.000	0.878	0.939	1.000	1.000	0.857	0.976	0.000	1.000	0.000	0.122	0.367	1.000	0.857
IoT	0.980	0.208	0.000	1.000	0.898	0.898	0.000	0.000	0.000	1.000	0.000	0.143	0.367	1.000	0.918
SL	1.000	0.938	0.204	0.918	0.918	0.918	0.816	0.952	0.000	0.143	0.833	0.000	1.000	0.347	0.857
CS	0.918	0.833	0.000	0.204	0.857	0.857	0.000	0.000	0.000	0.143	0.000	0.000	0.020	0.102	0.939
IC	0.918	0.833	0.816	1.000	0.204	0.816	1.000	0.952	0.000	0.347	0.854	0.082	0.000	0.306	1.000
DM	0.000	0.917	0.918	0.980	0.816	0.204	0.898	1.000	0.000	1.000	0.000	0.020	0.184	0.306	0.857
3PL	0.000	0.875	0.694	0.837	0.816	0.816	0.204	0.952	0.000	1.000	0.000	0.122	0.143	1.000	0.694
VR	0.694	0.688	1.000	0.857	0.816	0.816	1.000	0.238	0.000	1.000	0.875	0.143	0.306	1.000	0.694
UC	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	1.000	0.000	1.000	1.000	1.000	0.000
FC	0.000	0.000	0.000	0.837	0.000	0.000	0.000	0.000	0.000	0.796	1.000	1.000	0.000	0.000	1.000
FK	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.208	1.000	0.082	0.184	0.918
CBA	0.000	0.896	0.918	0.939	1.000	1.000	0.898	1.000	0.000	0.327	0.000	0.796	0.122	0.327	1.000
GW	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	1.000	0.796	0.000	0.878
CE	0.000	0.000	0.000	0.694	0.000	0.000	0.000	0.000	0.000	0.082	0.000	1.000	0.041	0.796	1.000
EU	0.000	0.875	0.000	0.816	0.000	0.000	0.000	0.000	0.000	0.184	0.938	0.122	0.041	0.000	0.204

$$N_{ij} = \frac{a_{j,\max} - a_{ij}}{a_{j,\max} - a_{j,\min}} \tag{22.7}$$

Equation 22.6 is used to find the normalized values for beneficial criteria, whereas the normalization of non-beneficial criteria is carried out by Eq. 22.7.

Step 2: Calculating the evaluating differences of an enabler with other enablers

Evaluative difference (D) is the difference between the normalized values of an enabler with respect to all other enablers. For example, evaluative differences are D(RFID-IoT), D(RFID-SL), D(RFID-CS) — D(RFID-EU) for the enabler RFID. This process is carried out for all the enablers.

Step 3: Calculating the preference function

The preference function $P(x, y)$ represents the amount of preference of a criteria x with respect to criteria y and is valued between 0 and 1. The preference value is zero if the evaluative difference is negative. If the evaluative difference is positive, the preference function takes the difference value.

$$P(x,\ y)\ =\ 0,\ if\ D(x,\ y)\ is\ negative \tag{22.8}$$

$$P(x,\ y)\ =\ D(x,\ y),\ if\ D(x,\ y)\ is\ positive \tag{22.9}$$

The preference function values are hence calculated for all the obtained matrices in the above step.

Step 4: Calculating aggregated preference (AP) values

The weights obtained using the AHP method and the preference values obtained in the previous step. $P(x, y)$ are used to calculate the AP for each row as follows:

$$\prod(x, y) = \sum_{j=1}^{n} w_i P_j(x, y) \tag{22.10}$$

$$\prod(y, x) = \sum_{i=1}^{n} w_i P_j(y, x) \tag{22.11}$$

The calculated values are written in the form of a $n \times n$ matrix with diagonal elements as zeroes.

Step 5: Calculating leaving and entering flows

The entering (in) flow (ψ_+) is calculated for rows as:

$$\psi_{i,+} = \frac{\prod_i (x, y)}{n - 1} \tag{22.12}$$

The leaving (out) flow value (ψ_-) is calculated for columns as:

$$\psi_{j,-} = \frac{\prod_j (y, x)}{n - 1} \tag{22.13}$$

The net flow is:

$$\psi_{n,net} = \psi_{n,+} - \psi_{n,-} \tag{22.14}$$

This net flow value is calculated for all the enablers (Table 22.8). Based on these net flow values, ranking is given to all the enablers. Higher the net flow value, the greater the rank.

The final rankings from the integrated AHP-PROMETHEE II method are represented in Table 22.9.

22.4 Findings and Discussions

The PFSC enablers are ranked based on two MCDM methods DEMATEL and integrated AHP - PROMETHEE II, and key enablers have been identified. After assessing the enablers, it is relatively simple to handle PFSC management. This study analyzes various PFSC enablers and cause-and-effect relationships and ranks them based on MCDM methods. The relationship matrix for the fifteen enablers based on the experts' ratings has been identified from the literature. The ranking results are shown in Table 22.10.

The significance of the enablers has been determined by DEMATEL using ($K_j + L_i$) values. The DEMATEL results show that SL, RFID, VR, and IC have been key enablers for PFSC, prioritizing rank of first, second, third, and fourth, respectively. Shelf Life is the most key influential enabler with the highest ($K_j + L_i$) value, 5.3254, and capacity has the least importance with ($K_j + L_i$) value of 0.0362. The previous study observed that SL, RFID, VR, and IC are the critical enablers for PFSC, quantifying based on the TISM approach at the highest level. The integrated AHP—PROMETHEE II results show that VR, RFID, SL, and IC have been key enablers for PFSC, prioritizing rank of first, second, third, and fourth, respectively. Comparing DEMATEL and integrated AHP—PROMETHEE II, there is an interchange between

Table 22.8 Enablers and their inflow, outflow and net flow values

Enabler	RFID	IoT	SL	CS	IC	DM	3PL	VR	UC	FC	FK	CBA	GW	CE	EU	$\psi_{i,+}$	ψ_{net}
RFID	0.000	0.315	0.166	0.394	0.176	0.138	0.151	0.122	0.627	0.569	0.689	0.103	0.634	0.597	0.495	0.370	0.298
IoT	0.111	0.000	0.077	0.183	0.138	0.218	0.173	0.073	0.423	0.364	0.484	0.198	0.429	0.391	0.385	0.261	0.080
SL	0.162	0.276	0.000	0.295	0.127	0.257	0.266	0.160	0.668	0.568	0.680	0.204	0.662	0.609	0.462	0.385	0.294
CS	0.097	0.089	0.002	0.000	0.057	0.173	0.134	0.063	0.415	0.373	0.398	0.122	0.395	0.372	0.274	0.212	−0.021
IC	0.148	0.313	0.103	0.326	0.000	0.213	0.238	0.107	0.666	0.526	0.655	0.170	0.652	0.581	0.439	0.367	0.268
DM	0.012	0.295	0.135	0.344	0.115	0.000	0.090	0.085	0.530	0.446	0.566	0.044	0.517	0.487	0.372	0.288	0.163
3PL	0.000	0.226	0.119	0.281	0.115	0.065	0.000	0.060	0.484	0.422	0.542	0.054	0.494	0.450	0.344	0.261	0.136
VR	0.123	0.278	0.165	0.361	0.136	0.213	0.213	0.000	0.630	0.538	0.685	0.202	0.641	0.602	0.486	0.377	0.299
UC	0.098	0.097	0.142	0.182	0.164	0.127	0.106	0.099	0.000	0.122	0.154	0.138	0.084	0.134	0.179	0.131	−0.400
FC	0.073	0.070	0.075	0.174	0.058	0.076	0.077	0.040	0.154	0.000	0.163	0.072	0.132	0.093	0.084	0.096	−0.234
FK	0.038	0.035	0.033	0.044	0.031	0.041	0.042	0.032	0.032	0.008	0.000	0.015	0.015	0.010	0.052	0.031	−0.400
CBA	0.032	0.330	0.137	0.348	0.128	0.099	0.134	0.130	0.597	0.498	0.596	0.000	0.584	0.514	0.424	0.325	0.222
GW	0.043	0.041	0.075	0.101	0.089	0.052	0.054	0.048	0.022	0.038	0.075	0.064	0.000	0.073	0.112	0.063	−0.335
CE	0.031	0.029	0.047	0.104	0.044	0.048	0.035	0.035	0.098	0.024	0.096	0.020	0.098	0.000	0.070	0.056	−0.308
EU	0.039	0.133	0.010	0.115	0.012	0.042	0.039	0.029	0.253	0.125	0.248	0.039	0.247	0.180	0.000	0.108	−0.190
$\psi_{j,-}$	0.072	0.181	0.092	0.232	0.099	0.126	0.125	0.077	0.400	0.330	0.431	0.103	0.399	0.364	0.298		

Table 22.9 Ranking of enablers from AHP-PROMETHEE II method

Rank	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Enabler	VR	RFID	SL	IC	CBA	DM	3PL	IoT	CS	EU	FC	CE	GW	UC	FK

the SL and VR. For perishable products, SL and VR are both significantly high priority to deliver the perishable products within the shelf life and optimize the vehicle route. Based on the DEMATEL results, CBA, DM, and 3PL prioritized ranks five, six, and seven. The PROMETHEE II results show that CBA, DM, and 3PL prioritize rank five, six, and seven. Comparing the CBA, DM, and 3PL with the above approaches, cost benefit analysis is the fifth priority in PFSC to optimize the transportation and operational cost throughout the SC. In addition to the ranking, from the DEMATEL analysis, RFID, IoT, CS, IC, DM, 3PL, VR, CBA, and EU are categorized as the causes, and the remaining enablers are effects.

22.5 Conclusions

Identifying the key enablers helps in maximizing the PFSC efficiency and promotes a better decision making process. In this chapter, we have considered the enablers related to PFSC from the previous literature conducted by the authors. To understand the key enablers importance and their relevance with respect to current technologies in the context of PFSC a MCDM has been adopted. The better approach leads to good performance solutions, hence, a hybrid MCDM approach with integrated AHP-PROMETHEE II has been proposed in which the AHP method has evaluated the weights of various enablers where PROMETHEE II is used to rank the enablers using the weights obtained from the AHP. Later, another prominent MCDM method i.e., DEMATEL is employed in this chapter that helps in ranking the enablers and categorizes them into cause and effects. Both the proposed and adopted methods are used in this chapter to compare and analyze the results for validation. It has been found that the enabler shelf life is the most key influential enabler with the highest ($K_j + L_i$) value 5.3254, followed by 4.937 (RFID), 4.653 (VR), 4.292 (IC), 4.064 (CBA), and 3.9512 (DM). From the integrated AHP—PROMETHEE II method, it is evident that VR, RFID, SL, IC, and CBA enablers have the highest ranking. Combining both the results, the proposed ranking is as follows: SL > RFID > VR > IC > CBA > DM > 3PL > IoT > CS > EU > FC > FK > CE > GW > UC. From the above results, SL, RFID, VR, IC, CBA, and DM are the most influential enablers in the context of PFSC. RFID helps in digitalizing the PFSC, thereby improving its performance and efficiency. Similarly, vehicle routing and inventory control directly impact the most significant characteristic of perishable goods i.e., shelf life. Lastly, the operational and transportation costs are important for transporting of the perishable products within their shelf life period in PFSC.

Table 22.10 Ranking of the enablers

Rank	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
DEMATEL	SL	RFID	VR	IC	CBA	DM	3PL	IoT	CS	EU	FC	FK	CE	GW	UC
PROMETHEE II	VR	RFID	SL	IC	CBA	DM	3PL	IoT	CS	EU	FC	CE	GW	UC	FK
Proposed	SL	RFID	VR	IC	CBA	DM	3PL	IoT	CS	EU	FC	FK	CE	GW	UC

References

1. Musavi M, Bozorgi-Amiri A (2017) A multi-objective sustainable hub location-scheduling problem for perishable food supply chain. Comput Ind Eng 113:766–778
2. Laniel M, Émond JP, Altunbas AE (2011) Effects of antenna position on readability of RFID tags in a refrigerated sea container of frozen bread at 433 and 915 MHz. Trans Res Part C: Emerg Technol 19(6):1071–1077
3. Jouzdani J, Govindan K (2021) On the sustainable perishable food supply chain network design. A dairy products case to achieve sustainable development goals.J Cleaner Prod 278
4. Karanam M, Krishnanand L, Manupati VK, Antosz K, Machado J (2022) Identification of the critical enablers for perishable food supply chain using deterministic assessment models. Appl Sci 12(9)
5. Kim K, Kim H, Kim SK, Jung JY (2016) i-RM: An intelligent risk management framework for context-aware ubiquitous cold chain logistics. Expert Syst Appl 46:463–473
6. Bogataj D, Bogataj M, Hudoklin D (2017) Reprint of "Mitigating risks of perishable products in the cyber-physical systems based on the extended MRP model." Int J Prod Econ 194:113–125
7. Singh AK, Subramanian N, Pawar KS, Bai R (2018) Cold chain configuration design: location-allocation decision-making using coordination, value deterioration, and big data approximation. Ann Oper Res 270(1):433–457
8. Chen MC, Lu CC, Liu YC (2018) Optimal consolidation of fresh agricultural products in a multi-temperature joint distribution system.Int J Log Manag
9. Stellingwerf HM, Kanellopoulos A, van der Vorst JG, Bloemhof JM (2018) Reducing CO_2 emissions in temperature-controlled road transportation using the LDVRP model. Transp Res Part D: Transp Environ 58:80–93
10. Bozorgi A (2016) Multi-product inventory model for cold items with cost and emission consideration. Int J Prod Econ 176:123–142
11. Joshi R, Banwet DK, Shankar R (2011) A Delphi-AHP-TOPSIS based benchmarking framework for performance improvement of a cold chain. Expert Syst Appl 38(8):10170–10182
12. Ma X, Wang J, Bai Q, Wang S (2020) Optimization of a three-echelon cold chain considering freshness-keeping efforts under cap-and-trade regulation in Industry 4.0. International Journal of Production Economics 220, (2020).
13. Song MX, Li JQ, Han YQ, Han YY, Liu LL, Sun Q (2020) Metaheuristics for solving the vehicle routing problem with the time windows and energy consumption in cold chain logistics. Applied Soft Computing 95, (2020).
14. Bozorgi A, Pazour J, Nazzal D (2014) A new inventory model for cold items that considers costs and emissions. Int J Prod Econ 155:114–125
15. Leng L, Zhang J, Zhang C, Zhao Y, Wang W, Li G (2020) Decomposition-based hyperheuristic approaches for the bi-objective cold chain considering environmental effects. Computers & Operations Research 123, (2020).
16. Meneghetti A, Ceschia S (2020) Energy-efficient frozen food transports: The refrigerated routing problem. Int J Prod Res 58(14):4164–4181
17. Cai X, Chen J, Xiao Y, Xu X (2010) Optimization and coordination of fresh product supply chains with freshness-keeping effort. Prod Oper Manag 19(3):261–278
18. Meneghetti A, Monti L (2015) Greening the food supply chain: an optimisation model for sustainable design of refrigerated automated warehouses. Int J Prod Res 53(21):6567–6587
19. Saif A, Elhedhli S (2016) Cold supply chain design with environmental considerations: A simulation-optimization approach. Eur J Oper Res 251(1):274–287
20. Zhang S, Chen N, Song X, Yang J (2019) Optimizing decision-making of regional cold chain logistics system in view of low-carbon economy. Trans Res Part A: Policy Prac 130:844–857
21. Bogataj D, Bogataj M, Hudoklin D (2017) Reprint of "Mitigating risks of perishable products in the cyber-physical systems based on the extended MRP model. Int J Prod Eco 194:113–125
22. Wu Z, Mohammed A, Harris I (2021) Food waste management in the catering industry: enablers and interrelationships. Ind Mark Manage 94:1–8

23. Kilic HS, Yalcin AS (2021) Comparison of municipalities considering environmental sustainability via neutrosophic DEMATEL based TOPSIS. Socio-Eco Plan Sci 75
24. Sharma M, Joshi S, Kumar A (2020) Assessing enablers of e-waste management in circular economy using DEMATEL method: an Indian perspective. Environ Sci Pollut Res 27:13325–13338
25. Khanzode AG, Sarma PR, Goswami M (2021) Modelling interactions of select enablers of Lean Six-Sigma considering sustainability implications: an integrated circular economy and Industry 4.0 perspective. Production Planning & Control, 1–7
26. Li P, Xu Z, Wei C, Bai Q, Liu J (2022) A novel PROMETHEE method based on GRA-DEMATEL for PLTSs and its application in selecting renewable energies. Inf Sci 589:142–161
27. Makan A, Fadili A (2020) Sustainability assessment of large-scale composting technologies using PROMETHEE method. J Cleaner Prod 261
28. Gupta N, Vrat P, Ojha R (2022) Prioritizing enablers for service quality in healthcare sector–a DEMATEL approach. J Health Org Manag
29. Tzeng G, Huang JJ (2011) Multiple attribute decision making: methods and applications. CRC press
30. Chakraborty K, Mondal S, Mukherjee K (2018) Developing a causal model to evaluate the critical issues in reverse supply chain implementation. Benchmark Int J
31. Şen CG, Çınar G (2010) Evaluation and pre-allocation of operators with multiple skills: a combined fuzzy AHP and max–min approach. Expert Syst Appl 37(3):2043–2053
32. Mousavi MM, Lin J (2020) The application of PROMETHEE multi-criteria decision aid in financial decision making: Case of distress prediction models evaluation. Expert Syst Appl 159

Chapter 23
Blockchain for Supply Chain for Perishable Goods

Anandu B. Ajith, Nikhil R, Pati Chandana, Vinod Pathari, and Vinay V. Panicker

Abstract Maintaining transparency in the perishable goods supply chain is critical to ensure that the items are not tampered with and that customers have faith in the products. This work aims to solve this problem utilizing blockchain technology, which is well-known for its decentralization, immutability, and trustless nature. The research goal was to develop an Android application that is both cost-free, easily deployable, and easy to use for recording product data and the supply chain. Before purchasing a product, a consumer might use this app to scan the QR (Quick Response) code to retrieve all of the product's details and supply chain to ensure that the product has not been tampered with and is still within its shelf life.

Keywords Supply chain · Perishable goods · Block chain · Transparency · Traceability · Android application

23.1 Introduction

Perishable goods have a short shelf life and are unmarketable after their expiry. However, they are being tampered with because of various reasons. The expiration dates on the product packages may be worn off, or counterfeit products may be sold under the names of well-known brands. Tampering with perishable foods, in particular, could expose consumers to dangerous food-borne infections [1]. Therefore, consumers must inspect and ensure that the product's box is not ripped, torn,

A. B. Ajith · R. Nikhil · P. Chandana (✉) · V. Pathari · V. V. Panicker
National Institute of Technology Calicut, Calicut 673601, India
e-mail: mail@chandana.dev

R. Nikhil
e-mail: mail@nikhilrajesh.com

V. Pathari
e-mail: pathari@nitc.ac.in

V. V. Panicker
e-mail: vinay@nitc.ac.in

M. K. Tiwari et al. (eds.), *Applications of Emerging Technologies and AI/ML Algorithms*, Asset Analytics, https://doi.org/10.1007/978-981-99-1019-9_23

or damaged and verify the expiration date [1]. However, the expiration date may be changed to extend the shelf-life of the products illegally, and the packaging could be reused to fill the faulty product with the genuine company's label. This tampering can be done very easily and quickly and would lead to customer skepticism in buying the products. Therefore, a need for a trusted source of product data that cannot be tampered with is necessary. This could be achieved through the application of blockchain technology.

Currently, a central database is used to store, retrieve, and modify the data in most industries, and a central system is often responsible for securing this data. However, this implementation has its disadvantages.

It is hard to retrieve the data once lost, there is a lack of transparency and security, and these centralized databases will not serve the requirement for sectors where decentralization, transparency, and immutability are much required. In this context, blockchain technology can be applied.

Blockchain is a decentralized ledger made up of a chain of blocks, where each block consists of a cryptographic hash of the previous block and several transactions and other details [2]. This distributed ledger technology ensures the following characteristics.

1. ***Decentralization***: Every node in the network owns a copy of the blockchain, ensuring that transparency is maintained and that there is no central entity solely responsible for the blockchain.
2. ***Security***: The transactions would be signed by the sender using the digital signature formed from the user's private key, and the other nodes in the network could verify the transaction's validity with the sender's public key, ensuring the transaction has not been tampered with. It would be impossible to tamper with the blockchain since it is stored across the network of devices, ensuring its security.
3. ***Immutability***: Even if a block of data in a blockchain is being modified, to make the blockchain valid, one should recalculate the cryptographic hashes of the entire chain from that block, which is not practically effortless; therefore, it cannot be modified or deleted once the blocks are added, making the chain practically immutable.
4. ***Consensus-based***: In order to agree that a transaction is valid, all the nodes in the network must agree on a consensus algorithm.

As a result, supply chain entities can store data in the blockchain and benefit from its qualities, while customers can look into the blockchain to check product details and ensure that the product is authentic before purchasing.

This paper is organized into six sections. Section 23.2 presents a literature review, followed by the problem described in Sect. 23.3. Section 23.4 describes the solution approach adopted in this research. Section 23.5 reports the results of our work, and the last section presents our conclusions.

23.2 Literature Review

Maintaining transparency in the perishable products supply chain is critical for consumer trust. If the product details such as manufacturing date, location, temperature of the product when manufactured and while transporting are recorded and made publicly accessible, the consumers can access the particular product's data and purchase it after making sure of its authenticity. But, storing the data in centralized servers would imply placing huge burden of trust and the responsibility of safeguarding the data on a single authority. Therefore, the collected data could be stored in a public block chain ledger so that the stored data is immutable once added to it and as the ledger is distributed, anyone can access the data anytime [3]. As the data is public, immutable, and distributed, prospective customers can access the data and verify the product's authenticity before purchasing it. This would boost the customer's trust in the manufacturing company [3]. Especially in the case of perishable food products, customers would often be worried about the temperature and other conditions maintained during the product's manufacturing, storage and transportation phases. To resolve it, data could be recorded by IoT enabled devices and stored in the public block chain ledger, so that customers can access the real-time and historical information about the products [4].

23.3 Problem Description

Due to the shorter shelf life of perishable goods, they are more vulnerable to tampering such as selling expired products or reusing the packaging of previously used products by stuffing the fake product with the genuine company's label. There is also a risk that the expiry date on the product has visibly deteriorated, which could cause customers to be confused about the product's expiry date. Customers' trust in the company would be eroded under such circumstances, and consuming products past their shelf life, particularly food products, could result in food poisoning and other food-borne diseases, which could be harmful. As a result, complete openness of all product data is essential so that consumers can access it prior to purchasing the product to assess its validity.

23.4 Solution Approach

We propose a solution based on blockchain technology to prevent the possibility of tampering of data in perishable goods. The product supply chain data in the blockchain is safe, tamper-proof, and readily available to all users. The idea is to store the lot number, product key, manufactured date, the expiry date, the manufactured location, and other product details in the blockchain. Date and location would also be added to the blockchain after the product is transacted in the supply chain. Consumers can view the entire supply chain and product details by scanning the QR code label on the product's outer cover.

When a customer purchases a product, the seller scans the QR code and marks it as purchased. If another consumer scans the QR code of a sold item, they see that the product has been purchased in the product supply chain and realize that there is an unacceptable reuse of the packaging (QR code) and avoid buying it.

A low-tech technique with which each product would be tracked is by using a combination of "Lot Number" and "Product Number," which, when combined, uniquely identifies the item. The product packaging will contain these identifiers as QR codes. Each participant in the supply chain has a unique private key and public key, which generates a unique address that is used to identify themselves.

Initially, the manufacturer constructs the blockchain, and the application adds the manufacturer as the network's authority node, which is responsible for generating blocks and delivering them across the network after confirming transactions from other nodes (wholesalers, retailers, etc.).

The manufacturer first creates the lot, which represents a set of similar products, and then records all of the details of the lot, such as Lot number, individual unique product keys of the products, manufacturing date, expiry date, and manufacturing location in a "CREATE LOT" transaction and transmits it to all nodes in the network. Except for manufacturing (first step) and purchasing of the product (last step), all intermediate transactions like wholesaler purchasing from the manufacturer or a retailer purchasing from a wholesaler would be recorded on the blockchain as a "PURCHASE" transaction. The transaction will contain details of when and where the transaction occurred. It is assumed that when a product is handed over, the address that sells the product adds the transaction.

"END PURCHASE" transaction is used to keep track of products purchased by end-users and products that had faults during distribution and are to be withdrawn from the market.

The product's route from manufacturer to retailer can be tracked with the above configuration. When a customer wants to examine the details of a product before buying it, they can use the app to scan the QR code of the product they want to buy, and the product's trip through the blockchain will be displayed. The end consumer can check the product's expiration date and ensure that the supply chain does not include an end purchase transaction before purchasing it. The end consumer can now make an informed choice of the legitimacy of the products and purchase the same. When a customer buys a product, the retailer adds an "END PURCHASE" transaction, which includes the date and location of the purchase, and marks the conclusion of the product's journey.

If a seller receives a tampered product, the seller can return it to the wholesaler and include an "END PURCHASE" with the wholesaler's address and the date and location the goods were returned and mark the product as purchased so that it cannot be sold again.

With the above configuration, the life-cycle of a product can be monitored transparently on the blockchain from producer to end consumer and, this can prevent tampering with perishable goods and boost customer confidence in the items and company.

In the traditional approach, storing all this information would require a lot of storage and the need to place trust in a centralized entity. By utilizing a distributed blockchain, with low powered Android devices, where there is no need to completely trust any entity, we aim to overcome the above limitations.

23.5 Results and Discussion

A Proof-of-Concept native android application that implements the above flow has been developed using our version of the blockchain library. The application can be used by.

- Producer to add CREATE LOT transactions, and, PURCHASE transactions while selling products to wholesalers.
- Wholesaler to add PURCHASE transactions while selling the products to retailers.
- Retailer to add END PURCHASE transactions while selling the product to the end consumer marking the product as purchased.
- End Consumer, can use the app to query by LOT ID, PRODUCT ID to get product details and supply chain of individual products.
- Everyone can view the information of all the transactions.

This application currently considers Producers as the Authority nodes, as it is their need that quality products reach the end consumer. Hence Producers are validators creating blocks. The schematic of the Android application is explained using the Use Case Diagram Fig. 23.1. Corresponding to each of these use cases, our blockchain undergoes an immutable change and this is depicted in Table 23.1. Screenshots of the proposed application are given in Fig. 23.2.

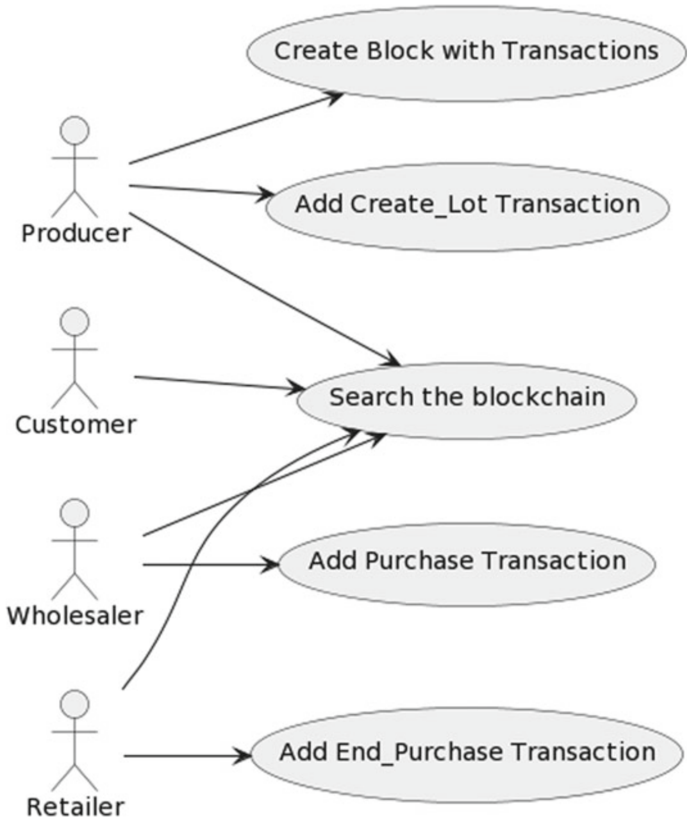


Fig. 23.1 Use case diagram of the application

23.6 Conclusion

Maintaining openness in the perishable foods supply chain has become increasingly critical with rising customer concerns about product authenticity. In order to accomplish this, a low-cost, quickly deployed android application is designed which is implemented utilizing blockchain technology at its core. This will ensure that the supply chain data stored in it is immutable, safe, and decentralized. The consumer can download the Android app, scan the QR code on the product, and inspect the complete product supply chain, ensuring that the goods have not been tampered with or re-sold at any point in the supply chain.

Table 23.1 Table with actions and activity on block chain by various actors

Actors	Action	Effect on the block chain
Producer	Add CREATE_ LOT Transaction	A CREATE_LOT transaction with lot number, product keys and other relevant metadata is created and broadcast on the Network
Producer	Create Block with Transactions	Periodically the Producer (which is the Authority node in our case) takes the valid transactions from the queue and produces a block. This block is then broadcast to all nodes
Wholesaler	Add PURCHASE Transaction	A transaction with product keys, lot number and other metadata is made, and broadcast on the network. Clients wait until this block is added to a block [i.e., recorded on the chain]
Wholesaler, Retailer	Add END_ PURCHASE Transaction	A transaction with product key, lot number is made and broadcast over the network. Clients will wait until a block with this transaction is added to the block chain
Producer, Wholesaler, Retailer, Customer	Search	The local copy of the block chain is searched using the parameters given [e.g.: lot number, product keys etc.]

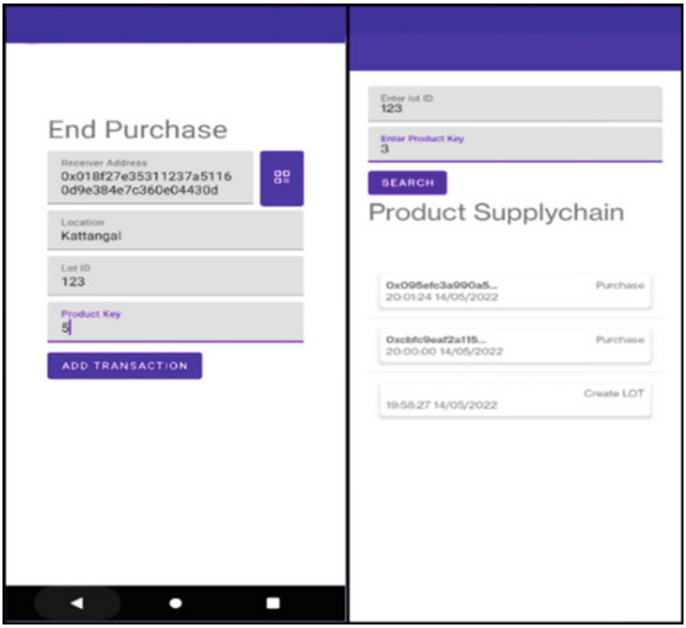


Fig. 23.2 Screenshots of the application

References

1. Food Tampering: An Extra Ounce of Caution. https://www.fda.gov/food/buy-store-serve-safe-food/food-tampering-extra-ounce-caution. Last accessed 12 May 2022
2. Gupta W (2017) Blockchain for dummies. IBM Limited Edition. Wiley, NJ USA
3. Ahamed NN, Thivakaran TK, Karthikeyan P (2021) Perishable food products contains safe in cold supply chain management using blockchain technology. In: 2021 7th international conference on advanced computing and communication systems (ICACCS), pp 167–172
4. Mondragon AEC, Coronado CE, Coronado ES (2019) Investigating the applicability of distributed ledger/blockchain technology in manufacturing and perishable goods supply chains. In: 2019 IEEE 6th international conference on industrial engineering and applications (ICIEA), pp 728–732

Chapter 24
A Novel Linear Mathematical Model Based Heuristic for a Class of Classification Problem with Non-linearly Separable Data

Anushee Jain and Chandrasekharan Rajendran

Abstract Classification is a supervised machine learning technique that is used to predict the class or category of a new observation based on training data. Classification techniques can be broadly categorized into logic-based (decision trees and rule-based classifier, perceptron-based (neural networks), and statistical techniques. In this study, we focus on developing a mathematical model based heuristic for binary classification problems. The application of Operation Research (OR) techniques is quite rare in the literature on classification. We propose a novel Linear Programming (LP) based model for binary classification. A very popular dataset available in the literature, Iris, has been used in our analysis. It is a multi-variate dataset first used by Fisher in [1], which consists of three classes of fifty instances each, where each class refers to a type of Iris plant. The classes are as follows: Iris Setosa, Iris Versicolor, and Iris Virginica. The proposed Linear Programming model consists of additional variables to show the interaction effect between the predictor variables along with the actual predictor variables so that non-linearly separable data points can also be classified by using the proposed LP based binary classifier. The values of the decision variables, obtained from building the classification model using LP on training data set, are used to predict the class of observations on the test data set. The overall accuracy of the proposed 2-stage classifier is found to be comparable with the results reported in the literature.

A. Jain · C. Rajendran (✉)
Indian Institute of Technology Madras, Chennai, India
e-mail: craj@iitm.ac.in

A. Jain
e-mail: anushee27@gmail.com

M. K. Tiwari et al. (eds.), *Applications of Emerging Technologies and AI/ML Algorithms*, Asset Analytics, https://doi.org/10.1007/978-981-99-1019-9_24

24.1 Introduction

One of the most common applications of machine learning is in the area of predictive analytics. Supervised machine learning forms an integral part of predictive analytics, which deals with algorithms to process labeled data. When the observations in the data are unlabeled, it is referred to as unsupervised machine learning [2]. Classification models in predictive analytics predict the class or category of a new observation while training from a given set of observations. The response variable in a classification problem is qualitative compared to a regression problem where the response variable is quantitative in nature [3]. The selection of any statistical technique depends on whether the response or dependent variable is qualitative or quantitative. Linear Regression is used for quantitative variables and logistic regression is used for modeling qualitative variables. However, most statistical learning techniques such as Naïve Bayes, tree-based techniques, and perceptron-based neural networks can be used for classification irrespective of the nature of the response variable after proper coding of the qualitative variables. A model which is able to classify two classes at a time is known as a binary classifier. A multi-class classifier can classify more than three classes. The main objective of the study is to develop a mathematical model-based heuristic for classification problems. A popular dataset used in classification problems, the IRIS dataset [1], has been used in our analysis.

24.2 Literature Review

In this section, we give a brief overview of the various applications of machine learning techniques for solving classification problems. We also give a review of the Linear Programming based models for classification found in the literature.

Classification is a supervised machine learning technique that is used to predict the class or category of a new observation based on training data. Classification techniques can be broadly categorized into logic-based (decision trees and rule-based classifier, perceptron-based (neural networks), and statistical techniques [4]. Application of ML-based techniques for classification has been found in intelligent decision-making, especially in the area of credit card fraud detection and identification of criminals. Another area where machine learning techniques find huge application is inventory management [5], optimizing logistics and warehousing costs. Healthcare is one of the prominent areas where machine learning techniques have been able to devise algorithms for the prognosis and detection of diseases like cancer, diabetes, and patient management [6–8].

Operations Research-based techniques using linear programming for classification were discussed by Freed and Glover [9]. The objective in their LP model maximizes the gap between the two class decision boundaries instead of minimizing the deviation from each class. Madankumar et al. [10] proposed an LP-based classifier by considering the interaction effect between the features and the contribution of features from their higher order polynomial degrees.

Pinto et al. [11] used various techniques such as Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and logistic regression on the Iris dataset and found the best accuracy using SVM. Thirunavukkarasu et al. [12] also used KNN in their study. Both studies used the scikit-learn library in python for their analysis. Wu et al. [13] study the application of Random Forests and Boosted trees on the Iris dataset. The study also compares the performance of the two techniques with popular algorithms KNN, SMO, and simple Classification and Regression Trees (CART).

The main research gap identified from the literature review is Operations Research based techniques for binary classification in data analytics have been less found in the literature. Therefore, our motivation is to integrate Operations Research and analytics. The remainder of the paper is organized as follows; Sect. 24.3 discusses the proposed Linear Programming model for classification for the Iris dataset, Sect. 24.4 discusses the numerical illustration and results, and Sect. 24.5 concludes the study along with key contributions, limitation, and future scope. It is to be noted that the proposed Linear Programming (LP) model consists of additional variables to show the interaction effect between the predictor variables along with the actual predictor variables so that non-linearly separable data points can also be classified by using the prepared LP-based classifier.

24.3 The Proposed Binary Classification Model

24.3.1 The Iris Dataset

It is a very popular dataset in the literature, and various techniques from data analytics and machine learning have been used to model it. It is a multi-variate dataset first used by Fisher in [1]. It is also published at UCI Machine Learning Repository. The data set consists of 3 classes of 50 instances each, where each class refers to a type of Iris plant. The classes are as follows: Iris Setosa, Iris Versicolor, and Iris Virginica. The attributes of the plant are given by sepal length (in cm), sepal width (in cm), petal length (in cm) and petal width (in cm).

24.3.2 The Linear Programming Model for Binary Classification

The dataset is partitioned into training and test sets, and the model building is done using the training set. We divide the observations by the square root of the sum of squares of every feature for normalization in our LP based heuristic model for classification. It is done by taking the square of every feature followed by taking the square root of the sum of the squares. Every attribute is then divided by its respective square root of the column sum. Test set is normalized by dividing the square root of the column sums of the training set. This is done because in real world size of test set is not known and therefore the square root of the column sums cannot be calculated a priori.

$$\text{normalize}(a_{i,j}) = \frac{a_{i,j}}{\sqrt{\sum_{i'=1}^{N} a_{i'j}^2}} \tag{24.1}$$

Linear Programming Model Formulation. We develop a model consisting of interaction effect terms C_r^P, where $r = 2$ along with P features for main effect. The concept of interaction effect between two predictors has been inspired by the design of experiments (DOE) where all the possible (full or partial factorial) combinations of the factors are hypothesized to affect the response variable. We also develop the general model consisting of P main features alone and compare the accuracy of the two. The parameters, decision variables and constraints for Iris dataset have been discussed below.

Parameters

P	Total number of features or predictors in each observation
N	Total number of observations in the training data set
j, j'	Index for the features; $j, j' \in 1, 2, 3, \ldots, P\}$
i	Index denoting an observation; $i \in 1, 2, 3, \ldots, N\}$
$a_{i,j}$	Value of feature j for observation i
$C1$	Set of observations belonging to Class 1
$C2$	Set of observations belonging to Class 2
k	Index denoting an observation from Class 1; $k \in C1$
l	Index denoting an observation from Class 2; $l \in C2$

Decision Variables

Intercept	A surrogate constant
x_j	An unrestricted variable to capture the coefficient for the term $a_{i,j}$
$x'_{j,j'}$	An unrestricted variable to capture the coefficient of the interaction term with respect to $\sqrt{(a_{i,j} \times a_{i,j'})}$
u_1	An unrestricted variable to denote the upper limit for observations belonging to class 1
l_2	An unrestricted variable to denote the lower limit for observations belonging to class 2
Gap	A positive variable which takes the difference of the lower limit of class 2 and upper limit of class 1

Objective

$$\text{Maximize } Z = \text{gap} \tag{24.2}$$

subject to

Constraints

$$u_1 \geq \text{intercept} + \sum_{j=1}^{P} x_j \times a_{k,j} + \sum_{j=1}^{P-1} \sum_{j'=j+1}^{P} x'_{j,j'} \times \sqrt{(a_{k,j} \times a_{k,j'})} \forall k \in C1 \tag{24.3}$$

$$l_2 \leq \text{intercept} + \sum_{j=1}^{P} x_j \times a_{l,j} + \sum_{j=1}^{P-1} \sum_{j'=j+1}^{P} x'_{j,j'} \times \sqrt{(a_{l,j} \times a_{l,j'})} \forall l \in C2 \tag{24.4}$$

$$\text{gap} = l_2 - u_1 \tag{24.5}$$

$$\begin{aligned} &\text{gap} \leq P \\ &\text{gap} \geq 0, \text{ and} \\ &\text{intercept}, \; x_j, \; x'_{j,j'}, \; l_2 \text{ and } u_1 \text{ unrestricted in sign.} \end{aligned} \tag{24.6}$$

We have established two decision boundaries of u_1 and l_2 for the two classes of our problem. This implies that any observation in the training data set less than u_1 belongs to $C1$ and any observation greater than l_2 belongs to $C2$. Constraint (24.3) calculates the highest value among observations from $C1$ as a function of the predictor variables and their interaction effects. Similarly, constraint (24.4) calculates the lowest value among observations from $C2$ as a function of the predictor variables and their interaction effects. Constraint (24.5) calculates the difference of l_2 and u_1. As *gap* is a positive variable, constraint (24.5) also imposes the lower limit l_2 to be higher than the upper limit u_2. Constraint (24.6) is used to give an upper limit to the variable *gap* because of the maximizing nature of the objective function. We have restricted the upper limit of gap to be equal to the number of features in the model (i.e., P).

Testing the Proposed Model. The values of the intercept and coefficients x_j, $x'_{j,j'}$ obtained from building the model on training data set are used to predict the class of observations on the test data set. The region between u_1 and l_2 is a grey zone and any observation lying in the grey zone is assigned to the class which is nearest to it in terms of the classification function with respect to data points:

$$f(i) = \text{intercept} + \sum_{j=1}^{P} x_j \times a_{i,j} + \sum \sum x'_{j,j'} \times \sqrt{(a_{i,j} \times a_{i,j'})} \tag{24.7}$$

In general, if $|f(i) - u_1| \leq |f(i) - l_2|$, then assign datapoint i to Class 1; else, assign datapoint i to Class 2.

24.4 Numerical Illustration and Results

Considering the Iris dataset, we have set $P = 4$ and taking the C_2^P interaction effects for the proposed model. The LP model is solved using ILOG CPLEX 20.1.0.0 on a PC with INTEL® Xeon® CPU E5-2620 v3@ 2.40 GHz 2 Processor, 64.0 GB RAM.

The steps followed for the binary classification of Iris dataset is given below in the flow chart (Fig. 24.1).

We follow a proportionate sampling approach to generate the training and test dataset. A sample of training set consists of equal observations from each class. We have used caret package from R library to create 5 equal folds of data for modeling. As our LP model is a binary classifier, only two classes are given as input to the model. In stage 1, a binary classifier is trained to classify species Setosa and Non-Setosa (versicolor and virginica). In stage 2, another binary classifier is trained over the remaining observations from Non Setosa class i.e., Versicolor and Virginica. It is evident that we have two binary classifiers, one for classifying into Setosa and Non Setosa, and another binary classifier for classifying into Versicolor and Virginica datapoints, from among Non Setosa data points.

24.4.1 *Accuracy of the Proposed Heuristic*

In general, accuracy of a classifier can be computed using a confusion matrix as:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{FN} + \text{TN}}, \tag{24.8}$$

where TP: True positive, TN: True Negative, FP: False Positive, and FN: False Negative.

Fig. 24.1 Flow chart to show the different stages in Iris classification

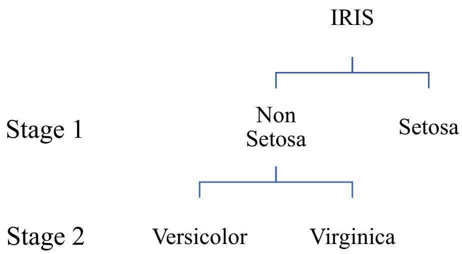


Table 24.1 Test accuracy for iris dataset using proposed LP based heuristic

Number of folds	Minimum accuracy (%)	Mean accuracy (%)	Maximum accuracy (%)
Five-folds	90	98	100

Table 24.2 Mean Accuracy of classification of Iris dataset as reported by [11]

	Support vector machine (SVM) (%)	K-nearest neighbor (KNN) (%)	Logistic regression (%)
Mean cross-validated accuracy	98	97	96
Mean accuracy (without cross validation)	96	93	91

Pinto et al. [11] used various techniques from machine learning for the classification of Iris dataset and found SVM to give the best test accuracy. We benchmark the accuracy reported by the authors in this work and report the overall accuracy of stage 1 and stage 2. We report the accuracy in terms of maximum, mean and minimum using the 80:20 training and test data split. For the 80:20 split, the data is divided into five equal folds. We also found that the model consisting of P features alone achieves 100% accuracy in the first stage but fails to give good results in the second stage, which indicates the presence of non-linearity in the Iris dataset. Table 24.1 gives the test accuracy for the model with interaction effects and Table 24.2 presents the results of techniques from literature.

In this work, we have proposed a new Linear Programming model for the classification problem. We observe that the accuracy of the proposed LP based heuristic for five-folds is comparable to the cross-validated accuracy reported by [11] using SVM (98%) in Table 24.2. We have also reported the minimum accuracy and the maximum accuracy which have not been reported in the literature. We also validated our dataset using the LP-model proposed by [9] and observe that the model cannot classify the given dataset, possibly due to a single decision boundary without the consideration of *gap* and without the consideration of interaction effect between the variables.

24.5 Conclusion and Future Work

Some of the key contributions of this study are as follows.

1. The LP model for binary classification is easy to implement and reproducible, while many of the machine learning techniques are not.

2. The normalization function in Eq. (24.1) is unique in our work. It takes care of the limitation of the min–max normalization technique in which some of the values in the test set may turn negative, if minimum value of training set is less than any of the observation in the test set.
3. The results of the LP-based classifier are comparable to that of the machine learning techniques reported in the literature.

The application of the proposed LP based heuristic is not only limited to Iris dataset, but also be used for other classification related datasets. Popular datasets from UCI Machine Learning repository like WDBC Breast Cancer dataset can be analyzed in future studies, and different aspects of supply chain such as classification and selection of suppliers can also be explored. The LP model for classification is binary and needs to be solved in more than one step for datasets with more than two classes. Future work can focus on experimenting with different values of the gap variable. There is also a scope of adding more features to bring additional non-linearity to the model or using different objective functions.

