



A multivariate quantitative approach for sustainability performance assessment: An upstream oil and gas company

Navid Salmanzadeh-Meydani^{1,2} · S. M. T. Fatemi Ghomi¹ ·
Seyedhamidreza Shahabi Haghighi¹ · Kannan Govindan²

Received: 21 August 2021 / Accepted: 31 December 2021 / Published online: 12 February 2022
© The Author(s), under exclusive licence to Springer Nature B.V. 2022

Abstract

This paper presents a quantitative approach to evaluate the sustainability performance (SP) of an organization. A methodology is proposed based on principal component analysis (PCA), numerical taxonomy (NT), statistical, and cluster analysis. Related factors are determined considering sustainability dimensions, including social, economic, and environmental ones. A unique procedure is presented to commensurate the monetary factors. PCA is developed for multivariate analysis, which can rank annual records and determine factors' importance degree. NT is developed to verify and validate the results of PCA. The multivariate approach is able to rank the alternatives and simultaneously determine the importance degree of factors. An upstream oil and gas company is given as a case study. The statistical analysis showed direct relationships between the results of different analyses. In addition, the factors are categorized into two and four partitions to reduce dimensionality. The outcome-related factors are found to be of great importance for organizational SP. Furthermore, a descending SP in recent years was identified. To improve potential SP, operational planning for enhancing the performance of personnel is recommended. This paper provides intuition on the clusters obtained and why these factors affect the organizations' performance. The approach of this paper can be utilized for managers in the other industrial sectors who have access to company performance data and wish to analyze which factors are vital for company performance or hurting the company performance using the methodology presented in this paper.

Keywords Sustainability performance · Quantitative assessment · Sustainable multivariate analysis · Performance measurement · Organizational sustainability

1 Introduction

Nowadays, sustainability is an inseparable topic for companies and organizations, especially in their survival and competition. Poor sustainability performance (SP) deprives managers of the stability they need for decision-making and disruption creation (Ding

✉ S. M. T. Fatemi Ghomi
fatemi@aut.ac.ir

Extended author information available on the last page of the article

et al., 2016). This causes the organization to be unbalanced during a recession and takes it out of normalcy. Predetermined and anticipated decisions of managers to deal with such an atmosphere in a competitive world is of great importance. Organizational and management decisions play important roles in the overall performance of engineering, procurement, construction, and installation systems (Toutounchian et al., 2018). The ability to recover and return quickly after downturns is associated with the ability to cope with vulnerabilities. This ability will be enhanced if decision-making tools are used to deal with the effects of negative shocks.

Businesses need to identify the factors which affect their SP. More importantly, companies need to identify which elements to prioritize and accordingly manage their limited resources. Studying the key performance indicators (KPIs) of an organization in the past can help examine the SP. On the other hand, it will enable predictions about the organizational system behavior by assessing the past SP of an organization (Dev et al., 2020). Moreover, the past performance of the organization can also indicate its attention to KPIs. Based on an organization's past performance, it is possible to obtain comprehensive information about programs, identify challenges and make plans for the future. Hence, an organization's performance in the past plays a significant role in its current and future performance.

Since the strategic definition for individual KPIs may disrupt the organization or have high costs, they can be clustered. The combination of KPIs also helps develop strategies to improve the organization's SP and inspire new thinking toward the organization's goals. Accordingly, practical operational, tactical, and strategic planning can be associated with each cluster instead of planning for each KPI. Therefore, it is essential to propose an integrated approach to continuously assess and improve the SP based on its previous performance. For this purpose, the identification and investigation of KPIs affecting the improvement of the SP should be attained. To specify a set of factors for measuring SP, their equilibrium to ensure various dimensions is necessary.

This paper aims to assess the SP of an organization through associated KPIs to improve its functional operating level and return to its normal condition during probable social, economic, and environmental downturns and crises. Dealing with sustainability under social, economic, and environmental dimensions, the research objectives are to evaluate annual SP ranking and to realize the importance degree of factors to be monitored.

With this background in mind, this paper examines these research questions:

1. How can an integrated multivariate quantitative approach be proposed to assess the SP of an organization using annual historical data?
2. What is the importance degree of effective KPIs in SP?
3. What is the relationship between sustainability dimensions in terms of KPIs' classification?