References

1. Fisher RA (1936) The use of multiple measurements in taxonomic problems. Ann eugen 7(2):179–188. https://doi.org/10.1111/j.1469-1809.1936.tb02137.x
2. Jain AK, Murty MN, Flynn PJ (1999) Data clustering: a review. In: ACM computing surveys, vol 31, issue 3. https://doi.org/10.4018/978-1-5225-1776-4.ch002
3. James G, Witten D, Hastie T, Tibshirani R (2013) An introduction to statistical learning, vol 7, issue 10. Springer, pp 995–1039. https://doi.org/10.1007/978-1-4614-7138-7
4. Kotsiantis SB, Zaharakis ID, Pintelas PE (2006) Machine learning: a review of classification and combining techniques. Artif Intell Rev 26(3):159–190. https://doi.org/10.1007/s10462-007-9052-3
5. Partovi FY, Anandarajan M (2002) Classifying inventory using an artificial neural network approach. Comput Ind Eng 41(4):389–404. https://doi.org/10.1016/S0360-8352(01)00064-X
6. Fatima M, Pasha M (2017) Survey of machine learning algorithms for disease diagnostic. J Intell Learn Syst Appl 09(01):1–16. https://doi.org/10.4236/jilsa.2017.91001
7. Nilashi M, Ibrahim O, Ahmadi H, Shahmoradi L (2017) An analytical method for diseases prediction using machine learning techniques. Comput Chem Eng 106. https://doi.org/10.1016/j.compchemeng.2017.06.011
8. Silahtaroğlu G, Yılmaztürk N (2019) Data analysis in health and big data: a machine learning medical diagnosis model based on patients' complaints. Commun Stat Theory Methods 50:1–10. https://doi.org/10.1080/03610926.2019.1622728
9. Freed N, Glover F (1986) Evaluating alternative linear programming models to solve the two-group discriminant problem. Decis Sci 17:151–162
10. Madankumar S, Navya P, Rajendran C, Gupta NS, Valarmathi B (2017) Linear programming (LP)-based two-phase classifier for solving a classification problem with multiple objectives. In: Big data analytics using multiple criteria decision-making models, pp 87–112. https://doi.org/10.1201/9781315152653
11. Pinto JP, Kelur S, Shetty J (2018) Iris flower species identification using machine learning approach. In: 2018 4th international conference for convergence in technology, I2CT 2018, pp 1–4. https://doi.org/10.1109/I2CT42659.2018.9057891

12. Thirunavukkarasu K, Singh AS, Rai P, Gupta S (2018) Classification of IRIS dataset using classification based KNN algorithm in supervised learning. In: 2018 4th international conference on computing communication and automation, ICCCA 2018, pp 5–8. https://doi.org/10.1109/CCAA.2018.8777643
13. Wu Y, He J, Ji Y, Huang G, Yao H, Zhang P, Xu W, Guo M, Li Y (2019) Enhanced classification models for iris dataset. Procedia Comput Sci 162:946–954. https://doi.org/10.1016/j.procs.2019.12.072

Chapter 25
Intermittent Demand Forecasting for Handtools in Forging Industries: A Svm Model

Harsora Karn and Ajay Gupta

Abstract One of the most challenging tasks is forecasting intermittent demand, yet the most important activity in the forging industry since it serves as a foundation for production and inventory level planning. It is likely to be the most difficult task in handtools manufacturing as well. While working with these kinds of demands, exponential smoothing is frequently utilised in practice. More improved approaches, such as the Croston method, SBA, MA, ARIMA, SARIMA model, and so on, have been researched based on the exponential smoothing method. Demand unpredictability and intermittency offer obstacles in accurate demand forecasting using traditional or better methods. Support vector machine (SVM) models have been found to outperform previous models in terms of accuracy. However, there are certain drawbacks to basic SVM models, such as the fact that the computation takes more time and does not result in a statistically significant gain in accuracy, and there are a few reasons for model resilience in demand forecasting. To anticipate intermittent demand, we used an adaptive univariate SVM (AUSVM) model. Real-world data from the forging (handtool) sector indicates its performance when compared to current models. In terms of computational time, AUSVM clearly surpasses basic SVM. The computational findings of the forging industry handtool scenario show that for a lot of non-smooth demand series, AUSVM provides an analytically important increase in accuracy and best inventory performance when compared to very famous parametric models. The paper concludes with an explanation of why AUSVM is better for forecasting demand and inventory control in the forging industry, in general, and for a handtool manufacturer in particular.

Keywords AUSVM · Intermittent demand · Support vector regression · Forecasting · SVM

H. Karn (✉) · A. Gupta
Dr. B R Ambedkar National Institute of Technology Jalandhar, Jalandhar, Punjab, India
e-mail: harsorakarn@gmail.com

M. K. Tiwari et al. (eds.), *Applications of Emerging Technologies and AI/ML Algorithms*, Asset Analytics, https://doi.org/10.1007/978-981-99-1019-9_25

25.1 Introduction

Many businesses are concerned about poor demand forecasts. It serves as the foundation for all aspects of inventory management. Various forecasting models have been created to account for various demand components such as trend and seasonality. Intermittent demand is a type of need that emerges at random, with many spans of time with little demand. Furthermore, when a demand emerges, it is occasionally for more than one unit. This is particularly true for service components and capital products. In this work, the word fluctuating demand refers to patterns that are smooth, intermittent, irregular, or lumpy. Intermittency, demand unpredictability, and the lack of explanatory factors all provide issues for the forging industry's intermittent demand forecasting (a handtool manufacturer). A recent achievement in predicting accuracy has been obtained by infrequent demand forecasting models relying on machine learning technology. The objective of this research is to see if, under certain conditions, a machine learning system with suitable carve may increase both forecasting veracity and stockpile effectiveness in the forging business (a handtool maker).

Over the last decade, a multitude of strategies for anticipating intermittent needs have been established in scholastic articles. The following forecasting methodologies have been discussed: parametric prediction, non-parametric bootstrapping, and machine learning models. The parametric model relies on a demand dissemination theory, but the inverse model does not require a distributional assumption to estimate demand. Because data gathering on sporadic needs in the production business is usually a challenging undertaking due to business exclusivity, researchers employ a variety of datasets.

Reference [2–9] contains the literature, which offers to some extent modestly a variety of possible planning approaches addressing the problem of demand uncertainty. Croston's [1] method underpins the majority of the research effort in the domain of intermittent demand in terms of the parametric forecasting methodology. This is the conventional strategy that is commonly used in practise and has been extensively researched in the academic literature [1]. Later, in 2001, Syntetos [3] publicised a key flaw in Croston's work, namely that it is skewed toward positive demands. And in 2005, the Syntetos and Boylan technique (SBA method) was introduced for improved accuracy, although it was subsequently discovered to be biased for low intermittent data. Several more approaches, such as the Bootstrapping techniques, have recently entered the intermittent demand forecasting space [4]. They examined six models for bootstrapping, the most important of which were the Willemain–Smart–Schwarz model (WSS) and the Viswanathan–Zhou model [4]. Croston and SBA concepts were updated to enhance later models, including Teunter–Syntetos–-Babai (TSB) [4], a seasonal autoregressive integrated moving average (ARIMA) [5].

Mobarakeh et al. (2017) used convoluted data with 36 periods to build a sliding window bootstrapping (SWB) model to anticipate next-period demand. Machine learning research publications include neural network and support vector regression (SVR). A few additional machine learning algorithms were also invented, such as XGBoost and Random Forest, which function similarly to a previously stated bagging approach [6, 7]. created a hyper parameterised SVM regression model set via a cumbersome grid search. Kourentzes [8] introduced two feed-forward neural network (FNN) models with 36 training measurements. Lolli [9] created an FNN as well as a recurrent neural network (RNN). Jiang [10] developed an adaptive autoregressive SVM model for spare component demand forecasting. Because of a powerful requirement on temporal series that are autocorrelated, its usefulness for intermittent demand is restricted.

This paper is divided into five sections. Section 2 examines the most extensively used technique for estimating intermittent demand. Section 3 offer an experiment containing specific techniques for using SVM regression, such as data set selection, performance criteria, and experiment results. The final portion contains conclusions and suggestions for additional research.

25.2 Demand Estimation Methods

Other forecasting models, based on successful models found in the literature and freshly built models, are discussed in this chapter. The models are validated using real-world demand data from the firm. Multiple markers of performance are used to contrast the various models.

25.2.1 Support Vector Regression (SVR)

In both theory and implementation, SVR is closely connected to SVM classifiers. A distance gauge must be added to the loss function. Regression can be both linear and nonlinear. A non-linear transformation such as the non-linear SVC technique, can be utilised to map the data into a high-dimensional feature space where linear regression is conducted. In the error loss function, Vapnik (1999) added the e-insensitive zone. This zone indicates the degree of accuracy at which the constraints on generalisation ability apply from a theoretical standpoint. Training vectors that fall inside the zone are considered accurate, but those that fall outside the zone are considered erroneous and help with the error correction function. These erroneous vectors are transformed into support vectors as shown in Fig. 25.1. Vectors that lie above and extreme of the dotted lines represent the support vectors, but those not outside the e-insensitive zone have no bearing on the regression function. Only support vectors may then be used to determine the regression surface.

SVR is fundamentally linear regression in the feature space. Despite being simplistic and inapplicable in real-world settings, it serves as a foundation for understanding sophisticated SVRs. The purpose of SVR is to develop an f (x) function that differs from the targets y_k by no more than ε for all of the training data while remaining as flat as feasible. Let the linear function f (x) have the following form:

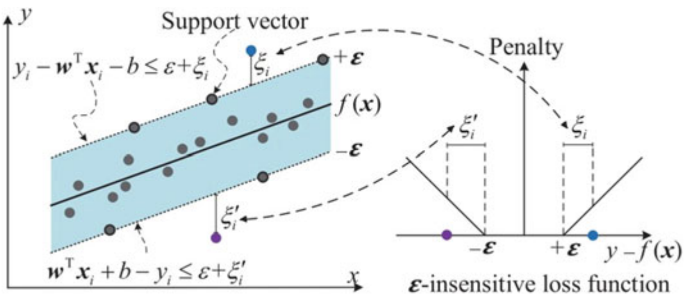


Fig. 25.1 One-dimensional non-linear regression with epsilon intensive band

$$f(x) = w^T x + b$$

In the fundamental weight space, the optimisation problem becomes

$$minimize\ \frac{1}{2} w^T w + C \sum \left(\xi^+ + \xi^-\right)$$

$$subject\ to: \begin{cases} y_i - \left(w^T \phi(x_i) + b\right) \le \varepsilon + \xi^+ \\ \left(w^T \phi(x_i) + b\right) - y_i \le \varepsilon + \xi^- \\ \xi^+, \xi^- \ge 0 \end{cases}$$

The primal Lagrangian form is used to solve this restricted optimisation problem:

$$\frac{\partial L}{\partial w} = 0 \to w = \sum\nolimits_{i=1}^{N} \left(\alpha_i - \alpha_i^*\right)\phi(x_i)$$

$$\frac{\partial L}{\partial b} = 0 \to w = \sum_{i=1}^{N} \left(\alpha_i - \alpha_i^*\right) = 0$$

$$\frac{\partial L}{\partial \xi_i} = 0 \to c - \alpha_i - n_i = 0$$

$$\frac{\partial L}{\partial \alpha_i} = 0 \to c - \alpha_i^* - n_i^* = 0$$

The kernel method was used here with $K(x, x_i) = \phi(x_k)^T \phi(x_l)$ or k, l = 1,…,N. The kernel trick is a technique for writing a nonlinear operator as a linear one in a higher dimension space. The model's dual form is the following equation, where $\alpha_{k,}\alpha_k^*$ are the QP problem solutions.

$$f(x) = \sum_{\dot{\mathrm{k}}=1}^{N} \left(\alpha_{\mathrm{k}} - \alpha_{\mathrm{k}}^*\right) K(x, x_{\mathrm{k}}) + b$$

25.2.2 Adaptive Univariate Svm (Support Vector Machine) Regression

Jiang (2021) created the adaptive univariate support vector machine (AUSVM) as a variation of the original support vector machine (SVM). The basic AUSVM is used for adaptive hyperparameter optimization, which replaces the usual grid search for classification and regression. Forecasting of demand for as long as lead time L and the review period R is required for a handtool firm using a system of periodic order-to-level inventory. In practise, the corporation can ideally decide between L and R. The following parameters are defined for modelling:

k = Lead Time + Review Period.

k = L + R (k = 1,2,3,….,K).

Training Period = 1,2,3,…..,T_1.

Validation Period = $T_1 + 1, T_1 + 2, T_1 + 3, \ldots.., T_1 + V$.

Testing Period = $T_1 + V + 1$, $T_1 + V + 2$, $T_1 + V + 3$,….., $T_1 + V + T_2$

$d_s^t = demand\ for\ hand\ tools\ s(s = 1, 2, 3, \ldots.., S)\ in\ a\ period\ t(t = 1, 2, 3, \ldots.., T)$

where t is defined as a weekly or monthly term depending on the circumstances

$$D_s^{t,k} = \sum\nolimits_t^{t+k-1} d_s^t$$

where $D_s^{t,k}$ is represents the demand of a handtool s over k period starting from t

$$w = Coefficient\ vector$$

$$b = intercept$$

$$C = Regularization\ constant$$

$$\varepsilon = Loss\ function\ thresold$$

$$\gamma = Kernal\ parameter$$

$$d_s^{t-1} = demand\ in\ periodt - 1$$

τ_s^{t-1} = number of zero demand periods between the last two non − zero demands

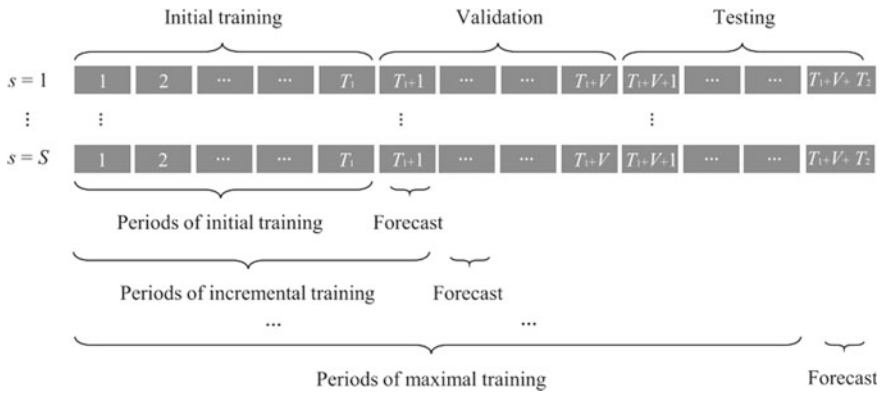


Fig. 25.2 Dynamic training, validation and testing process

δ_s^{t-1} = number of successive periods with zero demands at the end of t − 1 period
x_s^t = Attribute Vector for item s in period t

$$x_s^t = [d_s^{t-1}, \tau_s^{t-1}, \delta_s^{t-1}]$$

AUSVM combines the procedures of undulate validation, training and testing (Fig. 25.2). The total training sample capacity grows during the voluble training process (i.e. from T1 to T1 + V + T2 1), where T1, V and T2 signify the training, validation and testing set's time periods. The model is first created during the initial training period, and then forecasted during the first validation phase. The training period is then increased to the first validation period, and the forecast is performed again on the following validation period. In this manner, the model gets trained while also being enhanced. The validation procedure tries to fine-tune important hyperparameters utilised in testing. The testing periods are used to assess the model's efficacy.

The validation procedure is separated into three stages: main problem, subproblem, and dual-subproblem. When the hyperparameter is obtained from the Main problem, the Dual-subproblem and Subproblem are utilised to construct the best-suited SVM regression problem. The main issue is with the hyperparameter (e.g., regularisation constant, an error threshold in loss function and kernel parameters).

Main Problem:

$$\theta_k^* = \left(C_k^*, \varepsilon_k^*, \gamma_k^*\right)$$

$$= arg\left(\frac{1}{SV}\sum_{s=1}^{S}\sum_{t=T_1+1}^{T_1+V}\left(\frac{\left|D_s^{t,k} - F_s^{t,k}\right|}{\left(\frac{1}{V-1}\sum_{q=T_1+2}^{T_1+V}\left|D_s^{q,k} - D_s^{q-1,k}\right|\right)}\right)\right), \forall k \quad (25.1)$$

$$\text{subjectto}: F_s^{t,k}(\theta_k) = \sum_{i=1}^{t-k}\left(a_{i,s}^{t,k}(\theta_k) - a_{is}^{\prime tk}(\theta_k)\right)\pi_k\left(x_s^i, x_s^t\right) + b_s^{t,k}(\theta)_k \quad (25.2)$$

$$\pi_k\left(x_s^i x_s^t\right) = \begin{cases} \left(x_s^i\right)^T x_s^t + r, r \geq 0, if linear \forall s, \forall t, \forall i \\ \left[\gamma_k \left(x_s^i\right)^T x_s^t + r\right]^d, r \geq 0, \gamma_k \geq 0, d \in N, if polynomial \forall s, \forall t, \forall i, \forall k \\ \exp\left(-\gamma_k \|x_s^i - x_s^t\|^2\right), \gamma_k \geq 0, if radial basis \forall s, \forall t, \forall i, \forall k \\ \tanh\left(\gamma_k \left(x_s^i\right)^T x_s^t + r\right), r \geq 0, \gamma_k \neq 0, if sigmoid \forall s, \forall t, \forall i, \forall k \end{cases} \tag{25.3}$$

Subproblem:

$$w_s^{t,k}, b_s^{t,k}, \xi_s^{t,k}, \xi_s^{'t,k}$$

$$= arg \min_{w_s^{t,k}, b_s^{t,k}, \xi_s^{t,k}, \xi_s^{'t,k}} \frac{1}{2}\left(w_s^{t,k}\right)^T w_s^{t,k} + C_k \sum_{i=1}^{t-k}\left(\xi_s^{t,k} - \xi_s^{'t,k}\right), \forall s, \forall t, \forall k \tag{25.4}$$

$$\text{subjectto}: \begin{cases} D_s^{i,k} - \left(w_s^{t,k}\right)^T \phi\left(x_s^i\right) - b_s^{t,k} \leq \varepsilon + \xi_{i,s}^{t,k}, \forall k, \forall i, \forall s, \forall t \\ \left(w_s^{t,k}\right)^T \phi\left(x_s^i\right) + b_s^{t,k} - D_s^{i,k} \leq \varepsilon + \xi_{i,s}^{'t,k}, \forall i, \forall k, \forall s, \forall t \\ \xi_s^{t,k}, \xi_s^{'t,k} \geq 0, \forall s, \forall i, \forall t, \forall k \\ C_k \varepsilon_k > 0, \forall k \end{cases} \tag{25.5}$$

Dual-subproblem:

$$\begin{aligned} \left\{a_s^{t,k} a_s^{'t,k}\right\} = arg \max_{a_s^{t,k} a_s^{'t,k}} - \varepsilon_k \sum_{i=1}^{t-k}\left(a_{\text{i},s}^{t,k} + a_{\text{i},s}^{'t,k}\right) + \sum_{i=1}^{t-k}\left(a_{\text{i},s}^{t,k} - a_{\text{i},s}^{'t,k}\right) D_S^{i,k} \\ - \frac{1}{2}\sum_{i=1}^{t-k}\sum_{j=1}^{t-k}\left(a_{\text{i},s}^{t,k} - a_{\text{i},s}^{'t,k}\right)\left(a_{j,s}^{t,k} - a_{j,s}^{'t,k}\right) \times \pi\left(x_s^i, x_s^j\right), \forall s, \forall t, \forall k \end{aligned} \tag{25.6}$$

$$\text{subject to}: \begin{cases} \sum_{i=1}^{t-k}\left(a_{\text{i},s}^{t,k} - a_{\text{i},s}^{'t,k}\right) = 0, \forall s, \forall t, \forall k \\ 0 \leq a_{\text{i},s}^{t,k}, a_{\text{i},s}^{'t,k}, a_{j,s}^{t,k}, a_{j,s}^{'t,k} \leq C_k, \forall i, \forall j, \forall s, \forall t, \forall k \end{cases} \tag{25.7}$$

Equation 25.1 is an objective function of the main issue, which is to minimise the MSAE and optimise the hyperparameters. The Lagrange multiplier is the solution to the Dual-subproblem generated from Eqs. 25.6–25.7. These Lagrange multipliers are utilised in Eqs. 25.2–25.3 to calculate the forecast value, which is then applied to the goal function. $b_s^{t,k}$ signifies the regression function's intercept, whereas $\pi_k\left(x_{s,}^i x_s^t\right)$ denotes the kernel function derived from a linear, polynomial, or radial basis, and sigmoidal kernels. The least validation error criterion may be used to choose the best kernel function.

To tune the hyperparameter, the basic SVM employs the PSO (particle swarm optimisation) method. [8] created the new enhanced APSO (adaptive particle swarm optimisation) method, which has supremacy in comparison of speedy preventing prematurity, convergence, and robustness. This APSO method is included in the

Main problem to automatically jingle the AUSVM aggressive bounds. In the adaptive process, the MASE in Eq. 25.6 is used as the vigour value. Changes in {Ck, εk, γk} (i.e., θk). Update outcome Subproblem variables and their effects twin counterparts. For simplicity in Subproblem and Dual-subproblem θ_k is removed from it. Equation 25.4 optimises the model with a given set of hyperparameters from the Main issue by generating a flat regression model and dealing with demand uncertainty (Under a skewed distribution, the long tail of demand is extremely important). Constraints 1–2 of Eq. 25.5 allow the output function to estimate with k accuracy all attribute-response pairings $(x_s^i,\ D_s^{i,k})$. Constraints 3–4 of Eq. 25.5 define the boundaries.

We obtain the intercept b as follows from the optimisation results of constraint 1–2 of Eq. 25.5 based on the Lagrange and KKT conditions:

$$b_s^{t,k} = -\frac{1}{2}\sum_{i=1}^{t-k}\left(a_{\mathrm{i},s}^{t,k} - a_{\mathrm{i},s}^{\prime t,k}\right)\pi\left(x_s^i,\ x_s^{\varsigma}\right) + \pi\left(x_s^i,\ x_s^{\varsigma'}\right),\quad \forall s, \forall t, \forall k$$

whereas the final forecast value is given by

$$F_s^{t,k}\left(\theta_k^*\right) = \sum_{i=1}^{t-k}\left(a_{\mathrm{i},s}^{t,k}\left(\theta_k^*\right) - a_{\mathrm{is}}^{\prime tk}\left(\theta_k^*\right)\right)\pi_k\left(x_s^i,\ x_s^t\right) + b_s^{t,k}\left(\theta_k^*\right)$$

It is worth noting that subproblem and its consequences twin are changed from ε-SVM. They may be found using the Libsvm platform [6] by inserting demand and characteristics structured as shown in Eqs. 25.3, 25.5, 25.6.

25.3 Experimental Setting

25.3.1 Datasets

The dataset was compiled using firm data, namely sales data of ordered handtools in inventory from 2016 to 2019. Data is gathered on a weekly basis, including item codes, item descriptions, and ordered/shipped amounts. All goods are then classified into categories based on their inter-demand intervals and degrees of idiosyncrasy and repetitiveness. The average inter-demand interval (also known as the average demand interval) and the squared coefficient of variation were used to determine the amount of intermittence and erraticness of each item.

Following that, products with intermittent demand are divided into four categories based on the kind of Erratic, Lumpy, Smooth, and Intermittent demand patterns such that two cut-off values of 0.49 and 1.32, respectively.

25.3.2 Performance Criteria

The accuracy of the forecast is measured using the following demographic metrics: normalised mean squared error (NMSE), mean absolute error (MAE), and mean absolute scaled error (MASE). The NMSE and MAE are measurements of the difference between actual and projected values. The lower the NMSE and MAE values, the closer the anticipated time series values are to the actual values. In MASE, amplification of a loss relies mostly on the naïve method's in-sample MAE. The mean absolute error (MAE) and mean square error (MSE), which are commonly used to evaluate time series forecasting outcomes, are as follows:

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|y_i - \hat{y}_i\right|$$

$$MSE = \frac{1}{N}\sum_{i=1}^{N}\left(y_i - \hat{y}_i\right)^2$$

In modelling and forecasting, the criterion for determining the optimal model is the comparatively modest MAE and MSE.

25.3.3 Experimental Results

Using eminent real datasets from the handtool manufacturer, the prediction skills of the recommended model are compared with artificial neural networks (ANNs), autoregressive integrated moving average (ARIMA), and feed-forward neural networks (FNN) models. In this study, the mean absolute error (MAE) and mean squared error (MSE), as well as the mean absolute scaled error (MASE), which are frequently used for evaluating the results of time series forecasting, were tilized as performance measures. Table 25.1 summarises the MAE and MSE values for the models' latest observations. The bold numbers in each row show the finest veracity per stream model, while the highlighted by-line value denotes the best among all models.

Table 25.1 A comparison of predicting precision and calculation time of several models

Errors	Croston	ARIMA	ANN	SVM	AUSVM
MAE	37.3	12.68	**9.89**	14.51	14.19
MSE	42.49	19.54	21.05	22.15	**21.03**
MASE	0.839	0.869	0.997	0.826	**0.817**
Time (s)	**25.19**	189.11	986	768	681.66

25.4 Conclusions

Machine learning models outperform time series approaches, generating more accurate predictions that are closer to the real data, as we discovered in the literature.

In the forging business, AUSVM has been shown to be useful at forecasting intermittent demand (a handtool manufacturer). AUSVM has attained substantially shorter reckoning time among machine learning models, and a numerically symbolic accuracy yield Across the models of parametric, bootstrapping, and neural networks when tested on real-world data with 160 handtools. The following are the causes for these findings: The linear kernel function has been validated as a preferable alternative for AUSVM over the difficult RBF kernel utilised in SVM. Because of the linear kernel function, an adaptive particle swarm optimisation technique, rather than a cumbersome grid search, tunes fewer hyperparameters for numerous spare parts. AUSVM can recognise historically comparable scenarios by combining the three features, which aids in capturing demand dynamics. Simultaneous hyperparameter tuning aids in the prevention of overfitting and the reduction of the number of hyperparameters by the resulting linear kernel function.

References

1. Croston JD (1977) Operational Research Society is collaborating with JSTOR to digitize, preserve, and extend access to Operational Research Quarterly (1970–1977). ® www.jstor.org. Oper Res Q 28(3):663–670
2. Vapnik V (1999) The Nature of Statistical Learning Theory. Springer, New York
3. Syntetos AA, Boylan JE (2001) On the bias of intermittent demand estimates. Int J Prod Econ 71(1–3):457–466. https://doi.org/10.1016/S0925-5273(00)00143-2
4. Syntetos AA, Boylan JE, Croston JD (2005) On the categorization of demand patterns. Journal of the Operational Research Society 56(5):495–503. https://doi.org/10.1057/palgrave.jors.2601841
5. Zhang, R., Bao, Y. and Zhang, J. (2010) 'Forecasting erratic demand by support vector machines with ensemble empirical mode decomposition', *Proceedings - 3rd International Conference on Information Sciences and Interaction Sciences, ICIS 2010*, pp. 567–571. doi:https://doi.org/10.1109/ICICIS.2010.5534762.
6. Chang, C.C. and Lin, C.J. (2011) 'LIBSVM: A Library for support vector machines', *ACM Transactions on Intelligent Systems and Technology*, 2(3). doi:https://doi.org/10.1145/1961189.1961199.
7. Zhou C, Viswanathan S (2011) Comparison of a new bootstrapping method with parametric approaches for safety stock determination in service parts inventory systems. Int J Prod Econ 133(1):481–485. https://doi.org/10.1016/j.ijpe.2010.09.021
8. Jiang P, Liu X, Shoemaker CA (2017) An adaptive particle swarm algorithm for unconstrained global optimization of multimodal functions. ACM International Conference Proceeding Series, Part F1283:221–226. https://doi.org/10.1145/3055635.3056639
9. Lolli F et al. (November 2016) (2017) 'Single-hidden layer neural networks for forecasting intermittent demand.' Int J Prod Econ 183:116–128. https://doi.org/10.1016/j.ijpe.2016.10.021
10. Jiang P, Huang Y, Liu X (2021) Intermittent demand forecasting for spare parts in the heavy-duty vehicle industry: a support vector machine model. Int J Prod Res 59(24):7423–7440. https://doi.org/10.1080/00207543.2020.1842936

Chapter 26
Electric Vehicle and Charging Infrastructure Development: A Comprehensives Review Using Science Mapping and Thematic Analysis

D. V. Pendam and T. M. Rofin

Abstract This research study will conduct a literature review on the development of electric vehicles and charging station infrastructure. We used the four-phase method to review the literature. The review process includes bibliometric search, descriptive analysis, scientometric analysis, and citation network analysis. In phase I, the 957 articles retrieved from Scopus and Web of Science from 2008 to 2022 were systematically screened. The final selected articles were then subjected to descriptive analysis to identify the most influential authors, articles, keywords, and countries in the EV research domain. The research concepts/themes and methods were then classified using thematic analysis. Numerous discoveries have been made in the development of electric vehicles and charging infrastructure as a result of this review. China, the United States, and Germany are the leading countries in all areas of EV and EVC research. The research gap and issues of EV and EVCS are highlighted at the end of the review, as is the scope for future discussion.

Keywords Electric vehicle · EVs · Charging station · Charging infrastructure

26.1 Introduction

Greenhouse gas emissions (GHGs) from the burning of fossil fuels and climate change (CC) are critical issues in the world [1–3]. One of the reasons for this is the continuous usage of nonrenewable sources in conventional vehicles such as petrol and diesel which not only depletes the available natural resources but also pollutes the environment [4]. This leads to a negative impact on human lives as well as the

D. V. Pendam (✉) · T. M. Rofin
National Institute of Industrial Engineering Mumbai, Mumbai, India
e-mail: duryodhan.pendam.2020@nitie.ac.in

T. M. Rofin
e-mail: rofintm@nitie.ac.in

M. K. Tiwari et al. (eds.), *Applications of Emerging Technologies and AI/ML Algorithms*, Asset Analytics, https://doi.org/10.1007/978-981-99-1019-9_26

environment [5, 6]. To address GHG and CC issues, 193 countries around the world came together at Paris Summit 2015 and decided to reduce GHG emissions and limit the global limit up to 2 °C and further reduce to 1.5 °C [7]. With this commitment, participant countries (develop and developing) agreed for the net zero emission target by 2050 to achieve the sustainable development goal. This is the global motivation for the research community, scientists, academicians, and policymakers to reduce carbon emission. Adoption of EVs and electrification of transportation will help in reducing carbon emission. If the carbon emission keeps increasing will create a more worst situation for human lives and the environment. To overcome this issue, there is a need and necessity for the alternative sources to traditional vehicles. The adoption of the electric vehicle (EV) seems to be one of the best solutions in front of the government and the automobile industry to resolve the issue of GHG and CC [8, 9]. But, inadequate availability of charging facilities or infrastructure hinder the adoption of electric vehicle in developed as well as developing countries[10–12].

Now, developed countries have adopted the electric vehicle and started setting up the charging infrastructure to charge the electric vehicle without interruption. The amount of charge in an EV will decide how much it can run. So, it is good if the time battery will have enough amount of charging before starting the trip, otherwise, it will create problems while traveling long distances. Since COVID-19, most people prefer to travel using their own car to avoid infections. According to IEA, it is shown that some people prefer the electric vehicle over traditional cars and still some people like traditional vehicles. But the government has taken the bold decision to scrap the old conventional vehicle or convert it into EV. According to EV global outlook (2022), it is expected that to be 360 million by the end of this decade. In addition to this, the automobile industry has started making the EVs including Plug-in-hybrid electric vehicle (PHEV), Battery electric vehicle (BEV), and Electric vehicle (EV). These vehicle charge using the connector which are classified based on the type of charge connector used such as slow charging (level 1), high charging (level 2), and fast charging (level 3). The current status of electric vehicles and charging stations shows the increasing trends in EV and EVCS. Also, to increase connectivity, charging stations must have a global standard of charging connectors and connect all types of electric vehicles. But inadequate infrastructure, long charging time, frequent charging, and rollout of electric vehicle on the road are some technical issues [14–16]. To resolve these technical issues, the automobile industry is shifting toward electric mobility to improve the solution for battery range, charging time, EV-related issues, helps in reducing the carbon emission for sustainable development.

During the early stages of EV research, researchers, scientist, and policymakers identified the problems associated with EVs and charging infrastructure. It also has substantial barriers to adopting EVs and EVCS including technological, social, and economic challenges. For example, there exists too much market pressure for future technology requirements and how the EV industry will respond to it [21]. However, [22] emphasized the importance of charging infrastructure based on EV travel patterns and how EV infrastructure can address the issues of EV charging on distances (daily travel). In this regard, one of the critical issues is the design of a charging network [14, 23] mentioned the type of charging such as fast or slow charging, charging station planning [24], and discussed the scheduling of multiple charging points for charging EVs [25]. Although most research on EV design, planning, scheduling, and infrastructure management is taking place before 2020, the number of charging points and electric vehicles remains limited.

Current research from 2021–2022 continues to highlight issues such as charging infrastructure, charging type, charging location, charging facilities, and emerging technologies such as battery storage and wireless charging, among others. For example, [24] discussed the various types of EVs and their infrastructure, but they emphasized that the problem of charging infrastructure remains complicated. [25] stated that "the proportion of electric vehicles in transportation fleets is increasing, but widespread adoption will be impossible without adequate charging infrastructure." The deployment of such infrastructure should be guided by a strategy that takes into account both the environment in which it is installed and the behavior patterns of electric vehicle users. If these factors are not considered, there is a risk of failing to meet the needs of users. [26] raised the issue of an inadequate EV management system. [27] focuses on EV security and the EV Ecosystem. [28] concentrate on the development of EV infrastructure to meet the energy storage needs of e-mobility. This reflects the growing interest in EV and EVCS research among the research community to address the issues of carbon emissions.

Considering the above thoughts of the researchers and their directions, it is also very important to understand the current or as-is status of electric vehicles and charging infrastructure in developed and developing countries for achieving the net zero target [24–26]. To track and monitor the progress of EV and EVCS, our primary focus is to fill the identified research gap. Here, we have used science mapping and citation network analysis methods using RStudio and VOSviewer software to find the outcome of selected articles on EV and EVCS domain. This software aids in the analysis of data retrieved from Scopus and Web of Science for this research study. The objective of this research study is to conduct a thorough evaluation and determine the stages of electric vehicle and charging infrastructure development in developed and developing countries in the EV research domain. The following research questions will be addressed in the comprehensive review of EV and EVCS research domain:

RQ1: What are the current research trends in EV and EVCS?

RQ2: Which emerging authors, institutions, countries, and journal articles have had the greatest influence on the EV and EVCS domains?

RQ3: What academic structure does the EV and EVCS research knowledge base have?

Finally, the foregoing is the outline for the paper. Section 2 demonstrates the research methodology. Section 3 summarizes the results of the bibliometric, scientometric, and citation network analyses. Section 4 discusses the major key findings of EV and EVCS research with a discussion followed by the future scope, including new research directions.

26.2 Review Methodology

We used the four-phase review scheme [26] in this method, which is clarified and displayed in Fig. 26.1. The main objective of this study is to understand the current state of EV and charging infrastructure in developed and developing countries. Then, investigate the various emerging fields of research in EVCS research, classify a summary research field, and identify future research prospects in each area. The review process is organized into four levels: (a) bibliometric search, (b) descriptive analysis, (c) scientometric analysis, and (d) citation network analysis.

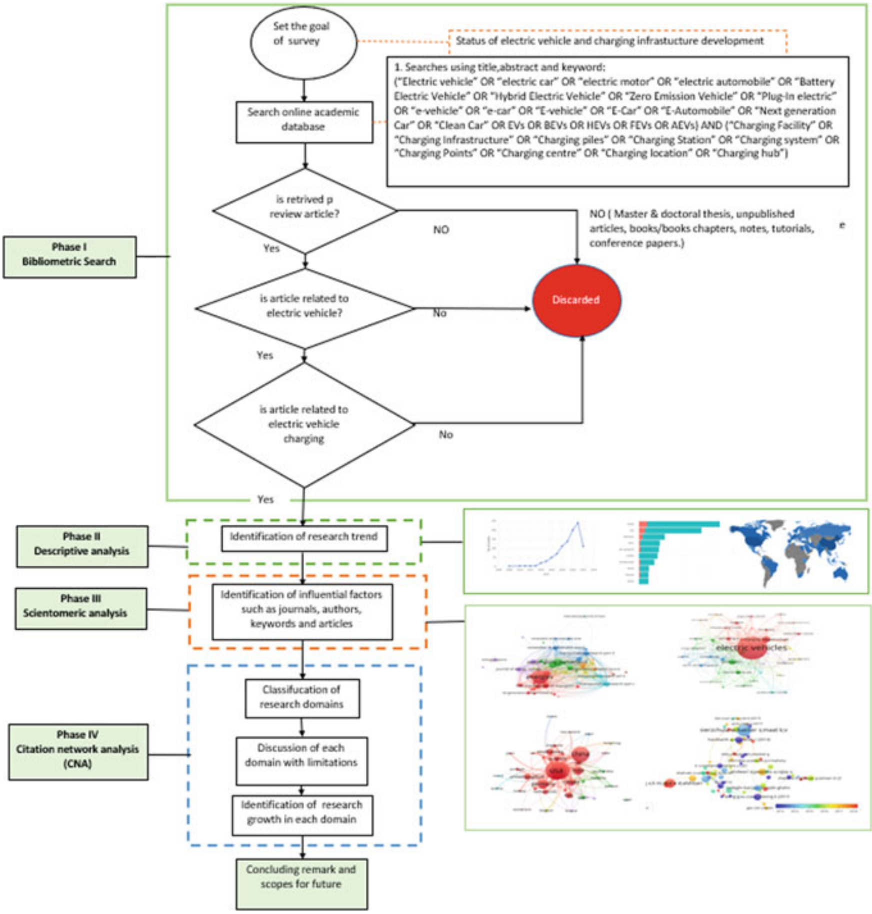


Fig. 26.1 Flowchart for four-phase literature review process

26.2.1 Phase I: Bibliometric Search

Time period

The period for the review studies does not give any limit cap to cover maximum coverage of the research field, but the momentum of research starts from the year 2008 onwards. So, we take all the articles from the 2008 to 2022 period.

Database selection

The articles were regained from the Scopus database for the literature survey because it has comprehensive coverage and covers almost all journals and publications associated with other sources. However, we took additional articles from the science web to not miss out on any paper from Scopus.

Journal selection

In this research study, the journals focusing on an electric vehicle, charging infrastructure, and its relevant domains include energy, environmental science, decision sciences, business management, transportation, and transportation science.

Article selection

For the selection of articles, we follow the keyword search function such as Abstract, Title, Keyword ((“Electric vehicle” OR “electric car” OR “electric motor” OR “electric automobile” OR “Battery Electric Vehicle” OR “Hybrid Electric Vehicle” OR “Zero Emission Vehicle” OR “Plug-In electric” OR “e-vehicle” OR “e-car” OR “E-vehicle” OR “E-Car” OR “E-Automobile” OR “Next-generation Car” OR “Clean Car” OR EVs OR BEVs OR HEVs OR FEVs OR AEVs) AND (“Charging Facility” OR “Charging Infrastructure” OR “Charging piles” OR “Charging Station” OR “Charging system” OR “Charging Points” OR “Charging Centre” OR “Charging location” OR “Charging hub”)). Initially, 10,616 total articles were found after that keyword function was used to keep the article in the line research field of interest. Lastly, the articles published in peer-reviewed journals in the English language are well-thought-out for the final study and exclude other articles. After that, the article and review article filter are used, selecting the journal publication for the last pieces. Finally, 939 articles were considered for the definitive research study. Here, we have used the R studio functions to merge all the data file from Scopus and Web of Science and convert them into CSV files. After that, we used the biblioshiny for further analysis.

26.2.2 Phase II: Descriptive Analysis

Following phase I process, a descriptive analysis is carried out with r studio and the biblioshiny function to identify:

1. Main information about data.
2. Article distribution by year of publication.
3. Journals' contributions.
4. The authors' influence.
5. Country contribution.
6. Institution contributions.

26.2.3 Phase III: Scientometric Analysis

After phase II is finished, a scientometric study will be conducted using the VOSviewer, a software tool (van Eck and Waltman, 2009). This tool can improve the visual interpretation of networks based on distance, with each separation between two nodes indicating the degree of proximity between them. This method identifies the effects/influences of articles, journals, authors, and keywords on our topic of interest.

26.2.4 Phase IV: Citation Network Analysis

Following the scientometric analysis, the research domain is classified using citation network analysis (CNA). CNA features include the ability to categorize research areas and aid in research growth or paradigm shifts. The references and citations of the collected articles are first investigated in this research, and then binary matrices are created. Citation networks in the form of clusters are created using these matrices. A modularity-based clustering strategy is used [27].

26.3 Results & Discussion

This section discusses the outcomes of descriptive analysis, scientometric analysis, and citation network analysis. The discussion of electric vehicles, charging stations, and newly developing fields follows. Finally, a number of significant study findings are presented.

Table 26.1 Main information used in research

Main information about the data	Results
Timespan	2008:2022
Sources (Journals)	155
Documents	957
Average Years from Publication	2.66
Average Citations Per Documents	16.68
Average Citations Per Year Per Doc	3.48
References	39,615
Keywords Plus (Id)	4229
Author's Keywords (De)	2735
Authors Of Single-Authored Documents	29
Authors Of Multi-Authored Documents	2723
Documents Per Author	0.35
Authors Per Document	2.88
Co-Authors Per Documents	3.98
Collaboration Index	2.93

26.3.1 Descriptive Analysis

The descriptive analysis uses quantitative data to highlight the primary contributions of publications, authors, journals, keywords, nations, and institutions, which aids in identifying patterns and trends in the field of EV and EVCS research. The explanation of it emphasizes the primary contribution in various fields.

26.3.1.1 Main Information About the Data

The information was gathered between 2008 and 2022 from the databases of Scopus and Web of Science. The several research field tags utilized for research analysis, include time period, sources, keywords, authors, documents, and references biblioshiny's analysis of 957 articles yielded the noteworthy findings displayed in Table 26.1 after taking into account 957 publications. The table's collaboration index, which shows the expansion of research in diverse fields, is 2.93, which is regarded as an excellent number in the research area. However, the author to document ratios of 2.88 and 3.98, respectively, also represent the growth of research.

26.3.1.2 The Distribution of Number of the Articles by years

The standard of publications is a crucial indicator of how scientific fields are evolving in light of trends. The quantity of citations for a given research piece is very significant

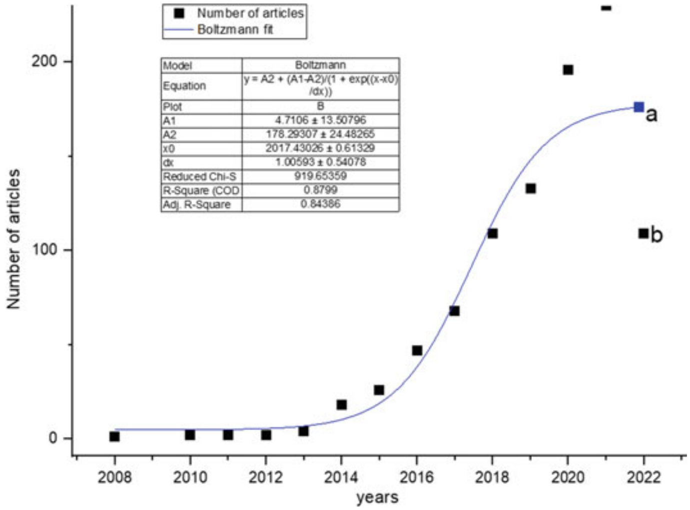


Fig. 26.2 Annual publication

because it indicates the excellence of the research. The study trends were recorded using the Boltzmann fit equation in this case. Figure 26.2's output for the Boltzmann equation shows that y stands for the prediction variable and x for the response variable. The Boltzmann fit line is represented by the little letter a, colored blue. For 2008 through 2022, the small letter b, colored black, reflects the number of papers published that year. The summary of the distribution of articles by year created using the biblioshiny function is given in Table 26.2, which provides evidence for the number of times an article is cited annually, the mean average number of times an item is cited annually, and the citable years. Considering the Boltzmann equation's nonlinear fit, it is discovered that R2 = 0.8799 and modified R-square = 0.84386. The anticipated value for 2022, estimated at 218, demonstrates that there will be an increase in future research articles on EV and EVCS areas.

26.3.1.3 Distribution of the Number of Articles by Journals

The nu.mber of studies published in various journals from 2008 to 2022 is examined using the Scopus and WoS databases, and a summary of the distribution of articles by the top 20 sources is shown in Table 26.2. The number of documents, citations, h index, and citation per document are used to analyze sources. The citation distribution yearly aids in assessing the quality of research in a given year, as shown in Table 26.2. According to Table 26.2, "Energies" took first place with 303 publications out of 957 articles with 3030 citations and a h index of 27, indicating the quality of research articles in the field of EV and EVCS. The second position is for sustainability, and the third position is for sustainability (Switzerland) with 81 and 47 articles, respectively.

Table 26.2 A summary of the distribution of the number of articles by years

Year	N	MeanTCperArt	MeanTCperYear	CitableYears
2008	1	107.00	7.64	14
2009	0	0.00	0.00	0
2010	2	33.50	2.79	12
2011	2	3.50	0.32	11
2012	2	92.50	9.25	10
2013	4	176.75	19.64	9
2014	18	68.11	8.51	8
2015	26	33.65	4.81	7
2016	47	54.11	9.02	6
2017	68	31.18	6.24	5
2018	109	27.50	6.87	4
2019	133	19.83	6.61	3
2020	196	8.89	4.45	2
2021	230	3.00	3.00	1
2022	109	0.47		0

MeanTCperArt- Mean Times Cited per Article, MeanTCperYear- Mean Times Cited per Year

Transportation Research Part D-Transport and Environment take fourth place with 35 articles. The number of articles produced by the sources is used to rank the sources in this case. When we consider other factors, such as citations, the position of sources can change. One of the most intriguing facts about Table 26.2 is that the number of citations/documents for Transportation Research Part A: policy and practice is 50.42 and for applied energy is 43.50, despite the fact that their publication dates begin in 2015 and 2014, respectively. It clearly demonstrates that in scientific research, quality is always important for good citation.

26.3.1.4 Distribution of the Number of Articles by Authors

The biblioshiny function in bibliometrics analysis has been used to evaluate the individual contributions of authors and provide the evaluation's results. Table 26.3 shows the annual distribution of articles by number in the year numbers column and the total number of citations for each year in brackets. Figure 26.3 shows the distribution of articles by author, with dark colors reflecting rising citation counts and light colors denoting declining citation counts. The number of articles in Fig. 26.3 is, however, indicated by the size of the bubble. Table 26.3 shows that Lin Z. is in first place with 14 articles and 760 citations, Zhang Y. is in second place with 12 articles and 154 citations, and Liu C. is in third place with 11 articles (674 citations). The interesting fact in Table 26.3 is just the number of citations for the research articles. Zhang Z only produced 9 articles, but they received fewer citations, while Chen W produced 6 articles, receiving more citations than Zhang Z, indicating that the quality of the research articles matters more than the quantity of articles when it comes to receiving citations. The dark bubble in Fig. 26.3 indicates that Lin Z. and Liu. C. have received more citations, and the other bubbles on the line indicate that there are more publications being published. Liu C. has the most citations per publication, which also reflects the value of his articles.

26.3.1.5 Distribution of the Number of Articles by Countries/regions

The results of the examination of nations' contributions for the years 2008 to 2022 are shown in Table 26.4. The table clearly shows that China holds the top spot in the EV and EVCS domain, contributing 267 articles with 2940 citations and an average citation count of 14.55. The United States, in second place, contributed 246 articles with 4414 total citations, which is a very high number compared to other nations. The input from the nations of origin was extracted into an Excel heat map, which showed the country of origin's geographic origin. Figure 26.4 depicts the geographic head of the institution contributing to the research scope of the current analysis using the extracted data and the "Geographic Heat Map" add-in in Excel. A created geographic heat map that uses color grades from deep red to deep green to show areas of low density (0–50) and high density (250–300) of published research papers is an interactive technique to show each nation's contribution. In the United States, China, and the United Kingdom, the contributing country has a more significant density.

26.3.1.6 Distribution of the Number of Articles by Different Institutes

The role played by several institutes in the field of EV and EVCS is also examined. Figure 26.5 presents the investigation's findings. The contribution made by various institutions is depicted in this image. The top university on the list is "North China Electric Power University, China," and only 4 of the top 5 universities are located in China. However, in terms of EV and EVCS research, Tsinghua University Beijing, China, holds second place, and the University of California, United States, holds third place. Figure 26.5 provides proof that the top spike in the graph indicates the article's contribution to the EV and EVCS research. The quantitative data about the article's % contribution is shown in Table 26.5, which highlights its tremendous impact on this field of study. For example, North China Electric Power University in China contributes 18%, and Tsinghua University in Beijing contributes 17%. However, the contribution of the University of California, USA is 15%. That is why these institutions are at the forefront of EV and EVCS research.

Table 26.3 A summary of distribution of articles by sources (Journals)

Source Title (Journal)	No of article	Citations	H index	Citation per Documents	Yearly contribution (frequency)										
					< 2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
Energies	303	3070	27	10.13	1	1	3	7	23	43	78	129	202	275	303
Sustainability	81	592	13	7.31	0	0	4	4	6	11	19	30	43	71	81
Sustainability (Switzerland)	47	540	13	11.49	0	0	1	1	2	4	8	16	32	42	47
Transportation Research Part D-Transport and Environment	35	1268	16	36.23	0	1	2	2	5	6	15	20	29	34	35
Applied Energy	28	1206	16	43.07	0	0	0	4	7	8	11	14	20	23	28
Energy Reports	21	39	3	1.86	0	0	0	0	0	0	0	0	1	12	21
Energy	20	739	13	36.95	0	0	0	0	1	6	8	10	12	14	20
Energy Policy	16	574	12	35.88	5	1	2	2	3	8	12	12	13	14	16
IEEE Transactions on Intelligent Transportation Systems	16	252	7	15.75	0	0	0	1	2	3	3	4	6	9	9
IET Generation Transmission and Distribution	13	118	4	9.08	0	0	0	0	0	1	4	4	6	10	13

(continued)

Table 26.3 (continued)

Source Title (Journal)	No of article	Citations	H index	Citation per Documents	Yearly contribution (frequency)										
					< 2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
Transportation Research Part C-Emerging Technologies	13	819	11	63.00	0	0	1	1	3	5	9	11	11	13	13
Journal of Modern Power Systems and Clean Energy	12	193	8	16.08	0	0	0	3	3	3	4	5	9	11	12
Transportation Research Part A-Policy and Practice	12	605	8	50.42	0	0	1	2	4	4	4	7	9	10	12
IEEE Transactions on Transportation Electrification	11	217	6	19.73	0	0	0	0	0	1	2	6	6	9	11
Journal of Advanced Transportation	11	67	4	6.09	0	0	0	0	0	3	5	6	10	11	11
IET Electrical Systems in Transportation	10	56	4	5.60	0	0	0	0	0	0	1	2	5	9	10

(continued)

Table 26.3 (continued)

Source Title (Journal)	No of article	Citations	H index	Citation per Documents	Yearly contribution (frequency)										
					< 2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
Renewable and Sustainable Energy Reviews	10	401	8	40.10	0	0	0	0	1	1	5	7	8	8	10
Environmental Research Letters	9	246	6	27.33	0	0	0	0	1	1	3	5	8	9	9
IEEE Transactions on Vehicular Technology	9	288	7	32.00	0	1	2	2	2	2	4	6	8	9	9

Table 26.4 A summary of author's contributions

Authors	DF	TA	SA	MAA	FAA	RBA	RBDF
Zhang Y	0.29	17	0	17	5	1	4
Zhang X	0.35	14	0	14	5	2	3
Liu C	0.07	14	0	14	1	2	10
Li Y	0.25	12	0	12	3	4	5
Li X	0.5	10	0	10	5	5	1
Wang Y	0.4	10	0	10	4	5	2
Wang J	0.11	10	1	9	1	5	6
Chen Z	0.10	10	0	10	1	5	7
Huang X	0.10	10	0	10	1	5	7
Li J	0.10	10	0	10	1	5	7

DF = Dominance Factor; TA—Total Articles; SA—Single Author; MAA—Multi Authors Articles; FAA—First Author Articles; RBA—Rank by Articles; RBDF—Rank by Dominance Factor

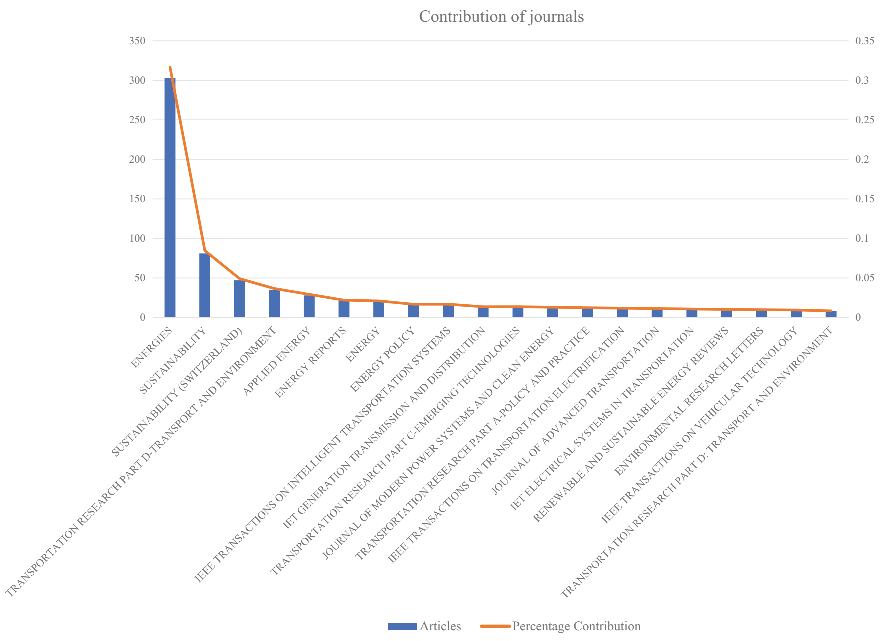


Fig. 26.3 Journals contribution

26.3.2 *Scientometric Analysis*

To identify the most influencing author, document, countries, organizations, and sources scientometric analysis was carried out on the 957 articles from the period

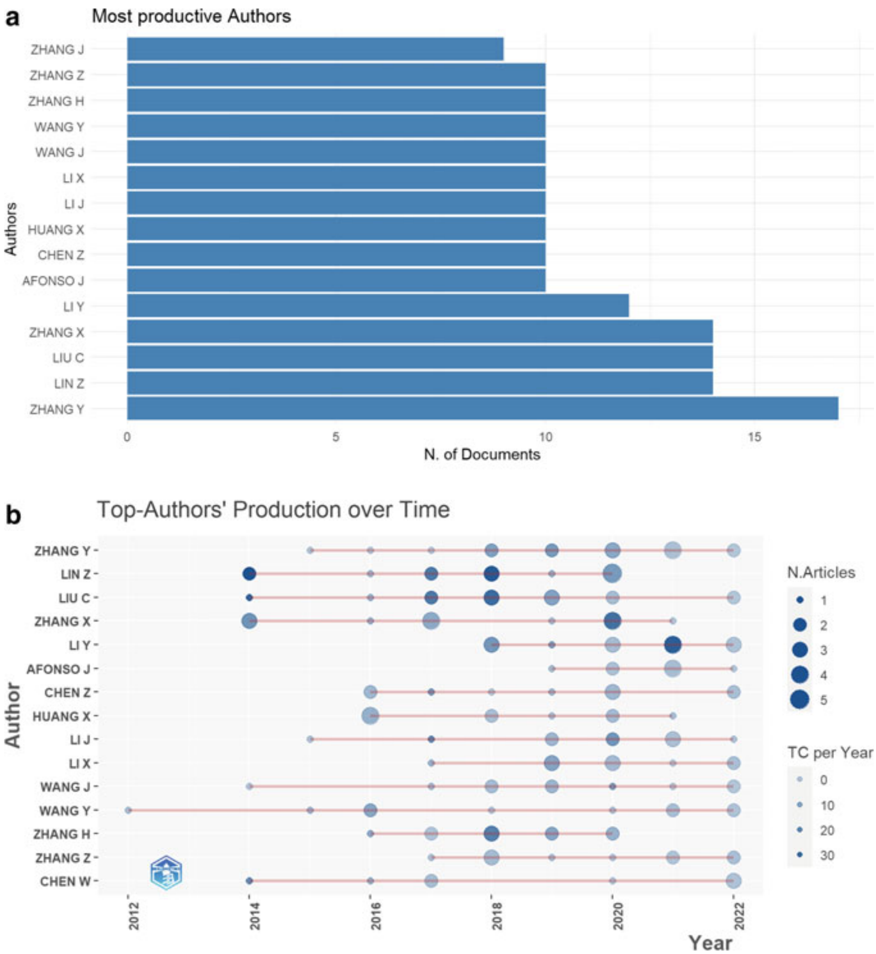


Fig. 26.4 **a** Most productive authors. **b** Top authors' production over time

2008–2022 using VOSviewer 1.6.17 version software and biblioshiny function in bibliometrics software. The investigation of findings shows the impact of various attributes in the field of EV and EVCS research area.

26.3.2.1 Source Impacts

The source impact was performed with the VOSviewer 1.6.17 software version. The minimum number of source documents and citations per source is kept at five. 33 of the 153 sources meet the criteria. The layout is generated using fractionalization in this normalization method. There are 6 clusters of 24 items. Energies, IEEE Transaction on Intelligent Transportation System, IEEE Transaction on Vehicular

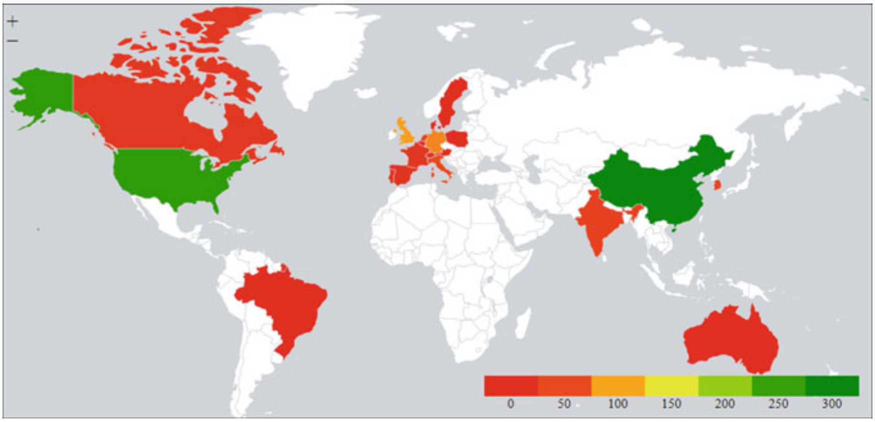


Fig. 26.5 Geographical heat map for country contributions

Table 26.5 Number of articles by countries

Country	Articles	TC	Avg. article citation
China	267	2940	14.55
United States	246	4414	33.19
United Kingdom	95	1135	25.8
Germany	83	1251	22.34
The Netherlands	51	634	18.11
Italy	49	368	9.65
South Korea	46	386	9.65
India	42	266	13.3
Canada	34	290	18.12
France	30	266	14
Spain	30	329	14.95
Portugal	29	59	3.28
Australia	23	174	14.5
Sweden	22	277	19.79
Denmark	21	228	15.2
Austria	20	187	13.36
Switzerland	20	186	14.31
Poland	19	132	6.95
Belgium	18	164	12.62
Brazil	17	59	6.56

Technology, International Journal of Sustainable Transportation, Journal of Modern Power System and Clean Energies, Sustainability, and Sustainability are among the seven items highlighted in red (Switzerland).

Cluster 2 is highlighted in dark green and contains 5 sources, Cluster 3 is highlighted in light blue and contains 4 sources, Cluster 4 is highlighted in yellow and contains 3 sources, Cluster 5 is highlighted in pink and contains 3 sources, and Cluster 6 is highlighted in faint green and contains 2 sources as applied energy and environmental research letter. The size of the bubble indicated the source citation. Because citations are weighted in this case, the larger the bubble, the more citations there are. Figure 26.5 also indicates that the size of the bubble is larger for energies, which takes first place with 3070 citations among all sources, and second place goes to transport research part d with 1268 citations. Figure 26.6 illustrates the evolution of new sources in the field of EV and EVCS research. Figure 26.6's critical observation reveals new sources of publishing the research article in EV and EVCs research, which is highlighted in yellow. Table 26.5 shows a summary of the impact of various sources using the biblioshiny function in bibliometrics software to generate the h index, g index, m index, total citations (TC), number production (NP), and publication year start (PY Start). The h, m, and g index-es are used to compare the quality of field research. The higher the index, the higher the research quality. Table 26.5 shows that energy sources have a higher h, m, g index, TC, and NP, proving their first place in the EV and EVCS research domain. Energies, for example, has an h index of 27, an m index of 2.45, and a g index of 39, ranking first among all sources with 3070 total citations and 256 publications. Transportation Research Part D: Transport and Environment, on the other hand, has an h index of 16, a g index of 32, and an m index of 1.6, with 1268 total citations and 32 publications. Both journals' publication years begin in 2012 and 2013, respectively.

26.3.2.2 Authors Impact

We first set the minimum number of articles per author to 1 and the minimum number of citations per author to 15 in VOSviewer 1.6.17. As a threshold, it filters 254 out of 902 authors. In this case, the fractionalization normalization method generates the layout. Filtrations revealed that 56 authors out of 902 consider by software and formed 9 clusters. In Cluster 1, chen t, dong j, eptimiuou d, marmaras c, morton c, namdeo a, pagany r, tu w, xi x, yi z, and zhao x are among the eleven authors with strong collaboration between them which is highlighted in red. Funke s has the most link strength (12) in the group, with 11 connections in Cluster 2 which is highlighted in green color having eight authors.

In comparison, Cluster 3 is highlighted in blue and has seven authors, Cluster 4 is highlighted in yellow and has six authors, Cluster 5 is highlighted in dark pink and has five authors, Cluster 6 is highlighted in light blue and has five authors, and Clusters 7 also have five authors. However, Clusters 8, and 9 all have four authors each. Still, their cluster colors are different, as shown in Fig. 26.7. The number of authors cited is indicated by the size of the node in the visualization network, and the

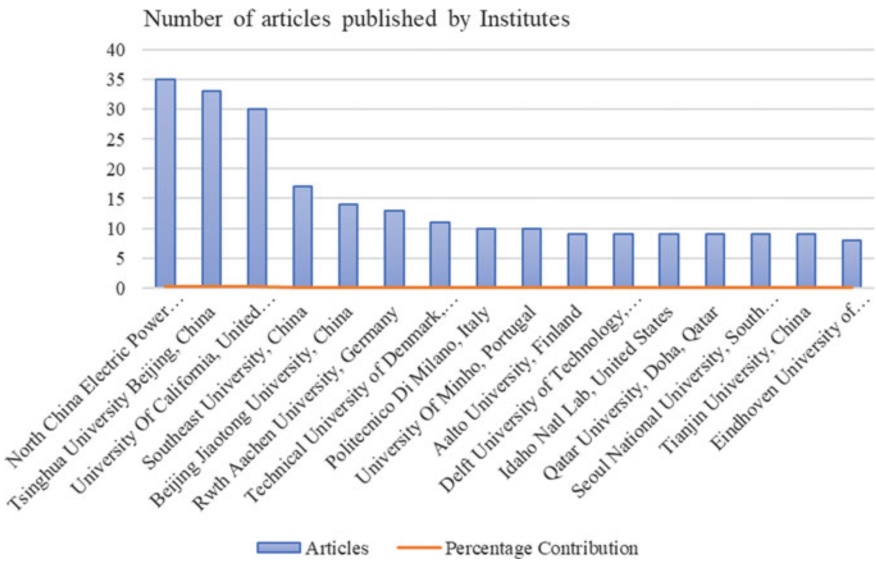


Fig. 26.6 Number of articles published by institutions

link displays the strength of the network between the nodes. The greater the power, the shorter the distance between two nodes. In Fig. 26.8, new authors in the field of EV and EVCS are highlighted in yellow. Table 26.6 summarizes the quantitative explanations for the research field's exploration. Dong j and Chen t have the most citations, with 292 and 278, respectively, according to the table observation in the citation column. Dong j, Chen t, Mersky a, and Hardman s have the most citations in this area, and they are also the most productive and active authors in the EV and EVCS.

26.3.2.3 Documents Impact

We first set the minimum number of citations to 20 in VOSviewer. It filters 244 documents out of 957 papers that meet the criteria. The total cluster generated by VOSviwer records is 9. The number 9 has 43 items with 58 links, with the most citations going to Dong j et al. Figure 26.6 depicts the most influential documents or articles in EV and EVCS research in terms of font size and highlighted text. Figure 26.7 depicts the progression of the EV and EVCS research domains, with bi z and song l, dek r, mic, and keolei as new emerging authors. Figure 26.8 depicts an emerging document in the EV research domain (Table 26.9).

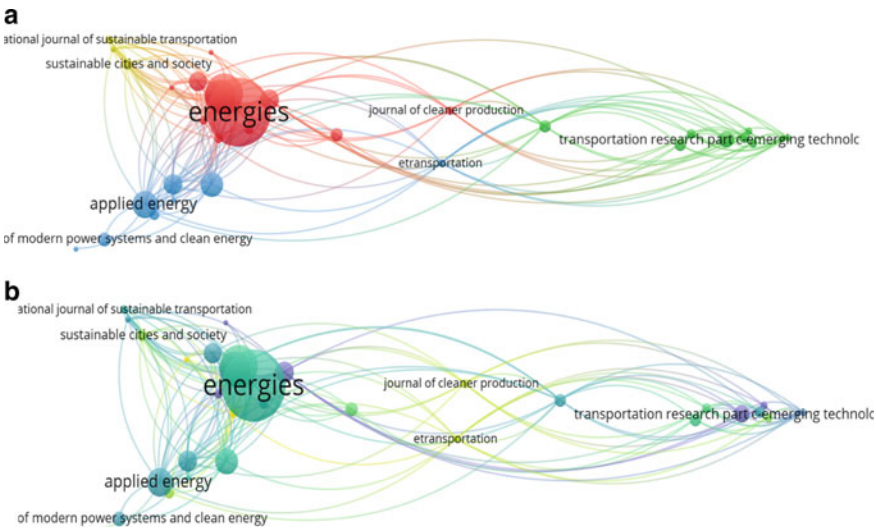


Fig. 26.7 **a** Visualization of sources (Network). **b** Visualization of journal sources (Overlay)

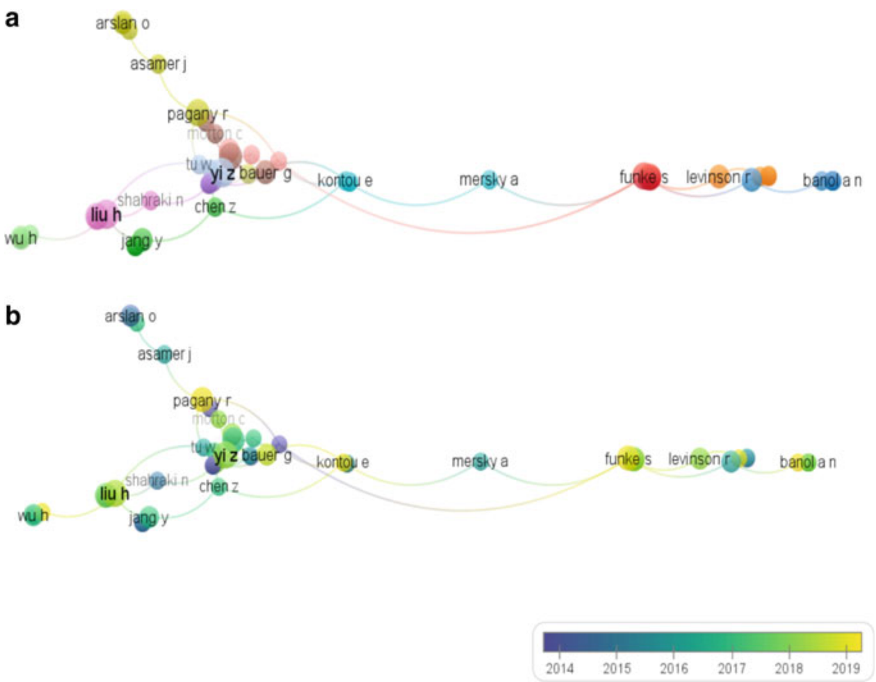


Fig. 26.8 **a** Visualization network of authors. **b** Overlay visualization of authors

Table 26.6 Contribution of different institutes in percentage

Institutes	Articles	Contribution in Percentage (%)
North China Electric Power University, China	35	18
Tsinghua University Beijing, China	33	17
University Of California, United States	30	15
Southeast University, China	17	9
Beijing Jiaotong University, China	14	7
Rwth Aachen University, Germany	13	7
Technical University of Denmark, Denmark	11	6
Politecnico Di Milano, Italy	10	5
University Of Minho, Portugal	10	5
Aalto University, Finland	9	5
Delft University of Technology, Netherland	9	5
Idaho Natl Lab, United States	9	5
Qatar University, Doha, Qatar	9	5
Seoul National University, South Korea	9	5
Tianjin University, China	9	5
Eindhoven University of Technology, Netherland	8	4

26.3.2.4 Keywords Impact

We first set the minimum number of keyword occurrences to 20 in VOSviewer. It filters 40 keywords out of 2758 that meet the criteria. Figure 26.8 portrays the EV and EVCS research network with various research areas. The majority of the investigation focused on electric vehicle concepts and charging station infrastructure. However, the type of charger and charging system, as well as energy storage, are emerging research using alternative energy sources. Overlay network Fig. 26.9 clearly shows that smart charging and energy storage are new areas of research in EV and EVCS (Table 26.10).

26.3.2.5 Country Impact

Using VOSviewer, we first set the minimum number of country documents to be five, and the minimum citation is 5. It filters 46 countries out of 253 countries meeting the thresholds. Figure 26.10 shows a network of countries working on EV and EVCS research areas. China, the United States, and the United Kingdom are the most influential countries. It is evident from Fig. 26.11 that the United States and China

Table 26.7 A summary of top sources' impact

Sources	h index	g index	m index	TC	NP	PY start
Energies	27	39	2.45	3070	256	2012
Transportation research part D-transport and environment	16	32	1.60	1268	32	2013
Applied energy	16	24	2.00	1206	24	2015
Transportation research part C-emerging technologies	11	13	1.22	819	13	2014
Energy	13	17	1.86	739	17	2016
Transportation research part a-policy and practice	8	10	0.89	605	10	2014
Sustainability	13	21	1.44	592	63	2014
Energy policy	12	15	0.80	574	15	2008
Sustainability (Switzerland)	13	22	1.44	540	40	2014
Renewable and sustainable energy reviews	8	10	1.14	401	10	2016
IEEE transaction on vehicular technology	7	8	0.70	288	8	2013
IEEE transaction on power delivery	1	1	0.10	258	1	2013
IEEE transaction on intelligent transportation system	7	10		252	10	
Environmental research letters	6	8	0.86	246	8	2016
Journal of power sources	5	5	0.56	230	5	2014
IEEE transactions on transportation electrification	6	8	1.00	217	8	2017
Journal of modern power system and clean energy	8	11	1.00	193	11	2015
International journal of sustainable transportation	4	4	0.67	181	4	2017
Electric power system research	5	6	0.63	176	6	2015

began the research on EV and EVCS. India, Denmark, Belgium, and Poland are new emerging countries in this research domain (Tables 26.7 and 26.8).

26.3.3 Citation Network Analysis/Thematic Analysis

Thematic analysis is performed on 957 articles from 2008 to 2022 using the biblioshiny function in RStudio's bibliometric analysis. It aids in detecting the themes of the research areas for a given time period and separating the fields in the thematic map based on the centrality and density measures. X-axis represents the centrality and y-axis represent density. Callon centrality expresses the role of a research topic's relevance in the domain, whereas callon density expresses topic development. Figure 26.1

Table 26.8 A summary exploring the impacts of the authors

Author	Documents	Citations	Total link strength	Normalized citations	Average publication year	Average citation	Average normalize citations
Aarslan O	2	62	1	1.43	2014	31	4.87
Levinson R	2	40	1	1.36	2018	20	0.68
Pagany R	2	34	5	2.49	2019	52	2.49
Dong J	1	292	6	4.08	2014	292	4.08
Chen T	1	278	2	4.87	2015	278	4.87
Mersky A	1	201	1	3.52	216	201	3.52
Hardman S	1	190	1	6.47	2018	190	6.47
Cai H	1	168	1	2.35	2014	168	2.35
Shahraki N	1	159	2	4.59	2015	159	4.59
Tu W	1	149	7	2.61	2016	149	2.61
Lieven T	1	115	0	3.32	2015	115	3.32
Bi Z	1	91	1	2.63	2015	91	2.63
Chen Z	1	86	4	2.62	2017	86	2.62
Liu H	1	47	4	2.25	2019	47	2.25

Note Citations, total link strength, normalized citation, average publication year, average citation, and average normalized citation retrieved from the VOSviwer software

Table 26.9 A summary showing the article with the highest number of citations

Document	Citations	Norm Citations
dong j; et al. (2014)	292	4.08
chen t; et al. (2016)	278	4.87
mersky a; et al. (2016)	201	3.52
hardman s; et al. (2018)	190	6.47
xi x; et al. (2013)	181	0.98
cai h; et al. (2014)	168	2.35
shahraki n; et al. (2015)	159	4.59
tu w; et al. (2016)	149	2.61
gnann t; et al. (2018)	134	4.56
bi z; et al. (2015)	91	2.63

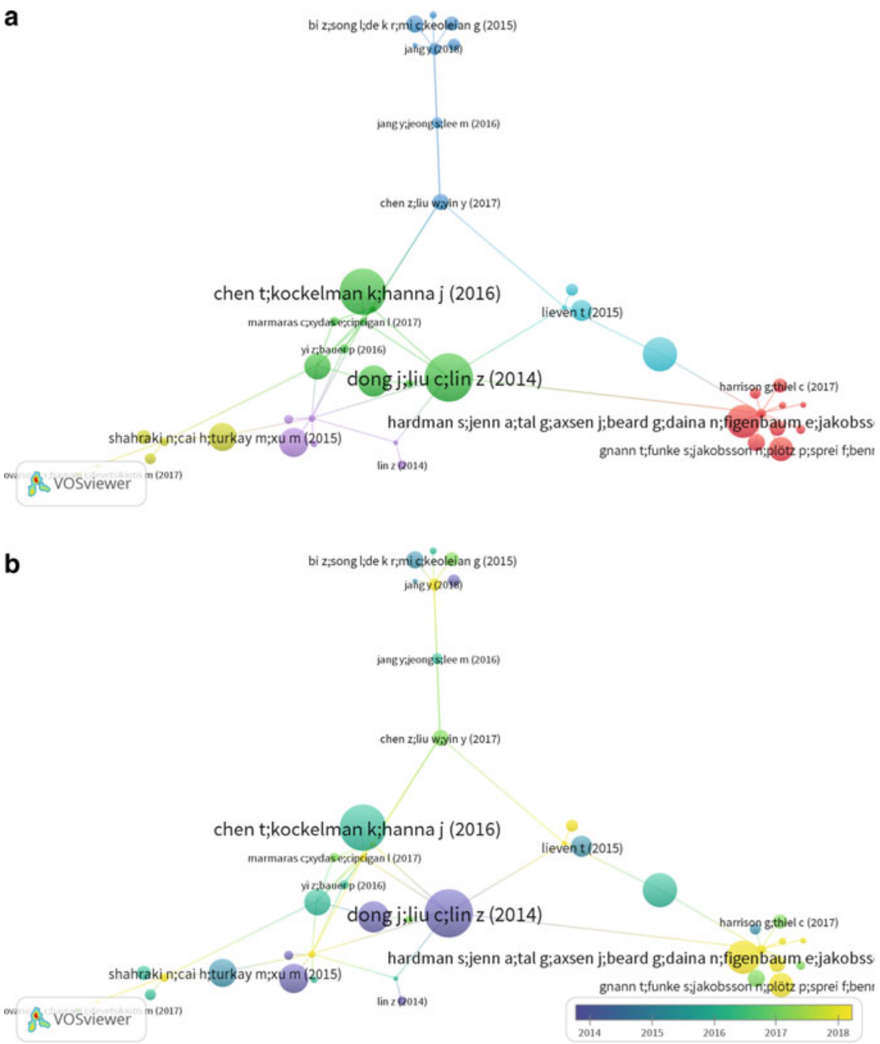


Fig. 26.9 **a** Document visualization (network). **b.** Document visualization (Overlay)

depicts a thematic map with four research theme categories: motor, niche, basic, and emerging or declining. In each theme, the color of the bubbles indicates the clusters and the size of the bubbles indicates the occurrence of keywords. Figure 26.1 shows the motor theme in the first (I) quadrant, which has a higher value of centrality and density and is well-developed and relevant for structuring the conceptual framework of the domains. The basic theme, significant for the domain and cross-cutting to its different areas, is shown in the second (II) quadrant, which has a higher value of centrality and a lower value of density. The niche theme is defined by the third quadrant (III), which has a lower value of centrality and a higher value of density, in

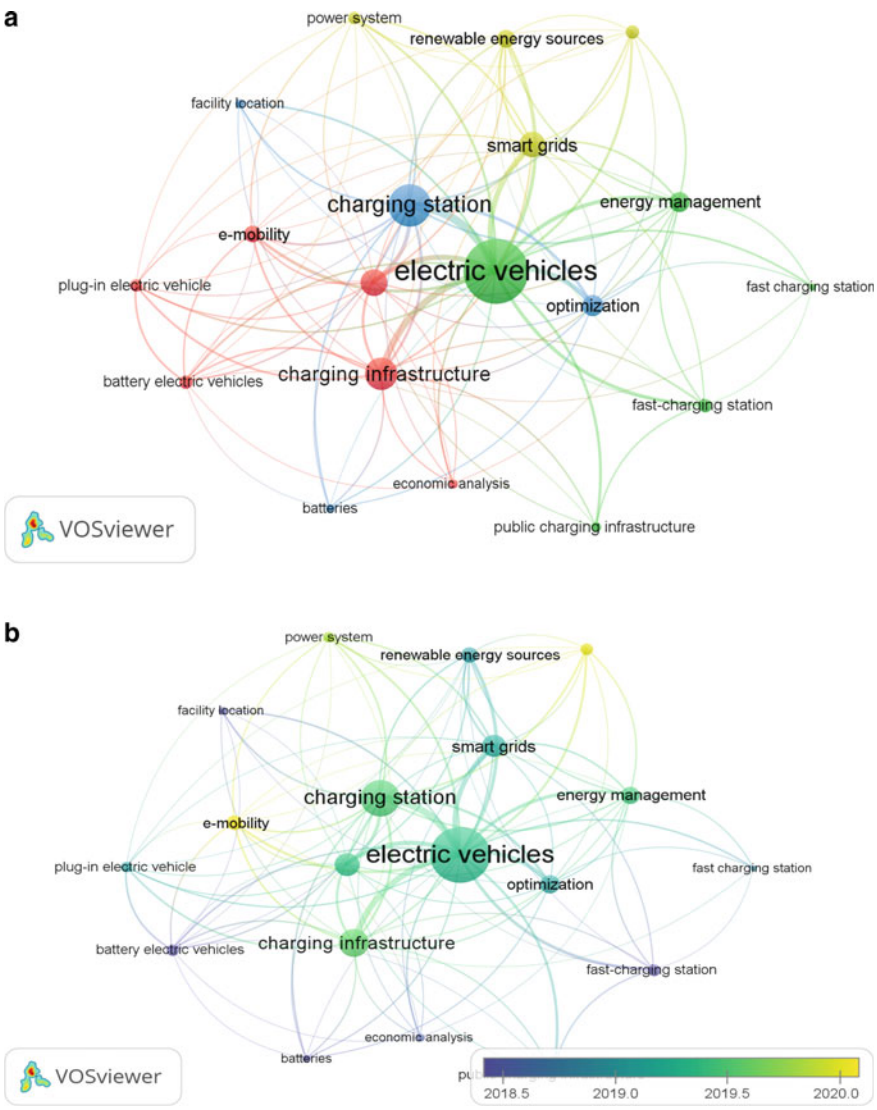


Fig. 26.10 **a** Keywords visualization (Network). **b** Keywords visualization (Overlay)

which the topics are well-developed but still marginal for the domain under investigation. The fourth (IV) quadrant has a lower centrality and density value, indicating an emerging or declining theme that is undeveloped and only marginally interesting for the investigation.

From 2008 to 2022, thematic evaluation (Fig. 26.2) was tried to carry out in RStudio using the "keyword plus" function with the biblioshiny function, because actual article production on EV and EVCS research began in 2012. The thematic

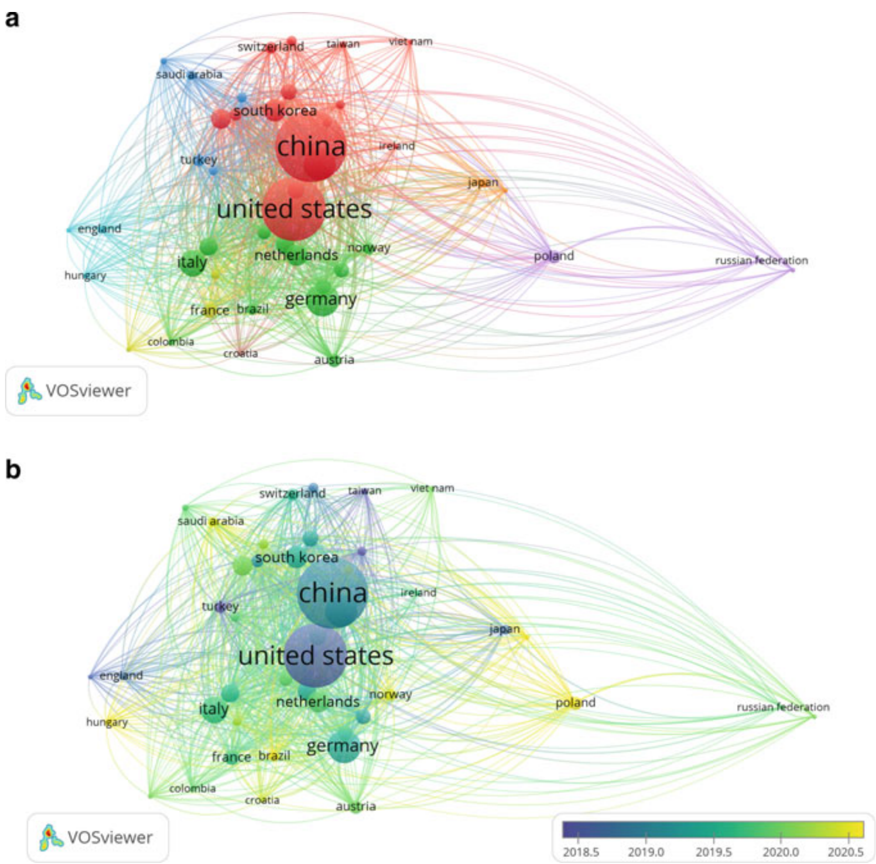


Fig. 26.11 **a** Country visualization (Network). **b** Country visualization (Overlay)

map depicts the research topic and patterns established by existing researchers using clustering techniques to extract the different topics of the specific domain based on the keyword network. The years 2012, 2014, 2016, 2018, and 2020 are chosen as five slicer-cutting selection based on the timeline of the published paper. Figure 26.13 depicts the complete thematic evaluation (alluvial diagram), which indicates the flow of the research topic and pattern for the selected period, and the individual thematic map, which depicts the detailing of various research theme evolution in Fig. 26.13a, b, c, d, e, respectively. Keyword such as “transportation infrastructure” and “Optimization” have been introduced in the beginning of the time phase 2008–2012 which serve as a unique theme. Figure 26.13a depicts the thematic map for time slice 2008–2012, which only contains emerging and motor themes. There is no distinct niche and basic theme. The size of the bubble in the emerging theme indicates that transportation infrastructure is the emerging theme at the beginning of the EV and EVCS domain. The motor theme, on the other hand, shows a higher

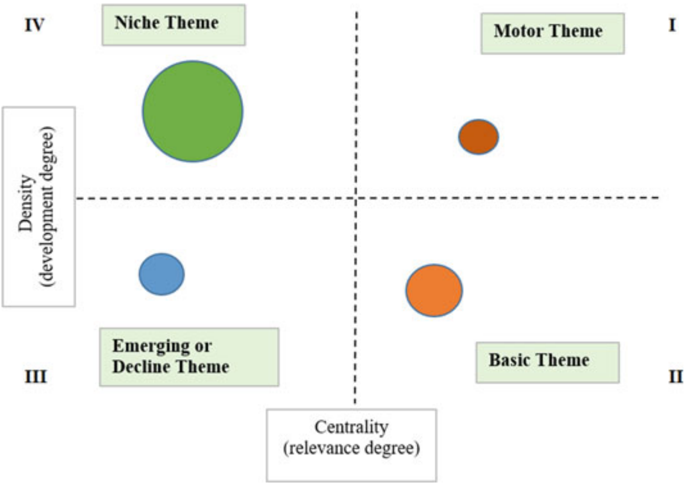


Fig. 26.12 Thematic map

value of centrality and density, as evidenced by the size of the bubble indicator and the red color, which also indicates the cluster of the thematic map. The shift in theme from 2008–2012 to 2013–2014 clearly demonstrates that the subject matter has changed in the new time slice 2013–2014 depicted in Fig. 26.13b. During this phase, transportation infrastructure was transformed into alternative fuel charging infrastructure, and batteries were transformed into charging(batteries), and charging infrastructure, as well as two new topics of EV and EVCS such as charging station, investment added to the thematic map. This reflects the fact that EV and EVCS research is rapidly expanding since all these topics are lies in the motor theme. The measurement of the callon centrality, callon density, centrality rank, density rank, and cluster are listed in Table 26.11.

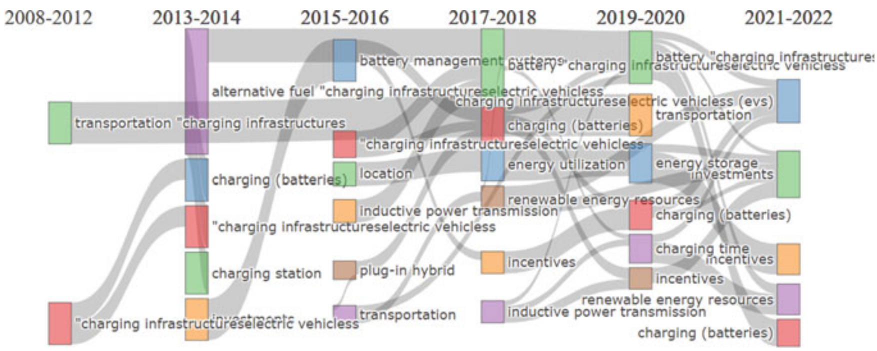


Fig. 26.13 Thematic evolution (2008–2022)

Table 26.10 A summary of influencing keywords

Keyword	Occurrences	Total link strength	Average publication year	Average citation	Average normalize citation
Electric vehicles	271	94	2019	18.03	1.02
Charging infrastructure	71	44	2019	21.65	1.07
Optimization	42	22	2019	20.95	1.13
Smart charging	32	23	2020	10.66	1.35
Electric mobility	22	13	2020	8.32	0.56
Renewable energy	22	15	2018	19.68	0.91
Smart grid	22	13	2020	21.05	0.72
Electric vehicle charging station	21	9	2020	13.95	1.08
Battery electric vehicles	20	7	2018	22.4	1.22
Fast charging	19	15	2017	32.79	1.22
Transportation	19	10	2020	18.89	0.91
Sustainability	18	19	2019	11.44	0.89
Energy storage	15	13	2019	26.93	1.31
Wireless charging	15	3	2018	23.33	0.79

Only the investment topic flows from 2014 to 2015 in the time phase 2015–2016, as do new research topics such as plug-in hybrid, battery management system, inductive power transmission, location, and transportation. The size of the bubbles represents the maturity, stages, or level of the research field. According to Fig. 26.13c, plug-in hybrid electric vehicle research is an emerging field, charging infrastructure and battery management are basic themes, location, and inductive power transmission are niche themes, and transportation is a motor theme.

The flow of research topics from 2016 to 2017 is depicted in an alluvial diagram (Fig. 26.13), and the research topics are investment, battery charging infrastructure vehicles, charging batteries, inductive power transmission, renewable energy resources, and energy utilization in Fig. 26.13d. In this phase, incentive and battery infrastructure are emerging themes, and their small bubble sizes show a clear picture of the new theme for the specified period. The basic theme is charging batteries, but the size of the bubble indicates the scope of the research. However, inductive power transmission is a niche theme, and the size of the bubble indicates that it requires more investigation for further development of the research field.

The research topic for 2017–2018 is time change, energy storage, battery charging electric vehicle infrastructure, incentive, transportation, and battery charging. The flow of the research theme for this period is shown in Fig. 26.13e, where charging time is extremely important in the battery management system of an electric vehicle

Table 26.11 Summary of countries/regions having higher publication impacts based on average normalized citations of articles

Country	Documents	Citations	Link strength	Norm citations	Avg. publication per year	average citation	Avg norm citation	Avg norm citation	Avg citation per year
China	202	3115	29	165.91	2018	33.19	0.82	0.82	15.42
United States	151	2083	73	190.5	2018	14.55	1.26	1.26	13.79
Germany	76	1706	84	90.87	2019	25.80	1.20	1.2	22.45
Switzerland	18	932	5	19.87	2019	14.31	1.10	1.1	51.78
Netherlands	43	882	30	58.28	2019	18.11	1.36	1.36	20.51
Australia	27	739	19	28.3	2019	21.75	1.05	1.05	27.37
Canada	33	702	12	36.41	2019	18.12	1.10	1.1	21.27
Italy	56	652	10	57.58	2019	9.65	1.03	1.03	11.64
Poland	22	589	2	29.25	2020	14.50	1.33	1.33	26.77
Denmark	29	557	28	40.03	2019	13.36	1.38	1.38	19.21
France	26	516	14	49.74	2019	19.79	1.91	1.91	19.85
United Kingdom	74	512	29	121.43	2019	22.34	1.64	1.64	6.92
Spain	34	487	4	24.68	2019	14.00	0.73	0.73	14.32
Sweden	25	466	43	43.8	2019	15.20	1.75	1.75	18.64
Hong Kong	15	459	7	17.54	2018	30.60	1.17	1.17	30.60
India	38	446	15	37.89	2020	14.95	1.00	1	11.74
Belgium	21	406	14	21.94	2019	12.62	1.04	1.04	19.33
Norway	18	341	29	22.66	2020	18.94	1.26	1.26	18.94
Qatar	12	324	5	11.34	2019	12.92	0.95	0.95	27.00

Table 26.12 A Summary of the clusters in the thematic analysis

Cluster 1	Collon centrality	Collon density	Rank centrality	Rank density	Cluster color
Charging infrastructure electric vehicles	50	1115.38	3.5	3	Red
Batteries	0	456.25	1.5	1	Blue
Transportation charging infrastructures	0	711.54	1.5	2	Green
Cluster 2					
Charging infrastructure vehicles	12.54	832.36	8	5	Red
Charging (Batteries)	14.23	1454.68	9	9	Blue
Charging station	9.81	1355	6‘	7	Green
Alternative fuel charging infrastructure vehicle	11.25	1436.11	7	8	Pink
Investments	22.95	1602.94	10	10	Orange
Cluster 3					
Charging infrastructure vehicles	26.77	239.29	6	2	Red
Battery management systems	16.88	333.23	4	4	Blue
Location	14.50	430.55	2	5	Green
Transportation	24.76	474.26	5	6	Pink
Inductive power transmission	15.36	415.50	3	4	Orange
Plug-In-Hybrid	1.16	169.44	1	1	Light Grey
Cluster 4					
Charging (Batteries)	23.36	77.98	6	3	Red
Energy utilization	18.42	152.04	5	5	Blue
Battery charging infrastructure electric vehicle	9.90	71.01	3	2	Green
Inductive power transmission	4.84	91.86	2	4	Pink
Inductives	0.41	59.15	1	1	Orange
Renewable energy resources	16.98	190.97	4	6	Light grey
Cluster 5					
Charging (Batteries)	12.04	46.93	6	5	Red
Energy storage	7.75	42.81	4	2	Blue

(continued)

Table 26.12 (continued)

Cluster 1	Collon centrality	Collon density	Rank centrality	Rank density	Cluster color
Battery charging infrastructure electric vehicle	7.92	42.49	5	1	Green
Charging time	4.85	42.95	2	3	Pink
Transportation	6.74	70.93	3	6	Orange
Inductives	1.22	43.15	1	4	Light grey

because maximum research is being done by researchers, industries, and stakeholders of EV and EVCS to reduce charging time. As a result of this, charging time has emerged as an emerging theme for this period. Energy storage and battery charging infrastructure for electric vehicles are also becoming increasingly important for electric vehicle management, and these fields fall under the basic theme, indicating that much more research is required to explore the field of EV and EVCS, which has piqued the interest of researchers and practitioners. Inductive power transmission is a niche theme that attracts research to explore the field, but charging batteries is a motor theme that shows that a lot of research is happening with charging batteries and new materials are coming for battery charging and it is still a challenging and evolving research theme. The size of each theme's bubble represents the expansion of various research topics in the EV and EVCS domains.