This paper has addressed the aforementioned issue by introducing a simple-to-follow quantitative methodology with five phases and demonstrated a use case for an upstream oil and gas company. In the first phase, KPIs are determined concerning the sustainability dimensions. The second phase implies data gathering and scaling. A unique procedure is proposed for the monetary factors based on annual inflation rates (AIRs) and calculating consumer price indices (CPIs). Phase three develops principal component analysis (PCA) to achieve annual SP ranking and importance degree of KPIs simultaneously. In phase four, numerical taxonomy (NT) is developed to verify and validate the results of the developed PCA. Finally, phase five analyzes the results. Two quantitative methodologies are

presented, for which the description of steps of data preparation, computation, and inference are provided. The main contributions of this paper can be highlighted and summarized as follows:

- Proposing a simple-to-follow multivariate quantitative approach considering sustainability dimensions by synthesizing PCA, NT, clustering, and statistical analysis to assess SP;
- Proposing a unique procedure to commensurate the monetary factors based on AIRs and CPIs;
- Developing PCA which is capable of ranking annual SP and KPIs by defining utility functions;
- Developing NT which can examine both positive and negative ideal values and integrate them for simultaneously evaluating the annual SP and KPIs;
- Validating and implementing the proposed approach through application in a real-world case study.

The remainder of this paper is organized as follows. Section 2 reviews the related literature. Section 3 describes the proposed methodology. Section 4 deals with the case study and computational results. Verification and validation of the results are investigated in Sect. 5. Section 6 provides discussion and managerial insights. The conclusion is summarized in Sect. 7.

2 Literature review

Although many studies and efforts have been made to improve the SP of organizations, adverse events still occur due to recession and various crises (Azadeh et al., 2017). In times of crisis, organizations face significant threats to their financial, competitive, and survival potential (Pal et al., 2014). Azadeh et al. (2017) studied the performance of a production plant in crises considering management decisions using data envelopment analysis. Chen et al. (2019) proposed a mathematical formulation to capture favorable economic outcomes and environmental performance. Battisti et al. (2019) proposed an integrated approach to specify the role of learning in the performance of small enterprises using datasets in the period of financial crises. According to Alnajem et al. (2021), circular economy research exponentially increases with a 47.1% growth rate per year, and the sustainability topic has been one of the hotspots in recent years.

Many of the studies have tended to specify the organizational KPIs to assess SP. Hayami et al. (2015) found that the suppliers with fewer waste products (WP) desire to have better SP. Khan and Qianli (2017) examined the relationships between environmental and economic factors of the UK using time series data. Sabogal-De La Pava et al. (2021) considered the market value-added (VA) as a criterion to maximize the SP. Shan et al. (2021) studied the effect of economic and energy KPIs in limiting carbon emissions to evaluate SP. Pashapour et al. (2019) specified resilience-based factors to improve the performance of a petrochemical plant in terms of economics. Schwab et al. (2019) examined the effect of economic KPIs on the financial sustainability of a company in Swiss within business growth using quantitative and qualitative data. Ahmad, Muslija, et al. (2021) investigated the relationship between environmental sustainability and economic prosperity in 11 developing economies. D'Inverno et al. (2020) evaluated the performance of water utilities in

Italy considering financial and economic factors, service quality, and environmental sustainability. D'Ambra et al. (2020) assessed the economic and social effects of carbon dioxide on sustainability. Singh and Misra (2021) proposed a framework to evaluate the safety performance of laborers by identifying and ranking key safety indicators in the Indian construction industry. Ahmad, Chandio, et al. (2021) examined dynamic interactions among sustainable development, air pollution, and energy investment in China using panel data in 27 provinces of China.

Some scholars have used annual data to evaluate the SP. Ahmad, Jan, et al. (2021) provided a model accompanying sustainability challenges by considering financial-based factors as inputs of production growth during 1997–2017. Brandenburg (2016) evaluated the trends of cost efficiency and SP using 8 years of data in European automotive industries. Ganda and Milondzo (2018) investigated the effect of carbon emissions on the SP of some African organizations. Salmanzadeh-Meydani and Fatemi Ghomi (2019) examined the causalities between economic growth and capital stock in Iran by the vector autoregression model. Forteza et al. (2017) examined the relationship among economic performance, accident rate, and site risk of Spanish companies using annual data. Hussain et al. (2021) evaluated financial development and economic openness on SP using annual data of Pakistan and performing time series methods.

Various studies have assessed the SP in the supply chain. Huo et al. (2018) investigated an organization's ability to improve economic and operational performance through the flexibility of the supply chain. Sabbaghnia et al. (2019) investigated reducing carbon emissions on the members' profit in a supply chain. Das (2018) examined all three sustainability dimensions to determine their impacts on organizational performance in India. Arıkan et al. (2014) investigated the causalities between economic and environmental performance. Furthermore, most previous studies have focused on the economic dimension, while the social and environmental dimensions have still received less attention (Elfarouk et al., 2021). Moreover, the knowledge system that shapes companies' engineering, procurement, construction, and installation and encompasses their broad concepts and numerous complimentary ingredients have been neglected (Tseng et al., 2021). Even though there are many studies in terms of SP evaluation in the supply chain, the intra-organizational studies regarding the assessment of SP are limited and require more attention.

The literature determines that multiple-criteria decision-making method (MCDM) techniques considerably help to address and assess SP. Moreover, the application of MCDMs is closely interlinked to enable the managers of organizations to make better decisions. Solangi et al. (2020) applied Delphi and fuzzy analytical hierarchy process methods to evaluate the SP of electricity generation in Turkey. Although the application of PCA and NT exists in the literature (Azadeh et al., 2009, 2011; Pashapour et al., 2019; Shirali et al., 2013, 2016), the development to simultaneously ranking the alternatives and determining importance degree of variables has not taken adequate attention. Therefore, this paper contributes to the existing literature by developing PCA, and NT methodologies for the assessment in the context of SP, analyzing the results by utilizing statistical and clustering, and the application in a real-world case study.

3 Methodology description

Figure 1 demonstrates the methodology proposed in this paper.

The nomenclatures for sets and indices are defined as follows:

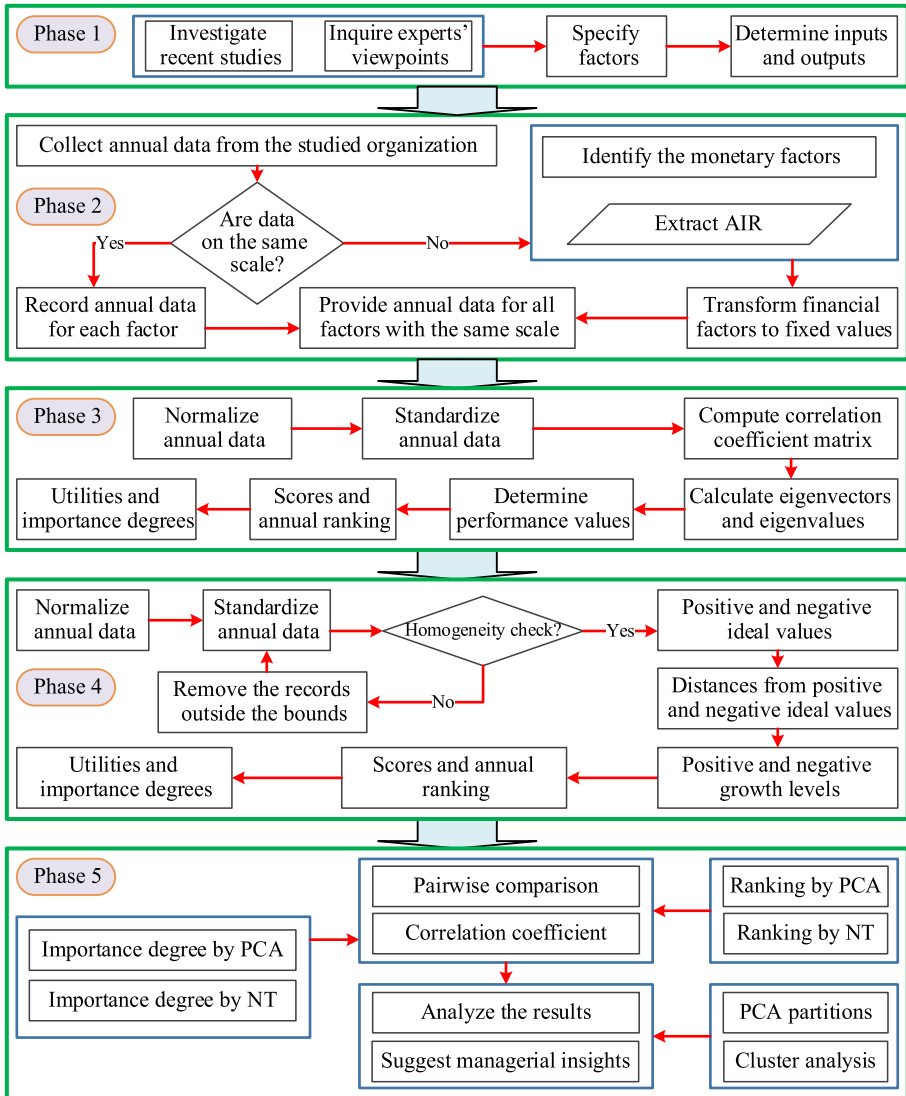


Fig. 1 Proposed methodology

I Set of input factors indexed by $i \in \{1, 2, \dots, a\}$.

O Set of output factors indexed by $o \in \{1, 2, \dots, b\}$.

F Set of financial-based factors indexed by $f \in \{1, 2, \dots, c\}$.

U Set of non-financial-based factors indexed by $u \in \{1, 2, \dots, d\}$.

V Set of all variables including input and output factors indexed by $v \in \{1, 2, \dots, m\}$ where $V = I \cup O = U \cup F$ and $m = a + b = c + d$.

L Set of the reduced dimension of principal components (PCs) indexed by $l \in \{1, 2, \dots, p\}$.

T Set of annual records indexed by $t \in \{1, 2, \dots, n\}$.