Electric vehicles, investments, energy storage, batteries, and renewable energy sources are the evolving fields of research in the EV and EVCS domain for 2021–2022. The stakeholder's investment in EVs and EVCS is critical to driving the electric vehicle market. In addition, to charge the batteries, electric energy is required, which can be obtained from renewable sources such as solar, wind, or other sources. To accelerate the adoption of EVs and EVCS, the government must make a bold decision on incentives for EV and EVCS users and manufacturers. However, it is critical because conventional vehicles remain on the road (Table 26.12).

To fully comprehend the contextual structure of the screen extent literature, biblioshiny must perform a factorial analysis to multiple correspondence analysis (MCA), which results in the classification of distinct clusters, as shown in Fig. 26.14. This analysis also categorized the study's sample, which was then combined with the results of keyword co-occurrence analysis, as shown in Fig. 26.8. Each cluster/group of publications' documents was ready to be analyzed using VOSviewer software's "Bibliographic coupling of the document," which aids in identifying the connection between the documents based on citation and text similarities. Each cluster was named to represent the domain of EV and EVCS research after cross-referencing with research articles identified using the factorial analysis approach (biblioshiny) and the bibliographic coupling (VOSviewer). As stated in the previous paragraph, biblioshiny used a factorial analysis with the auto-generation option to identify different clusters and their conceptual structure. The current study identifies three dominant macrostructures in the proposed field of research based on the observation of bibliographic results obtained using both software. These are the following categories: (1) electric vehicles, (2) charging infrastructures, and (3) energy storage. The associated dendrogram (shown in Fig. 26.14) generated by the results of the factorial analysis (i.e., MCA-Multiple correspondence analysis) techniques in biblioshiny revealed that these identified clusters shared some similarities and overlaps. Figure 26.7 depicts a cluster dendrogram that was generated and optimized with the 48 most commonly used keywords to improve readability due to visual constraints of small fonts and the generation of complicated tree structures when all keywords were included, as shown in Fig. 26.8. The keywords listed in cluster 1 in Fig. 26.7 are "infrastructure," "model," "energy," "design," and "adoption," all of which are related to electric vehicles. In contrast, 24 keywords in cluster 2 (blue color) are associated with charging infrastructure (shown in Fig. 26.7). The remaining keywords are primarily concerned with energy utilization, power distribution, renewable energy, and so on. The topic dendrogram confirms the various dominant clusters discovered by co-occurrence and factorial analysis (Figs. 26.15 and 26.16).

26.4 Discussion

For the analysis of literature, the four-stage review process applied on 957 articles which are taken from the Scopus and Web of Science database from 2008 to 2022(up to May) using bibliometric software such as RStudio and VOSviewer. The results of this comprehensive review in the area of EV and EVCS are listed below:

(a) According to descriptive analysis, (1) there is an increasing trend in EV and EVCS research during 2008–2022.
(b) The top three journals publishing the most articles on EV and EVCS are "Energies," "Sustainability," and "Transport Research Part D". contributing the most to the field. China, the United States, and the United Kingdom have the most influence on EV research. However, new countries participating in EV research include India, Belgium, and Poland.
(c) According to scientometric analysis, new countries are collaborating with dominant countries, and some authors have begun collaborating between the groups.
(d) The thematic map depicts the evolution of new emerging research themes in the field of EV and EVCS all around the world, and all nations now have begun their mission to reduce carbon emissions and achieve a net-zero target. The charging system, charging time, and energy storage are all new fields of study in EV and EVCS.

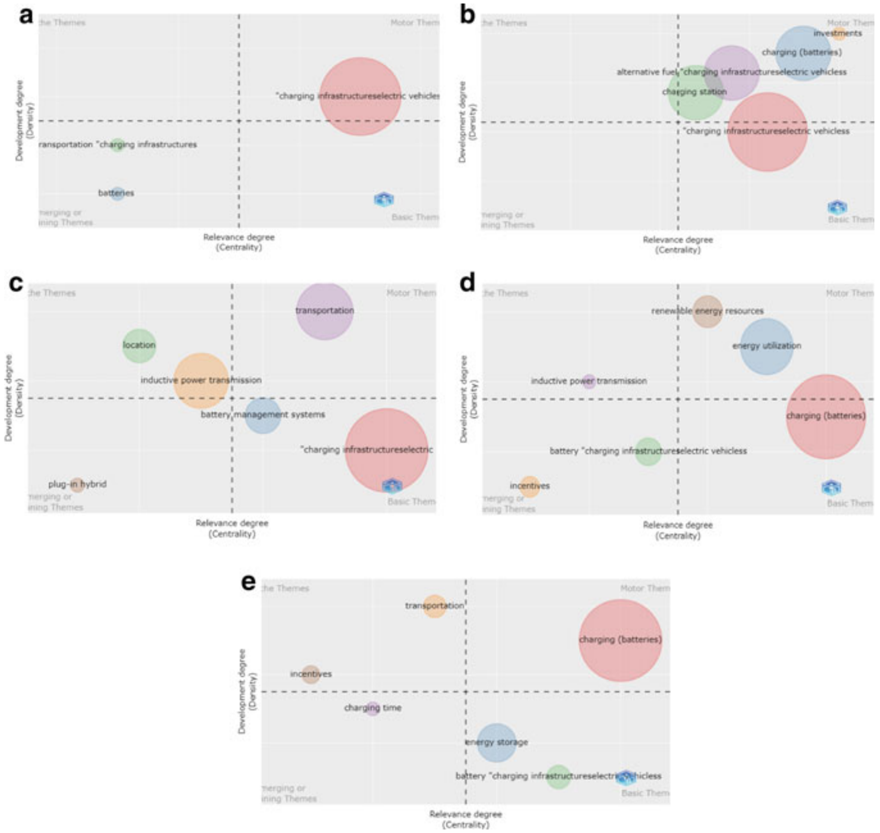


Fig. 26.14 **a** Time Slice I—(2008–2012). **b** Time slice II—(2013–2014). **c** Time Slice III—(2015–2016). **d** Time Slice IV—(2017–2018). **e** Time Slice V—(2019–2020)

26.5 Conclusion

The four-phase review process on EV and EVCS charging infrastructure development is used in this research study, which includes bibliometric search, bibliometric analysis, scientometric analysis, and citation network analysis. From 2008 to 2022, 939 articles were chosen for the bibliometric study. The bibliometric analysis is performed on these articles to investigate trends in authors, documents, sources, keywords, countries, and organizations. Furthermore, a scientific metric analysis is performed to investigate the impact of the EV and EVCS research domains. Finally, the citation network analysis employs R studio and a biblioshiny function to track the growth of emerging research themes in the development of EV and EVCS infrastructure. Overall, China is the leading country in the EV and EVCS domains.

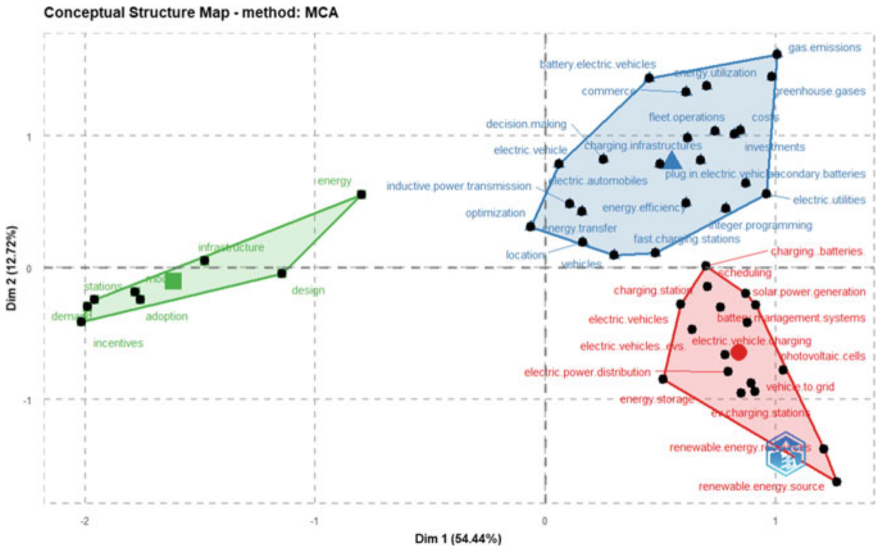


Fig. 26.15 Conceptual structure

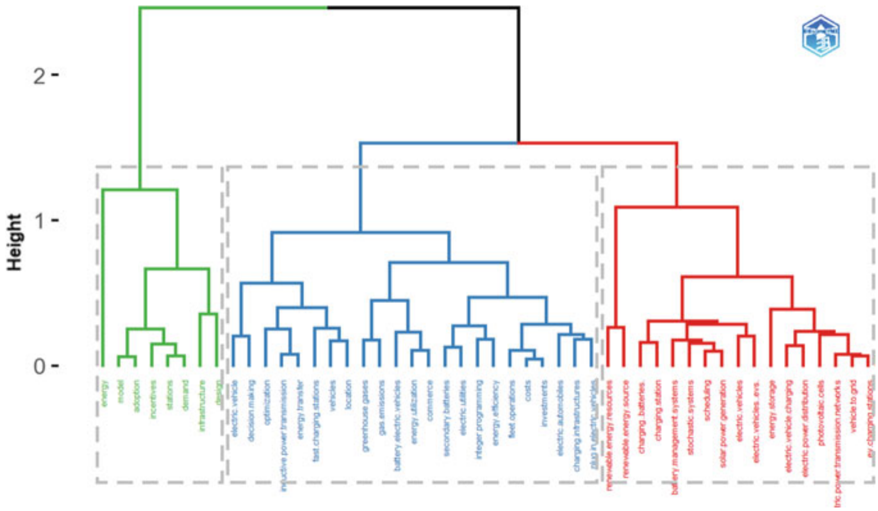


Fig. 26.16 Topic dendrogram

References

1. Zuo X, Xiao Y, You M, Kaku I, Xu Y (2019) A new formulation of the electric vehicle routing problem with time windows considering concave nonlinear charging function. J Clean Prod 236. https://doi.org/10.1016/j.jclepro.2019.117687
2. Bauer GS, Phadke A, Greenblatt JB, Rajagopal D (2019) Electrifying urban ridesourcing fleets at no added cost through efficient use of charging infrastructure. Transp Res Part C Emerg Technol 105:385–404. https://doi.org/10.1016/j.trc.2019.05.041
3. Erickson LE (2017) Reducing greenhouse gas emissions and improving air quality: two global challenges. Environ Prog Sustain Energy 36(4):982–988. https://doi.org/10.1002/ep.12665
4. Deb S, Tammi K, Kalita K, Mahanta P (2018) Review of recent trends in charging infrastructure planning for electric vehicles. Wiley Interdiscip Rev Environ 7(6):e306. https://doi.org/10.1002/wene.306
5. Noshadravan A, Cheah L, Roth R, Freire F, Dias L, Gregory J (2015) Stochastic comparative assessment of life-cycle greenhouse gas emissions from conventional and electric vehicles. Int J Life Cycle Assess 20(6):854–864. https://doi.org/10.1007/s11367-015-0866-y
6. Hoarau Q, Perez Y (2018) Interactions between electric mobility and photovoltaic generation: a review. Renew Sustain Energy Rev 94:510–522. https://doi.org/10.1016/j.rser.2018.06.039
7. Wicki M, Brückmann G, Bernauer T (2022) How to accelerate the uptake of electric cars? Insights from a choice experiment. J Clean Prod 355. https://doi.org/10.1016/j.jclepro.2022.131774
8. Chaudry M et al. (2022) The implications of ambitious decarbonisation of heat and road transport for Britain's net zero carbon energy systems. Appl Energy 305. https://doi.org/10.1016/j.apenergy.2021.117905
9. Ghosh A (2020) Possibilities and challenges for the inclusion of the electric vehicle (EV) to reduce the carbon footprint in the transport sector: a review. Energies 13(10). https://doi.org/10.3390/en13102602
10. Mo T et al. (2022) Commercialization of electric vehicles in Hong Kong. Energies 15(3). https://doi.org/10.3390/en15030942
11. Sommer S, Vance C (2021) Do more chargers mean more electric cars? Environ Res Lett 16(6). https://doi.org/10.1088/1748-9326/ac05f0
12. Alp O, Tan T, Udenio M (2022) Transitioning to sustainable freight transportation by integrating fleet replacement and charging infrastructure decisions. Omega (United Kingdom) 109. https://doi.org/10.1016/j.omega.2022.102595
13. Mersky AC, Sprei F, Samaras C, Qian ZS (2016) Effectiveness of incentives on electric vehicle adoption in Norway. Transp Res Part D Transp Environ 46:56–68. https://doi.org/10.1016/j.trd.2016.03.011
14. Huang K, Kanaroglou P, Zhang X (2016) The design of electric vehicle charging network. Transp Res Part D Transp Environ 49:1–17. https://doi.org/10.1016/j.trd.2016.08.028
15. Taylor M Fisher, Blair Farley K, Gao Y, Bai H, Tse ZTH (2014) Electric vehicle wireless charging technology: a state-of-the-art review of magnetic coupling systems. Wirel Power Transf 1(2): 87–96. https://doi.org/10.1017/wpt.2014.8
16. Harrison G, Thiel C (2017) An exploratory policy analysis of electric vehicle sales competition and sensitivity to infrastructure in Europe. Technol Forecast Soc Change 114:165–178. https://doi.org/10.1016/j.techfore.2016.08.007
17. Kirpes B, Danner P, Basmadjian R, Meer H, Becker C (2019) E-mobility systems architecture: a model-based framework for managing complexity and interoperability. Energy Infor 2(1). https://doi.org/10.1186/s42162-019-0072-4
18. Garruto R, Longo M, Yaïci W, Foiadelli F (2020) Connecting parking facilities to the electric grid: a vehicle-to-grid feasibility study in a railway station's car park. Energies 13(12). https://doi.org/10.3390/en13123083
19. Corinaldesi C, Lettner G, Schwabeneder D, Ajanovic A, Auer H (2020) Impact of different charging strategies for electric vehicles in an austrian office site. Energies 13(22). https://doi.org/10.3390/en13225858

20. IEA, Global EV Outlook. [Online]. Available: http://www.iea.org
21. Zhang X, Rao R, Xie J, Liang Y (2014) The current dilemma and future path of China's electric vehicles. Sustain 6(3):1567–1593. https://doi.org/10.3390/su6031567
22. Dong J, Liu C, Lin Z (2014) Charging infrastructure planning for promoting battery electric vehicles: an activity-based approach using multiday travel data. Transp Res Part C-Emerging Technol 38:44–55. https://doi.org/10.1016/j.trc.2013.11.001
23. Zhang J et al (2019) Design scheme for fast charging station for electric vehicles with distributed photovoltaic power generation. Glob Energy Interc 2(2):150–159. https://doi.org/10.1016/j.gloei.2019.07.003
24. Deb S, Tammi K, Kalita K, Mahanta P (2018) Impact of electric vehicle charging station load on distribution network. Energies 11(1):1–25. https://doi.org/10.3390/en11010178
25. Tu W, Li Q, Fang Z, lung Shaw S, Zhou B, Chang X (2016) Optimizing the locations of electric taxi charging stations: a spatial–temporal demand coverage approach. Transp Res Part C Emerg Technol 65(3688): 172–189. https://doi.org/10.1016/j.trc.2015.10.004
26. Sarkar S, Jhareswar M (2020) Machine learning in occupational accident analysis: a review using science mapping approach with citation network analysis. Saf Sci 131: 2020. https://doi.org/10.1016/j.ssci.2020.104900
27. Ludo W (2010) A unified approach to mapping and clustering of bibliometric networks. J Inf 4(4):629–635. https://doi.org/10.1016/j.joi.2010.07.002

Chapter 27
Contract Price Negotiation Using an AI-Based Chatbot

Divya Ramachandran, Anupam Keshari, and Manoj Kumar Tiwari

Abstract The contract management process is a tedious and often manually performed task in the procurement process in the industry. It includes requesting bids, negotiating, drafting the contract, etc. Within this purview, there is ample opportunity for smart automation. In this paper, we present a general framework for a negotiating bot that can bargain with the vendor for an agreement on the price in the contract negotiation process. Based on the data provided by the vendor, the bot determines a fair price as per market standards, using which the user (procurement staff) can further negotiate for a better deal as per the systematic steps described by the decision bot. This method can benefit the companies by saving a significant amount of time and money for the organization and reducing human dependency on such routine tasks.

27.1 Introduction

Contract management is an integral step in the procurement cycle. It involves requesting quotes from suppliers, compiling the bids, cost negotiation, and coming to an agreement on various KPIs such as quality standards, supplier lead time, fulfillment rate, etc. [1, 2]. This is a protracted process in the industry usually done manually by the procurement staff. It is essential to ensure efficient contract management for a sound supplier relationship, minimized supply cost, standardization, and an optimum trade-off between cost and quality. The long-drawn nature of this process hints at the need for a smart automated system that can bring substantial value addition by performing routine tasks without much human intervention.

D. Ramachandran (✉) · A. Keshari · M. K. Tiwari
NITIE, Mumbai, India
e-mail: divya97.r@gmail.com

A. Keshari
e-mail: anupam.keshari@nitie.ac.in

M. K. Tiwari et al. (eds.), *Applications of Emerging Technologies and AI/ML Algorithms*, Asset Analytics, https://doi.org/10.1007/978-981-99-1019-9_27

27.1.1 Contract Management Process

Contract Management for procurement involves several steps as described in the literature [2–4], i.e., Step-1: Request for a new contract or review of the existing contract, Step-2: Authoring the contract, Step-3: Negotiation and Approval, Step-4: Contract performance and Analysis, Step-5: Amendment and Expiry/Renewal.

Contract management poses challenges due to its time-consuming nature with a non-standardized process and increasing complexities in the regulations.

In this paper, we focus on the negotiation stage of the contract management process (step-3), presenting the automation of price negotiation with the support of an AI-based Chatbot. There are two major issues to deal with. First the functioning of the AI-based Chatbot, Second the activities/steps involved in the negotiation.

27.1.2 AI-Based Chatbot

Artificial intelligence (AI) is used to build a smart machine that can mimic human behavior. It carries out many decision-making tasks autonomously and reduces human intervention [5].

Similarly, a chatbot or chatterbot is a software application that can simulate human conversation. Chatbots can be broadly classified into three categories. 1. *Simple Chatbot*—That is rule-based and task-specific where the bot poses predefined questions, and the user has to choose from the options provided. 2. *Smart Chatbots*—It can have a free-flowing conversation, understand the sentiment, and learn from the interactions. 3. *Hybrid Chatbots*—It is a combination of simple and smart chatbots and is a balanced tool in terms of complexity and usability.

With consumers demanding relentless service, chatbots are setting foot in the industry as many brands are using this chatbot technology to streamline their processes involving basic customer services or enabling product discovery.

The presented chatbot carries out price inquiries from vendors and negotiates based on the recommended rate patterns.

27.1.3 Negotiation Process

While agreeing on the price, contractual terms, and services in the procurement, negotiation is a paramount skill.

BATNA (Best Alternative to Negotiated Agreement) gives the best option a party has if the negotiation fails. It can be determined by identifying all the alternative courses of action and evaluating these alternatives [6].

A strong BATNA ensures a solid stand during negotiation and provides a good backup plan. A stronger BATNA also increases the chances of the opposite party settling for a mutual agreement. Our proposed negotiation process is based on the recommended price of the item. For each vendor, a recommended price is determined using the vendors' record (fill rate, product quality, company's goodwill), and the performance category of the vendor. The calculated recommendation remains the reference price to reach the final deal.

27.2 Literature Survey

While there has been a significant study related to the usage of machine learning for supplier selection and usage of chatbots. This section first surveys the effectiveness of an AI-based chatbot in the procurement process and then discusses the existing natural language processing techniques that are used to create the chatbot algorithm.

27.2.1 Impact of AI in Procurement Using a Chatbot

The usage of AI chatbot is being established as a well-grounded solution for buyer–seller, supplier-manufacturer, etc. Previous studies have explored the effect on the seller's attitude and discrimination while interacting with a chatbot while providing a quotation. Cui et al. [7] were able to run a field experiment on an online platform of a trading company and compare the quotes across the chatbot, and female and male human buyers. The study considered the initial price quote for the experiment as it reflects the buyer's disposition to pay, suppliers could lose customers based on the initial price offered, and it is not affected by bargaining on negotiation.

They concluded that a simple chatbot without smart technology receives a higher quote price than human buyers. However, automation smartness simultaneously delivers the most value.

27.2.2 ML-Based Price Negotiation

Previous research has explored the areas of electronic negotiation [8], and the work done by Carbonneau et al. [8] presented a neural network-based model for predicting the opponent's proposal while negotiating. Various what-if scenarios optimization cases were studied, using which the model could exhibit interesting negotiation strategies.

Another research by Liu and Zheng [9] presented a negotiation bot that gave a fair market price based on used car data and provided suggestions to the negotiator based on the market price. They used a decision tree model that could simulate a step-by-step process to assess the offer given by the seller and negotiate further.

27.2.3 Natural Language Processing: An Interactive Chatbot

In 2019, Flipkart launched its first haggling bot that enabled customers to bargain for better deals during their online sales. Google Dialog Flow is an artificial intelligence software powered by Google AI, which is NLP based and can handle conversations for text-based and voice-based chatbots. Similarly, Haptik is another enterprise conversational AI-based platform. They built a WhatsApp bot during COVID-19 which was used by 21 million users as a government helpline.

In this paper, we would be using Natural Language Processing for creating the bot interface and sentiment analysis. We would be creating our own data dictionary for linking the vendor responses to the bot response.

27.3 Methodology

The paper presents a general framework that can be used in price negotiation for the procurement process, it integrates a decision tree for negotiating under different scenarios with an NLP-based Chatbot.

It is a hybrid chatbot (illustrated in the introduction section), for the balance of complexity in the implementation and usability. The hybrid chatbot in this case performs both price negotiation and interaction based on a decision criterion (Fig. 27.2).

First, the requirement of the product/raw material to be procured is identified. Following this, the vendor's bid data and historical information are collected and various KPIs are calculated. Then, the recommended price is predicted, and the negotiation process starts as per the Decision bot explained in (Fig. 27.2). The negotiation loop continues till the user and the seller reach an agreement or rejection as shown below (Fig. 27.1).

27.3.1 Recommended Cost Determination

The supplier bid information provided by the user is used to predict the recommended cost of the product and the performance category. Based on the historical information of the suppliers, the supplier KPI data [1, 2] is input into the model. We have used the following Key performance indicators of the vendor:

1. On-time Delivery (OTD)
2. Fill Rate
3. Defect Rate
4. Compliance Rate
5. Vendor Availability

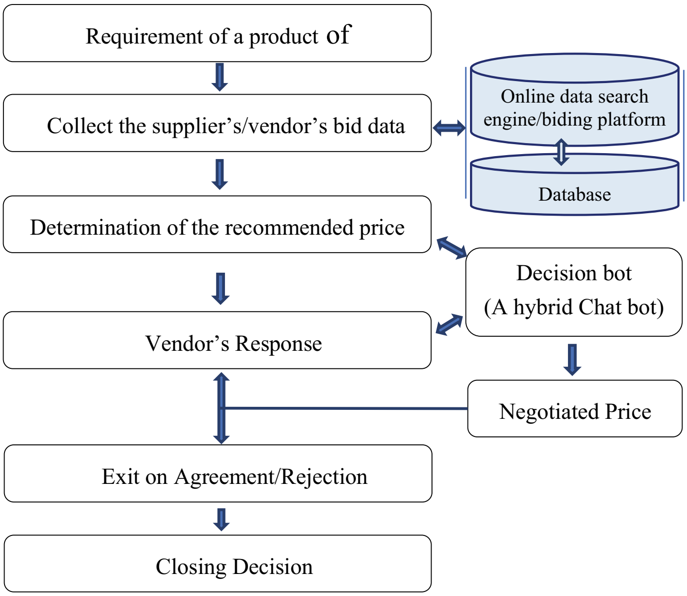


Fig. 27.1 Activities involved in price negotiation: a general framework to implement smart-automated negotiation in procurement

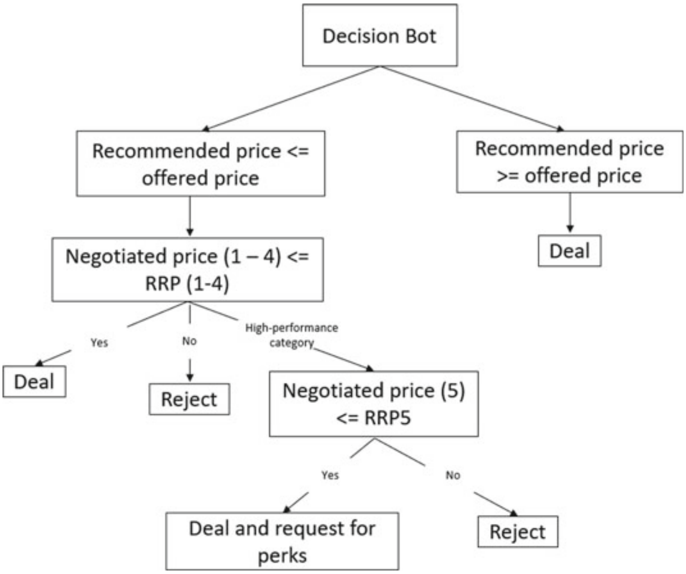


Fig. 27.2 Decision tree methodology for deal selection or rejection

Using the above information of the supplier, recommended price for the vendor is predicted by the bot, and the supplier is classified into different performance categories—High, Medium, and Low.

27.3.2 Decision Bot

The decision bot helps to arrive at a negotiated price systematically. If the recommended price is less than or equal to the quoted price, the deal is closed. Else, the negotiation process is carried out. Since we have proposed a hybrid model for the chatbot, the seller would have predefined options—Agree, Reject, and Tradeoff.

Based on this response, the negotiated prices are presented to the seller step by step to arrive at a final decision (Fig. 27.2).

After obtaining the recommended price and category from the previous step in (Sect. 27.3.1), the input-offered price is taken from the supplier. If the offered price is less than or equal to the recommended price, the deal is closed.

If the offered price is higher, then the bot guides the user through step-by-step negotiation. In [9], the authors have used the formula: 2 (recommended—offered) to calculate the aggressive price (lowest negotiated price), whereas, we have created a logic to calculate N.F and use it as an incrementing factor to calculate the revised recommended price (RRP).

First, the lowest price in the recommended category is calculated. Negotiating factor is ((Recommended price—Lowest price)/4)/Recommended price) if the recommended price is less than or equal to the lowest price, N.F is assumed to be 0.025. RRP1 is calculated by incrementing the lowest price by the negotiating factor (N.F) and subsequently, the RRP [2–5] is calculated by multiplying the N.F by the preceding RRPs.

If the negotiated price 1 stated by the seller after quoting RRP1 lies between RRP1—RRP4, then the deal is accepted, else further negotiation is carried out by quoting subsequent RRP. The upper limit for the accepted price is the recommended price. In the case of the high-performance category, RRP5 is considered, and additional perks are requested. The deal is rejected if the best price offered by the seller is greater than the last RRP.

27.3.3 Interface of Negotiating Bot

The chatbot interface can be built using the natural language processing technique [10]. A data dictionary will be used and created for intent recognition which will contain a set of tags, patterns, and responses. The chatbot responds to the seller as per the tag from the data dictionary and a random response is generated. The inputs from the seller are preprocessed—case normalized (lowercasing all the letters), tokenized (sentence broken to individual words), and stemmed (reducing words to stem words, e.g., Agree, Agreed, Agreeing would be reduced to the root word 'agree').

For simplicity, we have limited the input from the seller to a numeric entry for quoting and negotiating. The entered number is POS Tagged as 'Cardinal Digit'. The integer version of the digit is assigned to the input and compared with the recommended price and RRPs in the subsequent iterations. The responses are generated based on the predefined response options in the dictionary and the calculated RRP costs. The deal is accepted/rejected as per the conditions explained in (Sect. 27.3.2) and the chatbot is closed.

27.4 Conclusions and Future Work

The proposed methodology is a general framework that can be used to implement smart-automated negotiation in procurement. In the proposed methodology, the price negotiation process is referenced with a recommended price. For a product, the recommended price will be unique for each vendor, which will be depending on vendors' records, i.e., fill rate, OTD, defect rate, compliance rate, vendor availability, and average price associated with the vendors' region. Referencing the recommended price, a decision tree-based scheme demonstrates to settle the right deal for a procurement.

Future works could extend this methodology for negotiation from price to considering additional features in the contract process such as contract renewal dates, delivery time, tradeoffs to arrive at the final decision, document detection, and compliance to enable an end-to-end contract management process.

Here, the decision bot logic considers only a limited number of iterations for bargaining. The iterations can be made dynamic as per the product and business requirements. Apart from this, deep learning models can be incorporated to enhance the performance and interface experience of the chatbot.

References

1. Morris PWG, Pinto JK (2007) The Wiley Guide to project technology, supply chain & procurement management. Wiley, Inc., Hoboken, New Jersey
2. Chae BK (2009) Developing key performance indicators for supply chain: an industry perspective. Supply Chain Manage 14(6):422–428
3. Erhun F (2014) Contract complexity and performance. In: INFORMS TutORials in Operations Research
4. Kalkanci B, Chen K-Y, Erhun F (2013) Complexity as a contract design factor: a human-to-human experimental study. Prod Oper Manage
5. Adamopoulou E, Moussiades L (2020) An overview of chatbot technology. In: Maglogiannis I, Iliadis L, Pimenidis E (eds) Artificial intelligence applications and innovations. AIAI. IFIP advances in information and communication technology, vol 584. Springer (2020)

6. Brett JF, Pinkley RL, Jackofsky EF (1996) Alternatives to having a BATNA in dyadic negotiation: the influence of goals, self-efficacy, and alternatives on negotiated outcomes. Int J Confl Manag 7(2):121–138
7. Cui R, Li M, Zhang S (2022) AI and procurement. Manufacturing & Service Operations Management
8. Carbonneau R, Kersten GE, Vahidov R (2008) Predicting opponent's moves in electronic negotiations using neural networks. Expert Syst Appl Int J 34(2)
9. Liu T, Zheng Z (2020) Negotiation assistant bot of pricing prediction based on machine learning. Int J Int Sci 10:9–21
10. Tuarob S, Tucker CS (2015) A product feature inference model for mining implicit customer preferences within large-scale social media networks. In: 35th computers and information in engineering conference (Proceedings of the ASME design engineering technical conference, Vol. 1B-2015). American Society of Mechanical Engineers (ASME)

Chapter 28
A Deep Learning-Based Reverse Logistics Model for Recycling Construction and Demolition Waste

Sanjeev Sinha, Subodh Srivastava, Bal Krishan Sahay, and Abhinav Kumar

Abstract The outcome of construction activities leads to the production of large amounts of solid waste, primarily known as construction and demolition (C&D) waste. The reutilizing of C&D wastes plays a vital role in the sustainable growth of the environment, economy, and public health. The existing recycling methods have limitations, such as cost, human intervention, unstable identification process for recycling, on-site sorting techniques, irregular landfill events, and a lack of an effective waste tracking system. The paper proposes an end-to-end improved convolutional neural network (EEI-CNN) based reverse logistics model for recycling C&D waste to overcome these issues. The EEI-CNN is a customized convolutional neural network that performs the classification of C&D waste aggregates. The refine the efficacy of EEI-CNN, a preprocessed image is used. The effectiveness of the proposed method is judged for an openly available C&D waste image dataset. The evaluation metrics like accuracy, precision, true positive rate, true negative rate, and F-score are estimated. The proposed method outperforms existing methods based on comparative analysis.

Keywords Construction and demolition · Convolutional neural network · Landfill · Reverse logistics

S. Sinha
Department of Civil Engineering, National Institute of Technology, Patna 800005, India
e-mail: sanjeev@nitp.ac.in

S. Srivastava (✉) · A. Kumar
Department of Electronics and Communication Engineering, National Institute of Technology, Patna 800005, India
e-mail: subodh@nitp.ac.in

A. Kumar
e-mail: abhinav.ec19@nitp.ac.in

B. K. Sahay
Department of Bihar State Disaster Management Authority, Patna 800005, India
e-mail: bksahay.15@gmail.com

M. K. Tiwari et al. (eds.), *Applications of Emerging Technologies and AI/ML Algorithms*, Asset Analytics, https://doi.org/10.1007/978-981-99-1019-9_28

28.1 Introduction

Construction engineering is a complex process involving planning and managing large structures. It is a process of generating large quantities of solid waste known as C&D Waste [1]. Currently, the management of C&D waste is a global challenge and issue. Its proper management can help in natural aggregate restoration for concrete production with the environment and public health. It enables a recognition of the large construction wastes required for landfills. It also aids in avoiding a mix-up with municipal waste. Recycling is one of the most active strategies to minimize waste and achieve sustainable construction waste management. The three Rs method, which stands for reduce, reuse, and recycle, is the most common perspective to lessen the probable effect of waste [2]. While recycling waste has several advantages, it would be cumbersome as it takes time and human interference to separate the trash into predefined classes. Additionally, the high cost, human requirement, uneven sources for recycling, on-site sorting, unregulated landfill activities, and a lack of an effective waste tracking system [3] are the major drawbacks of existing C&D waste management techniques. To overcome such difficulties in recycling construction waste, technologies based on artificial intelligence have emerged as alternatives. This research aims to develop a deep-learning-based reverse logistic model for recycling C&D waste.

The rest of the manuscript is prearranged as follows; in Sect. 28.2, a literature study of the existing method of C&D waste is presented. Section 28.3 describes the methodology and materials for the present work. Section 28.4 illustrates the description of evaluation metrics. Section 28.5 is about the result and discussion of the proposed method. And Sect. 28.6 presents a conclusion.

28.2 Related Work

This segment presents navigation about the existing work. Akanbi et al. [4] presented a deep neural network (DNN) for waste classification in a globular economy. This model can predict the quantities of materials used after the demolition of the building. The numeric value of the average R-squared was 0.97, and the mean absolute error (MAE) ranged from 17.93 to 19.04. Majchrowska et al. [5] demonstrated a deep learning-based framework for waste recognition in natural and municipal surroundings. The suggested method achieved an average accuracy of up to 70% in waste identification and a classification accuracy of around 75% in the test data set. Na et al. [2] developed a framework based on transfer learning using artificial intelligence (AI) to recognize construction waste. The proposed model was developed based on an image dataset. There was a limited dataset, so the data augmentation increased the dataset. With the help of data augmentation and transfer learning, the mAP was increased by 16%. Chu et al. [7] have designed a multilayer hybrid deep learning method for waste classification and recycling. A convolutional neural network (CNN) extracts the features from the image dataset and combines them with other elements to make it a fusion feature. The obtained accuracy was 90%. Davis et al. [6] have used a deep convolutional neural network to classify construction waste material. The performance evaluation metric used was the F1 score and accuracy. The method achieved an accuracy of 94%. Awe et al. [7] depicted a faster regional convolutional neural network (F-RCNN) model to classify waste. It was an experimental project that categorizes waste into paper, recycling, and landfill. The value of mean precision was 68%. Thung and Yang [8] employed a support vector machine (SVM) with CNN to classify waste into six groups. It has an accuracy of 63%.

Based on the literature study, it has been noticed that very few performance evaluation parameters are estimated to judge the efficacy of the developed model. And also, the existing method has limited accuracy (Table 28.1).

Table 28.1 A detailed summary of the reported work as a literature

Author	Year	Dataset	Method	Parameters	Remarks
Akanbi et al. [4]	2022	C&D image dataset	DNN	Average R-squared = 0.97, MAE = 17.90–19.04	Less number of parameters are estimated
Majchrowska et al. [5]	2020	Open litter image dataset	Deep learning model	Precision = 75%, accuracy = 70%	Accuracy is limited
Na et al. [2]	2022	Self-acquired image dataset	YOLACT	Accuracy = 97%	Less number of parameters are estimated.
Chu et al. [9]	2018	Self-acquired image dataset	Transfer learning with AI	mAP increased by 16%	No other parameters are estimated, only mAP
Davis et al. [6]	2021	Self-acquired image dataset	Deep CNN	Accuracy = 94%	Less evaluation metrics are estimated
Awe et al. [7]	2107	Image dataset	F-RCNN	Precision 68%	Less number of parameters are determined
Thung and Yang [8]	2016	Image dataset	SVM and CNN	Accuracy = 63%	Less evaluation metrics are assessed

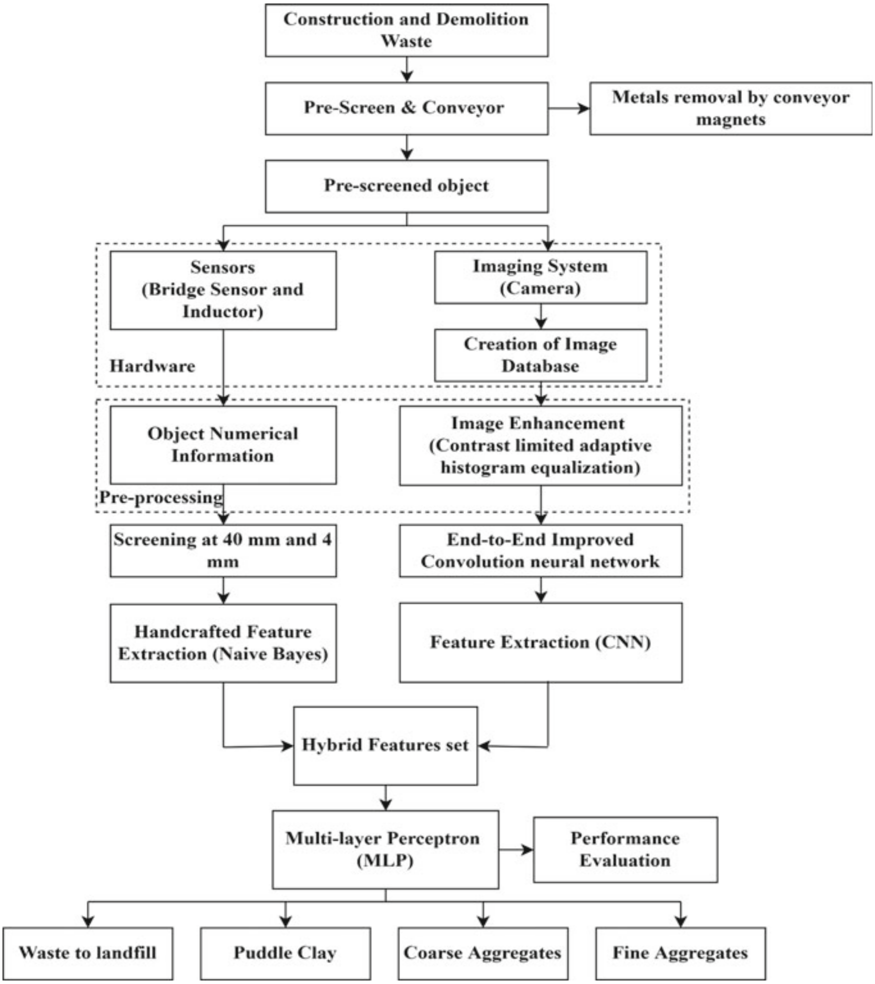


Fig. 28.1 Proposed flowchart of deep learning-based reverse logistic model for recycling C&D waste

28.3 Method and Models

Figure 28.1 shows the proposed flowchart for recycling C&D waste for real-time application. In the proposed flowchart, there are two ways to collect the data. The first way uses the sensors to acquire the numeric information of the pre-screened object. Sensors [9] like bridges measure the weight of the C&D solid waste. And inductors differentiate between metallic and non-metallic C&D waste. Further, a Naive Bayes

Table 28.2 Complete enlightenment of used datasets for the proposed work

Attributes	Description
Number of images	1231
Number of classes	11
Image size	Variable in nature
Image file format	Joint photographic experts group
Accessibility	Publicly

classifier extracts the features based on numerical information. The second way uses an imaging system (Cameras) to acquire images of waste. This process results in the formation of an image database. Further, image enhancement is performed with the help of contrast-limited adaptive histogram equalization (CLAHE) [10]. After that, the preprocessed image data is categorized as model training, validation, and test data. The train and validation data are fed into an end-to-end improved CNN. It is used for feature abstraction. The features extracted from both ways are merged, and a hybrid feature set is created. Further, the hybrid feature set is given to a multilayer perceptron neural network that classifies the C&D waste as landfill, Puddle Clay, Coarse aggregates, and Fine aggregates. It is the working explanation of the proposed method of recycling C&D waste for real-time application. For the present paper, the proposed flowchart works as an assumption where only an end-to-end improved CNN has been developed for the available image dataset.

28.3.1 Dataset

The proposed method utilizes an open-source dataset [11]. It consists of a total of 11 groups of solid and C&D waste. For easy working, concrete blocks and bricks are labeled as waste to landfill class. Soil vegetation is labeled as puddle clay. The groups of gravels and stones are labeled coarse aggregates. Similarly, paving and sands are marked for fine aggregates. The complete explanation of the dataset is presented in Table 28.2.

28.3.2 Pre-processing

In the present paper, contrast-limited adaptive histogram equalization [10] is accomplished for the quality improvement of C&D waste images. It is an essential step for enhancing image quality, and the result works as an input for accurate classification. It improves the contrast of low-quality images globally. For CLAHE, all necessary information such as the number of rows and columns, image bit or dynamic range, clip limit, and distribution parameter type are collected. After that, a mapping of the gray level with a clipped histogram is generated. It gives the average number of pixels over a contextual region. It mathematically reads as

$$M_{\text{Avg}} = \frac{M_{CR} - X_p * M_{CR} - Y_p}{M_{\text{gray}}} \tag{28.1}$$

where M_{Avg} is the average number of pixels, M_{gray} is the number of gray levels, M_{CR-X_p} and M_{CR-Y_p} are the number of pixels in the contextual region's X and Y direction, respectively. Using Eq. (28.1) and histogram clip information, CLAHE is determined. It is given as

$$M_{CL} = M_{\text{CLIP}} * M_{\text{Avg}} \tag{28.2}$$

28.3.3 Proposed an End-to-End Improved Convolutional Neural Network (EEI-CNN) Based Reverse Logistics Model for Recycling Construction and Demolition Waste

EEI-CNN is a customized convolution neural network [12] that uses four convolution layers to extract the features. CNN is a deep learning architecture that processes structured array data, commonly called images. It consists of three layers: convolution named as Conv2D, pooling, and fully connected layers. The initial two layers, i.e., Conv2D and pooling, are used for feature abstraction. The third, fully connected layer is used to map the extracted features into the final output, such as classification with the help of SoftMax. Only the convolutional and pooling layer is cast to extract the low-level and high-level features. The low-level features reflect the minor details of the image, whereas the high-level features refer to objects or instances present in the image. The proposed method uses two activation functions, i.e., Rectified Linear Unit (ReLU) and sigmoid functions. ReLU can be read as

$$f(x)_{ReLU} = \max(x, 0) \tag{28.3}$$

ReLU gives a zero value for negative numeric pixel values and returns the x for positive numeric pixel values. The sigmoid activation is a nonlinear analysis of $f(x)$ that indicates the changes with the assessment of x. It reads as

$$f(x)_{\text{Sigmoid}} = \frac{1}{1 + \exp(-x)} \tag{28.4}$$

Figure 28.2 shows the architecture of the proposed EEI-CNN. It is designed and developed for the publicly available C&D image dataset. After preprocessing by CLAHE, the dataset is alienated into train, validation, and test image data of

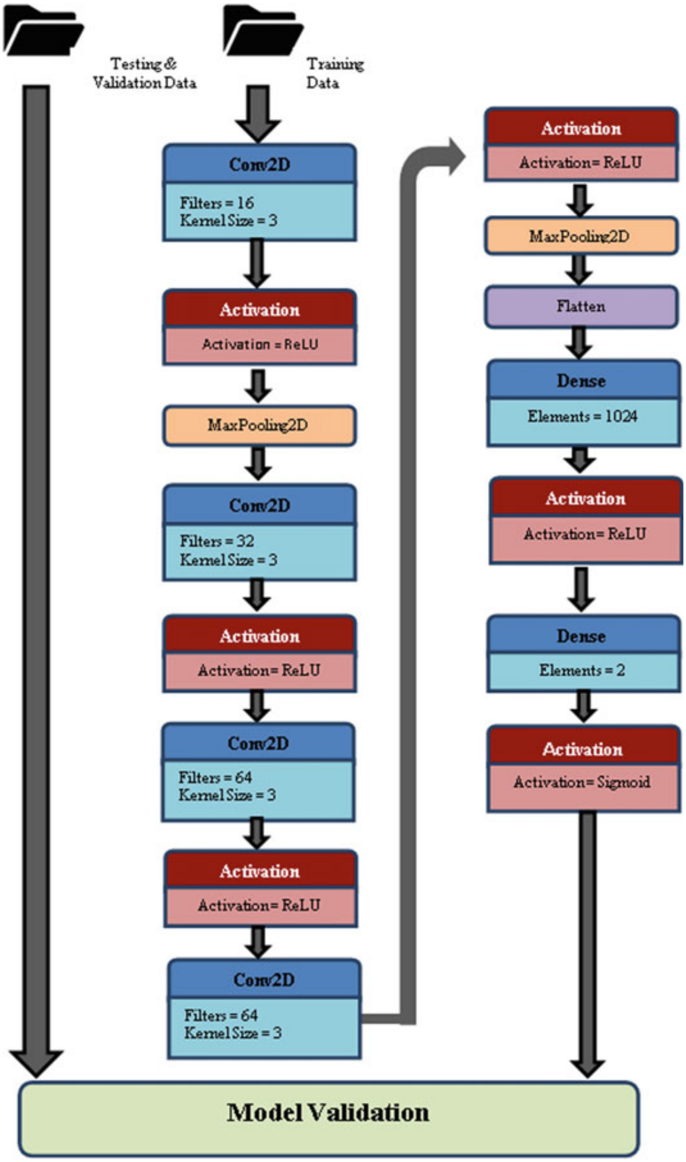


Fig. 28.2 The proposed architecture of the EEI-CNN

ratio 8:1:1. It comprises four convolutional operations named Conv2D, followed by activation and a pooling operation. The four consecutive Conv2D uses filter sizes 16, 32, 64, and 64. The abstracted features are used to make a set of features in 1D vector form. Then the feature set is fed to the flattening layer. Further, the output of flatten layer is processed through the SoftMax layer that providing a probability-based classification.