3.1 Phase 1: Determining factors

The factors are categorized based on the sustainability dimensions. For the social dimension, the factors affecting the performance of personnel are considered. For the economic dimension, the outcome-based factors are specified, which mainly affect the sales and income. For the environmental dimension, pollution-based factors like different emissions are investigated.

Table 1 provides the extracted factors and their related dimensions in sustainability. Also, their input- and output-oriented nature and unit of measurements are given in this table. Output-oriented factors are positive criteria where the great values are favorable. In contrast, input-oriented factors are negative criteria where the small values are favorable.

3.2 Phase 2: Collecting and scaling data

The steps of this phase are as follows.

Step 1 Collecting annual data.

Raw annual data are collected from the studied organization as matrix $D = [d_{iv}]_{n \times m}$.

Step 2 Investigating commensurability of factors.

If factors are on the same scale, they are recorded without any changes. Otherwise, they are transformed to the same scale values. The unit of measurement for IV, EV, and VA are reported as million IRR. Therefore, the time value of money is considered, which is necessitated by the requirement of importance degree calculation.

Step 3 Transforming non-scale factors.

The AIR and CPI are applied for transforming monetary factors. CPI is a tool to measure the financial level of products and services to examine price changes and calculates as Eq. (1).

$$CPI_t = CPI_{t-1} (1 + AIR_t) \quad \forall t \in T \tag{1}$$

To apply Eq. (1), the CPI is considered 100 for the last period. Concerning a backward computation, Eq. (2) calculates the CPI for all periods.

$$CPI_t = \begin{cases} \frac{CPI_{t+1}}{(1+AIR_{t+1})} & \forall t \in T - \{n\} \\ 100 & \forall t = n \end{cases} \tag{2}$$

Step 4 Determining commensurate matrix.

Equation (3) calculates the fixed value of financial factor f in period t (FV_{tf}) based on its original value in period t (OV_{tf}). Matrix D is transformed to the commensurate matrix $X = [x_{iv}]_{n \times m}$ by Eq. (4).

$$FV_{tf} = \frac{CPI_n}{CPI_t} OV_{tf} \quad \forall t \in T, f \in F \tag{3}$$

$$x_{iv} = \begin{cases} d_{iu} & \forall t \in T, u \in U \\ \frac{CPI_n}{CPI_t} d_{if} & \forall t \in T, f \in F \end{cases} \tag{4}$$

Table 1 Extracted factors

Sustainability dimension	Factor	Type	Code	Unit of measurement	References
Social	Number of personnel	Input	NOP	Person	Ahmad, Jan, et al. (2021), Azadeh et al. (2009), Goyal et al. (2021), Nishitani and Kokubu (2020) and Pashapour et al. (2019)
	Number of accidents	Input	NOA	Numeral	D'Inverno et al. (2020), Forteza et al. (2017), Pashapour et al. (2019) and Toutouchian et al. (2018)
	Number of injuries	Input	NOI	Numeral	Ahi et al. (2016), Forteza et al. (2017), Pashapour et al. (2019) and Toutouchian et al. (2018)
	Number of lost days	Input	NOL	Day	Ahi et al. (2016), Toutouchian et al. (2018) and Tsao et al. (2018)
	Waste products	Input	WP	Ton	Ahi et al. (2016), Al-Tuwajri et al. (2004), Chen et al. (2019), D'Inverno et al. (2020) and Dal Pozzo et al. (2018)
Economic	Imports value	Input	IV	Million IRR	Ling-zhi et al. (2018) and Pal et al. (2014)
	Production rate	Output	PR	Ton	Ahmad, Jan, et al. (2021), Nishitani and Kokubu (2020), Pashapour et al. (2019) and Sabogal-De La Pava et al. (2021)
	Exports value	Output	EV	Million IRR	Ling-zhi et al. (2018), Nishitani (2011), Pal et al. (2014) and Taherdangkoo et al. (2017)
	Return on investment	Output	ROI	Percent	Brandenburg (2016) and D'Ambra et al. (2020)
	Value-added	Output	VA	Million IRR	Ahi et al. (2016), Anser et al. (2021), Ling-zhi et al. (2018), Nishitani (2011) and Nishitani and Kokubu (2020)
Environmental	Emission of CO ₂	Input	ECO	AQI	Arikan et al. (2014), D'Ambra et al. (2020), D'Inverno et al. (2020), Nishitani and Kokubu (2020), Rahimi et al. (2019) and Sabogal-De La Pava et al. (2021)
	Emission of SO ₂	Input	ESO	AQI	Dal Pozzo et al. (2018), Król-Badziak et al. (2021) and Rahimi et al. (2019)
	Emission of NO _x	Input	ENO	AQI	D'Inverno et al. (2020), Dal Pozzo et al. (2018), Nishitani and Kokubu (2020) and Rahimi et al. (2019)
	Environmental noise	Input	EN	AQI	Alves et al. (2015), Hammad et al. (2017) and Zhong and Wu (2015)

3.3 Phase 3: Developing PCA

In PCA, multidimensional factors are summarized into non-correlated components, linear combinations of the main factors. It has the following steps.

Step 1 Normalizing annual dataset.

Since the unit of measurement for various factors is different, they are normalized. Equations (5) and (6) apply for normalizing the input- and output-oriented factors, respectively. Matrix $R = [r_{iv}]_{n \times m}$ is obtained as a normalized matrix.

$$r_{ii} = \frac{\frac{1}{x_{ii}}}{\max_{i \in T} \left\{ \frac{1}{x_{ii}} \right\}} = \frac{\min_{i \in I} x_{ii}}{x_{ii}} \quad \forall t \in T, i \in I \tag{5}$$

$$r_{to} = \frac{x_{to}}{\max_{i \in T} \{x_{io}\}} \quad \forall t \in T, o \in O \tag{6}$$

Step 2 Standardizing annual dataset.

For homogenizing factors, the standard data are considered with mean zero and unit variance as Eq. (7). Matrix $S = [s_{iv}]_{n \times m} = [S_1, S_2, \dots, S_m]$ is a standardized matrix.

$$s_{iv} = \frac{r_{iv} - \text{mean}_{v \in V} \{r_{iv}\}}{\text{StDev}_{v \in V} \{r_{iv}\}} \quad \forall t \in T, v \in V \tag{7}$$

Step 3 Computing correlation coefficient matrix.

The covariance between factors j and k is calculated as Eq. (8). Covariance depends on the unit of measurements, while the correlation coefficient is a unit-free index. The Pearson's correlation coefficient is calculated as Eq. (9). Since S_1, S_2, \dots, S_m are standard, we have Eq. (10). Pearson's correlation coefficient matrix is asymmetrical as Eq. (11). The main diagonal of the matrix is +1 because $\text{Corr}(S_j, S_j) = E[S_j S_j] = E[S_j^2] = \text{Var}(S_j) = 1$.

$$\text{Cov}(S_j, S_k) = E[(S_j - E[S_j])(S_k - E[S_k])] = E[S_j S_k] - E[S_j]E[S_k] \quad \forall j, k \in V \tag{8}$$

$$\text{Corr}(S_j, S_k) = \frac{\text{Cov}(S_j, S_k)}{[\text{Var}(S_j)\text{Var}(S_k)]^{1/2}} \quad \forall j, k \in V \tag{9}$$

$$\text{Cov}(S_j, S_k) = \text{Corr}(S_j, S_k) = E[S_j S_k] \quad \forall j, k \in V \tag{10}$$

$$C = [c_{jk}]_{m \times m} = [\text{Corr}(S_j, S_k)]_{m \times m} = \begin{bmatrix} +1 & E[S_1 S_2] & \dots & E[S_1 S_m] \\ E[S_2 S_1] & +1 & \dots & E[S_2 S_m] \\ \vdots & \vdots & \ddots & \vdots \\ E[S_m S_1] & E[S_m S_2] & \dots & +1 \end{bmatrix} \tag{11}$$

Step 4 Calculating eigenvalues.

Equation (12) calculates eigenvalues from the matrix C . Solving Eq. (12), m eigenvalues are obtained. They are arranged in decreasing order, $\lambda_1, \lambda_2, \dots, \lambda_m$ so that $\sum_{v=1}^m \lambda_v = m$.

$$\det [C - \lambda I] = \begin{vmatrix} 1 - \lambda & E[S_1S_2] & \dots & E[S_1S_m] \\ E[S_2S_1] & 1 - \lambda & \dots & E[S_2S_m] \\ \vdots & \vdots & \ddots & \vdots \\ E[S_mS_1] & E[S_mS_2] & \dots & 1 - \lambda \end{vmatrix} = 0 \tag{12}$$

Step 5 Calculating eigenvectors and PCs.

Equation (13) calculates eigenvectors in which E_v is an $m \times 1$ matrix as Eq. (14)

To avoid multiple solutions, a unit vector is considered by Euclidean norm as Eq. (15). These unit vectors are PCs in this methodology: PC_1, PC_2, \dots, PC_m .

$$CE_v = \lambda_v E_v \quad \forall v \in V \tag{13}$$

$$E = [e_{jk}]_{m \times m} = \begin{bmatrix} e_{11} & e_{12} & \dots & e_{1m} \\ e_{21} & e_{22} & \dots & e_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ e_{m1} & e_{m2} & \dots & e_{mm} \end{bmatrix} = [E_1, E_2, \dots, E_m] \tag{14}$$

$$E_v = E_v^T E_v = e_{1v}^2 + e_{2v}^2 + \dots + e_{mv}^2 = 1 \quad \forall v \in V \tag{15}$$

The first extracted PC accounts for the enormous possible variance in the dataset. The correlation between the two PCs is zero, and they are orthogonal. Generally, Eq. (16) is established among the coefficients of PCs.

$$E_j^T E_k = e_{1j}e_{1k} + e_{2j}e_{2k} + \dots + e_{mj}e_{mk} = 0 \quad \forall j, k \in V \quad \text{and} \quad j \neq k \tag{16}$$

Step 6 Reducing the number of PCs.