28.4 Evaluation Metrics

Some evaluation metrics [13] are predefined to measure the efficacy and a comparative study. For the proposed work, classification evaluation metrics are used. The metrics are rooted in the confusion matrix, which is shown in Table 28.3.

$$\mathrm{Accuracy(AC)} = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN}} \tag{28.5}$$

$$\mathrm{Precision(PC)} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} \tag{28.6}$$

$$\text{True positive rate(TPR)} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \tag{28.7}$$

$$\text{True negative rate(TNR)} = \frac{\mathrm{TN}}{\mathrm{TN} + \mathrm{FP}} \tag{28.8}$$

$$\mathrm{F} - \mathrm{score} = \frac{2 * \mathrm{Precision} * \mathrm{TPR}}{\mathrm{Precision} + \mathrm{TPR}} \tag{28.9}$$

$$\text{Balanced classification rate(BCR)} = \frac{1}{2}[\mathrm{TPR} + \mathrm{TNR}] \tag{28.10}$$

$$\text{Youden's index (YI)} = \mathrm{TPR} - (1 - \mathrm{TNR}) \tag{28.11}$$

Table 28.3 Confusion matrix

	Real positive	Real negative
Anticipated positive	TP	FP
Anticipated negative	FN	TN

28.5 Result and Discussion

Figure 28.1 shows the proposed EEI-CNN for the real-time application. Assuming Fig. 28.1 as a core concept, an EEI-CNN is developed for existing datasets. Figure 28.2 shows the architecture network of the proposed EEI-CNN. It uses four consecutive layers of convolutional operations trailed by a ReLU with a pooling operation. The last Conv2D uses the sigmoid activation function. Here, the first layer abstracts the low-level image features while the last layer provides the high-level features information of the input image. Figures 28.3 and 28.4 are the proposed method training-validation, accuracy, and loss curves. It is a visualization plot of the learning proportion of accuracy and loss in accordance with numeral of epochs executed. A total of 50 epochs are run to form a training validation model of the proposed method. As shown in Fig. 28.3, the accuracy of the proposed method increases as the number of epochs increases. Similarly, the loss of the proposed EEI-CNN of Fig. 28.4 decreases as the number of epochs increases.

Table 28.4 is a comparative study of the proposed method with the existing CNN method, such as Akanbi et al. [1], Davis et al. [5], Chu et al. [4], and Majchrowska et al. [2]. The numeric entity of AC, TPR, TNR, PC, F-score, BCR, and YI are 95.46%, 97.61%, 93.02%, 0.9401, 0.9577, 0.9532, and 0.9063, respectively. Figures 28.5 and 28.6 are comparative graphical plots of existing CNN and the proposed EEI-CNN.

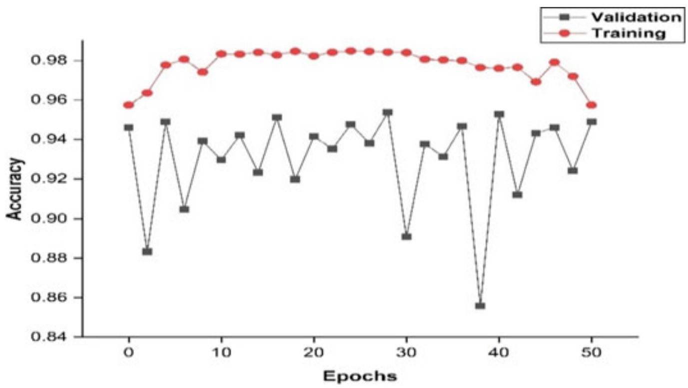


Fig. 28.3 Accuracy curve of training and validation of the proposed EEI-CNN

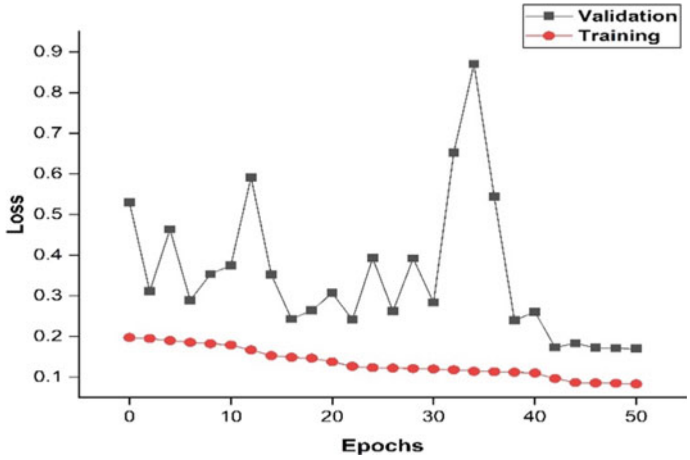


Fig. 28.4 Loss graph of training and validation of the proposed EEI-CNN

Table 28.4 A comparative study of existing CNN and the proposed EEI-CNN

CNN model	AC (%)	TPR (%)	TNR (%)	PC	F-score	BCR	YI
Akanbi et al. [1]	93.16	93.53	92.78	0.9307	0.9330	0.9316	0.8631
Davis et al. [5]	94.00	86.00	85.00	0.8601	0.8602	0.8602	0.7103
Chu et al. [4]	91.00	91.00	92.78	0.9286	0.9192	0.9189	0.8378
Majchrowska et al. [2]	87.20	87.20	84.20	0.7740	0.8601	0.8202	0.8101
Proposed model	95.46	97.61	93.02	0.9401	0.9577	0.9532	0.9063

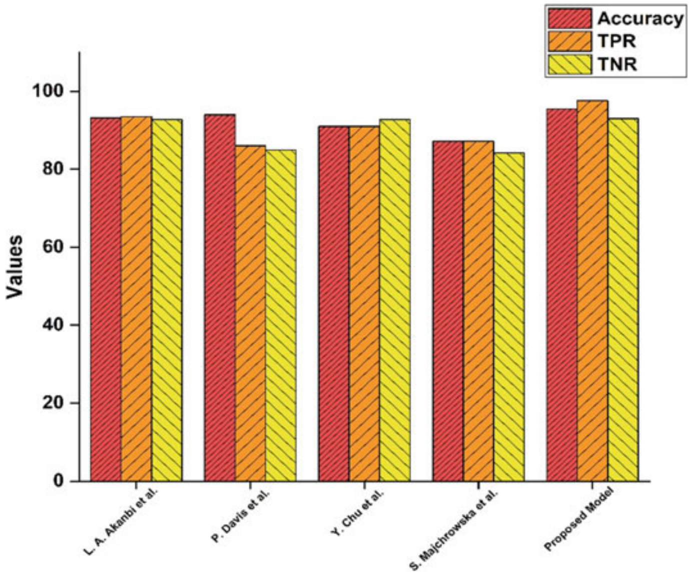


Fig. 28.5 Comparative graphic plot between existing CNN and the proposed EEI-CNN in terms of accuracy, TPR, and TNR

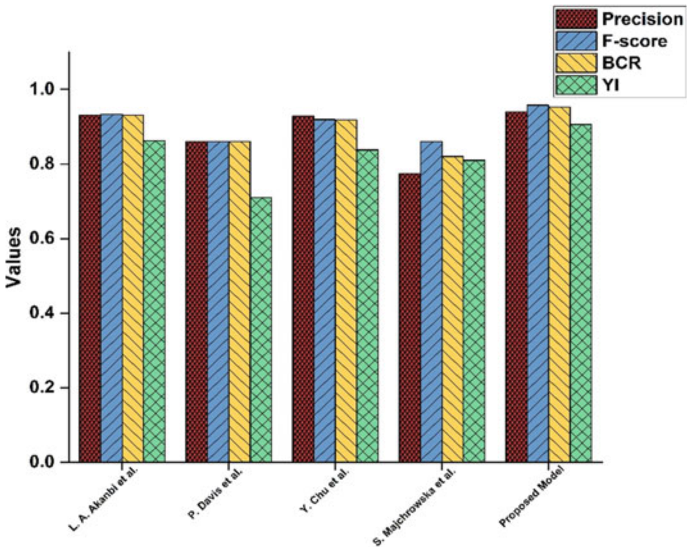


Fig. 28.6 Comparative visual plot between existing CNN and the proposed EEI-CNN in terms of precision, F-score, BCR and YI

28.6 Conclusion

In this paper, an EEI-CNN-based reverse logistic model was proposed to recycle C&D waste. It was developed for the openly available image dataset. It was based on the fundamental concept of image processing with deep learning. In terms of image processing, CLAHE was used to improve image quality. Further, a customized CNN model that comprises four Conv2D layers was designed. It was used to classify the image datasets as waste to landfill class, puddle clay, coarse aggregates, and fine aggregates. The corresponding numeric entity of AC, TPR, TNR, PC, F-score, BCR, and YI was 95.46%, 97.61%, 93.02%, 0.9401, 0.9577, 0.9532, and 0.9063, respectively. Based on the comparative study, it can be concluded that the proposed method is performing well. The proposed method helps in reducing the cost required for the recycling process with the automatic classification of C&D waste. It also facilitates a waste tracing system that results in support for on-site sorting of waste.

Acknowledgements The authors did not receive support from any organization for the submitted work.

Declarations There is no Conflict of Interest and this work did not involve human subjects or animals in its research.

References

1. Sinha S, Shankar R, Taneerananon P (2010) Modelling and case study of reverse logistics for construction aggregates. Int J Logist Syst Manag 6:39–59. https://doi.org/10.1504/IJLSM.2010.029720
2. Na S, Heo S, Han S et al (2022) Development of an artificial intelligence model to recognise construction waste by applying image data augmentation and transfer learning. Buildings 12.https://doi.org/10.3390/buildings12020175
3. Ma M, Tam VWY, Le KN, Li W (2020) Challenges in current construction and demolition waste recycling: a China study. Waste Manag 118:610–625. https://doi.org/10.1016/j.wasman.2020.09.030
4. Akanbi LA, Oyedele AO, Oyedele LO, Salami RO (2020) Deep learning model for demolition waste prediction in a circular economy. J Clean Prod 274.https://doi.org/10.1016/j.jclepro.2020.122843
5. Majchrowska S, Mikołajczyk A, Ferlin M et al (2022) Deep learning-based waste detection in natural and urban environments. Waste Manag 138:274–284. https://doi.org/10.1016/j.wasman.2021.12.001
6. Davis P, Aziz F, Newaz MT et al (2021) The classification of construction waste material using a deep convolutional neural network. Autom Constr 122:103481. https://doi.org/10.1016/j.autcon.2020.103481
7. Awe O, Mengistu R, Sreedhar V (2017) Smart trash net: waste localization and classification. CS229 Proj Rep 1–6
8. Yang M, Thung G (2016) Classification of trash for recyclability status. CS229 Proj Rep 1–6
9. Chu Y, Huang C, Xie X et al (2018) Multilayer hybrid deep-learning method for waste classification and recycling. Comput Intell Neurosci 2018.https://doi.org/10.1155/2018/5060857

10. Yadav G, Maheshwari S, Agarwal A (2014) Contrast limited adaptive histogram equalization based enhancement for real time video system. In: Proceedings of 2014 international conference on advances in computing, communications and informatics, ICACCI 2014, pp 2392–2397. https://doi.org/10.1109/ICACCI.2014.6968381
11. Ghassemi N, Mahami H, Darbandi MT et al (2020) Material recognition for automated progress monitoring using deep learning methods
12. Kumar P, Srivastava S, Mishra RK, Sai YP (2022) End-to-end improved convolutional neural network model for breast cancer detection using mammographic data. J Def Model Simul 19(3):375–384
13. Kumar P, Kumar A, Srivastava S, Padma Sai Y (2022) A novel bi-modal extended Huber loss function based refined mask RCNN approach for automatic multi instance detection and localization of breast cancer. Proc Inst Mech Eng Part H: J Eng Med 236(7):1036–1053

Chapter 29
Supplier Prioritization and Risk Management in Procurement

Virendra Kumar Verma

Abstract Today's global business is evolving into a one-economy and one culture. As a result, the production sector is shifting toward high-value-added operations, and supplier and risk management needs may change significantly. The current overview addresses the issues of finding suitable and viable suppliers, prioritization, and risk management in procurement. The main types of supplier selection have been determined at this point. Companies utilize qualifying criteria, selection criteria, and other considerations to categorize criteria during the selection process and risk management. The SLR technique is utilized to establish the relative relevance of supplier prioritization and risk management in procurement. In the context of the Indian manufacturing business, this helps to determine purchaser preferences in prioritizing suppliers. A review was conducted for this aim to define supplier prioritization and proactive risk management in procurement. This review paper's contribution is to examine the research findings and future research directions. The SLR study helps find and contribute to viable suppliers based on delivery, quality, cost, added value, flexibility, risk, service, ecological ways, and social accountability, as well as operational and disruption risks in risk management in procurement. Researchers can use an innovative approach to operations research to solve supplier prioritization and risk issues in procurement uncertainty concerns.

Keywords Supplier · Prioritization · Risk management · Procurement · Review · Manufacturing industry

V. K. Verma (✉)
Operations and SCM, National Institute of Industrial Engineering (NITIE), Mumbai, India
e-mail: virendra.verma.2018@nitie.ac.in

KIIT School of Management, KIIT University, Bhubaneswar, India

M. K. Tiwari et al. (eds.), *Applications of Emerging Technologies and AI/ML Algorithms*, Asset Analytics, https://doi.org/10.1007/978-981-99-1019-9_29

29.1 Introduction

Today's manufacturing industry faces strong competition in terms of pricing, quality, and logistics [1]. Various manufacturing organizations begin to subcontract business that is outside their core strength in demand to carry on with the industry [2]. Outsourcing has made the supply chain more complicated and challenging to handle. If businesses do not understand the risk's implications and prevalence, they may undergo unfavorable outcomes. In the manufacturing company, the risk is well described as the inconsistency of return, and risk can have a harmful effect on forecasted yield [3]. Risk management is not only a critical concern in the financial sector; it is also a developing worry in the supply chain, as the unpredictability of an incident happening in one part of the supply chain might result in an unpleasant string effect for the supply chain channel [4, 5]. Several possible risk factors, such as supplier selection, supply yield, pricing, and demand should obviously be considered while making optimal decisions in risk management in procurement under uncertainty. Researchers have conducted much research in this field. Campuzano-Bolarin et al. [6] explored the issue of stochastic demand procurement planning [6]. In the aspect of unpredictable order and amount in the open market. Federgruen and Yang [7] established a model to construct resource support and determine demand distribution with suppliers in the existence of demand fluctuations and yield [7]. Nevertheless, before applying procurement risk management solutions, it is essential to recognize the source of unpredictability that causes procurement risk. As there is no organized approach to categorizing procurement risk, the manager's experience and intuition play a big role in how they manage it. After an unexpected risk happens, a short-term risk management strategy is put in place. There are a lot of literature review articles on supply chain risk, but according to the author, there hasn't been a review paper on risk management in procurement. We also reviewed and discussed supplier prioritization criteria such as delivery, quality, cost, added value, flexibility, service, risk, environmental practices, and social responsibility in Sect. 29.3.1. This brief review study gives information about supplier prioritization and procurement risk for researchers and practitioners. Section 29.1 delivers a short introduction to the background of supplier prioritization and risk management in procurement. Section 29.2 explains the literature review. Section 29.3 supports an overview of supplier prioritization. Section 29.4 is about the overview of risk management in procurement. Section 29.5 shows opportunities and challenges. The conclusions are described in Sect. 29.6.

29.2 Literature Review

The review approach is established on SLR [8], a literature review methodology. Figure 29.1 shows a flow chart of the literature search method to give a strong image of the review activity and findings. The articles chosen were mostly published between 2000 and 2022, with "Scopus, Science Direct, Web of Science, and Google

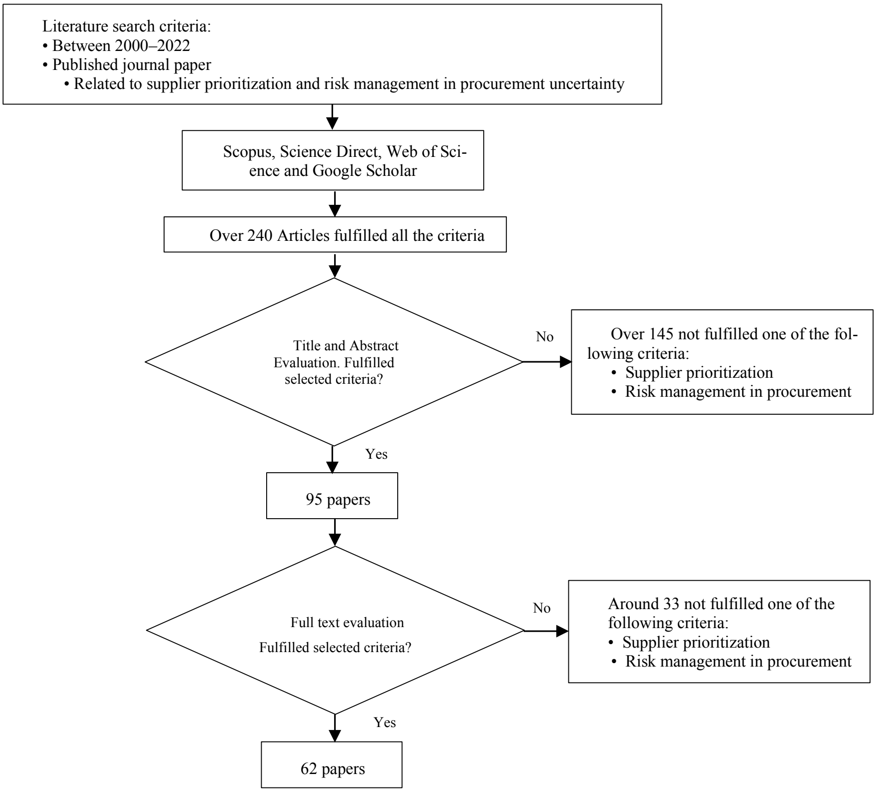


Fig. 29.1 Literature review search flowchart

Scholar" as the primary electronic databases. Supplier prioritization, Risk management, procurement, sourcing, supply risk, supply chain, and contract were among the keywords searched to find relevant publications. The papers are further sorted based on their topical relevancy, citation count, and publication year. Over 62 papers from top reputed journals such as the International Journal of Production Economics, the International Journal of Production Research, and many more were chosen, as shown in Fig. 29.1.

29.3 Overview of Supplier Prioritization

Suppliers are a highly substantial element in any supply chain, so procurement is one of the most significant decisions to make during the planning stage. For any firm, the necessity to attain international supply-edge competitiveness has expanded significantly over the previous era [9, 11, 14]. Suppliers play a significant role in achieving business competitiveness in new procurement and operational strategies.

As a result, choosing the correct suppliers is an important part of these strategies. In reality, the uncertainty and inaccuracy of the problem's goals, limitations, and parameters make decision-making difficult. As a result, the challenge of supplier selection has gotten a lot of attention in the area of supply chain and main strategies look at the issue through the lens of selection criteria. In reality, a company's decision to choose a supplier may be affected by a number of factors, such as the price quoted, the performance of the product, on-time shipment, the after-sales support, the supplier's locality, and their economic stability [10, 12, 14, 56, 59]. Supplier selection appears to be a multi-benchmarks challenge involving together qualitative and quantitative aspects. When looking for a suitable provider, you must make a deal between these direct and indirect factors.

29.3.1 Supplier Prioritization Criteria

Table 29.1 shows a detailed description of the supplier prioritization criteria.

29.4 Overview of Risk Management in Procurement

According to Tang [26], managing supply chain risk involves coordination or cooperation between supply chain players to ensure competitiveness and consistency [26–28]. In addition, in Sect. 29.4.1, we identify common procurement concerns to help us define the scope of our evaluation. Procurement risks and approaches are discussed in Sect. 29.4.2.

29.4.1 Risk Management in Procurement

Procurement is a set of processes that a company uses to get resources such as materials, expertise, capacity, and services to carry out its primary firm endeavors [29–31]. HP was a pioneer in developing a price, demand, and availability uncertainty measurement approach based on scenario analysis [29, 47, 50, 60]. Following the realization of these uncertainties, the contract is set up and its value is determined. Numerous studies have been conducted, and various purchasing portfolios have been developed, taking into account operational costs, product availability, and demand information, to achieve optimal control over supply while also minimizing risks or variances [28, 32, 33]. Procurement risk management is the method of decreasing risk and uncertainty in prices, lead-times, and demands to assure a continuous supply flow (materials, expertise, capacity, and services) with the smallest amount of interruption [34, 35, 46]. Although procurement is a part of the supply chain, the implications of insufficient risk management in procurement might be comparable to or the same as

Table 29.1 Description of the supplier prioritization criteria

Supplier prioritization criteria	Descriptions
Delivery	Integration with logistics services can help manufacturing supply chains save money and increase overall effectiveness. Decreased transport costs, value-added services, and shorter shipment times can all help the supply chain play more efficiently [9, 11, 13, 14]
Quality	Quality must be taught to all players in the channel as a streamlined way of doing business. Quality has become very significant because supply chains that use just-in-time production with optimal inventory levels have little safeguards to protect against quality problems [10, 12, 14, 54, 57]
Cost	Direct labor, supplies, rework, and scrap are all areas where you can save money on the inside. Other possibilities can be identified through decreasing expenses in novel approaches. Non-value-added expenses can be minimized by utilizing methods including activity-based costing (ABC), just-in-time manufacturing, supplier-managed inventories, and lean manufacturing. While cost cuts are crucial, more effort will be necessary to sustain a competitive benefit and meet the challenge of progressively more proficient international competitors [12, 13, 15]
Added value	Value-added services like low-cost storage, fast responses to warranty issues, easy access to replacement components, and improved logistics can give you an edge over your competitors [14, 23–25]
Flexibility	Production flexibility, variety flexibility, demand flexibility, and service flexibility are all aspects that influence the level of flexibility required from a supplier. Due to a lack of statistical data, precise flexibility needs, such as a specific upward flexibility percentage, are now impossible to achieve in this study [13, 21, 22, 52]
Service	Customer beliefs for prompt service before, during, and after a sale are continuing to rise, and service providers are responding to these demands, aided by the Internet and innovative transportation methods. Maintenance manuals, service bulletins, and answers to frequently asked inquiries are all available on the Internet. Customers can place orders every day, including holidays, from somewhere in the globe using e-commerce. In the US, replacement parts can be supplied instantly, and essentially anywhere else in a few days [9, 13, 14, 51]
Risk	Maintaining continuous supply network flows is essential for a supply chain's success in the market. However, each of these flows has associated risks that require appropriate mitigation techniques. The theme of supply chain risks has increased in prominence as a result of the realization that supply network breakdowns are lethal to the presence of all supply chain participants. This is due to a lack of resources and proper planning to address supply chain risks [5, 7, 11, 13, 14, 53]

(continued)

Table 29.1 (continued)

Supplier prioritization criteria	Descriptions
Environmental practices	Incorporating environmental considerations into purchase decisions introduces a new set of trade-offs, confounding the decision-making method with both quantitative and qualitative aspects. Even so, not many businesses use any kind of organized study to evaluate suppliers based on how they affect the environment. Environmental factors are supplier environmental evaluation, hazardous waste management rules, ISO 14000 certified, hazardous material list, management of toxic waste pollution, eco-friendly product packaging, and recyclable content [16, 17, 22, 55]
Social responsibility	The ability to handle stakeholders was viewed as a social responsibility. Many academics have tried to define social sustainability in their study. Social sustainability is a human rule of behavior that must be attained fairly, inclusively, and responsibly. The function of procurement managers in the area of social responsibility in the supply chain, as well as how carrying out these responsibilities can lead to more trust and participation from suppliers through procurement social responsibility and logistics social responsibility [18–20, 58, 61, 62]

those of a lack of risk management in the supply chain [36, 37, 48]. The authors make an effort to differentiate between supply chain risk and procurement risk by emphasizing that procurement risk is mainly focused on the risks concerned with supply interruption and the agreement that exists between the purchasing organization and the suppliers.

29.4.2 Risk Management Approaches in Procurement

To be more specific, in today's unpredictable and competitive market, procurement is subjected to a wide range of uncertainties, including fluctuating lead times, unpredictable demand, and unstable pricing. Most of these unpredictabilities and threats can be divided into two types: disruption and operational risks, which are discussed in Table 29.2. The risk of potential financial loss that can be attributed to various partners in the supply chain not effectively collaborating with one another is known as operational risk. Natural disasters, strikes, political uncertainty, and equipment failure are all examples of disruption risk [27, 40–45, 49]. Table 29.2 summarizes all potential procurement uncertainties following a thorough analysis of the research paper. As a result, it is essential to first identify potential risks before selecting and implementing risk management in procurement techniques. The usual procurement risks and accompanying risks management in procurement approaches are depicted in Fig. 29.2.

Table 29.2 Risks management in procurement

Operational risks	Disruption risks
Variable lead time; uncertainty on the supply yield	Disruption of information systems
Informational system with inadequate data consistency and effectiveness	Civil war or political crisis
Financial constraints; interest rate variability; price uncertainty	Fire or poor weather conditions
Incorrect inventory records; changing currency rates	Legal threats
Demand uncertainty; essential manpower constraint	Supplier failure, machine failure

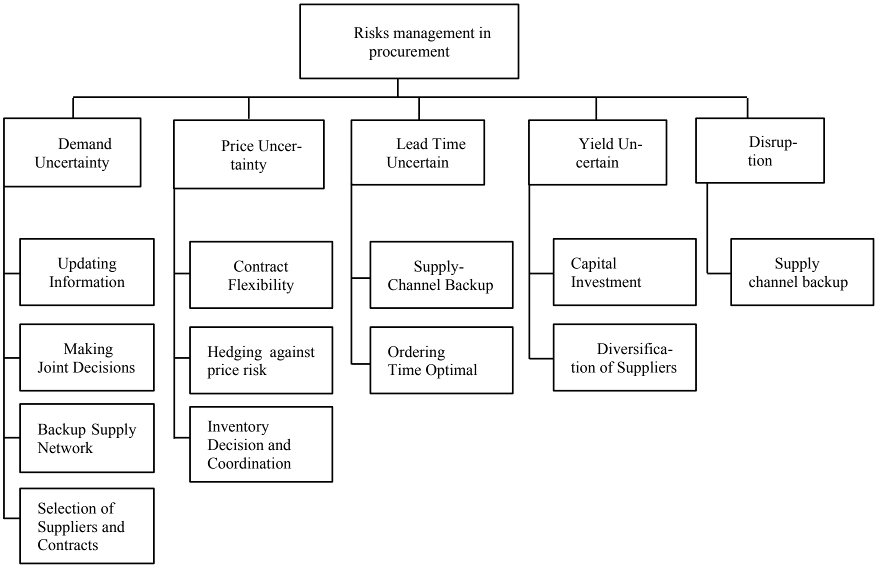


Fig. 29.2 Risk management approaches in procurement flowchart

29.5 Opportunities and Challenges

Opportunities and challenges for suppliers are ensuring quality, reducing risk, fostering corporate social responsibility, satisfying regulatory needs, controlling expenses, producing value for the company, and driving innovation. Supplier channels are more complicated than ever; therefore, it is time to shift away from telephonic conversations, mail, and spreadsheets and incorporate new approaches to communicating with suppliers. This will boost your capacity to evaluate the performance of suppliers and organize tasks that require access to numerous organizations. Smart Enterprise work management on a digital platform will help conquer today's largest supplier relations difficulties and improve suppliers' performances across the KPIs that matter most to the company.

29.6 Conclusions

A variety of procurement uncertainties are discussed in this brief review study. Uncertain demand, multi-uncertainties, disruption risks, pricing, yield, and lead time are all extensively explored. This work also includes an assessment of the most recent uncertainty management strategies. The genuinely significant factors are flexibility, delivery, quality, and cost, while risk, service, social responsibility, and environmental practices are highlighted as desired attributes for supplier prioritization criteria. It covers the aspects of future study in the area of supplier prioritization and risk management in procurement. The manager's experience shows a big responsibility in controlling unexpected risk during a short-term risk management strategy and supplier prioritization strategy. These are the major areas of improvement and future research direction.

References

1. Pereira CR, Christopher M, Da Silva AL (2014) Achieving supply chain resilience: the role of procurement. Supply Chain Manag Intern J 19(5/6):626–642
2. Kremic T, Tukel OI, Rom WO (2006) Outsourcing decision support: a survey of benefits, risks, and decision factors. Supply Chain Manag Int J 11(6):467–482
3. Guo H, Whitelaw RF (2006) Uncovering the risk–return relation in the stock market. J Financ 61(3):1433–1463
4. Finch P (2004) Supply chain risk management. Supply Chain Manag Int J 9(2):183–196
5. Tummala R, Schoenherr T (2011) Assessing and managing risks using the supply chain risk management process (SCRMP). Supply Chain Manag Int J 16(6):474–483
6. Campuzano-Bolarin F, Mula J, Díaz-Madroñero M, Legaz-Aparicio ÁG (2020) A rolling horizon simulation approach for managing demand with lead time variability. Int J Prod Res 58(12):3800–3820
7. Federgruen A, Yang N (2008) Selecting a portfolio of suppliers under demand and supply risks. Oper Res 56(4):916–936
8. Onwuegbuzie AJ, Frels R (2016) Seven steps to a comprehensive literature review: a multimodal and cultural approach. Sage Publication, London
9. Jafarnezhad Chaghooshi A, Kazemi A, Arab A (2016) Identification and prioritization of supplier's resiliency evaluation criteria based on BWM. J Ind Manag Persp 6(3):159–186
10. Jain R, Singh AR, Mishra PK (2013) Prioritization of supplier selection criteria: A fuzzy-AHP approach. MIT Int J Mech Eng 3(1):34–42
11. Susanty A, Puspitasari D, Hartini S, Sugi G (2014) Prioritization of supplier selection criteria in batik industry: a fuzzy-AHP approach. Int J Anal Hiera Proc 29:1–7
12. Zeng S, Peng X, Baležentis T, Streimikiene D (2019) Prioritization of low-carbon suppliers based on Pythagorean fuzzy group decision making with self-confidence level. Eco Res Ekonomska istraživanja 32(1):1073–1087
13. Sharmin I, Nafisa U, Das A, Adnan TM (2018) Supplier prioritization to manage logistics risk factors using AHP tool: an empirical case analysis. J Ind Mech 3(3)

14. Kahraman C, Onar SC, Öztayşi B (2022) A novel spherical fuzzy CRITIC method and its application to prioritization of supplier selection criteria. J Intell Fuzzy Syst (Preprint)1–8
15. Kumar SK, Tiwari MK, Babiceanu RF (2010) Minimisation of supply chain cost with embedded risk using computational intelligence approaches. Int J Prod Res 48(13):3717–3739
16. Foerstl K, Reuter C, Hartmann E, Blome C (2010) Managing supplier sustainability risks in a dynamically changing environment—Sustainable supplier management in the chemical industry. J Purch Supply Manag 16(2):118–130
17. Hallikas J, Virolainen VM (2004) Risk management in supplier relationships and networks. Supply Chain Risk 43–65
18. Hallikas J, Karvonen I, Pulkkinen U, Virolainen VM, Tuominen M (2004) Risk management processes in supplier networks. Int J Prod Econ 90(1):47–58
19. Hallikas J, Puumalainen K, Vesterinen T, Virolainen VM (2005) Risk-based classification of supplier relationships. J Purch Supply Manag 11(2–3):72–82
20. Buzdoğan-Lindenmayr E, Güray KA, Selcuk AS (2019) Assessment of supplier risk for copper procurement. Commun Faculty Sci Univ Ankara Series A1 Mathem Stat 68(1): 1045–1060
21. Barbarosoglu G, Yazgac T (1997) An application of the analytic hierarchy process to the supplier selection problem. Prod Invent Manag J 38(1):14
22. Yadegari D, Darestani SA (2021) Supplier evaluation with order allocation in mega-projects. Manag Res Rev 44(8):1157–1181
23. Chan FT, Kumar N, Tiwari MK, Lau HC, Choy K (2008) Global supplier selection: a fuzzy-AHP approach. Int J Prod Res 46(14):3825–3857
24. Ghadge A, Jena SK, Kamble S, Misra D, Tiwari MK (2021) Impact of financial risk on supply chains: a manufacturer-supplier relational perspective. Int J Prod Res 59(23):7090–7105
25. Igoulalene I, Benyoucef L, Tiwari MK (2015) Novel fuzzy hybrid multi-criteria group decision making approaches for the strategic supplier selection problem. Expert Syst Appl 42(7):3342–3356
26. Tang CS (2006) Perspectives in supply chain risk management. Int J Prod Econ 103(2):451–488
27. Kleindorfer PR, Saad GH (2005) Managing disruption risks in supply chains. Prod Oper Manag 14(1):53–68
28. Babich V, Li H, Ritchken P, Wang Y (2012) Contracting with asymmetric demand information in supply chains. Eur J Oper Res 217(2):333–341
29. Nagali V, Hwang J, Sanghera D, Gaskins M, Pridgen M, Thurston T, Mackenroth P, Branvold D, Scholler P, Shoemaker G (2008) Procurement risk management (PRM) at Hewlett-Packard company. Interfaces 38(1):51–60
30. Hong Z, Lee CK, Zhang L (2018) Procurement risk management under uncertainty: a review. Ind Manag Data Syst 118(7):1547–1574
31. Manea I, Popa I (2010) Risk management in public procurement process. Stud Sci Res Eco Edit 15(15)
32. Aouam T, Rardin R, Abrache J (2010) Robust strategies for natural gas procurement. Eur J Oper Res 205(1):151–158
33. Arnold J, Minner S, Eidam B (2009) Raw material procurement with fluctuating prices. Int J Prod Econ 121(2):353–364
34. Hong Z, Lee CK (2013) A decision support system for procurement risk management in the presence of spot market. Decis Support Syst 55(1):67–78
35. Wong JT (2020) Dynamic procurement risk management with supplier portfolio selection and order allocation under green market segmentation. J Clean Prod 253:119835
36. Ji X, Huang S, Grossmann IE (2015) Integrated operational and financial hedging for risk management in crude oil procurement. Ind Eng Chem Res 54(37):9191–9201
37. Lenderink B, Halman JI, Boes J, Voordijk H, Dorée AG (2022) Procurement and innovation risk management: how a public client managed to realize a radical green innovation in a civil engineering project. J Purch Supply Manag 18:100747
38. Nahavandi N, Tavakoli P (2022) Risk management of procurement processes in automotive supply chain; Bahman motor company. Int J Ind Eng 29(1)

39. Murray JG (2013) Debate: the need for procurement risk management. Public Money Manag 33(4):285–287
40. Jafarnejad A, Ebrahimi M, Abbaszadeh MA, Abtahi SM (2014) Risk management in supply chain using consistent fuzzy preference relations. Int J Acad Res Bus Soc Sci 4(1):77
41. Sharma SK, Singh R, Matai R (2018) Force field analysis of Indian automotive strategic sourcing risk management enablers and barriers. Meas Bus Excell 22(3):258–275
42. Zsidisin GA, Panelli A, Upton R (2000) Purchasing organization involvement in risk assessments, contingency plans, and risk management: an exploratory study. Supply Chain Manag Int J 5(4): 187–198
43. Faisal M (2009) Prioritization of risks in supply chains. Managing supply chain risk and vulnerability, Springer, London, pp 41–66
44. Kern D, Moser R, Hartmann E, Moder M (2012) Supply risk management: model development and empirical analysis. Int J Phys Distrib Logist Manag 42(1):60–82
45. Elmsalmi M, Hachicha W (2013) Risks prioritization in global supply networks using MICMAC method: a real case study. 2013 international conference on advanced logistics and transport, IEEE, pp 394–399
46. Osipova E, Eriksson PE (2011) How procurement options influence risk management in construction projects. Constr Manag Econ 29(11):1149–1158
47. Kalvet T, Lember V (2010) Risk management in public procurement for innovation: the case of Nordic–Baltic Sea cities. Innov Eur J Soc Sci Res 23(3): 241–262
48. Federgruen A, Yang N (2011) Procurement strategies with unreliable suppliers. Oper Res 59(4):1033–1039
49. Lee CY, Li X, Xie Y (2013) Procurement risk management using capacitated option contracts with fixed ordering costs. IIE Trans 45(8):845–864
50. Dohale V, Akarte M, Gupta S, Verma V (2021) Additive manufacturing process selection using MCDM. Advances in Mechanical Engineering, Springer, Singapore, pp 601–609
51. Matook S, Lasch R, Tamaschke R (2009) Supplier development with benchmarking as part of a comprehensive supplier risk management framework. Int J Oper Prod Manag 29(3):241–267
52. Kuo TC, Chu CH (2013) Risk management of hazardous substances in selection of green suppliers. Int J Precis Eng Manuf 14(6):1057–1063
53. Belhadi A, Kamble SS, Mani V, Venkatesh VG, Shi Y (2021) Behavioral mechanisms influencing sustainable supply chain governance decision-making from a dyadic buyer-supplier perspective. Int J Prod Econ 236:108136
54. Raut RD, Bhasin HV, Kamble SS, Banerjee S (2011) An integrated fuzzy-AHP-LP (FAHLP) approach for supplier selection and purchasing decisions. Int J Serv Operat Manag 10(4):400–425
55. Swarnakar V, Bagherian A, Singh AR (2022) Prioritization of critical success factors for sustainable Lean Six Sigma implementation in Indian healthcare organizations using best-worst-method. TQM J (ahead-of-print). https://doi.org/10.1108/TQM-07-2021-0199
56. Jagtap M, Kamble SS, Raut RD (2017) An empirical examination of the procurement-led supply chain in the construction industry. Int J Proc Manag 10(1):1–20
57. Verma VK, Kamble SS, Ganapathy L, Belhadi A, Gupta S (2022) 3D Printing for sustainable food supply chains: modelling the implementation barriers. Int J Log Res Appl 25(2):1–27
58. Pellegrino R, Costantino N, Tauro D (2020) The role of risk management in buyer-supplier relationships with a preferred customer status for total quality management. TQM J 32(5):959–981
59. Kaur H, Singh SP (2021) Multi-stage hybrid model for supplier selection and order allocation considering disruption risks and disruptive technologies. Int J Prod Econ 231:107830
60. Hallikas J, Lintukangas K, Kahkonen AK (2020) The effects of sustainability practices on the performance of risk management and purchasing. J Clean Prod 263:121579
61. Giannakis M, Dubey R, Vlachos I, Ju Y (2020) Supplier sustainability performance evaluation using the analytic network process. J Clean Prod 247:119439
62. Elmsalmi M, Hachicha W, Aljuaid AM (2021) Prioritization of the best sustainable supply chain risk management practices using a structural analysis based-approach. Sustainability 13(9):4608

Chapter 30
Development of an Integrated Customer Relationship Management Tool for Predictive Analytics in Supply Chain Management

S. N. Dhisale, V. B. Nikumbhe, P. P. Kerkar, H. P. Pinge, A. D. Revgade, and U. A. Dabade

Abstract Most businesses nowadays are designing their products and services with the customer in mind. Many businesses throughout the world are expected to shift from a product-centric to a customer-centric attitude. Customer relationships, experiences, and happiness are therefore crucial for any business's long-term existence, sustainability, and profitability in any industry, yet small businesses lack the resources and skills to succeed. An integrated decision-making framework is built in this present study, integrating diverse data mining methodologies from many fields. The primary goal of this research is to improve the application of predictive analytics in small and medium-sized organizations. The decision-making framework in the form of a Customer Relationship Management (CRM) tool for an online retail sector is the solution offered as part of this study. An integrated decision-making framework tool is built in the pretense of a predictive analytical CRM system, with seven core characteristics and more than thirty sub-activities. The seven features created as part of this study are Data Visualization and Analysis, Customer Segmentation, Customer Classification, Product Recommendation, Customer Linked Predictions, Sales Forecasting, and Forensic Analysis, since they are frequently requested by CRM tool users. It is built using a variety of data mining and machine learning methods. The tool is then made available as a real-time online application. This tool, which consists of a frontend and a backend application, is essentially designed to give users a complete picture of the data. Aside from that, there are several enhancements that can be made to this tool.

S. N. Dhisale (✉) · V. B. Nikumbhe · P. P. Kerkar · H. P. Pinge · A. D. Revgade
Bachelor of Technology in Mechanical Engineering, WCE, Sangli 416415, India
e-mail: dhisaleshreyas82@gmail.com

U. A. Dabade
Professor in Department of Mechanical Engineering, WCE, Sangli 416415, India
e-mail: uday.dabade@walchandsangli.ac.in

M. K. Tiwari et al. (eds.), *Applications of Emerging Technologies and AI/ML Algorithms*, Asset Analytics, https://doi.org/10.1007/978-981-99-1019-9_30

Keywords Customer relationship management · Predictive analytics · Data mining · Machine learning · Online retail · Supply chain management

30.1 Introduction

As the digital developments in the domains of Artificial Intelligence (AI), Internet of Things (IoT) are having an impact on different industries globally, putting too much emphasis on them can cause a company to lose sight of what matters most to them: their customers. American Express found that 33% of customers will consider switching companies after just one instance of poor customer service. As a response to this concern, nowadays, most of the companies are shifting from a product-focused view of the world to a customer-focused one. Companies are trying hard to devise strategies for providing customized services to their customers. Daqing et al. presented various simple data mining techniques that can be employed in relation to customer-centric business intelligence [1]. And this is the major driving force behind the usage of CRM and CRM systems by various industries.

CRM is a business strategy containing a set of practices that organizations adopt to maintain and increase their customer base and thus develop mutually profitable customer–supplier relationships as stated by Bin-Nashwan et al. [2]. According to research, it costs about 50% more to acquire new customers than it does to retain existing ones, and customer acquisition costs have climbed by nearly 60% in only five years. In short, customer relations can bring profits or losses as per the steps a company puts forth. Hence, a CRM approach and a dedicated system are of utmost importance to thrive in the market as mentioned by Srivastava [3]. Any CRM-based IT system can be regarded as a centralized database where all the crucial information related to customers of one's company is present.

But there are certain challenging issues related to CRM and systems based on it. CRM technology has yet to achieve the influence that it deserves in the industry. Around 22% of salespeople say they have no idea what a CRM system is, and 40% of businesses don't use one at all. Also, there exist other problems like lack of support, less local vendors, siloed departments, and most importantly improper data management and analysis. Therefore, Mishra et al. [4] proposed various data mining techniques which can be used for building a predictive analytics model. Moreover, CRM has only been perceived as a database and contact management tool with a complex interface to date. Despite many existing CRM systems, each system emphasizes different features to manage customer relationships and most of these systems can be afforded by big industries and firms only. This leads to the question, "How can a CRM system encompassing all the features to manage customer relationships be affordably developed?".

To address this question, an integrated decision-making framework tool is constructed in the guise of a predictive analytical model. This framework is beneficial for small- and medium-scale industries, specifically for the ones in the domain of online retail as the related dataset is considered while developing the model.

This paper is organized in different sections such as, section two includes the theoretical background of the CRM terminologies and the key features of the CRM system. Section three includes the methodology to develop a decision-making framework. In section four, experimental analysis and results have been discussed by considering a case study. Lastly, section five discusses the comprehensive summary consisting of research contributions and future scope.

30.2 Literature Review

30.2.1 CRM: Customer Relationship Management

CRM has become a critical topic among scholars since its inception as a study field in the 1990s [5]. A CRM system is a commercial technique for establishing mutually beneficial customer–supplier connections. Many studies suggest that firms may improve customer satisfaction [6], customer retention [7], customer loyalty [8], and company success [9] by implementing CRM. It is an industrial management technique that assists in the creation, development, and enhancement of relationships with precisely chosen customers to maximize their value to the firm and improve profits.

30.2.2 Supply Chain Management: Integration with CRM

Companies seldom combine Supply Chain Management (SCM) and CRM, according to Kracklauer et al. [10]. They claim that by integrating these business processes, firms may break through and reach a number of performance and financial metrics that would be impossible to attain with simply stand-alone CRM and SCM methodologies. The primary goal of CRM is to gain and keep lucrative long-term clients. Marketing, sales, and support workers, as well as top management, utilize CRM solutions to monitor and analyze crucial client connections [10]. Fuxiang and Yuhui [11] emphasized that enterprises must view their CRM operations through the lens of SCM.