The variance explained by each PC is calculated as Eq. (17), where w_v can be considered as a weight criterion of the v th PC.

$$w_v = \frac{\lambda_v}{\sum_{v=1}^m \lambda_v} = \frac{\lambda_v}{m} \forall v \in V \tag{17}$$

Three criteria are applied to reduce the number of PCs as follows.

- Scree test: The eigenvalues are plotted in decreasing order of their values. The point at which the last significant drop or break occurs determines the number of PCs (Ledesma et al., 2015).
- Kaiser rule: Those PCs whose eigenvalues are greater than one are considered.
- Explained variance: The Pareto principle or 20/80 rule considers p first PCs to define 80 percent of variance.

Considering these criteria, the number of PCs is reduced to p where $p < m$.

Step 7 Determining performance values.

Equation (18) calculates the performance value of record t for PC_j . n annual records regarding p PCs constitute matrix $Y = [y_{it}]_{n \times p}$.

$$y_{tl} = \sum_{v \in V} s_{tv} e_{vl} \quad \forall t \in T, l \in L \tag{18}$$

Step 8 Scores and annual ranking.

The total weighted performance of each annual record denotes the score of that record and is calculated as Eq. (19). The higher the z_t^{PCA} , the better the annual record is. The ranking of annual records is made via these values.

$$z_t^{PCA} = \sum_{l \in L} w_l y_{tl} \quad \forall t \in T \tag{19}$$

Step 9 Importance degree of factors.

The influential factors are those that lead to the most score for the annual records and the most standard value. Therefore, the aggregation of record scores and standard values is applied for comparative measurement of factors as Eq. (20).

$$A_v^{PCA} = \sum_{t \in T} z_t^{PCA} s_{tv} \quad \forall v \in V \tag{20}$$

A_v^{PCA} is linearly normalized between -1 and $+1$ as Eq. (21), named the utility of each factor. The higher magnitude of the U_v^{PCA} , the more influential the factor is.

$$U_v^{PCA} = \frac{\left(A_v^{PCA} - \max_{v \in V} \{ A_v^{PCA} \} \right) + \left(A_v^{PCA} - \min_{v \in V} \{ A_v^{PCA} \} \right)}{\max_{v \in V} \{ A_v^{PCA} \} - \min_{v \in V} \{ A_v^{PCA} \}} \quad \forall v \in V \tag{21}$$

3.4 Phase 4: Developing NT analysis

NT method is able to identify homogeneous states. In this method, factors are applied the same as the PCA. It has the following steps.

Step 1 Normalizing annual dataset.

Similar to step 1 of PCA analysis.

Step 2 Standardizing annual dataset.

Similar to step 2 of PCA analysis.

Step 3 Computing distance matrix.

The distance of every two annual records g and h for factor v is $|s_{gv} - s_{hv}|$. Therefore, Euclidean distance of annual record g from h can be computed as Eq. (22). This is done to homogenize the annual records. Computing pairwise distances, the symmetrical distance matrix is obtained as $Q = [q_{gh}]_{n \times n}$.

$$q_{gh} = \sqrt{\sum_{v \in V} (s_{gv} - s_{hv})^2} \quad \forall g, h \in T \quad \text{and} \quad g \neq h \tag{22}$$

Step 4 Calculating upper and lower bounds.

To calculate lower and upper bounds (LB, UB), vector $q = [q_t]_{n \times 1}$ is extracted from matrix Q , where q_t is the minimum amount of the t th row in matrix Q as Eq. (23).

$$q_t = \min_{h \in T} \{q_{th}\} = \min_{g \in T} \{q_{gt}\} \quad \forall t \in T \quad \text{and} \quad g, h \neq t \tag{23}$$

Equation (24) calculates the bounds. If all q_t are within the bounds, homogeneity is attained, and the next step is performed. Otherwise, cluster analysis should be made until all annual records become homogeneous. To do so, the records outside the bounds are removed. Again, the standard matrix is formed with these records removed. It repeats until the homogeneity is achieved.

$$\begin{cases} \text{LB} = \text{mean}_{t \in T} \{q_t\} - 2 \text{StDev}_{t \in T} \{q_t\} \\ \text{UB} = \text{mean}_{t \in T} \{q_t\} + 2 \text{StDev}_{t \in T} \{q_t\} \end{cases} \tag{24}$$

Step 5 Positive and negative ideal values.