30.2.3 Customer Segmentation and Classification: Grouping of Customers

Many strategies may be used to segment customers, but the most common are Recency-Frequency-Monetary (RFM) and k-means. According to Lin [12], RFM is a straightforward yet efficient approach for market segmentation that can be used to segment the customer base. RFM analysis increases market segmentation by looking at when (Recency), how often (Frequency), and how much money was spent on a given item or service (Monetary) [13]. The k-means technique is a non-hierarchical method that has become quite popular for classification due to its ease of setup and quick execution [14]. Market segmentation, pattern detection, and information retrieval are just a few of the applications [15].

30.2.4 Product Recommendation: Efficient Promotion of Products

Liu [16] developed a recommendation system for apparel fashion evaluation by proposing items using SVM algorithms. The filtering process employing the K-Nearest Neighbor (KNN) algorithm and matrix factorization was explored by Sharma et al. [17]. Between a collection of elements, association rules extract appropriate correlations, common patterns, relationships, or casual structures. Zhao et al. [18] used methodologies like Apriori, and Frequent Pattern-Tree algorithms to address classification and clustering in association rule mining techniques. Cheung et al. [19] proposed a distributed approach that generates a modest number of candidate sets and decreases the number of messages sent to the mining association.

30.2.5 Customer Linked Prediction

Net Profit Contribution Using machine learning models to estimate customer lifetime values (CLTV) and segmentation, Venkatakrishna et al. [20] explored marketing decision-making and future marketing strategies and plans. Using the Conway–Maxwell–Poisson (COM)-Poisson Distribution, Mohamed et al. [21] introduced the Pareto- Negative Binomial Distribution (NBD) and the Beta Geometric Negative Binomial Distribution (BG-NBD) models for simulating purchasing behavior. With BG-NBD and Pareto-NBD models, Jasek et al. [22] investigated the empirical statistical analysis and predicting capacities of chosen probabilistic Customer Lifetime Value (CLV) models.

Sales Forecasting The practice of sales forecasting allows a company to anticipate future revenues. Cash flow planning, inventory planning, production scheduling, and revenue projections, etc. all rely on forecast prediction [23]. Forecasting may be done using a variety of methodologies, including qualitative techniques, time series analysis and projection, and causal models [24]. Autoregressive integrated moving average (ARIMA) models, according to Box and David [25], are a type of statistical model that is used to analyze and forecast time series data. Many additional machine learning models, such as Long-Short-Term-Memory (LSTM), have also been widely used since they deal with non-linear data. Saxena et al. [26] did research

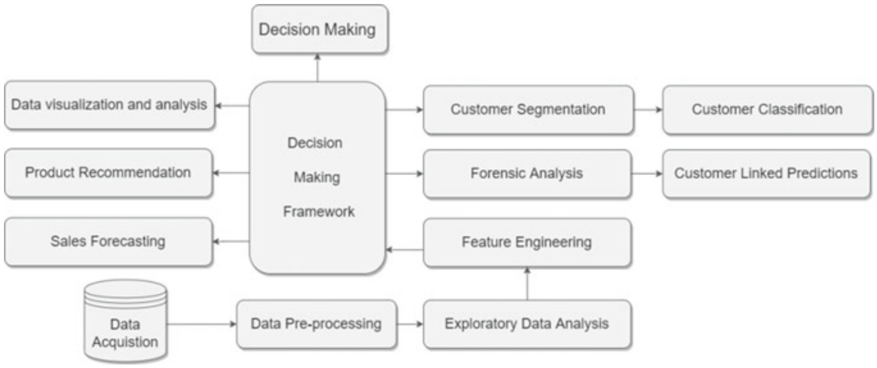


Fig. 30.1 Decision-making framework highlighting major key components

on deep learning-based models for anticipating future automobile sales trends and the results show that LSTM-Recurrent Neural Network (RNN) outperforms ARIMA for multivariate datasets.

30.3 Decision-Making Framework

30.3.1 Model Overview

The approach used in this study is described in this section, as shown in Fig. 30.1. The crucial seven components of an integrated decision-making framework tool are constructed which are data preparation and analysis, customer segmentation, customer classification, product recommendation, customer-linked predictions, sales forecasting, forensic analysis. The integrated decision-making framework tool uses seven major components and over 30 sub-components to make suitable judgments. Except for forensic analysis, all features are illustrated in the following sections.

30.3.2 Data Preparation

This study focuses on customers who have previously bought items to complete the seven critical components. Exploring dataset attributes, identifying and handling numerical and categorical variables, missing, sparse, duplicate, and inconsistent data types, as well as dealing with feature transformation, feature scaling, feature generation, and data visualization are all part of data preparation.

Table 30.1 Customer segmentation using hybrid modeling (K-means with RFM)

Method: K-means with RFM-based customer segmentation approach
Input: Customer ID, transaction date, quantity, and price of purchased product **Output**: Customer segmentation **Start** 1. To calculate Recency (amount of time elapsed since the last date of a customer's purchase), Frequency (total number of transactions), and Monetary (total amount paid up to the last date of purchase) and generate an RFM matrix for each customer 2. To create plots to infer the relationship between each of the RFM metrics 3. Generate RFM score and correlate it with labeled segmentation and then segment customers based on it 4. To re-create an RFM matrix suitable for k-means clustering 5. Find the optimum number of clusters using the Elbow Method and create plots to visualize the distribution of customers in labeled segments and clusters
End

30.3.3 Customer Segmentation and Classification

This section proposes a hybrid technique that combines RFM with k-means clustering. The customer segments are calculated using this approach, which combines the RFM score and the k-means algorithm. For example, the strategy outlined in Table 30.1 can be used to segment customers.

Support Vector Classifier (SVC), Logistic Regression, K-NN, Decision Tree, Random Forest, AdaBoost Classifier, and Gradient Boosting Classifier are some of the classifiers utilized in our decision-making system to categorize clients. It will also allow users to choose the best classifier based on its ability to forecast, quality of fit, and user needs.

To do so, first split the labeled dataset generated by segmentation into a train-test format, then define the associated classifiers, train and test them further, and then build the confusion matrix, which will be used to validate the findings of the various classifiers. After that, compute recall, precision, accuracy, and area under the receiver operating characteristic curve by viewing the classifiers' performance. Finally, utilize a soft classifier to select the best classifier based on predictive ability and fit quality.

30.3.4 Product Recommendation

The association rule is one such proposed collaborative filtering mechanism. The relationship between two or more items is described by association rules. The product suggestion using Apriori, for example, is shown in Table 30.2.

Understanding the importance of support, lift, and confidence may assist companies in determining which products will sell the best and so improve earnings.

Table 30.2 Product recommendation

Method: Association rules using the Apriori algorithm
Input: Identification number of product, label of the product **Output**: Association rules **Start** 1. To format the pre-processed dataset by retaining the attributes related to the Identification number of the product and the label of the product 2. To train the model using the Apriori algorithm and generate association rules 3. To compute the product's Support, Confidence, and Lift parameters 4. To create filtered association rules and product recommendation table
End

Table 30.3 Expected future transactions and company's net profit contribution

Method: BG-NBD model and GG model approach
Input: Customer ID, transaction date, quantity, and price of purchased product **Output**: Expected future transactions for a specific time period, expected average profit, CLV, customer segments based on scaled CLV **Start** 1. To compute Recency, Frequency, Monetary, and Time period (T) for each customer 2. To determine coefficients of BG-NBD and GG Model by computing equations for various values of variables 3. To predict expected transactions for a specific time period (1 week/1 month) by inserting the coefficient and variable values for each customer 4. To calculate the expected average profit for each customer by inserting the coefficient and variable values for each customer 5. To compute CLV, according to the standard equation for each customer 6. To scale the CLV in a fixed range (0–1) and segment the customer
End

30.3.5 Customer Linked Predictions

The BG-NBD model is used in the decision-making framework to forecast predicted future transactions that will be completed by consumers in a particular time period. The Gamma Geometric (GG) Model is used to anticipate a company's net profit contribution to its total future relationship with a client.

For example, the approach stated in Table 30.3 can be used for predicting the number of transactions for a given time period and to anticipate a company's net profit contribution throughout the course of a customer's lifetime.

30.3.6 Sales Forecasting

In this work, time series forecasting models are used to anticipate future sales based on previous sales trends found in the dataset. The ARIMA and LSTM models were

Table 30.4 Sales forecasting

Method: LSTM approach
Input: Customer ID, transaction date, total price of the purchased product **Output**: Forecasted values of future sales **Start** 1. To check the data for stationarity and seasonality by Augmented Dickey–Fuller Test (ADF Test) 2. To compute ADF value, P-Value, number of lags, and number of observations used for ADF Regression and comparing P values to analyze data 3. To split the dataset into train and test data, then train the model in LSTM 4. To forecast the future sales for a certain time period and compute performance metrics like prediction accuracy and RMSE score
End

utilized in this part, and after comparing the findings, the LSTM model was shown to be more accurate on bigger sequential data. Also, presented is a system for forecasting future sales. The approach presented in Table 30.4 can be used, for example, to forecast sales for a future time period.

30.3.7 Web Application Overview

Under the name *Grahak360,* the decision-making framework created as part of this project has been implemented as a web application interface as depicted in Fig. 30.2. It comes with a dashboard that graphically shows data and charts, as well as provides the necessary forecasts and analysis. It is also intended to be user-friendly and simple to comprehend.

Fig. 30.2 Home page of the deployed web-application interface *Grahak 360*

Table 30.5 Attribute description in the customer transaction dataset

Attribute	Description; meaning and values	Data type
Bill	A 6-digit unique bill number assigned to each transaction. (e.g., 540,026)	Categorical
MerchandiseID	A unique number assigned to each distinct product. (e.g., 21,519)	Categorical
Product	Name of the product	Categorical
Quota	Quantity of each product per transaction	Numerical
BillDate	Billing date of transaction. (e.g., 04/01/2019)	Numerical
Amount	Product price per unit in Pound sterling. (e.g., £2.85)	Numerical
CustomerID	A 5-digit unique number assigned to each customer. (e.g., 14,031)	Numerical
Country	Name of the country where the customer resides. (e.g., United Kingdom)	Categorical

30.4 Experimental Analysis and Results

Each of the seven characteristics yielded appropriate results. However, all analyses and outcomes are shown in the following section, apart from customer-related prediction, product suggestion, and forensic analysis.

30.4.1 A Case Study

Business Background The online shop in interest is a UK-based, non-store firm that has been registered since 1981 [1]. The merchant's customer transaction dataset has 8 variables and a total of 1,067,371 transactions. The transactions in the dataset are from December 1, 2017 to December 9, 2019**.** The 8 attributes in the dataset are elaborated in terms of descriptions and their data type as shown in Table 30.5.

30.4.2 Feature Engineering Analysis and Results

Feature Engineering The attributes *CustomerID* and *Product* both have missing data, accounting for 22% and 0.4% of the total data points, respectively. On average, about ten units of quota per product were sold. At a maximum price of £ 6.15, 75% of purchasers purchased less than or equal to 10 items. The cleansed data has 824,364 rows and 8 columns after missing values were deleted. The dataset had 15,487 duplicate rows, accounting for less than 2% of the total data points.

30.4.3 Customer Segmentation and Classification Results

Customer Segmentation Customer segmentation is done using RFM and Hybrid (K-means plus RFM) modeling approaches.

RFM. A scatter plot was used to infer the association between each of them. So, based on the relationship between recency and frequency, this graph indicates that the majority of newly recruited consumers buy regularly.

Customers are labeled in Table 30.6 based on the 10 segments indicated in Fig. 30.3.

Hybrid Customer Segmentation These scatter plots may be used to compare clusters such as recency versus frequency, monetary versus recency, and monetary versus frequency. The recency versus frequency Scatter figure below shows that the three separate groups are scattered throughout a wide variety of clients. The green clusters reflect customers that have a high recency but a low frequency. Customers with a high frequency but low recency are represented by the yellow cluster, whereas customers with a mid-range frequency and recency are represented by the purple cluster as inferred from Fig. 30.4.

Similarly, a 2D depiction of clusters can be shown in terms of monetary versus frequency.

Table 30.6 Customer IDs being labeled with 10 correlations of segmentations

Sr. No	Customer ID	Recency	Frequency	Monetary	R	F	M	RFM Score	Segment
0	12,346	326	47	452,198.66	2	3	5	235	At Risk
1	12,347	3	240	11,659.22	5	5	5	555	Champions
2	12,348	76	51	7447.40	3	3	5	335	Need attention
3	12,350	311	17	728.40	2	2	2	222	Hibernating

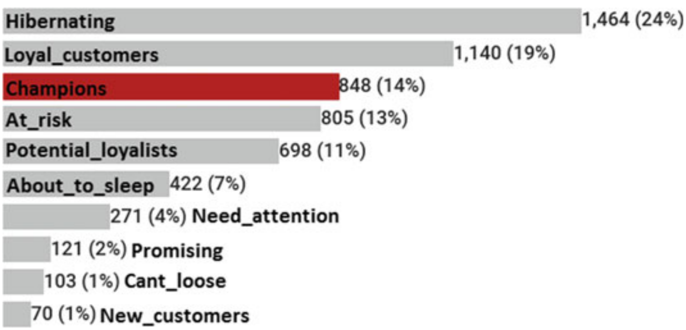


Fig. 30.3 Number of customers in each segment

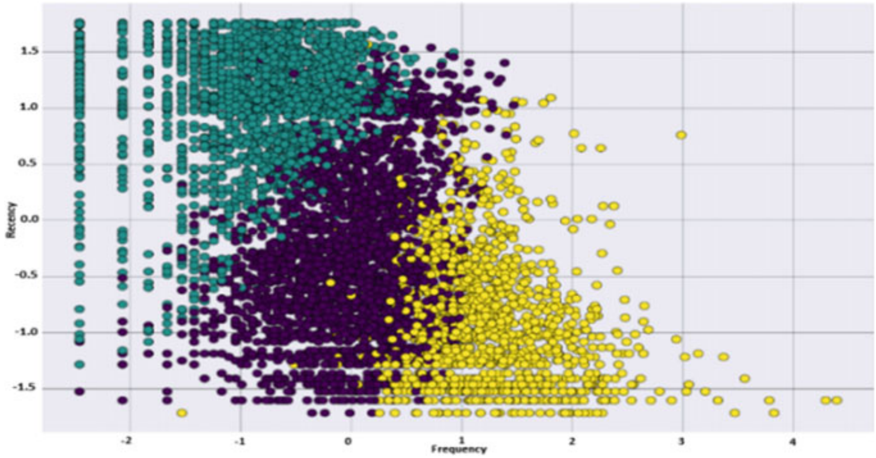


Fig. 30.4 The 2D representation of clusters in terms of recency versus frequency

Table 30.7 Precision of respective classifiers

Classifier	Precision	Classifier	Precision
Support vector machine	99.78%	Decision tree	99.23%
Logistic regression	99.90%	Random forest	99.58%
k-nearest neighbors	98.67%	Gradient boosting	99.63%

Customer Classification A train test is created from the labeled dataset obtained via segmentation. The classifier is then built using the necessary code. The data is then examined and validated using several classifiers, with the best classifier being chosen based on precision, accuracy, predictability, and quality of fit in learning curves.

Soft Vote Classifier Result All of the classifiers are evaluated in the analysis presented in Table 30.7 and the best classifier suited for the operation is chosen.

30.4.4 Sales Forecasting Results

The data is mostly fluctuating; therefore, time series analysis is done. Future Sales Prediction using ARIMA Model has been calculated. The ARIMA prediction accuracy of 19% (as calculated through RMSE metrics) denotes unsatisfactory results.

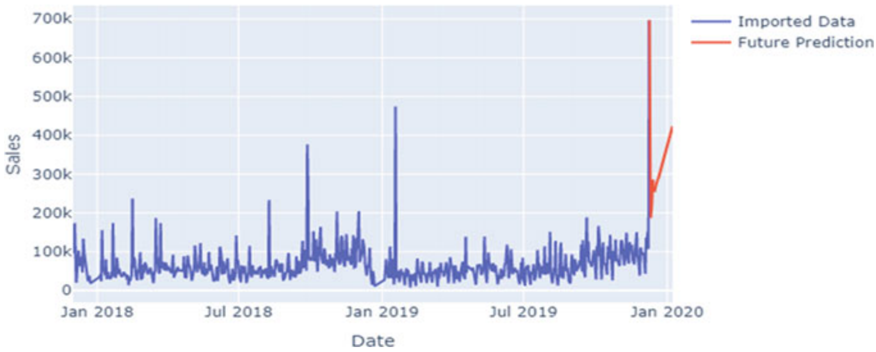


Fig. 30.5 Future sales prediction using LSTM model

Figure 30.5 indicates future sales prediction using the LSTM Model with a prediction accuracy of 77% (as calculated through RMSE metrics) denoting satisfactory results.

30.5 Conclusion

This paper demonstrates an integrated decision-making framework tool that is constructed in the guise of a predictive analytical model. Data visualization and analysis, customer segmentation, customer categorization, product suggestion, customer-linked predictions, sales forecasting, and forensic investigation are the seven essential characteristics of this platform. With all of these capabilities in place on the tool's backend, an operational interface is created on the front to assist decision-makers in making informed decisions.

30.5.1 Research Contributions

The proposed decision-making framework operates the tool such that it can guide one to get a panorama of data. It aids in customer segmentation and categorization, aids in the creation of association rules for product suggestions, aids in the prediction of future transactions and the contribution of one's company's net profit to an entire future connection with consumers, and much more. Our suggested framework-based tool is more comprehensive, configurable, and collaborative than previous studies by many scholars and researchers in this field.

30.5.2 Potential Future Research

Despite these important contributions, the suggested framework-based tool has notable shortcomings that serve as a springboard for future research. The input dataset must be in a well-defined format to run the tool effectively, and the time it takes to create results is dependent on the number of iterations and the size of the dataset.

The initially focused research area would be to include more capabilities to process datasets involving different organizational levels to assist decision-makers in gaining insights and creating transparency via proper decision-making. The second potential approach would be to develop an algorithm to change the tool's user interface based on the domain of usage and make it more flexible.

Acknowledgements The authors express their gratitude to the anonymous reviewers for their insightful remarks and ideas that helped improve the quality of this paper.

References

1. Chen D, Sain SL, Guo K (2012) Data mining for the online retail industry: a case study of RFM model-based customer segmentation using data mining. J Database Mark Cust Strategy Manag 19:197–208. https://doi.org/10.1057/dbm.2012.17
2. Nashwan SA, Hassan H (2017) Impact of customer relationship management (CRM) on customer satisfaction and loyalty: a systematic review. J Adv Res Bus Manag Stud 6:86–107
3. Srivastava M (2012) Customer relationship management (CRM): a technology driven tool. SAMVAD 4:14–25
4. Mishra N, Silakari S (2012) Predictive analytics: a survey, trends, applications, oppurtunities & challenges. Int J Comput Sci Inf Technol 3:4434–4438
5. Hannu S, Karjaluoto H (2015) Customer relationship management: the evolving role of customer data. In: Robinson L (ed) Marketing dynamism & sustainability: things change, things stay the same…. Springer International Publishing, Cham, pp. 505–515
6. Mithas S, Krishnan M, Fornell C (2006) Why do customer relationship management applications affect customer satisfaction? Inf Syst & Econ eJ 69. https://doi.org/10.1509/jmkg.2005.69.4.201
7. Verhoef PC (2003) Understanding the effect of customer relationship management efforts on customer retention and customer share development. J Mark 67:30–45. https://doi.org/10.1509/jmkg.67.4.30.18685
8. Reinartz WJ, Kumar V (2000) On the profitability of long-life customers in a noncontractual setting: an empirical investigation and implications for marketing. J Mark 64:17–35. https://doi.org/10.1509/jmkg.64.4.17.18077
9. Moorman C, Boulding B, Staelin R, Commentators R, Chandy M, Dekimpe A, Guha J, Francis W, Hoyer J, Irwin M, Lovett J, Lynch V, Mahajan A, Rangaswamy R, Rust R, Srivastava R, Varadarajan R (2005) Strategic firm commitments and rewards for customer relationship management in online retailing. J Market 69
10. Kracklauer AH, Mills DQ, Seifert D (eds) (2004) The integration of supply chain management and customer relationship management. In: Collaborative customer relationship management: taking CRM to the next level. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 57–69
11. Fuxiang LIU, Yuhui YOU (2011) Study and explores on CRM based on the supply chain integration. Manag Sci Eng 5:1–9

12. Wei J-T, Lin S-Y, Wu H-H (2010) A review of the application of RFM model. Afr J Bus Manag Dec Spec Rev 4:4199–4206
13. Yang AX (2004) How to develop new approaches to RFM segmentation. J Target Meas Anal Mark 13:50–60. https://doi.org/10.1057/palgrave.jt.5740131
14. Cheung Y (2002) k *-means—a generalized k -means clustering algorithm with unknown cluster number. Presented at the June
15. Wu H-H, Shieh J-I, Liao AYH, Lin S-Y (2008) An application of the generalised K-means algorithm in decision-making processes. Int J Operat Res 3:19–35
16. Zhang Y, Liu X, Shi Y, Guo Y, Xu C, Zhang E, Tang J, Fang Z (2017) Fashion evaluation method for clothing recommendation based on weak appearance feature. Sci Program 2017:8093057. https://doi.org/10.1155/2017/8093057
17. Sharma J, Sharma K, Garg K, Sharma AK (2021) Product recommendation system a comprehensive review. IOP Conf Ser: Mater Sci Eng 1022:12021. https://doi.org/10.1088/1757-899x/1022/1/012021
18. Zhao Q, Bhowmick S (2003) Association rule mining: a survey
19. Cheung DW, Han J, Ng VT, Fu AW, Fu Y (1996) A fast distributed algorithm for mining association rules. In: Fourth international conference on parallel and distributed information systems. pp 31–42
20. Venkatakrishna MR, Mishra MP, Tiwari MSP, Customer lifetime value prediction and segmentation using machine learning
21. ben Mzoughia M, Limam M (2014) An improved BG/NBD Approach for modeling purchasing behavior using COM-poisson distribution. Int J Model Optim 4. https://doi.org/10.7763/IJMO.2014.V4.362
22. Jašek P, Vraná L, Sperkova L, Smutny Z, Kobulsky M (2019) Comparative analysis of selected probabilistic customer lifetime value models in online shopping. J Bus Econ Manag 20:398–423. https://doi.org/10.3846/jbem.2019.9597
23. Vavliakis KN, Siailis A, Symeonidis AL (2021) Optimizing sales forecasting in e-commerce with ARIMA and LSTM models
24. Chambers JC, Mullick SK, Smith DD (1971) How to choose the right forecasting technique. Harvard University, Graduate School of Business Administration Cambridge, MA
25. Box GEP, Pierce DA (1970) Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. J Am Stat Assoc 65:1509–1526. https://doi.org/10.2307/2284333
26. Saxena P, Bahad P, Kamal R (2020) Long short-term memory-RNN based model for multivariate car sales forecasting. Int J Adv Sci Technol 29:4645–4656

Chapter 31
Predicting Top Companies Amid Changing Macro Environment—Optimal Sampling Imposing Restriction Filters

Selvan Simon and Hema Date

Abstract Predicting top-performing companies is an essential part of long-term business-to-business relationships. For example, supplier selection, sourcing risk management, enterprise loan customer selection, and stock selection. Moreover, long-term financial decisions involve analyzing the microeconomies of the companies. In other words, ratio analysis helps us to diagnose the health profiles of companies to select loan customers, equities, suppliers, vendors, producers, buyers, sellers, and others, for any business collaboration. However, the financial data by its nature is noisy, and the relationships among fundamental variables are dynamic due to the external influence of the changing macroeconomies. This study reviews optimal sampling techniques used in machine learning models for mitigating the problems caused by changing economic environments. It brings some immensely important sampling designs to develop models for ration analysis in decision support systems. Filtering noise due to external disturbances necessarily imposes constraints on the industry, size, consistency, history, performance, and the environment of the company. The moving window system allows us to retrain the models with more recent environments. It makes the model comparison difficult, an essential step of model validation. However, in real-world situations, models must prove to be more profitable rather than more accurate for usability.

Keywords Ratio analysis · Financial health · Long-term decisions · Classification of companies · Moving window system · Optimal training sample

31.1 Introduction

Machine Learning (ML) is considered a subfield of *Artificial Intelligence* (AI). It uses many algorithms for the development of *models* that enable computers to learn and perform tasks. For example, *classification* is a learning task where some training

S. Simon (✉) · H. Date
National Institute of Industrial Engineering, Mumbai, India
e-mail: selvansimon@outlook.com

M. K. Tiwari et al. (eds.), *Applications of Emerging Technologies and AI/ML Algorithms*, Asset Analytics, https://doi.org/10.1007/978-981-99-1019-9_31

samples with known classes or categories are provided for training a model, called *classifier*. Then, the model is used to *predict* the class they belong to for unseen new samples. The task may be a simple *binary* classification with exactly two classes or a complex multiclass classification with several classes. A classifier is said to *generalize* well if it can predict the correct classes well on a set of unseen test samples [25]. *Artificial Neural Network* (ANN) is the most dominant ML tool used in stock prediction [23, 24]. ML models employed ANN in *ratio analysis* for the financial health classification of companies for long-term decisions [22].

Predicting top-performing companies enhances the *return on investment* (ROI) in demand and supply models. For example, supply chain management is essentially chaining business to business in provider and procurer roles in various stages. To maintain the smooth functioning of business, it is essential to select partnering companies, at every stage. According to ROSMA (*Return on Supply Management Assets*) Performance Check Report 2016, ROSMA standard classifies the procurement profession as a top-tier group of standout performers, a middle tier performing well below the top tier, and a large group of bottom-quartile performers. The top-quartile performers deliver more than seven times their cost and investment base in procurement [20].

31.1.1 Predicting Top Companies to Win Market

Models applying *fundamental analysis* are based on the financial performance of the companies. They predict top-performing companies that are more likely to achieve higher ROI in the subsequent year [10]. They inspect the financial health of the companies using fundamental variables to classify them [1, 2, 6]. The portfolio managers can buy top-class stocks [9, 11, 18, 21, 26] for constructing portfolios to outperform the market in the long term.

31.1.2 Machine Learning in Predicting Top Companies

Unlike umpteen ML models for technical analysis, very few studies [3, 4, 7, 9–13, 17–19, 26, 28, 29] employed ML models for predicting top companies to enhance the stock selection. They applied fundamental analysis for improving the ROI of portfolios. The models predicted top-performing stocks from a set of listed companies for investment in the subsequent year or quarter.

31.1.3 Constrained Sample in Financial Health Prognosis

The amount of the macro environment's influence depends on how much of a company's business is dependent on the health of the overall economy. For instance, cyclical industries are heavily influenced, while basic staple industries are less influenced (for details read "Cyclical vs. Non-Cyclical Stocks: What's the Difference?"). Situations may worsen when the predictive analysis includes various industrial sectors which are affected by the environments to varying degrees. The weak relationships among financial variables in the noisy financial data display dynamic patterns in the changing market environment. Thus, studies may impose constraints on sampling based on industry type, market capitalization, data availability, historical period, performance slab, and economic environments. Filters on training and testing samples help us to optimize the search among fewer more relevant patterns.

31.2 Restriction Filters for Datasets

It is common to impose restrictions on the companies to be included in ratio analysis for financial health classifications. This section reviews the restriction filters on companies imposed by ML models, for all financial years.

31.2.1 Filter on Industry

Studies may impose filters on industry to maintain homogeneity in the financial characteristics of the underlying assets.

Yildiz [28] used a dataset comprising 958 firms listed on Istanbul Stock Exchange during 1992–1999. It included manufacturing, commercial, and service firms but excluded financial institutions, holdings, and transportation companies due to their different financial characteristics. Firms with different financial reporting periods were also deleted from the final dataset.

Becker et al. [3] developed stock selection models for S&P 500. They included monthly financial data and stock returns of 350 investable stocks. They excluded the financials and utilities sectors from the study to maintain the homogeneity of the underlying assets.

Liu et al. [14] included the financial data of 629 electronic firms listed on the Taiwan Stock Exchange (TSE) and Over the Counter (OTC) in 2004 and 2005 in their revenue growth rate forecasting model.

31.2.2 Filter on Capitalization

Studies may impose filters on the size of the companies to maintain consistency in the sample.

Olson and Mossman [17] obtained an initial sample of 4750 observations of firms traded on the Toronto Stock Exchange during 1976–1993 from databases. They imposed a set of restrictive filters on companies as there were many ratios and few returns missing for some smaller companies. Thus, the final sample included 2352 observations only from medium to large companies, with annual sample size ranging from 79 in 1976 to 213 in 1993.

Huang et al. [10] included only 200 companies with the largest market capitalizations listed on the TSE as investment universe.

31.2.3 Filter on Completeness

Studies may impose filters on the completeness of data in standard databases to reduce noise.

Fan and Palaniswami [7] used financial indicators of the stocks listed on the Australian Stock Exchange during 1992–2000. To reduce noise and maintain consistency, they included only data from annual reports and discarded the rows with more than one missing variable. The final dataset consisted of 273–537 rows each year.

Feng and Kong-lin [8] developed a model for enterprise credit risk classification. They selected the eastern large and medium-sized loan enterprises from a loan database. However, they deleted the firms whose financial data were not available.

Quah [18] used data from 1630 DJIA equities traded during 1995–2004. The author included the delisted companies to avoid survivor bias. However, the author deleted the entries with missing values for more than half of the features to reduce noise.

Yildiz and Yezegel [29] included all the firms traded in NYSE, AMEX, and Nasdaq exchanges over the fiscal years from 1962 to 2005. Their aim was to overcome the limitation of using a sample from a single stock exchange. Nevertheless, they excluded the companies with missing data from their experiments.

31.2.4 Filter on Multiple Constraints

Studies may impose multiple filters on different criteria to make the dataset more compact with fewer patterns.

Min and Lee [15] included all 944 non-bankrupt firms in 2002 and 944 bankrupt firms during 2000–2002. All those firms were from the Korean medium-size heavy industry.

Min et al. [16] included 614 Korean companies, among them 307 were bankrupt and filed for bankruptcy between 1999 and 2002. All the firms were from the medium-sized light industry.

Yeh et al. [27] used a sample consisting of information and electronic manufacturing firms listed on TSE. They initially considered a set of variables following literature and industry experts. However, they deleted those variables with missing values.

31.2.5 Filter on History

Obviously, including very old historical records may not be useful due to many new developments in technology. Therefore, it is normal that researchers include the historical data of the recent past 10 to 15 years in their experiments. Thus, applying the above restriction filters, we can obtain the datasets for 10 to 15 years. We denote them by $D0,\ D1,\ \ldots,\ D\mathrm{n}$. Where $D\mathrm{i}$ is the dataset of the ith financial year.

31.3 Restriction Filters for Training Sets

Adding a large training sample may not necessarily estimate more accurate models in case of noisy financial data. Therefore, to build a resilient model, it is necessary to impose some restriction filters on the training sample to reduce noise.

31.3.1 Filter on Performance

The fundamental analysts aim to buy (sell) stocks of companies with strong (weak) fundamentals like that of the top (bottom) performing companies. Essentially, the fundamental analysis is guided by extreme cases: identifying patterns from the historical records of companies that performed extremely well (top) to buy stocks and companies that performed extremely bad (bottom) to sell stocks. Therefore, for predicting top companies, we can include only companies with extreme performances as training examples from the dataset of each fiscal year. This process is known as *downsampling*. At the end of this process, we will have n sets of training samples obtained from the first n datasets. We denote them by $\tau 0,\ \tau 1,\ \ldots,\ \tau \mathrm{n}-1$. Where $\tau \mathrm{j}$ is the set of all training samples from the financial year j.

31.3.2 Filter on Environment

Stock selection models are required to classify the stocks correctly according to the future performances of their companies. However, the performance of the models may deteriorate in classifying the stocks when used in an unseen environment that is different from the training environment. Since the financial markets are extremely unstable the correlation between the data included in training and the future performance of a given stock may become weaker. Thus, a dynamic adaptation to the changing environments is very important in responding to market instability. Models may implement dynamic adaptation via retraining, using new data drawn from new environments. Ideally, the system should be retrained as and when an occurrence of a change in the market. However, it turns out to be very difficult to detect such points at which a market changes [9].

31.4 Moving Window System

Alternatively, for the purpose of constructing portfolios using annual rebalancing, stock selection models may employ *Moving Window System* which it retrains every year on the most recent n years of data. Using a moving window system, [19] assessed the generalization power of ANN using small sample drawn from recent environments. ANN showed its ability in deriving relationships in constrained environments through moving windows, making it recommended for stock selection by adapting to the changing market environments.

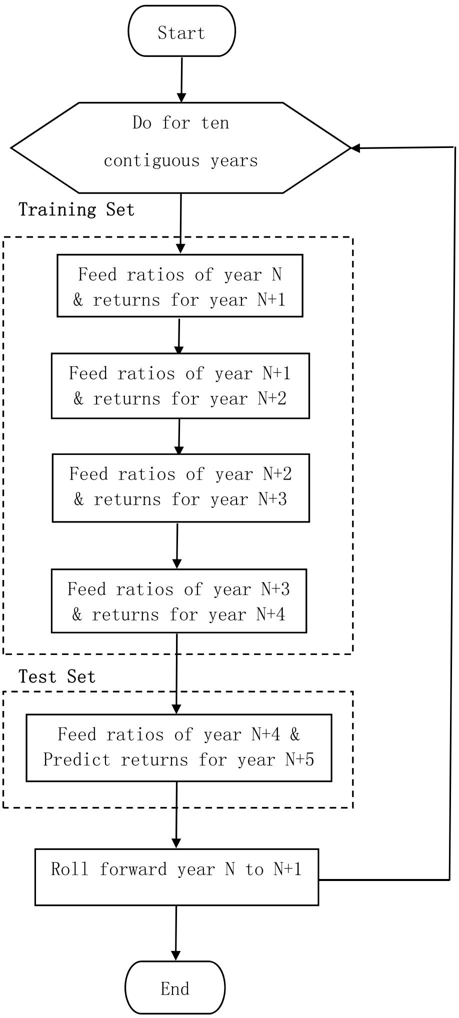
In the moving window system for annual stock selection with the size of training window $\eta = 1$, a training window, trailing with a testing window, recursively rolls forward in unison toward the successive target years. The financial ratios of year N and the stock returns of year $N + 1$ are used in retraining the model. The model is then applied to ratios of the year $N + 1$ to predict stock returns for the target year $N + 2$. Thus, a revised model is estimated for each testing year for a one-time application [12].

The training window recursively includes the historical records of recent past n-years until before the testing year in training sets. Thus, the retraining captures the latest relationship among the variables from their dynamic relationships amid the changing market environment [19].

31.4.1 Training and Testing Window

The training windows allow retraining of the models using recent past few years' historical records. Thus, the training set comprises training samples of the recent past few years obtained above and test samples of the subsequent year constitute the corresponding test set. For instance, in the case of the model illustrated above, the training samples of every contiguous recent past four years were included for retraining the model, and the dataset of the fifth year was used in testing them. This process is repeated ten times rolling forward 1 year each time to obtain a series of ten training sets for the model. Figures 31.1 and 31.2 give the flowchart and illustration of this process for a model with 4 years of records included in training sets ($\eta = 4$) with a prediction horizon of one year ($\theta = 1$).

Fig. 31.1 Training and test sets construction in a moving window system



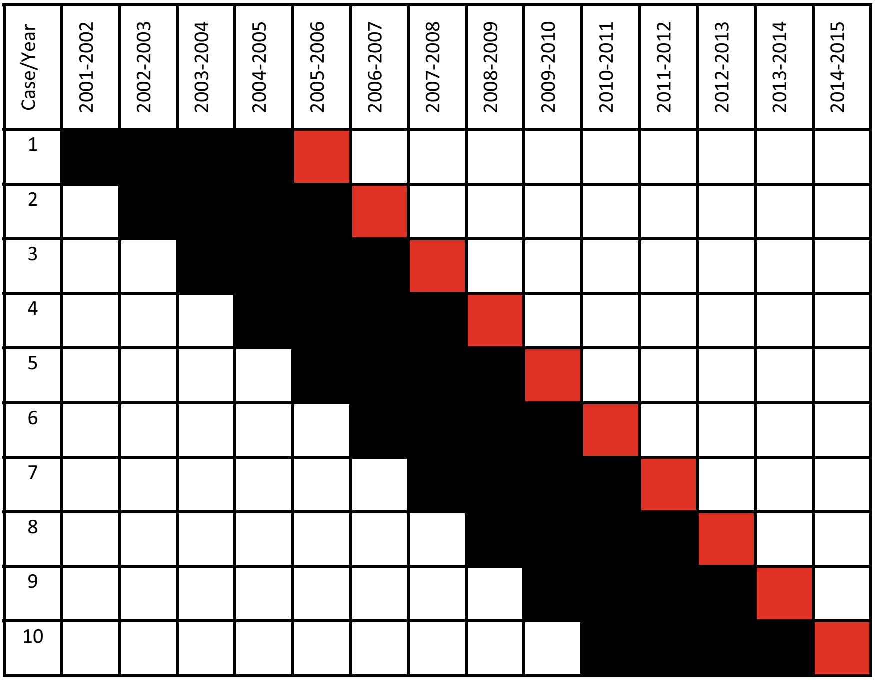


Fig. 31.2 Training and test sets in a moving window system

31.4.2 Top Companies Selection Window: Cases and Scenarios

The following studies on fundamental analysis by ML models adopted the moving window system**.** These models used the financial ratios of the listed companies as inputs and classified their stocks based on historical records. The models selected top-performing stocks from a set of companies based on their predicted returns for the subsequent financial year. The models classified a stock as Class-A (say) if it belongs to the top-performing minority category and as Class-B, otherwise. Where the Class-A stocks have the potential for higher annual returns.

Quah and Srinivasan [19] used the dataset consisted of quarterly stock prices and financial ratios from 1993 to 1996. The window allowed samples from the recent three quarters in training sets and samples from the subsequent quarter in test sets. Thus, they could select stocks to construct portfolios for a series of 13 testing quarters.

Fan and Palaniswami [7] used the data for consecutive three years to predict high-return stocks for the subsequent year. Thus, utilizing the data from 1992 to 2000, they could recursively experiment to make predictions for five years from 1995 to 1999.

Olson and Mossman [17] used the dataset comprising observations of firms during 1976–1993. The recent six years of data were recursively rolled forward each year to estimate returns for each of the 12 out-of-sample years 1983–1993. That is, the 6-year average in-sample parameters were used with accounting ratios at the beginning of the fiscal year to forecast 1-year-ahead.

Krishna Kumar et al. [12] used the financial ratios of the companies listed on the BSE 500 during 1995–2005. The model used the relationship among variables from the recent year to predict high-return stocks for the subsequent year. Moving forward through time, a new model was estimated for each successive year for constructing value portfolios for the years 1996–2006.

31.4.3 Training Window Optimization

Consider a moving window system with the training window size η years. The disadvantage of the moving window is that the first η-1 years of training samples are predominantly from old environments. Therefore, the system must include as fewer years' data as sufficient for retraining. What should be the value of η, the size of the training window? If η is too small, say $\eta = 1$, too little data will be available, because the training set comprises only observations of the recent past year. Thus, it may not be feasible to retrain on just a handful of data points. On the other extreme, if n is too large, say $\eta = 20$, the retraining may be literally ineffective as the training set is a bundle of data drawn from a wide variety of environments as the market has changed several times over the period of 20 years. Therefore, choosing the size of the training window is an optimization problem: For a given investment universe U with κ number of companies, minimize η, with η => 1. Note that when κ is decreasing, we have to increase η, to avoid the lack of data for retraining and vice versa.

31.4.4 Testing Window Optimization

In the moving window system for annual stock selection, it is common to set the size of the test window θ to 1 (see Sect. 31.4.2). Moreover, it is not necessary to apply downsampling on test sets [18]. Thus, we include the entire dataset of the corresponding year in the test set of a target year. However, we exclude a column in the test set, if that was either the response variable or used in deriving the response variable. Thus, we obtain a series of n test sets for the n target financial years Y_1, Y_2, …, Y_n. We denote them by S_1, S_2, …, S_n. Where S_k is the set of all test samples of the fiscal year Y_k.

31.4.5 Advantages of the Moving Window System

The moving window system helps to construct a series of training sets for training the model.

- It allows retraining the models using new data drawn from recent environments for fundamental analysis.
- It always includes constrained environments that reduce the number of data records by eliminating invalid data. Thus, it works as a data pre-processing step in modeling to enhance performance by *data reduction.*
- It provides a practical solution for the effect of changing macroeconomies.
- It can be used to construct and rebalance portfolios for long-term stock investments [7, 9, 12, 17, 19].
- Similarly, it can be adopted by any decision support system predicting the performance of companies, where the decisions are revised periodically.

31.5 Conclusions

The ML models explored in this study are intended to enhance equities screening to improve long-term ROI. However, the optimal sampling concepts of modeling discussed in this review remain applicable to any demand and supply model for long-term decisions. Such models select top-performing companies for improving long-term ROI. They predict the performance of companies by diagnosing their financial ratios to find matching patterns in historical records. However, due to changes in the economic environment, the patterns acquired from the training sample may have no matching pattern in the unseen test sample. The changing macroeconomies alter the patterns among microeconomic indicators dynamically. Therefore, a statistically proven, well-trained model may not yield profit in the long term. This conceptual review brings forth optimal sampling designs of noisy financial data for training and testing the models that minimize pattern distortions.

31.6 Limitations and Challenges

When we adopt the moving window system:

- We may have to forgo the recommended minimum sample size for training the ML algorithms.
- The models are trained for the one-time application: predicting subsequent target years. Thus, we must repeat all the steps of model estimation for each test case.
- It makes the model comparisons almost impossible to apply statistical tests, as the models are fitted in different environments every time. However, comparative studies are an essential part of the ML modeling.

31.7 Scope for Further Studies

ML models predicting top companies for improving long-term ROI are rare.

- Apply ratio analysis for predicting top-performing companies.
- Obtain optimal length of historical period to train such models, which are iterative. Such optima are application specific: they depend on the prediction target, ML algorithm used, number of companies to be classified, number of categories in classification, etc.
- Develop prediction models resilient to noisy financial data in changing macroeconomies.
- Compare prediction models on the basis of profits and not merely on statistical measurements.
- Apply *reinforcement learning*, a new area of ML that considers the environments for decision.

References

1. Afolabi M, Olude O (2007) Predicting stock prices using a hybrid Kohonen self organizing map (SOM). In: 40th annual hawaii international conference on system sciences (HICSS'07). IEEE, pp 1–8
2. Al-Qaheri H, Hassanien A, Abraham A (2008) Discovering stock price prediction rules using rough sets. Neural Netw World J
3. Becker Y, Fei P, Lester A (2007) Stock selection: an innovative application of genetic programming methodology. Genet Program Theory Pract IV 617:315–334
4. Chen S-S, Huang C-F, Hong T-P (2013) A multi-objective genetic model for stock selection. Kaigi.Org 2–6
5. Cyclical versus Non-Cyclical Stocks: What's the Difference? (n.d.). In: Investopedia. https://www.investopedia.com/articles/00/082800.asp
6. Esfahanipour A, Mardani P (2011) An ANFIS model for stock price prediction: The case of Tehran stock exchange. Int Symp Innov Intell Syst Appl (INISTA) 44–49
7. Fan A, Palaniswami M (2001) Stock selection using support vector machines. In: Proc IJCNN'01 Int Jt Conf Neural Netw 3:1793–1798
8. Feng X, Kong-lin K (2008) Five-category evaluation of commercial bank's loan by the integration of rough sets and neural network. Syst Eng Theory Pract 28(1):40–45
9. Hassan G, Clack C (2009) Robustness of multiple objective GP stock-picking in unstable financial markets. In: GECCO'09 proceedings of the 11th annual conference on genetic and evolutionary computation, pp 1513–1520
10. Huang C-F, Chang BR, Cheng D-W, Chang C-H (2012) Feature selection and parameter optimization of a fuzzy-based stock selection model using genetic algorithms. Int J Fuzzy Syst 14(1):65–75
11. Huang C-F, Hsieh T, Chang BR, Chang C (2011) A comparative study of stock scoring using regression and genetic-based linear models. IEEE Int Conf Granul Comput 2011:268–273
12. Krishna Kumar MS, Subramanian S, Rao US (2010) Enhancing stock selection in Indian stock market using value investment criteria: an application of artificial neural networks. IUP J Account Res Audit Pract 9(4):54–67
13. Lai K, Yu L, Wang S, Zhou C (2006) A double-stage genetic optimization algorithm for portfolio selection. In: 13th international conference on neural information processing, pp 928–937

14. Liu J-W, Cheng C-H, Chen Y-H, Chen T-L (2010) OWA rough set model for forecasting the revenues growth rate of the electronic industry. Expert Syst Appl 37(1):610–617
15. Min JH, Lee YC (2005) Bankruptcy prediction using support vector machine with optimal choice of kernel function parameters. Expert Syst Appl 28(4):603–614
16. Min SH, Lee J, Han I (2006) Hybrid genetic algorithms and support vector machines for bankruptcy prediction. Expert Syst Appl 31(3):652–660
17. Olson D, Mossman C (2003) Neural network forecasts of Canadian stock returns using accounting ratios. Int J Forecast 19(3):453–465
18. Quah T-S (2008) DJIA stock selection assisted by neural network. Expert Syst Appl 35(1–2):50–58
19. Quah T-S, Srinivasan B (1999) Improving returns on stock investment through neural network selection. Expert Syst Appl 17(4):295–301
20. ROSMA 2016. https://www.kearney.com/procurement/rosma/full-report
21. Samaras GD, Matsatsinis NF, Zopounidis C (2008) A multicriteria DSS for stock evaluation using fundamental analysis. Eur J Oper Res 187:1380–1401
22. Simon, S. (2019). Artificial neural networks in ratio analysis for financial health classifications. *JBIMS Spectrum*, *Vol. VII,*(1), 297–308.
23. Simon S, Raoot A (2012) Accuracy driven artificial neural networks in stock market prediction. Int J Soft Comput 3(2):35–43
24. Simon S, Raoot A (2012b) Competitive artificial neural network models in stock market prediction. Indore Manag J (Special Issue) 335–347
25. Turban E, Sharda R, Delen D (2011) Decision support and business intelligence systems (9th ed). Prentice Hall
26. Vanstone B, Finnie G, Tan C (2004) Enhancing security selection in the Australian stockmarket using fundamental analysis and neural networks. Bond University EPublications@bond
27. Yeh C-C, Chi D-J, Hsu M-F (2010) A hybrid approach of DEA, rough set and support vector machines for business failure prediction. Expert Syst Appl 37(2):1535–1541
28. Yildiz B (1999) Fundamental analysis with neuro-fuzzy technology: an experiment in Istanbul stock exchange. Sbd Ogu Edu Tr 8(2):25–42
29. Yildiz B, Yezegel A (2010) Fundamental analysis with artificial neural network. Int J Bus Financ Res 4(1):149–159

Chapter 32
Role of Artificial Intelligence in Green Public Procurement

(with Special Reference to European Economic Deal)

Bhakti Parashar and Amrita Chaurasia

Abstract In order to get the market to give the public sector products and services that have the least possible negative impact on the environment, green public procurement, or GPP, is now being implemented. Despite the fact that the scope varies from nation to nation, it is a widely acknowledged environmental policy tool. In the meanwhile, modifications in environmental legislation have led to states randomly purchasing goods, services, and works as well as their application. Green Public Procurement (GPP) has been shown to be an effective strategy for achieving the environmental policy objectives indicated in the communications from the European Commission. The field of artificial intelligence (AI) has expanded dramatically at the same time that businesses, governments, and society have greatly benefited from the intelligence of computers with machine learning capabilities. They also have an impact on broader trends in global sustainability. Key problems for sustainable manufacturing can be helped by artificial intelligence (e.g., optimization of energy resources, logistics, supply chain management, waste management, etc.). In this framework of smart production, there is a push to integrate AI into green manufacturing processes in order to abide by increasing environmental requirements. The government mandates a seamless integration of a complex system of laws, institutions, innovations, health care, nutrition, and education into all goods and services. The bulk of ecological and biodiversity study domains, as well as environmental and ecosystem management in general, are anticipated to benefit from advancements made feasible by artificial intelligence (AI) and related technologies like the Internet of Things. The essay aims to demonstrate how AI and ML could be applied in this situation to improve the efficacy and accuracy of the public procurement system.

Keywords Green · Sustainable · Procurement · Public · Artificial intelligence · Machine learning

B. Parashar (✉)
Vellore Institute of Technology, Bhopal University, Bhopal, India
e-mail: sharma.bhakti@gmail.com

A. Chaurasia
Christ University, Delhi-NCR, Ghaziabad, India

M. K. Tiwari et al. (eds.), *Applications of Emerging Technologies and AI/ML Algorithms*, Asset Analytics, https://doi.org/10.1007/978-981-99-1019-9_32

32.1 Introduction

Managing a product's life cycle's sustainability effect, raising consumer demand for better goods and more advanced manufacturing processes, and helping consumers make informed purchasing decisions are all crucial green economy activities [1]. However, the government's initiative is critical in starting and maintaining the green economy's virtuous circle [2]. Public purchases of goods, services, and projects, particularly green public procurement, must have the fewest possible negative effects on the environment (GPP). Before going on to procurement, where it is vital to select the best short-term remedies with long-term consequences, green supply chain management (GSCM) must also be developed [3]. It is crucial to acknowledge the rise in popularity of green and sustainable public procurement in order to address socioeconomic, environmental, and social issues through procurement [4].

Since the Single European Act took effect in 1986, the concept of green public procurement, or GPP, has been strongly related to the increasing significance of environmental protection in EU policy goals and aims. (SEA) [5] Given that the state of the environment is frightening on a global scale for a variety of reasons, environmental issues have risen relatively high on the political agenda of the EU over the past 20 years [6]. To address these issues and advance the achievement of local, regional, national, and EU environmental goals, the EU legislator envisioned a green shift in public procurement. Several legally required EU green public procurement structure features were consequently implemented. Sustainable public procurement (SPP), which is built on the three pillars of economic, social, and environmental responsibility, includes GPP as one of its components. Instead of emphasizing social and economic sustainability, this essay concentrates on the environmental aspects of public procurement [7].

Moreover, because they make it simpler and quicker to analyse massive volumes of data, which increases our knowledge and enables us to better understand and manage environmental challenges, artificial intelligence (AI) and other technologies have a lot of potential for green public procurement. Environmental effects and trends can be tracked more precisely, effectively, and quickly with the help of AI and Earth observation data fusion. It also enhances predicting skills by offering a new understanding of the dynamics that lead to environmental effects.

Significant environmental harm has been caused by the overuse of natural resources, the enormous volume of trash produced, and the incorrect disposal of dangerous chemicals, which have been paid for by corporate organizations. The hallmark of our corporate strategy for the past 20 years, lowering our global environmental impact has become increasingly important. The success of this endeavor depends on small and medium-sized businesses (SMEs). One of the corporate sectors that embraces change the most is supply chain management. The connection between the supply chain and environmental activities is made possible by the use of green supply chain management (GSCM). Thanks to green supply chain management, businesses can alter resource recycling, distribution, manufacturing, and consumption to have a less harmful influence on the environment (GSCM). Data analytics' significance in operations management as it relates to the management of operations has likewise rapidly increased.

32.1.1 Green Public Procurement

Public procurement is the practise of purchasing commodities, services, and creative works from private sector vendors by government agencies and state-owned businesses., has a lot of potentials to advance calls for improving the sustainability of goods and services, provided that various objectives can be balanced. The GPP is "the process by which public bodies try to buy goods, services, and performance with a low environmental effect throughout their life cycle", according to the European Commission [5] The GPP is a component of a larger paradigm for environmentally, socially, and economically sustainable public property procurement [6].

32.1.2 AI and Green Public Procurement

Artificial intelligence (AI) applications in sectors including energy, agriculture, housing, and mobility, which are commonly cited for their potential environmental impact. Environmental research could advance innovative applications, such as those in the energy or mobility sectors, with the usage of AI. Gathering, analyzing, and processing huge amounts of data also enhance the decision-making processes used in regulatory or political contexts. But just as AI has the potential to be a strong tool for achieving the objectives of a green transition, it also has the potential to accentuate unfavorable processes and generate fresh environmental dangers. The demands for energy and resources associated with digitization are strongly driven by AI systems [5].

AI-based techniques are already in use to better design transportation systems and infrastructure, increase the effectiveness of internal combustion engines, improve electric car charging, coordinate different modes of transportation, and regulate and manage railway systems. AI can also be used in use cases that boost the circular economy, such as improving the efficiency of eco-design or assisting in the inspection, sorting, separation, and disassembly processes that help materials circulate in the economy. AI-generated information may also help businesses and customers modify their behavior to be more ecologically friendly. The use of automated steering systems and technology Thanks to technology that makes predictive maintenance possible, infrastructure safety will continue to increase.

A number of critical sectors of a green transformation also show AI's promise. Electrical systems can benefit from the integration of renewable energy sources when artificial intelligence (AI) is utilized to monitor and optimize energy use. Applications in agriculture could help farmers use water, pesticides, and fertilizers more efficiently, minimizing any unfavorable environmental effects. AI systems are becoming more prevalent in all facets of our existence. Every day, a wider range of data-driven, linked, and automated applications help us communicate, shop, and acquire information. However, AI is not only becoming more commonplace in our daily lives in private: All facets of economics, research, and policymaking are already utilizing technology's ability to increase our knowledge, better manage operations, and ultimately control and influence our surroundings [5].

The method's benefits for the environment are frequently highlighted: Artificial intelligence (AI) has the ability to cut energy and resource consumption, hasten decarbonization, and support the circular economy through its applications in sectors such as energy, agriculture, housing, and mobility. By employing AI, environmental research could promote creative applications in the energy or mobility industries [9]. By collecting, evaluating, and processing huge amounts of data and helping to identify complex patterns in the environment or social interactions, AI has the potential to enhance the decision-making processes used in regulatory or political settings.

So far, despite the fact that AI has the potential to be a strong tool for achieving the objectives of a green transition, it also carries the risk of amplifying undesired processes and generating new environmental risks. As a result of their energy-hungry algorithms and extensive use of both old and new information and communication technology (ICT) devices and ICT infrastructure (i.e., data centres and data networks), artificial intelligence (AI) systems have a substantial impact on the energy and resource requirements of the digitalization process [9].

32.2 Application of AI in Green Public Procurement

Relatively few AI-based solutions are currently being developed to address environmental issues, and economic actors continue to be the primary motivators for the development of AI applications. Nonetheless, these applications are becoming more widespread and are employed, among other things, to keep track of environmental deterioration, improve resource utilisation, and increase our knowledge of the environment and climate. Growing amounts of research are being done on AI's potential to support a sustainable and green revolution. Being green should become the standard in society, and GPP is a government-centred practise that integrates social responsibility [10]. In other words, the transition from conventional to sustainable use depends heavily on the government [12].

Policymakers must develop initiatives to increase consumer understanding of sustainability, with a focus on what actions consumers should and shouldn't do [13]. The public authority can effectively serve as a role model and a mentor for its followers and individual customers because it is the most powerful consumer in the world [14].

Following are the broader area where artificial intelligence may apply to make green public procurement more effective (Fig. 32.1).

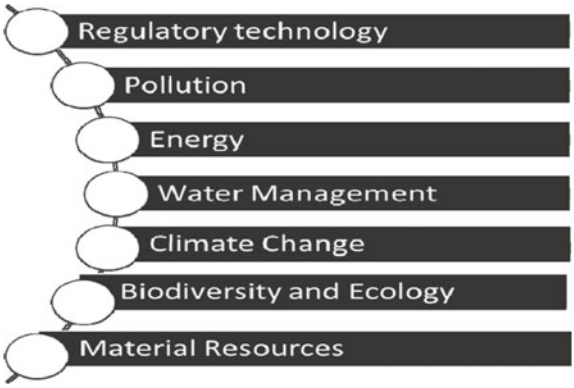


Fig. 32.1 Application of AI in green public procurement

32.2.1 Regulatory Technology

An attempt to "regulate (autonomously) governing systems" in line with socio-ecological goals might be seen as a green AI policy. From the perspective of environmental policy, answering two interconnected concerns is necessary to fully realise AI's potential for the green transition. How may the increase use of AI systems affect the environment and enduring problems like anthropogenic climate change, if at all? Determining how AI systems identify the tasks and activities they do is so important. Which rules and practises work best for controlling AI dynamics while accomplishing environmental objectives? The creation and use of regulatory technology is referred to as "RegTech", a phrase that emerged in the field of financial regulation. Deatherage [15] describes RegTech as only an expansion of the concepts of automating compliance and offering insight [15]. When traditional techniques of monitoring prove too time-consuming, AI may also be used to track specific offences [16]. As an illustration, a programme called Protection Assistant for Wildlife Security (PAWS) that uses predictive AI has been created to stop poaching. To forecast poachers' behaviour, it combines game theory and machine learning. Then, using a model of the poachers' behaviour, PAWS offers patrol routes [17]. In this regard, more AI tools have been created to find listings for animals for sale on well-known e-commerce sites [18].

32.2.2 Pollution

When it comes to urban governance and planning, for example, digital modelling of policy decisions can help governments make more sensible choices. IMEC and the Netherlands Centre for Applied Scientific Research unveiled a "interactive digital twin" of Antwerp in 2018. (TNO). A computerised 3D representation of the city is produced using data on noise pollution, real-time sensor data on air quality, and traffic. Its goal is to provide a current and future perspective of the city in order to forecast and assess the effects of proposed legislative actions [19].

Another indication of the potential of AI in urban management is the employment of sensors with built-in AI systems to covertly identify different road users at specific intersections and control traffic lights so that different forms of transportation can be given priority as and when necessary. In "smart junctions", which can give priority to pedestrians and cyclists when and where it is appropriate [20], there will be more bikes on the road as a result of COVID-19. For instance, [21] AI can be used to identify the most workable future land use and transportation plans.

AI-based solutions may be used to monitor, measure, forecast, and simulate pollution dispersion. Additionally, it can help to optimise and regulate the removal of pollutants from the environment. These applications may be found in the fields of pollution control for wastewater and solid waste, as well as air, soil, and water pollution, including marine contamination. For instance, AI apps can estimate pollution levels many days in advance when it comes to clean air. They are able to provide notifications regarding the air quality as a result. They can aid in the tracking and averting of air pollution since they provide real-time pollution monitoring and the location of air pollutant sources. 29 PE 662.906, "Artificial Intelligence and the European Green Deal", 23.

With the use of computer algorithm examples of untreated sewage discharges, daily effluent flow patterns, or historical spill events, the computer programme was taught to identify when a spill was likely to happen. The programme allegedly found almost 1,000 spills that had allegedly gone unreported for a period of eleven years [23].

32.2.3 Energy

For decarbonization, a shift to the production of renewable energy will be required. As a result, there will be a noticeable rise in the number of decentralised power producers, supply will be influenced by the climate and the length of the day, and flexible production and storage solutions will be necessary. Generation, storage, and consumption are all combined in intelligent power networks, or "smart grids". An AI-powered central control system balances out power variations, notably those brought on by erratic renewable energy sources in the grid, by efficiently coordinating between generation, storage, and consumption. Since AI is being used to handle load management, estimate supply and demand, and manage grid networks, it will be crucial in the future to achieve high levels of renewable energy integration and efficiency.

For example, VeriFone Energy leverages data from grid devices, energy demand, pricing, and weather forecasts to offer energy utilities adaptable AI solutions that automate forecasting processes in real time [24]. AI-based smart grid solutions are now being produced by numerous start-ups, and Siemens [25] has integrated AI into its suite of tools for designing, running, and managing smart grids. The production of renewable energy might be further optimised with AI technology, just as in other industries.

For instance, reinforcement learning has been used to maximise the amount of power produced by wind turbines or mobile solar panels [26]. It can decide where the turbines in wind farms should be placed [27]. In order to identify faults, automate issue diagnostics, and predict failures, wind turbines are also exposed to machine condition monitoring systems, particularly in remote and offshore areas [28]. This increases availability, productivity, and decreases the maintenance costs of wind farms. Other AI technologies can be used to find issues with rooftop solar panels.

AI may be used to find leaks, forecast methane emissions from gas pipelines and compressor stations, and make proactive pipeline management recommendations. By utilising AI algorithms and data from 23 satellites, the start-up Bluefield helped Florida authorities locate the source of multiple methane leaks and gives 120,000 firms in 130 countries with information on their CH4 emissions [29].

32.2.4 Water Management

AI is also used to assess soil moisture and other crucial factors that could help farmers use irrigation water more effectively. This is done by fusing data on precipitation volumes with satellite images [30]. Accurate meteorological data are also required for irrigation optimization because the weather has a direct impact on evapotranspiration. As a result, artificial intelligence (AI) applications have recently grown, even reaching the level of individual farms [31, 32].

32.2.5 Climate Change

Precision farming has the potential to reduce fuel usage of agricultural equipment right away and have favourable second-order effects on energy use. For instance, if inputs are utilised more effectively, less energy may be required to produce them [33]. AI may also aid in reducing nutrient runoff into the water and direct and indirect N_2O emissions from agricultural soils if it results in a drop in anthropogenic nitrogen inputs (from mineral or organic fertilisers into farms and grassland). The net EU GHG emissions from N_2O emissions will be 3.9% in 2018, according to current projections. [34]. The potential to boost yields and maintain soil health are two frequently mentioned effects of modern technology in agriculture [35]. Emissions from farming may decrease if this lessens the pressure to expand agriculture or frees up crops and grassland on organic soils. In 2018, organic soils were responsible for 30 Mt of the EU's CO_2 emissions. Its importance is questionable [36] because a variety of other factors also affect how farmers manage their land.

If artificial intelligence is successful in increasing the carbon content of agricultural soils, it may be conceivable for it to make a more significant direct contribution to climate change mitigation. To more precisely determine soil carbon, Cloud Agronomics uses deep learning and hyperspectral observations of the soil [37].

32.2.6 Biodiversity and Ecology

A wide range of options for biodiversity and conservation are presented by AI-based technology. They can support the creation of protected areas by monitoring ecosystems, recognising and tracking habitat deterioration, simulating interactions between animals and their environment, or estimating animal migration patterns, for example. For instance, the Biodiversity Observation Network (GEO BON) of the Group on Earth Observations (GEO) uses in-situ and remote sensing data to support the monitoring of changes to biodiversity and ecosystem services with the goal of providing high-quality observations, information, and data to scientists, policymakers, and the general public. GEO BON offers a map of the forest cover that illustrates the climatic change.

Invasive, hazardous, or threatened species can be found using artificial intelligence (AI) systems, which can also provide precise population and location estimates. They can also keep an eye on a species' health. It is well known that the majority of current monitoring methods, such as video traps and on-foot surveys, are time- and resource-consuming, incorrect, and imprecise.

AI provides a new method to more quickly analyse wider areas when combined with other technologies like unmanned aerial vehicles or tiny thermal imaging sensors [38]. Satellite tagging is used by programmes like "Whaletrack" to gather information on whale migration patterns. Machine learning may be used to analyse and assess acoustic sensor data in order to distinguish between various species in a region and research their behaviour [22].

32.2.7 Material Resources

The first direct impact on natural resources comes from the development of ICT and microelectronic components, which act as the foundation for the operation of the digital infrastructures needed for the application of AI. Manufacturing active semiconductor components consumes a lot of materials and energy. Metals like cobalt, palladium, tantalum, silver, gold, indium, copper, lithium, and aluminium stand out when it comes to the detrimental consequences that the extraction of raw materials for the construction of digital technology has on the environment and society.

Due to the strong demand for digital end-user devices, cobalt (9.4% of the world's primary output) and palladium (8.9% of the world's primary production) are in great demand. The incorporation of AI could increase demand for such materials, for example in the form of a customised neural network, given that such throw-away consumer devices would require the use of huge batteries. The inventory of data centres includes ICT components like servers and data storage in addition to supporting systems like cooling systems, uninterruptible power supplies, and physical buildings that all have an effect on resource utilisation.

Often, they are heavier than the server equipment. Under a variety of circumstances, IPOL | Policy Directorate for Economic, Scientific, and Quality of Life Policies PE 662.906 34. Steel, aluminium, and copper, which are common building materials and needs, make up the majority of them. The support systems' expected lifetime is frequently mentioned in decades. Servers, on the other hand, are composed of common ICT materials like copper, rare earths, precious metals, and semiconductors. Depending on performance needs, ICT for servers is updated every few years, up to ten years at most.

32.3 Findings and Suggestions

32.3.1 Findings

The main purpose of the European Green Deal can be advanced in a number of ways by using artificial intelligence (AI). The negative effects of AI on the environment, however, may compromise the achievement of these objectives. In this study, the potential of the environment is studied together with the characteristics and underlying causes of environmental dangers. Best practices and environmental policy measures are also covered. The study's recommendations in its conclusion emphasize the importance of passing legislation in order to harmonies AI development and application with the objectives of the European Green Deal.

32.3.2 Suggestions

While using AI systems and applications have many potential advantages for socio-ecological development, it also has substantial and pervasive environmental hazards. In order to provide clean, quality-controlled training datasets, convenient access to observation data, the creation of novel use cases, and prototype applications, researchers with competence in both AI and earth observation must develop their capabilities.

Theoretically, environmental policy and regulation can make use of a wide range of measures to support the judicious development of new technologies and innovations as well as the defence and alleviation of some sustainability-related issues. Supporting projects that aim to increase the number of open datasets and give the general public access to pertinent data is crucial when taking environmental issues into account.

Due to this, it is possible to include underserved nations and regions as well as thematic areas that may not have a solid business reason. According to the World Development Report [39], “An international agreement is needed to ensure that data are retained as a global public good and as a resource to help fair and sustainable development”, the UN said. To move the Green Deal’s main projects on climate.

change, the circular economy, zero-pollution, biodiversity, reforestation, and compliance assurance in this direction, the “GreenData4All” programme and the “Common Europe Green Deal Data Space” must be established.

In order to use AI to address environmental concerns in the sectors of research, business, government, and civil society, developing countries also need support in building the skills necessary for them to do so. Emerging and developing economies must address the problems that are causing the decline in global environmental resources, such as temperature and biodiversity. It is imperative that these nations have access to the promise of AI. There are still a lot of significant issues that need to be resolved, including the lack of sufficient local data, access to crucial data, a talent shortage, the relative unattractiveness of various business models, or governments’ limited capacity to encourage and regulate the development and use of AI.

References

1. [EC] European Commission Directive 2008/105/EC of the European Parliament and of the Council of 16 December 2008 on Environmental Quality Standards in the Field of Water Policy, Amending and Subsequently Repealing Council Directives 82/176/EEC, 83/513/EEC, 84/156/EEC, 84/491/EEC, 86/280/EEC and Amending Directive 2000/60/EC of the European Parliament and of the Council (2013)
2. Testa F, Grappio P, Gusmerotti NM, Iraldo F, Frey M (2016) Examining green public procurement using content analysis: existing difficulties for procurers and useful recommendations Environ. Dev Sustain 18(1):197–219
3. Roehrich JK, Hoejmose S, Overland V (2017) Driving green supply chain management performance through supplier selection andvalue internalization: a self-determination theory perspective. Int J Oper Prod Manag 37:489–509
4. Sönnichsen SD, Clement J (2020) Review of green and sustainable public procurement: towards circlar public procurement. J Cleaner Prod 245
5. Gailhofer P, Herold A, Schemmel JP, Scherf C-S, Urrutia C, Köhler AR, Braungardt S (2021) The role of artificial intelligence in the European green deal. Öko-Institut e.V. https://doi.org/10.13140/RG.2.2.26789.22244
6. Pouikli K (2021) Towards mandatory green public procurement (GPP) requirements under the EU green deal: reconsidering the role of public procurement as an environmental policy tool. ERA Forum 21:699–721
7. Commission E (2017) Strategic public procurement: facilitating green, inclusive and innovative growth. Eur Procure Public Priv Partnersh Law 3:219–223
8. Johnson PF (2020) Purchasing and supply management (sixteenth ed.). McGraw-Hill/Irwin, Burr Ridge, IL
9. Walker H, Brammer S (2012) The relationship between sustainable procurement and eprocurementin the public sector. Int J Prod Econ 140(1):256–268

10. Uttam K, Roos CL (2015) Competitive dialogue procedure for sustainable public procurement. J Clean Prod 86(1):403–416
11. Welsch H, Jan K (2018) How green self-image is related to subjective well-being: pro-environmental values as a social norm. Ecol Econ 149(7):105–119
12. Bala A, Munoz P, Rieradevall J, Ysern P (2008) Experiences with greeing suppliers the universitat autonoma de Barcelona. J Cleaner Produ 16(15):1610–1619
13. Pircher B (2019) EU public procurement policy: the economic crisis as trigger for enhanced harmonization. J Eur Integr 42:509–525
14. Paul J, Modi A, Patel J (2016) Predicting green product consumption using theory of planned behavior and reasoned action. J Retail Consum Serv 29(3):123–134
15. Deatherage P (2021) Environmental compliance through regulation technology. J Int Relat, Peace Stud Dev 6(1), Article 4
16. Purdy R (2010) Using Earth observation technologies for better regulatory compliance and enforcement of environmental laws. J Environ Law 22(1):59–87
17. Feng D et al (2020) Advances in plant nutrition diagnosis based on remote sensing and computer application. Neural Comput Appl 32(22):16833–16842
18. Rusch N, Arora P (2021) AI-based strategies to combat wildlife trafficking and wet markets in Asia. A critical review. Konrad-Adenauer-Stiftung
19. Imec (2018) Imex and TNO launch Digital Twin of the city of Antwerp, press release
20. Vivacity Labs (2020) TfGM chooses Vivacity Labs to deploy AI-controlled road junctions to optimise traffic networks. Vivacity Blog
21. Wu N, Silva EA (2010) Artificial intelligence solutions for Urban land dynamics: a review. J Plan Lit 24(3):246–265
22. Liu R, et al (2019) Impacts of the digital transformation on the environment and sustainability. Öko Institut,issue paper under Task 3 from the Service Contract on future EU environment policy
23. Hammond P, et al (2021) Detection of untreated sewage discharges to watercourses using machine learning. Clean Water 4(18)
24. Veritone (2021) Real-time predictive AI for smart grid optimisation and resilience
25. Siemens (2020) Energy intelligence—tapping the potential of a smart energy world
26. Abel D, Williams EC, Brawner S, Reif E, Littman ML (2018) Bandit-based solar panel control. In: Thirty-second AAAI conference on artificial intelligence
27. Dilkina B, et al (2015) Method for designing the layout of turbines in a windfarm. US Patent 9, 189, 570
28. Hastings M (2020) AI for enhanced wind-turbine monitoring. Wind Systems
29. Bluefield (2021) Independent ESG data via satellites and AI
30. HelioPas AI (2021) WaterFox feldgenau bewässern
31. Roach J (2019) Is drought on the horizon? Researchers turn to AI in a bid to improve forecasts
32. Fasal (2021) Company homepage. https://fasal.co/
33. Bocker T, et al (2019) An economic and environmental assessment of a glyphosate ban for the example of maize production. European Rev Agric Econ
34. European Environment Agency (EEA) (2020) Annual European Union greenhouse gas inventory 1990 2018 and inventory report 2020. Submission to the UNFCCC Secretariat
35. Lezoche M, et al (2020) Agri-food 4.0: a survey of the supply chains and technologies for the futureagriculture. Comput Ind 117:103187
36. Bottcher H, et al (2021) Exploratory analysis of an EU Sink and restoration targe
37. Cloud Agronomics (2021) Our solutions tackle the biggest challenges in sustainable agriculture
38. World Economic Forum (2018) Harnessing artificial intelligence for the earth. Fourth Industrial Revolution forthe Earth Series
39. World Bank (2021) Data for better lives. World Development Report

Chapter 33
Supplier Prioritization and Risk Management in Procurement

Shilpa Narayanswamy and Nikhil Ghantial

Abstract People analyze risks and task corrective actions to reduce the severity. But, in supply chain, risks cannot be eliminated, only reduced. Traditionally, supply chains involve three sub-sections, procurement, processing, and distribution. While processing is in-house; distribution and procurement are critical, and any issue there affects the whole supply chain. Therefore, resources are deployed to make them resilient and minimize impact. Procurement is defined as a process of buying goods and services. A procurement strategy needs to maintain a balance between the risks involved and create a win–win situation for both the organization and supplier. Since it involves decision-making at all levels and can affect every part of the organization, procurement strategy must prioritize risk mitigation to maximize its efficiency and efficacy. Hence, supply chain departments must realign their approaches to include evaluating suppliers, paying premium prices for raw materials, redefining product characteristics, identifying alternate suppliers globally, developing sources in the vicinity, changing logistics, outsourcing technology, and so on. This paper attempts to showcase how supply chain risks can be minimized using a structured approach, which includes supplier prioritization followed by risk management techniques, assessing impact, and creating a contingency plan to eliminate risks. A real case is presented in which the structured approach is discussed to mitigate risks in the nascent stage itself.

Keywords Supplier prioritization · Risk management · Procurement · Supply chain · Risk assessment

S. Narayanswamy · N. Ghantial (✉)
Operations Management Professional, Mumbai, India
e-mail: ghantialnikhil@gmail.com

S. Narayanswamy
e-mail: shilpa.n18@gmail.com

M. K. Tiwari et al. (eds.), *Applications of Emerging Technologies and AI/ML Algorithms*, Asset Analytics, https://doi.org/10.1007/978-981-99-1019-9_33

33.1 Introduction

The advent of COVID-19 has highlighted various loopholes in the supply chain which led to the restructuring of various strategies of procurement and distribution; as the market is facing a lot of raw material shortages, delayed shipments, and issues such as container shortages and increasing cost of reverse supply of containers, companies are highly interested in developing suppliers in the vicinity of their manufacturing locations. To give an example, a top automotive manufacturer in India sources 95% of the material required for assembly from local vendors. Also, on average, the production cost in manufacturing is more than 50–60% associated with raw material costs. In today's dynamic business climate, procurement's role is not limited to just getting the right goods and services at the best possible price, with the right volumes, at the right time. Their role has critically evolved to identify procurement risks, and with organizational involvement develop a risk mitigation strategy that reduces a company's vulnerability to potential supply management impacts.

The raw material cost is impacted by various factors since the advent of cross-border trade, impact on one country causes issues in other countries due to interlinked supply chains. One of the rising concerns contributing to the RM (raw material) cost is inflation, as per RBI data, the inflation was set to be in the range of 6 to 8% which now has crossed the threshold adding to it US economy is facing all-time high inflation these have a huge impact on world trade the primary reason has been US dollar being the default exchange currency in many countries. In a such high inflationary environment, the operating profit margin (OPM) is reduced significantly and the pressure is on the supply chain to reduce the cost by using some or other innovative ideas.

People working in the supply chain are under tremendous pressure since the supply chain is vulnerable to any disruption occurred in the economy hence, input cost increases, and their key result area (KRA) is directly linked to how they reduce the input cost, provide faster delivery, and reduce the cost of supply and distribution. Supply chains use various strategies to meet their KRAs and one of the critical strategies is supplier prioritization and risk management in procurement.

Supplier prioritization and supplier risk management are two different terms but in real-life scenarios, they are closely linked to each other. Supplier prioritization means the process by which a firm identifies, evaluates, and contracts with the supplier and provides the work to them, evaluation is done on various factors to set the context some of which are vicinity to plant, accepting the company policy, and using environment, health, and safety (EHS) norms at their premises.

Supplier risk management, on the other hand, is the process of identifying, assessing, and controlling threats to a project which might occur while the project is ongoing so the idea is that risk is identified using risk assessment techniques and use risk mitigation methods to avoid the risk or reduce the same. When talking the structured approach of risk management constitutes risk identification, risk assessment, risk mitigation contingency plan, and risk monitoring by internal stakeholders. For risk mitigation, company should understand the source and driver of risk.

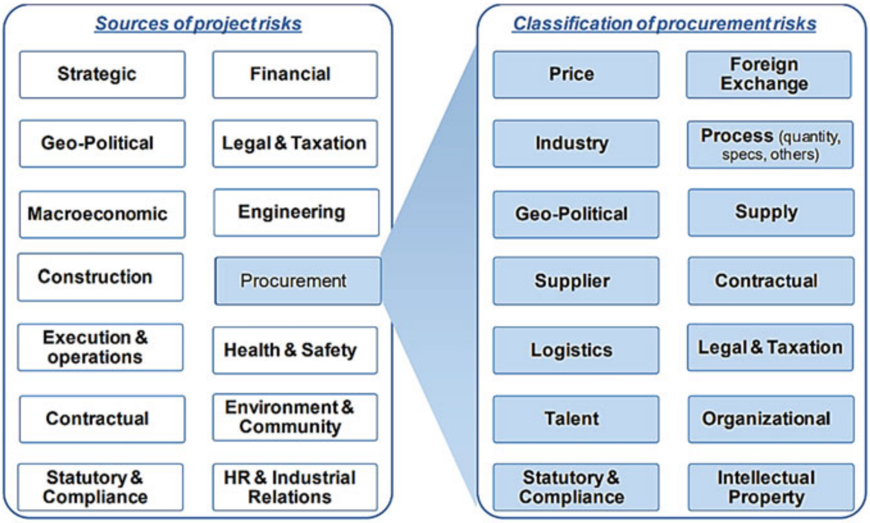


Fig. 33.1 Procurement risks

Both supplier prioritization and risk management are critical for the project's success.

A broad classification of procurement risks in a project is depicted in Fig. 33.1.

33.2 Literature Review

One of the most important parts of the supply chain is procurement and it is said that on average a company spends about 50 to 60% on raw material procurement, thus a major part of the revenue is consumed by raw material [1]. Procurement is very often existing in an uncertain business environment. This includes uncertain customer demands, market fluctuations, unavailability of raw materials, extended lead time, etc. Due to this, a major focus is on procurement and how the procurement cost viz a viz supplier can be identified such that minimum cost is spent on the same. the modern supply chain has changed from what used to be a resource optimization to a now so-called customer-centric approach, due to which having profit maximization and cost minimization are a concern, but people are going beyond it [2]. Risk mitigation is now more important and supply chain-related risks are now discussed at a higher level in supply chain literature [3]. When we say supply chain risk it is a very broad term and encapsulates various issues which can be related to disruption in production, inventory obsolesce, customer return of goods, seasonal-based trends of products, supplier failure to provide raw material on time, and so on [4]. To mitigate some of the issues in the supply chain companies are ferociously working to establish a reliable supplier, which can supply both in an uptrend of their product with more raw materials and can sustain the downtrend of seasonality, therefore, companies need to have a vision and reduce supply chain-related risk while supplier prioritization which means supplier need to be evaluated with the risk associated in our supply chain and thus trying to reduce the risk [1, 5].

Risk arises in the supply chain in two phases. One for manufacturing of the product from raw material and the other from moving the product from company premises and delivering to the customer and the reverse logistics in case of customer returns, so the probability of undesirable things happening is more and no one can predict the same but the consequence can be analyzed [6]. The risk is demand-driven and the result of disruption in the supply chain globally the risk can be categorized as operational risk and strategic risk [7].

With increasing globalization and cross-border trade, the risk associated with the supply chain is increased so supply chains need to have a plan in place if the current strategy does not work or at least be ready that's how can they mitigate the risk [8]. Firstly, network integration in the supply chain is important and the coordination analysis is used by several researchers for the same. in line with network analysis, risk management is important and various research is done researchers have identified some models that when used can help in reducing the risk for example one of the researchers used bibliometric and network analysis tools to identify key research clusters/topics, interrelationships, and generative research areas on quantitative models for managing supply chain risks [9] and research is also done in some service-based industries as well.

The main value addition comes into the picture when companies started using it and started developing their measures of risk assessment depending on the product and the nature of the business in which the company is operating.

33.3 Supplier Management

In a manufacturing enterprise, there is a value chain by which the product is manufactured, the value chain consists of various processes by which the raw material (RM) is converted into finished goods (FG) for which customers will be willing to disburse money from his pocket. The process of converting raw material to finished goods involves a set of operations and coordination between management and supplier, managing supplier become a big task as there need is at the upstream operation like raw material supply and any deviation caused upstream resonate till downstream and there might be a delay in supply of finished goods to the customer.

When we say managing the supplier it involves the quantity, quality, delivery time, and acceptance of payment terms although some of these are known, some are unknown; like in India various festivals are celebrated throughout the year. Suppose a supplier is in Gujarat and our plant location is in Maharashtra and due to the festive season in Gujarat supplier may not be able to supply the material due to manpower shortage in the plant. But, in Maharashtra, it is not a festive season and customer demand is intact; in such a scenario reaching a consensus by the procurement team and supplier before a festive season becomes important and due to unavoidable scenarios procurement might have to procure more before the festive and stock up their warehouses. Also, the supplier might have to add additional manpower and supply the requested quantity.

Supplier management becomes more critical because the increasing dependence on suppliers makes organizations highly exposed to supplier risks.

Many of the procurement risk causes could be linked to supply-side risks. Supply-side risks could be due to various reasons like physical damage to the facility, natural disasters, labor strikes, outsourcing, manufacturing delays, etc. Due to volatility and increased supply-side risks, there is usually a spillover in terms of increased production or overstocking.

A few factors that have a high impact on supplier management are supplier concentration, over-reliance on a supplier, global sourcing trends, trust towards suppliers, relationship with suppliers, information sharing with suppliers, etc. Understanding these risks in advance can enable the purchasing organization to take effective action in response to those risks.

Such scenarios might arise and hence managing suppliers becomes an important part of the whole supply chain such that production will continue, and the supply chain will not be hampered.

Hence organizations need to have full visibility of not only their first-tier supplier but also their second- and third-tier suppliers.

33.4 Supplier Prioritization

While selecting and evaluating suppliers' assessment basis, some pre-set criteria become an important part of the process. Here assessment of suppliers refers to certain measurable parameters such as procurement metrics to evaluate the credibility of the supplier.

One of the most important parameters is the product or raw material which the company needs from the supplier. Once this is finalized then other parameters like quality, quantity, and delivery are evaluated among different potential suppliers and the best is considered to supply the goods. These processes for the assessment of the supplier are part of the supplier prioritization practice. Certain components are part of the broad supplier prioritization practices globally. Some of the factors which are considered important are listed Table 33.1 and 33.2.

Table 33.1 Broad components of supplier prioritization

Component	Explanation
Goal-driven	Responds to organizational strategy, goals, and end-customer needs and wants
Resource driven	Prioritize interventions based on available resources
Market-driven	Responds to market understanding that we are selecting the right suppliers

Table 33.2 Key criteria for identifying potential suppliers

No	Criteria for potential suppliers
1	Revenue
2	Location
3	The segment of supplier-customer
4	Current capacity
5	No. of clients
6	Employee strength
7	ISO certification or other
8	Management of supplier
9	Feedback from supplier
10	Feedback from an existing client
11	Feedback from old client (4 to 5 years)
12	Acceptance to supply material on a credit basis

The above points are some of the key aspects to consider while prioritizing suppliers but one of the crucial aspects is the sector in which our supply chain is operating. A case in point is; if it is a pharma industry then despite meeting all the criteria there must be a supplier visit conducted to check their working standards, hygiene maintenance in the nearby vicinity, knowledge of the product which the supplier is supplying, following of current good manufacturing practices (CGMP), adhering to good documentation practices (GDP), etc.

33.5 Risk Management

Proactive risk management needs preparedness with suppliers. Preparedness in this context refers to a situation where if a contract was drafted for 6 months, then the supplier continues to provide us raw materials which are affected due to petrol prices at the same earlier rate despite the market fluctuations. Here, even if the rate is increased marginally, it will not reduce the profit margin for the industry. Like the role of project risk management in project management, supplier risk management is gaining prominence to mitigate risks of the supply chain.

Supplier risk assessment and management can be modeled using.

- Qualitative models
- Quantitative models
- Hybrid models

which include both qualitative and quantitative models.

Similarly, one-time advance payment is no more functional in many organizations, and nowadays agreements are in terms of 20% amount once the order is dispatched;

30% once the order is within industry premises, and the rest after the material is tested for quality.

In the case of resource-oriented industries like steel and aluminum, the material test report is one of the important criteria to check the quality and is compared on a point-to-point basis to the in-house testing report post which the payment is initiated.

A simple depiction of risk management in the form of a matrix, where risk assessment and management is mostly applied to the High and Medium importance category. The supplier risk assessment and management process can comprise sequential steps as follows (Table 33.3).

Also, once we identify various types of risk associated, we can categorically allocate them in a bucket of risk response matrix wherein each risk is assigned a category according to which action will be taken (Fig. 33.2).

- Avoid: Find a different approach to the activity that creates risk.
- Mitigate: Take steps to reduce the risk with appropriate control and monitoring measures.
- Transfer: Transfer all or some of the risk using insurance or other methods (Such as partnership)
- Accept: Accept the risk and all associated costs in case of exposure.

In general, the most preferred response should be to avoid risk such that in the first case, the risk will not affect any stakeholder, following these will be to mitigate risk such that no stakeholder will be affected by risk but the cost of mitigation may be shared collectively among them, in case of transfer of risk this response should be used only when avoid and mitigate doesn't work out, in transferring the risk at least one of the stakeholder will be impacted, the last one in risk response is the accept risk these is least preferred and directly it affects the cost and schedule.

Once we prioritize our actions, we can reduce the severity of risks and design emergency plans for them.

Table 33.3 Steps for supplier risk assessment

Step no:	Step description	Process
1	Identify	Understand potential risks for each supplier. Use data from portfolio analysis and supplier preferencing
2	Assess	Assess the likelihood of the risk being realized and the severity of the impact
3	Prioritize	Prioritize risks for action basis assessment done earlier and ensure the right risks are highlighted
4	Plan	Create either a mitigation action plan or a contingency action plan in response to priority risks. Develop an action plan
5	Manage	Manage and review the risk process on an ongoing basis and repeat the process on a loop

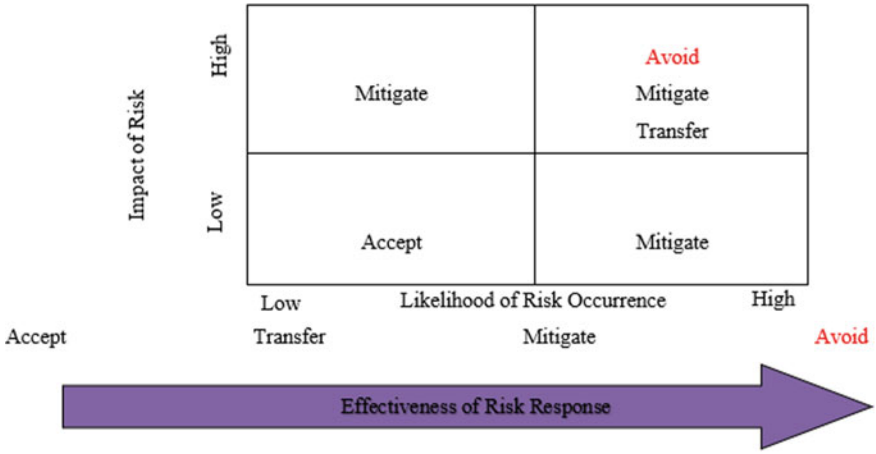


Fig. 33.2 Risk response matrix

Risk mitigation planning would include detailed contract planning and provision, supplier audits, maintaining rapport with suppliers, communication strategy, and maintaining policies and systems to name a few.

Risk contingency planning would include disaster recovery planning, switching suppliers, switching to substitute products, pausing operations, and creating readiness for alternatives to name a few.

Decisions taken in global supply chains are complex and reversing them in the short term becomes a difficult and costly affair. It is important from an organization's perspective to build robust programs for both known and unknown risks. Hence, understanding all possible risks and spending time in evaluation and assessment can be helpful and provide better results to the organization. Risk management is not only process management but also entails mindset and culture shifts.

33.5.1 Live Case

A live application of how risk management helps is explained below. The case study company was in the manufacturing sector. It supplied parts to domestic customers. The company produced a high variety of products and was planning to launch many new ones as well. The company followed a make-to-order strategy as per customer requirements and plan its upstream activities from raw material (RM) to work in progress (WIP).

A customer had a requirement for a new product (typically in the shape of a round plate) that required a pressing and heating process simultaneously. Hence the company floated a requirement for the machine specifically an special purpose machine (SPM) to the internal team handling such projects. The design of the machine was made by the internal team, and the manufacturing of the machine was to be outsourced since the company's core competency was in manufacturing the product, not the machine so the purchase team and project team came into the picture and took forward the task.

The top management had communicated the following 4 points to the purchase and project team as their evaluation criteria for the success of the project.

1. Cost
2. Delivery
3. Quality
4. Service

Historically supplier evaluation and assessment were based on traditional methods in the company. However, over time, the company started facing many supply risk problems as they neglected many suppliers-related risks in which the purchase team was well versed. The performance of the suppliers had an impact on the profitability and sustainability of the company. There was a need to reduce supplier-related risk by eliminating unreliable and risky suppliers right at the beginning of the procurement process. Collectively purchase and project teams worked out the approach for supplier identification and reducing risk for the same.

A framework was developed to measure the severity of the procurement risks. Accordingly, basis risk prioritization and the circumstances of the case, rank-based solutions were suggested to mitigate the severity of the procurement risks.

The case details the model followed for supplier evaluation and selection in the company.

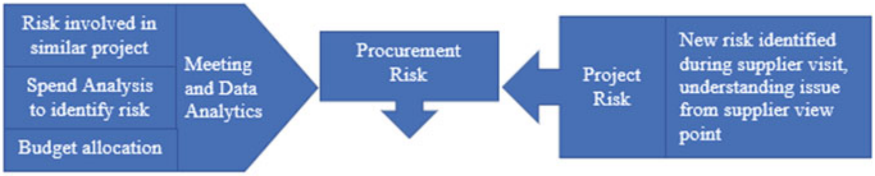
Identifying potential suppliers started with the following steps:

1. Filtering the registered suppliers based on their capacity and required bed size for manufacturing the special purpose machine (SPM).
2. Communicate the requirement to potential suppliers and ask for request for quotation (RFQ).
3. One-on-one meeting with the project team and suppliers to have clarity on the design and the supplier in build capacity to verify whether the required quality will be maintained or not.

Overall, the methodology followed for risk assessment is described in the following chart (Fig. 33.3).

The effectiveness of the proposed supplier evaluation method would largely depend on the right identification of supplier risk factors. After several discussions with key stakeholders in the case, the purchase team finalized the criteria for selection. Specific requirements of the company, the uniqueness of the sector, and the risk attitude of the company were also considered.

One of the points which were identified in the discussion was that the machine required some plates of which the length was more than 1400 mm and the width was 500 mm. Although the width can be made on any conventional vertical milling machine (VMC) it was difficult to find a computer numerical control (CNC) or a vertical milling machine (VMC), wherein the length of the plate can be manufactured



Risk Category	Risk Description
Contractual & Compliance	One of the potential supplier was not registered with the company (Registration will take 3 months)
Quality	In near by premises of the industry requirement quality won't be possible to maintain as desired capable supplier not identified
Installation and first piece check	One of the supplier not agree to ready to install machine in company premises and won't send his person (To be done by project team itself)
Price	Quoted price by supplier outside state was higher than the budget allocated

Fig. 33.3 Risk identification process

with the overall quality parameter of parallelism and tolerance need to be maintained less than ± 0.05 mm which was a difficult task.

It was difficult to find many suppliers because of the specific design requirement and finally, a supplier was shortlisted who had the required bed size, but the purchase team found that this single criterion would create a lot of dependency on the supplier removing any scope of controllable factors in our procurement and project team.

The issue was communicated to the design team and the project, purchase, and design had a lot of discussions after which it was finalized that the plate will be split into two pieces and will be assembled but in this case, the quality parameter of the tolerance was more critical as per the specification given by design.