The positive and negative ideal values for each factor are as Eqs. (25) and (26), respectively.

$$s_v^{\text{PIV}} = \max_{t \in T} \{s_{tv}\} \quad \forall v \in V \tag{25}$$

$$s_v^{\text{NIV}} = \min_{t \in T} \{s_{tv}\} \quad \forall v \in V \tag{26}$$

Step 6 Distance from positive and negative ideal values.

The positive and negative distances of factor v for annual record t in the standard matrix are $s_v^{\text{PIV}} - s_{tv}$ and $s_{tv} - s_v^{\text{NIV}}$. Therefore, Euclidean distance of annual record t for all factors from positive and negative ideal values is calculated as Eqs. (27) and (28), respectively.

$$c_t^{\text{PIV}} = \sqrt{\sum_{v \in V} (s_v^{\text{PIV}} - s_{tv})^2} \quad \forall t \in T \tag{27}$$

$$c_t^{\text{NIV}} = \sqrt{\sum_{v \in V} (s_{tv} - s_v^{\text{NIV}})^2} \quad \forall t \in T \tag{28}$$

Step 7 Positive and negative growth levels.

Each annual record's positive and negative growth levels are calculated as Eqs. (29) and (30), respectively.

$$f_t^{\text{PIV}} = \frac{c_t^{\text{PIV}}}{c^{\text{PIV}*}} \quad \forall t \in T \tag{29}$$

where $c^{\text{PIV}*} = \text{mean}_{t \in T} \{c_t^{\text{PIV}}\} + 2 \text{StDev}_{t \in T} \{c_t^{\text{PIV}}\}$

$$f_t^{\text{NIV}} = \frac{c_t^{\text{NIV}}}{c^{\text{NIV}*}} \quad \forall t \in T \tag{30}$$

where $c^{\text{NIV}*} = \text{mean}_{t \in T} \{c_t^{\text{NIV}}\} + 2 \text{StDev}_{t \in T} \{c_t^{\text{NIV}}\}$

Since $c_t^{\text{PIV}}, c_t^{\text{NIV}} \geq 0$, $\text{mean}_{t \in T} \{c_t^{\text{PIV}}\}, \text{mean}_{t \in T} \{c_t^{\text{NIV}}\} \geq 0$. Also, $\text{StDev}_{t \in T} \{c_t^{\text{PIV}}\}, \text{StDev}_{t \in T} \{c_t^{\text{NIV}}\} \geq 0$. Hence, $c^{\text{PIV}*}, c^{\text{NIV}*} \geq 0$. Therefore, $f_t^{\text{PIV}}, f_t^{\text{NIV}} \geq 0$. f_t^{PIV} and f_t^{NIV} indicate the distance of

annual record t from the best and the worst possible one, respectively. Therefore, the lower f_t^{PIV} and the higher f_t^{NIV} are desirable.

Step 8 Scores and annual ranking.

The scores of annual records are calculated as Eq. (31). The higher the z_t^{NT} , the better the annual record is. The ranking of annual records can be obtained via these values.

$$z_t^{NT} = \frac{f_t^{NIV}}{f_t^{PIV} + f_t^{NIV}}, \quad 0 \leq z_t^{NT} \leq 1 \quad \forall t \in T \tag{31}$$

Step 9 Importance degree of factors.

The aggregated comparative measurement and the utility of each factor in NT are obtained as Eqs. (32) and (33), respectively. The higher the magnitude of the U_v^{NT} , the more influential the factor is.

$$A_v^{NT} = \sum_{t \in T} z_t^{NT} s_{tv}, \quad \forall v \in V \tag{32}$$

$$U_v^{NT} = \frac{\left(A_v^{NT} - \max_{v \in V} \{ A_v^{NT} \} \right) + \left(A_v^{NT} - \min_{v \in V} \{ A_v^{NT} \} \right)}{\max_{v \in V} \{ A_v^{NT} \} - \min_{v \in V} \{ A_v^{NT} \}} \quad \forall v \in V \tag{33}$$

3.5 Phase 5: Analyzing, validating, and verifying the results

The results are discussed in four categories:

- First, the impact of dimension reduction in the proposed PCA is examined.
- Second, extracted factors are clustered in different partitions.
- Third, the annual SP is studied.
- And finally, the importance degree of factors is evaluated.

To verify and validate the results of PCA, the annual ranking and importance degree of factors of the PCA and NT are compared. For this purpose, first, the existence of differences in the results is examined by a pairwise comparison test. If there is no significant difference, then the compatibility of the results is carried out. If it is high, it is concluded that the results of PCA are valid.

3.5.1 Pairwise comparison test

The statistical hypothesis is as follows:

$$\begin{cases} H_0 : \mu_{PCA} = \mu_{NT} \\ H_1 : \text{Otherwise} \end{cases} \tag{34}$$

where μ_{PCA} and μ_{NT} are the mean of annual ranking by PCA and NT, respectively. Since the scales obtained from PCA and NT methods for annual records are different, they cannot be used directly for pairwise comparison. Therefore, the annual ranking of records is

evaluated. The Wilcoxon signed-rank nonparametric statistical test is applied with the following steps.