After due consultation with the design, purchase, and project teams some design changes were done which can be manufactured from a local registered supplier, and RFQs were floated.

The quotation came from many potential suppliers and finally based on the quality and cost, delivery, and service following apple-to-apple comparison was made with supplier 1 who was a local registered supplier, and supplier 2 who had the bed size and quality that was required (Table 33.4).

Based on the comparable purchase team decided to finalize supplier 1, one of the primary reasons for the same is supplier 2 was not accepting any clause of rectification post machining, while in the case of supplier 1 if there are some issues with the machine, supplier 1 will be in the nearby vicinity and communication would be faster in real-time with additional support from internal design and project team on any clarification regarding the machine.

But there was one thing the project team was concerned about which was some critical standard parts that are required and are used in the design of the machine and can be sourced by supplier 2 only, so the purchase team had no other option but to

Table 33.4 Comparison with supplier on the basis of the same characteristics

Criteria	Supplier 1	Supplier 2
The location of the organization's plant	2 km	200 km
Capacity	One Single machine with bed size (L = 700 mm)	4 machines with a bed size of (L = 1000 mm)
Cost	1,00,000 INR	1,00,000 INR
Quality assurance by the supplier	Within a ±0.05 mm tolerance range	Within a ±0.025 mm tolerance range
Material test report acceptance	Yes, will be provided	No
Corrosion Resistance Coating on the final product to reduce the erosion	Yes, will be provided	No
Acceptance to Install the product by the supplier in the organization plant	Yes	No
Payment terms: 20% After dispatch 30% After the material is received in the plant 50% After full functioning of the product	Yes	No. (Need full payment once the product is on organization premises)
Delivery	2 months	1 Month
Inspection weekly	Possible	Not possible as far away from the organization

buy it from them, yes since it was standard part available it was ready to dispatch and was in company premises in a matter of days.

Before finalizing the deal with supplier 1, various risk associated were identified and the following risk assessment was done and critical points were discussed with supplier 1 and a consensus was obtained such that some part of the risk was transferred to supplier 1, such that he will take responsibility to meet the desired result as expected. Risk in the case of supplier 1 cannot be avoided or mitigated since he was selected so the third step of transferring/sharing risk in our context was done (Table 33.5).

- Controllability score—is the extent to which a project can manage the risk
- Likelihood score—is the probability of risk to occur
- Impact score—is derived from schedule, cost and non-quantifiable impact
- Overall risk score—is the multiplication of impact and likelihood (likelihood score * impact score)

This table was designed to outline and prioritize supplier-side risks. This would help the organization to review and analyze all the potential characteristics of the suppliers and prioritize accordingly. The scoring was given on a scale of 0 to 1, with

Table 33.5 Illustrative risk matrix

Sr. no	Risk category	Risk description	Controllability score	Likelihood score	Impact score	Overall risk score
1	Quality of material	Material not meeting the desired tensile strength	0.4	0.3	0.55	0.165
2	Quality of machining	Not meeting the desired tolerance level	0.2	0.6	0.6	0.36
3	Delivery	Deviation from the committed date by the supplier	0.6	0.7	0.5	0.35
4	Installation	Fitment issue	0.5	0.68	0.6	0.408

0 being the lowest and 1 being the highest, risk which has a high impact and high likelihood of occurrence was given high priority, while in case the impact of risk is high but the likelihood of risk to occur is low in that case the priority given to the risk should be less, a risk having a median impact but the likelihood of occurrence is high is prioritize more as there is more relevant to the company.

To showcase a factor considering the above overall risk score is used which collectively uses the impact and likelihood and gives better insight on prioritization of risk.

To rank the risks by considering the probability and deductibility of the risks, failure mode effects analysis (FMEA) model was used as the reference base.

FMEA can be explained by an illustration, let us say a machine is broke down into smaller sub-assembly followed by parts, and for each part, potential failure mode and effects caused by them are noted the whole idea is to measure the effect in a numerical term such that better controllable measure could be taken and in turn, the output of controllable measure can also be measured, the technique is highly useful in the manufacturing industry and is applied in the service industry as well.

Overall FMEA provides a framework for problem-solving, in our case although we thought of all the risks that could occur with both suppliers and came down to a conclusion to go ahead with supplier 1, in real-life scenarios nothing is perfect and there are chances that in work allocated to supplier 1 something might go wrong.

FMEA helps in the following areas:

1. Helps to identify potential problem
2. Helps to identify preventive action
3. Helps to identify contingency action

Preventive action is the one to avoid the occurrence of the potential problem in the first case whereas the contingency plan is the one which is used in the case where the potential problem has occurred. The same ideology was used in developing Table 33.5. Wherein the risk is measured in numerical terms.

In Table 33.5, the value mentioned and the risk identified were mentioned by consensus of both the internal project and design teams as technical insight and supplier capability were observed by them.

Installation was identified as the risk which needs to be prioritized the factor behind that risk are

- Criticality of the design
- New machine one of its types
- The Fitment issue of all parts and movement of parts in motion

Quality of machining

- The tolerance level committed on paper by the supplier and maintained on the machine might differ
- Post-corrosion resistance coating chances of fitment may be tight since the tolerance level are close

Quality of material

- The chemical report does not meet the desired level
- The tensile report did not meet the desired level
- The machine will be used to heat a product in that case if material quality is not good then chances of sagging and breakage are high

Delivery

Sometimes suppliers did not assess the project properly and committee a date to supply the material with a thought to get the project in the first place, post start of working supplier gives various reason for not being able to comply with the date mentioned in the contract some of the reasons include.

- Unavailability of the operator of the machine
- Raw material unavailability
- Material available at a higher price compared to the price on the day of signing the contract
- Power cut issue

Note: The likelihood of an issue in the quality of the material was less as the material was tested at internal company premises.

Post prioritizing, the risk next on the basis of the matrix responsible, accountable, support, consult and inform (RASCI) risk mitigation and concerned responsibility table were developed where each risk identified was allocated to a concerned person who will make sure that the risk is reduced and in the first place should not occur (Table 33.6).

Also, post-risk assessment, a meeting was set up with the supplier, and the below questions were asked:

1. Date of delivery after the purchase order is released.
2. Mechanical and chemical test reports from a certified testing lab.
3. A timely update on the status of manufacturing.

Table 33.6 Risk mitigation and concerned responsibility

Risk	Responsibility	Risk mitigation measures
1. Quality of material	Project	Material test report duly checked before the start of manufacturing
2. Quality of machining	Design, project	Supplier visits on a weekly basis to assess the progress and help in any technical issue and if any bottleneck is identified try to solve it at the supplier premises
3. Delivery	Supplier	Written commitment and failure to adhere may have an impact on future new project orders to the supplier
4. Installation	Supplier, execution, project	At least 2 persons from supplier end for installation in company premises till the first ok piece comes, Internal execution team will hold the primary responsibility to assembly and the required assembly details will be provided by the project team

4. In-house assembly and no full payment till the first ok piece check.
5. Assembly post coating at supplier premises as well to check the fitment.
6. Service support till the next 6 months to rectify issues that might occur in the machine.

Finally, after due diligence, an agreement was signed with the supplier, and the order was released.

From the above case, we can see how critical risk management was and the vision required in the supply chain to eliminate critical risk at the start during supplier selection and evaluation.

33.6 Conclusion

The supply chain is tricky and will be hampered by any global issues, and the supply chain will be impacted more if risks associated with it are not assessed and reduced. Supplier prioritization and risk management go hand in hand and are closely related to each other use of project management orientation helps to better manage risk management and imbibe a culture in a company as in our case methods like risk assessment matrix, comparison matrix, and considering the issue at supplier end and redesign the product, collectively helped to reduce/mitigate the risk in our case.

In our scenario of supplier preferencing from the case study, we can see how the risk which might occur is considered.

Further, at a granular level, the risk associated is decomposed, and then associated prioritization was given Some of these risks are known risks, and some are unknown risks which are not visible to the purchase team. The severity of known risk is reduced by following the methodology of a risk matrix as highlighted in the case of the manufacturing industry and assigning a responsible department for the same such that they take it their prime focus area and measurement criteria.

In manufacturing, it is said that a bottleneck process, if it is at the start or end, can be controlled but if it is in the middle, it becomes difficult. So, to improve the bottleneck process most of the available resources need to be allocated to the bottleneck such that the bottleneck is improved and overall, the process is improved.

Similarly, before finalizing the supplier, internal resources should be allocated to supplier assessment and identifying the risk for the same such that in case of the project many times supplier might become a bottleneck the chances of the same are reduced.

This case implementation was specifically done in the manufacturing sector. Further research can be done in the said area of supplier prioritization and risk management and transfer the ideas to different sectors and a suitable methodology can be designed to help in supplier preferencing and risk mitigation. This will enable companies to proactively handle supply chain risks in their value chain. A global study can also be undertaken to give valuable insights to implement the solution in other countries and the applicability of the same due to cultural differences.

References

1. Hamdi F, Ghorbel A, Masmoudi F, Dupont L (2018) Optimization of a supply portfolio in the context of supply chain risk management: a literature review. J Intell Manuf 29:763–788
2. Wetzstein A, Hartmann E, Benton WC, Hohenstein NO (2016) A systematic assessment of supplier selection literature—state-of-the-art and future scope. Int J Prod Econ 182:304–323
3. Er Kara M, Oktay Fırat SÜ (2017) Supply chain risks literature review and a new categorization. Beykent Univ J Sci Eng 10:31–60
4. Manners-Bell J (2018) Supply chain risk management. In: Understanding emerging threats to global supply chains, 2nd ed. Kogan Page, London, UK. ISBN 978-0-7494-8015-8
5. Chen PS, Wu MT (2013) A modified failure mode and effects analysis method for supplier selection problems in the supply chain risk environment: a case study. Comput Ind Eng 66:634–642
6. Yu MC, Goh M (2014) A multi-objective approach to supply chain visibility and risk. Eur J Oper Res 233:125–130
7. Tapiero CS (2008) Analyse des risques et prise de décision dans la chaine d'approvisionnement. Rev Fr Gest 34:163–182
8. He B, Yang Y (2018) Mitigating supply risk: an approach with quantity flexibility procurement. Ann Oper Res 271(2):599–617
9. Fahimnia B, Tang CS, Davarzani H, Sarkis J (2015) Quantitative models for managing supply chain risks: a review. Eur J Oper Res 247(1):1–15

Chapter 34
An Empirical Study on Recruitment Management Systems: Start of a New Era

K. V. N. Lakshmi, Y. Fathima, R. Ravichandran, and N. Sathyanarayana

Abstract The importance and relevance of AI in current times cannot be overemphasized. As with evolving technologies, AI too has grown in size and scope and permeated across disciplines and industry sectors. Starting from making the highly mundane and process-oriented tasks, into automation mode, it also provides a higher level of accuracy, comfort and economic value to both the industry applying AI and the consumer who gets a standard quality product or service at reasonable costs. AI and machine learning play thus a very crucial role in different processes and help improve accuracy, enhance the quality of prediction and enable better decisions. As always these involve rule-based heuristics processes, interactive processes, which are constructed to fit an algorithm, that can clearly express the relationship and process the data to give the final output. This has the aspect of higher accuracy, greater automation and lesser human intervention. Typically, the scenario that happens is that AI algorithms are applied using Machine learning applications, which then coordinate together to give the desired output. Typical areas where applications of AI can be seen are related to a Human Resource Management System (HRMS)—social recruiting, automated computerized job advertisements, candidate identification, flagging potential issues, apportionment of timeslots in line with the supervisor's availability for job interviews as also automating the first level interview process. Over time it is expected that only selected positions will go for an interview by an HR manager and the rest may be done by the use of AI-ML integrated platforms, using bots. These have serious amplifications and consequences for the industry and how and what roles are likely to be taken up, depending on the convenience, payoff arising

K. V. N. Lakshmi (✉) · Y. Fathima · N. Sathyanarayana
JAIN (Deemed-to-Be University), Bangalore, India
e-mail: kvn.lakshmi@jainuniversity.ac.in

Y. Fathima
e-mail: y.fathima@jainuniversity.ac.in

N. Sathyanarayana
e-mail: n.sathyanarayana@jainuniversity.ac.in

R. Ravichandran
Kirtane & Pandit LLP Chartered Accountants, Pune, India
e-mail: ravi.ramadass@kirtanepandit.com

M. K. Tiwari et al. (eds.), *Applications of Emerging Technologies and AI/ML Algorithms*, Asset Analytics, https://doi.org/10.1007/978-981-99-1019-9_34

in the process, etc. This study examines the various aspects of these technologies and how they can help in the process of recruitment, accelerate the onboarding process and what factors help one to identify the process, examine the use of social media, social recruiting, virtual assessment, speeding up the hiring process and the like. Based on the specially designed questionnaire for this purpose, the authors propose an empirical study on a sample response of 281 individuals, across different levels from managers, senior executives and to examine how far the factors considered are responsible for explaining the impact of the changes happening in the application of AI and ML in the recruitment, selection and training process, and whether they have a significant role to play in the HRMS, and whether the use of AI and ML in HRMS is going to increase. Using factor analysis and regression analysis the authors studied the role played by the various factors and conclude that these factors have got a significant impact on the process of automating the HR process and application of various modern technologies including AI and ML in the system. The authors conclude that going forward, the HR process across industries is likely to be automated, with benefits in terms of quality, cost and efficiency flows accruing to the industry and final pass down to consumers in the form of reduced costs for products and services.

Keywords Artificial intelligence · HR automation · Machine learning · Deep learning · Algorithm based selection and recruitment process · HRMS · Social recruiting

34.1 Introduction

The authors propose here to review previous research in this respect, as well as how the rising involvement of AI in HR automation activities will influence society and the economy. The days of sifting through classified advertising for a job are long gone. Most corporate positions are now advertised online, and digital workplaces are becoming the standard. Companies want the recruiting process to be as virtual as possible. Unemployment is nearing a peak, and firms are receiving so many applications that it may be difficult to keep track of them all. So, how do you gather and evaluate applications while still fairly assessing all candidates? It is possible to accomplish so by utilising a Recruitment Management System (RMS).

RMS is a piece of HR software that helps streamline and control the hiring process inside an organisation. It facilitates a full-fledged recruitment cycle from first job advertising to final offer, including applicant management and interview scheduling. Most systems have a job posting template of a Job description and Job specifications like job title, location, employment type, education, level of skills, years of experience, etc.

Once the company enters the job in the recruitment management system, various stages of the recruitment process need to be set up. These stages are based on the company's hiring needs. RMS gathers all information submitted by an applicant when they apply for a job. It also allows recruiters to exclude prospects who do not satisfy the system's predefined criteria. To maintain track of each person's details, most systems allow recruiters to take extensive notes on their contacts with each prospect. Each recruiter has a personal account that is accessible via their business's email address, allowing the firm to reply directly to applicant applications within the recruitment management system and schedule interviews. When required, the organisation can even reject candidates with individualised comments. Nowadays, a company may complete all of the processes in the recruiting process from a single platform, making the recruitment management system a lifesaver for companies with ambitious hiring ambitions.

34.2 Review of Literature

Recruitment and selection are defined by Edwin Flippo as "Recruitment is the process of locating and attracting prospective employees to apply for openings inside an organisation."

In layman's terms, recruiting and selection are mutually exclusive procedures. They are very different from one another yet are critical components of the organisation. It aids in determining the potential and skills of applicants for anticipated or real organisational openings. It connects job seekers and employers.

Human resource management theorists like Korsten [15] and Jones et al. [13] extol the virtues of interviews, evaluations, and psychometric testing in the course of the employee selection process, emphasising the need for a thorough and methodical approach to recruitment and selection. They continued by saying that online applications and interviews might be part of an internal or external employment procedure. This system is often based on policies, job listings, and information as well as advertisement, the initial screening process, evaluation, decision-making, final selection, and training [15].

Jones et al. [13] state that the procedures involved in designing recruitment policies and defining management objectives may be gleaned from instances of recruiting policies in the healthcare, commercial, and industrial sectors.

Successful recruitment techniques involve things like reviewing the job description, the current labour market scenario/conditions, conducting interviews, and administering psychometric testing to candidates. Jones et al. [13] provide a number of other selection procedures, such as a variety of interviews, role plays, group discussions, and group projects.

Recruiting is at the heart of every management process, and failing to do so may create issues and undesirable barriers for any organisation, such as detrimental effects on profitability and inadequate staffing or employee skills [13].

According to Karade and Dsouza [14] author of HRM in a Business Context, the term "recruitment and selection" refers to the procedure of finding and attracting competent job candidates. He argues that finding and selecting the most qualified candidates is more of a complex task that calls for strategic management decision-making and comprehensive preparation. This would suggest that management places a premium on workers who are both competent and familiar with the needs of the position. In light of the fact that the ability to operate in a team is essential for every manager to succeed.

Human resource management (HRM) practises, HRM-organizational strategies, and organisational performance were all linked by Hiltrop [10] Further research by Hiltrop [10] confirmed that careful recruiting improved organisational performance, providing important takeaways for decision-makers in the business world. Furthermore, the process of hiring employees has become a major talking point. Successful businesses invest more in training, especially in soft skills like communication and cooperation, because of their thorough hiring procedures [11].

Human resource management strategies are used in any organisation with the end aim of enhancing the company's performance and revenues via the training and development of employees, as stated by Jackson et al. [12].

Bratton and Gold [5] state that in order to find employees who are a good fit for their company, businesses develop models of the kind of people they want to work there, and then use rigorous and precise selection procedures to narrow down the pool of applicants. However, the discussion between candidates and employers that will decide the nature of their working relationship doesn't begin and end with the selection and recruiting processes.

Recruitment and selection provide an opportunity to elaborate on applicants' preconceptions about the company's treatment of them [5, 12]. Realistic job previews (RJPs), which might include case studies of employees and their overall work and experiences, the chance to "cover" someone at work, employment samples, and films, are a method for constructing the perspective, as stated by Bratton and Gold [5]. The purpose of RJPs is to help job-seekers have more reasonable expectations.

Ability management was the main emphasis of Silzer et al. [21] research, and they were successful in addressing issues like the debate over whether or not talent is innate or acquired.

After considering the same issues in the BOU, Bizer et al. [4] suggested changing the recruiting process by making it HR's obligation to publicise the position directly rather than relying on any intermediary authority. The plan is to get rid of all the things that don't bring value.

34.3 Objectives of the Study

1. To identify the different components from the variables developed for the study.
2. To know the impact of the different components identified on the effectiveness in the decision-making.
3. To know the effect of predictor components on the effectiveness of the recruitment process of the organization.

34.4 Sample Design and Methodology

34.4.1 Method of Research

An empirical study on recruitment management systems: Start of a new era.

34.4.2 Sampling Technique

The data is collected using the purposive sampling technique to suit the research needs and specifically approached individuals with certain characteristics.

34.4.3 Sample Unit

The population selected for this study is managers, vice presidents, intrapreneurs, CMDs and team leaders.

34.4.4 Sample Size

In total, 281 respondents were interviewed who are from different demographic and professional backgrounds.

34.5 Limitations of Study

Through the study tried to cover all aspects of RMS, it is not out of limitations only 281 respondents are selected for the study time is another constraint with a restricted preview of the study.

34.6 Data Analysis

The KMO MSA test was performed to determine if the data is appropriate for factor analysis. Since the value is 0.788, this data is eligible for FA (Table 34.1).

The following table shows the test results of the factor loadings (Table 34.2).

Table 34.1 KMO and Bartlett's test

Kaiser–Meyer–Olkin measure of sampling adequacy		0.788
Bartlett's test of sphericity	Approx. Chi-square	11,140.789
	df	120
	Sig	0.000

Table 34.2 Rotated component matrix[a]

	Component		
	1	2	3
This will support optimizing job advertisements	−0.022	0.985	0.148
An automated purchasing space will be provided by software for advertising a vacancy through programmatic job advertisements	0.007	0.974	0.137
It empowers advertisers to laser-target their ideal demographics	−0.008	0.972	0.137
It will provide tangible results and a high Return on Investment	−0.040	0.979	0.144
It enables advertisers to easily access details on which stage a candidate is in the hiring process. This is one of the major benefits of ATS	0.139	0.131	0.967
It will allow the organization to clearly see gaps in the hiring process. This makes the hiring process more transparent	0.131	0.153	0.964
It provides ways to automate manual tasks	0.129	0.152	0.967
It enables speeding up hires without sacrificing quality	0.126	0.147	0.966
It provides enough sourcing of a candidate's resume	0.900	−0.051	0.058
Social media may give extra information about applicants for use in sourcing choices	0.887	−0.050	0.041
Through social recruiting, there is a less chance of losing out best candidates	0.901	−0.049	0.062
This will stimulate the sensory experiences of physically being there. This enables potential candidates to know exactly to know what it is like to be a member of the organization. This is more effective comparatively	0.920	0.020	0.133
The virtual application assessment process will streamline the assessment process comparatively	0.904	0.017	0.106
Virtual/video interviewing saves time, efforts, and travel cost	0.933	0.030	0.162
VR interviews will boost the effectiveness of recruiting managers in making decisions	0.935	0.026	0.163
Extraction Method: Principal Component Analysis Rotation Method: Varimax with Kaiser Normalization			

[a]Rotation converged in 5 iterations

Table 34.3 (VR interviews) - Model summary

Model	R	R Square	Adjusted R square	Std. error of the estimate
1	0.956[a]	0.914	0.913	0.270

[a]Predictors: (Constant), Application_Tracking_System, Social_Recruiting, Programmatic_Job_Advertisements

Table 34.4 (VR interviews) - ANOVA[a]

Model		Sum of Squares	Df	Mean Square	F	Sig
1	Regression	387.165	3	129.055	1773.159	0.000[b]
	Residual	36.610	503	0.073		
	Total	423.775	506			

[a]Dependent Variable: VR interviews will boost the effectiveness of recruiting managers in making decisions.
[b]Predictors: (Constant), Application_Tracking_System, Social_Recruiting, Programmatic_Job_Advertisements

Table 34.5 (VR interviews) - Coefficients[a]

Model	Unstandardized coefficients		Standardized coefficients		
	B	Std. error	Beta	t	Sig.
1 (Constant)	−0.632	0.092		−6.854	0.000
Application_Tracking_System	0.050	0.016	0.044	3.083	0.002
Programmatic_Job_Advertisements	0.024	0.016	0.021	1.555	0.121
Social_Recruiting	1.077	0.015	0.944	69.614	0.000

[a]Dependent Variable: VR interviews will boost the effectiveness of recruiting managers in making decisions

Table 34.6 (RMS results) - Model summary

Model	R	R square	Adjusted R square	Std. error of the estimate
1	0.701[a]	0.491	0.485	0.671

[a]Predictors: (Constant), Application_Tracking_System, Social_Recruiting, Programmatic_Job_Advertisements

Table 34.7 (RMS results) - ANOVA[a]

Model		Sum of squares	df	Mean square	F	Sig.
1	Regression	119.768	3	39.923	88.735	0.000[b]
	Residual	124.175	276	0.450		
	Total	243.943	279			

[a]Dependent Variable: RMS results in an effective recruitment process for an organization.
[b]Predictors: (Constant), Application_Tracking_System, Social_Recruiting, Programmatic_Job_Advertisements

Table 34.8 (RMS results) - Coefficients[a]

Model	Unstandardized coefficients		Standardized coefficients		
	B	Std. error	Beta	t	Sig.
1 (constant)	0.087	0.310		0.281	0.779
Social_Recruiting	0.813	0.052	0.696	15.680	0.000
Programmatic_Job_Advertisements	0.153	0.052	0.132	2.928	0.004
Application_Tracking_System	−0.024	0.054	−0.020	−0.441	0.659

[a]Dependent variable: RMS results in effective recruitment process for an organization

Table 34.9 Social_medias_Identified frequencies

		Response		Percent of cases (%)
		N	Percent (%)	
Social_medias_Identified[a]	Facebook	175	24.3	62.7
	LinkedIn	175	24.3	62.7
	Twitter handle	99	13.7	35.5
	WhatsApp	195	27.0	69.9
	Skype	77	10.7	27.6
Total		721	100.0	258.4

[a]Dichotomy group tabulated at value 1

Table 34.10 Software_For_RMS frequencies

		Responses		Per cent of cases (%)
		N	Percent (%)	
Software for RMS[a]	JazzHR	184	27.8	65.7
	RecruiterBox	149	22.5	53.2
	Merittracs	48	7.3	17.1
	Zoho Recruit	138	20.9	49.3
	Recruiterflow	60	9.1	21.4
	Recruit CRM	82	12.4	29.3
Total		661	100.0	236.1

[a]Dichotomy group tabulated at value 1

By combining many variables, the data were summed up in 3 dimensions. Factors were taken out with the help of the principal component. 15 different things were narrowed down to 3 factors. These three things were named:

- Social Recruiting
- Programmatic Job Advertisements

- Application Tracking System

Various variables are grouped into these components identified. The variables which are grouped under each component are as follows.

34.6.1 Social Recruiting: (SR)

- It provides enough sourcing of a candidate's resume.
- Social media may give extra information about applicants for use in sourcing choices.
- Through social recruiting, there is less chance of losing out best candidates.
- This will stimulate the sensory experiences of physically being there. This enables potential candidates to know exactly to know what it is like to be a member of the organization. This is more effective comparatively.
- The virtual application assessment process will streamline the assessment process comparatively.
- Virtual/video interviewing saves time, effort, and travel cost.

34.6.2 Programmatic Job Advertisements: (PJA)

- This will support optimizing job advertisements.
- An automated purchasing space will be provided by software for advertising a vacancy through programmatic job advertisements.
- It empowers advertisers to laser-target their ideal demographics.
- It will provide tangible results and a high return on investment.

34.6.3 Application Tracking System: (ATS)

- It enables advertisers to easily access details on which stage a candidate is in the hiring process. This is one of the major benefits of ATS.
- It will allow the organization to clearly see gaps in the hiring process. This makes the hiring process more transparent.
- It provides ways to automate manual tasks.
- It enables speeding up hires without sacrificing quality.

$\mathbf{H_0}$: There is no impact of ATS, SR, and PJA on 'RMS will increase the efficiency of decision-making hiring managers' (Table 34.3).

$\mathbf{H_1}$: There is an impact of ATS, SR, and PJA on 'RMS will increase the efficiency of decision-making hiring managers'.

Predictor variables namely, ATS, SR, and PJA were regressed on the dependent variable Whether RMS will increase the efficiency of decision-making hiring managers. The R2 is showing 0.914 which depicts that the model equation explains 91.4% of the variance in RMS decision-making (Table 34.4).

Overall the regression coefficient is significant is statistically significant with a p-value < 0.05 because of which the null hypothesis is rejected and concluded that there is an impact of ATS, SR and PJA on 'RMS will increase the efficiency of decision-making hiring managers' (Table 34.5).

Standardised coefficients Beta for SR is 0.944 which is influencing decision-making hiring managers to a greater extent comparatively whereas PJA is showing the least influence on the dependent variable.

$\mathbf{H_0}$: There is no impact of ATS, SR, and PJA on 'RMS result in the effective recruitment process for an organization'.

$\mathbf{H_1}$: There is an impact of ATS, SR, and PJA on 'RMS result in the effective recruitment process for an organization' (Table 34.6).

The predictor variables ATS, SR, and PJA is showing 0.491 variances in the Dependent variable 'RMS results in the effective recruitment process for an organization'. So the impact of independent variables is to the extent of 49.1% on the dependent variable (Table 34.7).

ANOVA being a mean squared test is applied to determine whether factors (treatments) are significant. Since the F value is high (88.375) and low significant value (0.000) it can be said that: there is an impact of ATS, SR, and PJA on 'RMS results in the effective recruitment process for an organization'. Thus null hypothesis is rejected (Table 34.8).

Standardized coefficients beta for SR is 0.696 which is again showing a higher impact on the dependent variable RMS results in an effective recruitment process for an organization. But ATS is showing a negative influence on the dependent variable.

The respondents were asked to identify the social media which the respondents identified most of the time. They were asked to select those sites in which they are actively participating. Multiple response analysis was conducted for the same (Table 34.9).

The majority of the respondents are actively participating in WhatsApp, and 175 respondents are active on Facebook and LinkedIn. Less percentage of respondents are involved through Twitter followed by it is skype.

RMS needs software to automate the recruitment process. The respondents were asked to identify the prominently used software for the recruitment process. Multiple response analysis was conducted for the same (Table 34.10).

JazzHR happens to be the prominent RMS Software.

34.7 Findings

The study was undertaken to study the RMS and its effectiveness in the recruitment process. 15 variables were developed to analyse the respondent's opinions on the same. Factor analysis was conducted for the same. KMO's Measure of Sampling Adequacy was conducted to know whether the sample is adequate or not. Since the value was greater than 0.06, FA was conducted. The results were generated from the rotated component matrix. The entire set of variables is divided into 3 components, i.e., Social Recruiting, Programmatic Job Advertisements, and Application Tracking systems. Further, to analyze the impact of these components regression analysis to conduct.

The dependent variable was whether RMS will increase the efficiency of the decision-making hiring managers and the dependent variables were SR, PJA, and ATS. Since the R square was high, it was concluded that there is a high level of impact of these predictor variables on the dependent variable. So, the null hypothesis was rejected. RMS will ease the job of the hiring managers enabling them to speed up the recruitment process due to which the efficiency of the managers will increase.

Another hypothesis was developed to test whether RMS will effective recruitment process for an organization. The same predictor variables were regressed on this dependent variable. The R square was higher again resulting in rejecting the null hypothesis.

Multiple response analysis was conducted to identify the commonly used social media platforms in which respondents are active. From the results, it was found that WhatsApp is the most commonly used social media platform. Followed to which are Facebook, LinkedIn, Twitter, and Skype. The commonly used recruitment software for RMS was JazzHR as opined by respondents.

34.8 Conclusion

The importance of the use of automation and AI in RMS is increasing day by day and is likely to be more pronounced due to various mixes of available technology solutions. These are likely to increase the productivity, effectiveness, and efficiency of the overall system. The different HR software which is popularly used in the market is a testimony to the same. However, it is not a mixed blessing, As with any AI initiative these are likely to face passive resistance from workers and employees, and may need persuasive skills of the highest order from middle and top management to address the fallout and make the initiative a success. Another possible dimension is the impersonal, cold approach practised in automation which may lack a sense of human touch and feel, and sensitivities of languages and semantics may come into play here especially if one single algorithm based software is used across geographies without thought for differences in culture too. These may create possible bias, and bring about a trust deficit which needs to be addressed carefully by the top management. AI and automation in HR have travelled much far and go into face recognition, and retina-based access to employees which may indicate a crossing line into the forbidden area of privacy and a delicate balance needs to be hit between the technology and human social traits. With the role of AI, and Robots in automation tasks going up right from recruitment it may be possible that special counsellors of a different sort are needed to alleviate the apprehensions of the employees in a free and frank discussion, which seldom is done in reality by management. This can compound the already fragile HR ecosystem which is getting increasingly dominated by AI and automation. These call for skills of a very high capability and grade so that we can get the benefit of both the human touch and sensitivity aspect as well as the improvement in efficiency and effectiveness in the HR process using an RMS. All in all one case certainly Automation and AI, as well as deep learning applications, linked with IoT devices promise a good future if used carefully by a skilled manager and this puts pressure on top management as well as middle management to be ready with change strategies and fathom out the resistance to change aspect. Although it is too early to speculative on which direction the future is likely to unfold towards us, it is certainly on the key aspect that: (i) AI and automation in HR tasks will accelerate over the next few years and (ii) IoT devices will increasingly play a part in recruitment training, simulation, development, and assessment process. (iii) Resistance to change will be the passive and deliberate outcome of management responses. (iv) Trust between employee and employer may become a causality in the process. (v) Continuous investment efforts in training and development are called for, with reskilling and upskilling which only can perhaps help sustain the employee in the job he holds and be productive in the process. (vi) People and Employees may have to make significant efforts in terms of time, money and resources to keep pace with developments in technology. The benefits however are worthy of a cause and may be adopted or looked for being internalised in a structured and deliberate manner.

References

1. Allen DG, Mahto R, Otondo RF (2007) Web-based recruitment: Effects of information, organizational brand, and attitudes toward a Web site on applicant attraction. J Appl Psychol 92:1696–1708
2. Alleyne C, Kakabadse A, Kakabadse N (2007) Using the HR intranet. Pers Rev 36:295–310
3. Bell BS, Lee SW, Yeung SK (2006) The impact of e-HR on professional competence in HRM: Implications for the development of HR professionals. Hum Resour Manag 45:295–308
4. Bizer C, Heese R, Mochol M, Oldakowski R, Tolksdorf R, Eckstein R (2005) The impact of semantic web technologies on job recruitment processes. In: Wirtschaftsinformatik 2005: eEconomy, eGovernment, eSociety. Physica-Verlag HD, pp 1367–1381
5. Bratton J, Gold J (1999) Human resource management: theory and practice. Macmillan, Basingstoke
6. Feldman DC, Klaas BS (2002) Internet job hunting: a field study of applicant experiences with on-line recruiting. Hum Resour Manage 41:175–192
7. Florkowski GW, Olivas-Lujan MR (2006) The diffusion of human-resource information technology innovations in US and non-US firms. Pers Rev 35:684–710

8. Grant D, Dery K, Hall R, Wailes N, Wiblen S (2009) Human resource information systems (HRIS): An unrealized potential. In: The Paper presented at the chartered institute of personnel and development, Nottingham, UK
9. Haines VY, Lafleur G (2008) Information technology usage and human resource roles and effectiveness. Hum Resour Manage 47:525–540
10. Hiltrop JM (1996) The impact of human resource management on organisational performance: theory and research. Eur Manag J 14(6):628–637
11. Hiltrop JM (1999) The quest for the best: human resource practices to attract and retain talent. Eur Manag J 17(4):422–430
12. Jackson DL, Gillaspy JA Jr, Purc-Stephenson R (2009) Reporting practices in confirmatory factor analysis: an overview and some recommendations. Psychol Methods 14(1):6
13. Jones DA, Shultz JW, Chapman DS (2006) Recruiting through job advertisements: the effects of cognitive elaboration on decision making. Int J Select Assess 14(2):167–179(13)
14. Karade MM, Dsouza MJ (2009) A study of recruitment and repudiation of offers: a case of daffodils systems pvt. limited, Pune. Emerg Trends Manag Technol 40
15. Korsten AD (2003) Developing a training plan to ensure employees keep up with the dynamics of facility management. J Facil Manag 1(4):365–379(15)
16. Kovach KA, Hughes AA, Fagan P, Maggitti PG (2002) Administrative and strategic advantages of HRIS. Employ Relat Today 29:43–48
17. Lievens F, De Corte W, Westerveld L (2015) Understanding the building blocks of selection procedures: effects of response fidelity on performance and validity. J Manag 41:1604–1627
18. Marler JH, E Parry (2015) Human resource management, strategic involvement and e-HRM technology. Int J Human Resour Manag 1–21. https://doi.org/10.1080/09585192.2015.1091980
19. Martin G, Reddington M (2010) Theorizing the links between e-HR and strategic HRM: a model, case illustration and reflections. Int J Hum Resour Manag 21:1553–1574
20. Ruta CD (2009) HR portal alignment for the creation and development of intellectual capital. Int J Hum Resour Manag 20:562–577
21. Silzer R, Church AH (2010) Identifying and assessing high-potential talent. Strat-Driven Talent Manag: LeadShip Imp 28:213–280
22. Svoboda M, Schröder S (2001) Transforming human resources in the new economy: developing the next generation of global HR managers at Deutsche Bank AG. Hum Resour Manage 40:261–273
23. Tansley C, Watson T (2000) Strategic exchange in the development of human resource information systems (HRIS). N Technol Work Employ 15:108–122
24. Wilson-Evered E, Hartel CEJ (2009) Measuring attitudes to HRIS implementation: a field study to inform implementation methodology. Asia Pac J Hum Resour 47:374–384

Chapter 35
Traceability of Unwitting Disclosure Using Explainable Correlation in Procurement and Supply Chain

Harish Vishnu Gunjal, Vaibhav Ingale, Shikhar Bhardwaj, and Rajendra M. Belokar

Abstract Many firms maintain highly complicated and worldwide supply chains in today's world, which results in tons of transactions. Because of the gap in record maintenance for transactions, there arises the possibility of an anomalous transaction, i.e., any suspicious transaction that does not follow the historical pattern of regular transactions, thus creating a demand for disclosure of procurement and supply chain transactions. In this paper, the authors define the method of segregation of transactions and obtain derived parameters from a study. This also establishes a model-based correlation between the derived parameters and the blockchain. Blockchain plays a critical role in creating immutable records. It will help to reduce human intervention and induced errors, maintain transparency in the transactions, and improve the verifiability and traceability of an anomalous transaction during an audit.

Keywords Blockchain · Traceability · Transactions · Transparency · Procurement · Supply chain

---

Vaibhav Ingale and Shikhar Bhardwaj: These authors contributed equally to this work.

H. V. Gunjal (✉) · R. M. Belokar
Department of Production and Industrial Engineering, Punjab Engineering College (Deemed to Be University), Chandigarh, Chandigarh 160012, India
e-mail: harishvishnugunjal.mtprod20@pec.edu.in

R. M. Belokar
e-mail: rmbelokar@pec.edu.in

V. Ingale (✉) · S. Bhardwaj (✉)
Department of Computer Science and Engineering, Punjab Engineering College (Deemed to Be University), Chandigarh, Chandigarh 160012, India
e-mail: vaibhavshamraoingale.phd21cs@pec.edu.in

S. Bhardwaj
e-mail: shikharbhardwaj.phd19cse@pec.edu.in

M. K. Tiwari et al. (eds.), *Applications of Emerging Technologies and AI/ML Algorithms*, Asset Analytics, https://doi.org/10.1007/978-981-99-1019-9_35

35.1 Purpose of Research

In the era of the modern world, demand for the supply chain is increasing, especially in developing countries like India. With this, tons of tangible and intangible transactions occur among companies, warehouses, stakeholders, and customers. Some companies keep a transaction record, while some omit non-desired transactions [1]. These transactions may or may not be anomalous. Thus, this research focuses on tracing anomalous transactions.

Tracing is essential because some measures or ACTs enforce the transparency and visibility between the transactions, yet, the measures are not effective for anomalous transactions. There is a need for some improvised strategy [2].

The authors emphasize establishing an explainable correlation of procurement and supply chain with blockchain technology, which develops a generalized model that can be used to trace anomalous transactions. The paper shall review numerous strategies to trace anomalous transactions. This paper consists of five sections; the first section briefly introduces procurement and supply chain transactions and dashboard studies and explains the blockchain concept. The second section enumerates the successful procurement and supply chain disclosure strategies. The third section uses syntactic analysis to classify transactions by type, construct a link with the blockchain, successfully implement strategies, and develop a generic model for tracing unwitting transactions in procurement and supply chains. Section four provides an idea about the benefits observed after a model implementation in the supply chain. The fifth section defines the implications of the model derived in the study and provides scope for future research in this area.

35.2 Methodology

This study presents a model-based blockchain solution for procurement and supply chain transactions, which will aid in maintaining transparency and tracing anomalous transactions. The research looks into the use of blockchain in the supply chain in-depth and breaks it down into three questions, then try to answer them.

1. How to sort out the transactions based on their types?
2. Gaps in transactions and how to categorize/differentiate anomalous transactions?
3. How to trace anomalous transactions with the use of blockchain in procurement and supply chains? (Fig. 35.1).

Transactions: The standard definitions of transactions for different scenarios could be given as follows:

1. A general transaction is a business occurrence that has a monetary impact on an entity's financial statements and is recorded in its accounting records as an entry [3].

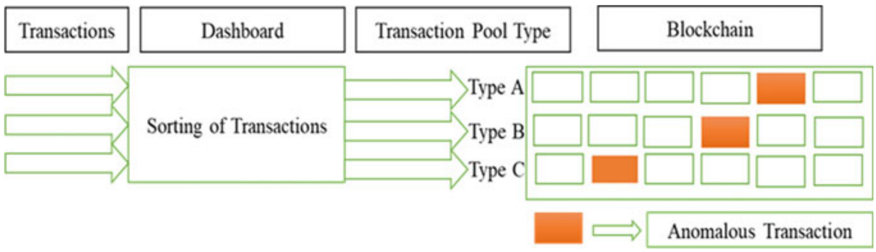


Fig. 35.1 Pipeline—tracing of anomalous transactions

2. A procurement transaction is a legal document through which a recipient or sub-recipient purchases the property or services required to carry out the project or programme specified in the award or sub-award.
3. A supply-chain transaction is a series of deliveries of the same commodities from one party to the next, all of which are transported directly from the first supplier to the ultimate customer. At least three subjects are engaged in chain transactions, although there are frequently more.
4. The lifecycle of a transaction begins with its creation, also known as origination. The transaction is then signed with one or more signatures, indicating that the money referenced by the transaction is authorized to be spent. The transaction is then broadcast on the bitcoin network, where it is validated and propagated by each network node (participant) until it reaches (nearly) every node in the network. Finally, a mining node verifies the transaction and adds it to a block of transactions that are stored on the blockchain. The transaction becomes a permanent part of the bitcoin ledger and is acknowledged as valid by all participants if it is published on the blockchain and confirmed by enough consecutive blocks (confirmations) [4].

Dashboard: It defines the sorting of the transactions criteria-wise (vary according to the company requirement). Sorting is any process of arranging items systematically, either arranging in a sequence or categorizing the items in similar groups. Sorting will help make groups that help the user identify the required transaction without any time-consuming, and it will also reduce the storage. Various techniques are available to sort the transactions like Quicksort, Bubble sort, Selection sort, Insertion sort, and many more [5].

In Blockchain, transactions are grouped into blocks to be efficiently verified and then synchronized with other computers on the network.

Transaction Pool Type: Basically, a transaction pool is a data structure containing a set of transactions that have not been mined but validated by a transaction miner. It stores the information on the transactions taken in procurement and supply chain, by suppliers, stakeholders, etc., according to the type generated by a company. For example, if the company procures the material, the pool type will be based on the amount of material or money required for procurement.

In a Blockchain, a transaction pool is stored on a special device, and its contents can be accessed or observed in real-time.

Blockchain: It is a system of recording information that makes it difficult or impossible to change, hack or cheat the system. A blockchain is essentially a digital ledger of transactions that is duplicated and distributed across the entire network of computer systems on the blockchain.

Blockchain in procurement and supply chain provides transparency at every stage of the product supply lifecycle by keeping a record of each transaction that occurred. Blockchain in procurement could benefit effectively by using smart contracts.

35.3 Example

Consider the Packaging Industry, in which multiple transactions are taken in different departments of the industry such as in the form of material and machinery procurement, supply of manpower in an industry like workers, managers, engineers, etc., products from different suppliers or companies for packaging, different vendors, and others.

Sorting of the transactions will be performed based on range. All the transactions are divided and classified into a range of money, assets, quantity, etc., according to company requirements. In this example, the authors decided to classify materials based on procurement in tons, manpower is classified based on class (which represents variable salary amount), and products received from different companies for packaging are classified based on quantities received in one transaction (Table 35.1).

Transaction Pool Type—After sorting transactions, a pool is generated, in which each transaction has three types—small, medium, and large. According to the transactions in each type (single transactions or group of transactions), blockchain technology generates the hash value of each type. Due to hash value, transactions will be secured and immutable. Once the hash value is generated for transactions, no one can change the transaction information stored in the block (without permission).

In this process of transactions to the blockchain, only at the initial step is there a chance of human intervention or error, i.e., a change in transaction information, which will automatically change the type of transaction (if the value lies in a different type). Eventually, it changes the hash value of the block.

Table 35.1 Sorting of transactions based on a range

Transactions	Small/low	Medium	Large/high
Material (price per ton)	0–100 ton	100–500 ton	500–1000 ton
Manpower (class)	Worker	Engineers	Managers
Products (lot size)	0–100	100–500	500–1000

Tracing of anomalous transactions—Suppose, after generating the hash value of transaction type, if anyone makes any change at the initial step or in the record, the hash value will automatically change. During an audit, per formed after some time (depending upon company policy), the auditors take the base ledger, divide the transactions into blocks as per the norms decided by the company/firm/organization, and verify the hash of the transactions stored in the block with the original transaction blocks [6]. Any change in block hash points towards the presence of anomalous transaction(s) is traceable and can be identified if the company wants any further investigation.

35.4 Implications

1. Implementation of the model will improve the transparency and traceability of both anomalous and non-anomalous transactions.
2. It will reduce the percentage of freight invoices containing inaccurate information, which are the principal source of conflicts.
3. The classification of transactions and tracing of anomalous transactions eases the audits.
4. The procurement and supply chain will be more efficient by reducing human interventions and errors.
5. It will help to maintain proper collaboration between companies.

35.5 Conclusion

Implementation of a derived model from the strategic blockchain, which will change the scenario of managing and tracing the transactions in procurement and supply chains among global companies. There are various benefits:

- Transparency and traceability,
- Security,
- The trusted relationship between stakeholders,
- A visible platform for bidding,
- Accurate flow of information and money among the companies through invoices,
- Reduction in the investment of money and time.

References

1. Dasaklis TK, Voutsinas TG, Tsoulfas GT, Casino F (2022) A systematic literature review of blockchain-enabled supply chain traceability implementations. Sustainability 14(4):2439
2. Omar IA, Debe M, Jayaraman R, Salah K, Omar M, Arshad J (2022) Blockchain-based supply chain traceability for covid-19 personal protective equipment. Comput Ind Eng 167:107,995
3. Raj PVRP, Jauhar SK, Ramkumar M, Pratap S (2022) Procurement, trace-ability and advance cash credit payment transactions in supply chain using blockchain smart contracts. Comput Ind Eng 167:108,038
4. Marshall D, McCarthy L, McGrath P, Harrigan F (2016) What's your strategy for supply chain disclosure? MIT Sloan Manag Rev 57(2):37–45
5. Akaba TI, Norta A, Udokwu C, Draheim D (2020) In: Conference on E-business, e-Services and e-Society. Springer, pp 3–14
6. Morishima S (2021) Scalable anomaly detection in blockchain using graphics processing unit. Comput Electr Eng 92:107,087