Step 1 Calculate $|R_t^{PCA} - R_t^{NT}|$ and $\text{sgn}(R_t^{PCA} - R_t^{NT})$ for $t = 1, \dots, n$, where R_t^{PCA} and R_t^{NT} are ranking of PCA and NT for annual record t and sgn indicates sign function.

Step 2 Omit $|R_t^{PCA} - R_t^{NT}| = 0$ annual records. Let n_r the number of the reduced sample size.

Step 3 Order nonzero $|R_t^{PCA} - R_t^{NT}|$ from smallest to largest.

Step 4 Rank so that the smallest $|R_t^{PCA} - R_t^{NT}|$ is 1. For equal values, insert average rank. Let R_t^W denote the ranking.

Step 5 Calculate Wilcoxon test statistic as follows which is the sum of signed ranks:

$$W = \sum_{t=1}^{n_r} [\text{sgn}(R_t^{PCA} - R_t^{NT}) \times R_t^W] \tag{35}$$

Step 6 Compare W with the critical value of the Wilcoxon signed-rank test (Wilcoxon et al., 1970). For a two-tailed test, H_0 rejects if $|W| > W_{\alpha, n_r}$.

3.5.2 Compatibility test

To evaluate the compatibility of PCA and NT results, Spearman’s rank correlation coefficient is examined as Eq. (36).

$$\rho = 1 - \frac{6 \sum_{t \in T} (R_t^{PCA} - R_t^{NT})^2}{n(n^2 - 1)} \tag{36}$$

4 Case study

The case study is an upstream oil and gas company in Iran. The main activities of such companies are contracting in engineering, procurement, construction, and installation in various projects for large offshore structures.

4.1 Data description

The data are reported based on Solar Hijri (SH) calendar in the organization. They cover 15 years from the first of 1382 SH (equals to March 21, 2003) to the end of 1396 SH (equals to March 20, 2018). Let AR_t denote annual record t . It should be noted that the total number of incidents is categorized into the NOI and the NOA. NOA implies intense incidents, and personnel needs more than four days’ absence and even for several months.

To commensurate IV, EV, and VA, the time value of money was considered. The AIR of Iran is applied according to the report of the central bank of Iran.¹ Table 2 gives the AIR and calculated CPI for 15 periods. Thereafter, normalizing and standardizing are performed, but the corresponding values are not reported because of brevity.

¹ https://www.cbi.ir/Inflation/Inflation_en.aspx.

Table 2 Annual AIR and CPI for considered periods

Annual record	AIR (%)	CPI
AR ₁	15.6	11.51
AR ₂	15.2	13.26
AR ₃	10.4	14.64
AR ₄	11.9	16.39
AR ₅	18.4	19.40
AR ₆	25.4	24.33
AR ₇	10.8	26.96
AR ₈	12.4	30.30
AR ₉	21.5	36.81
AR ₁₀	30.5	48.04
AR ₁₁	34.7	64.71
AR ₁₂	15.6	74.81
AR ₁₃	11.9	83.71
AR ₁₄	9.0	91.24
AR ₁₅	9.6	100

4.2 Dimension reduction

Table 3 gives the correlation coefficient matrix. The eigenvalues of the correlation matrix in relation to PCA are as shown in Table 4. The higher the eigenvalue, the higher the explained variance of each PC is. Figure 2 illustrates the scree plot of eigenvalues. The scree plot is able to display how much each PC captures from the dataset. According to Table 4 and Fig. 2, the first four PCs comprise about 80 percent of data variability. The magnitudes of eigenvalues in these PCs are more than one. Therefore, the other 10 PCs are ignored according to the three criteria described in the proposed methodology. Table 5 introduces the coefficients of PCs.

Figure 3 demonstrates the 2D loading plot of factors with PC1 and PC2. The value of a factor on each PC illustrates how much weight it has on each PC. EV and VA have the highest positive values and NOP, and IV have the highest negative values on PC1. The other point from the 2D loading plot is the angles between factors that specify how a factor is correlated. The factors are categorized into two groups according to their dispersion, schematically considering the first two PCs. Figure 4 displays the 3D loading plot of factors with PC1, PC2, and PC3. Like the 2D loading plot, schematically, considering the first three PCs, the factors are categorized into four groups according to their dispersion.

4.3 Performance values, annual ranking, and importance degree

Table 6 provides the performance values, scores, and annual ranking by PCA. Performance values of each annual record are obtained by multiplying standardized values of the dataset by eigenvectors of each PC. Consequently, the scores of PCA for each annual record are attained by multiplying weights of reduced PCs by performance values as Eq. (37). According to these scores, the annual records are ranked. Eventually, the utility and importance degree of factors are gained, as shown in Table 7.

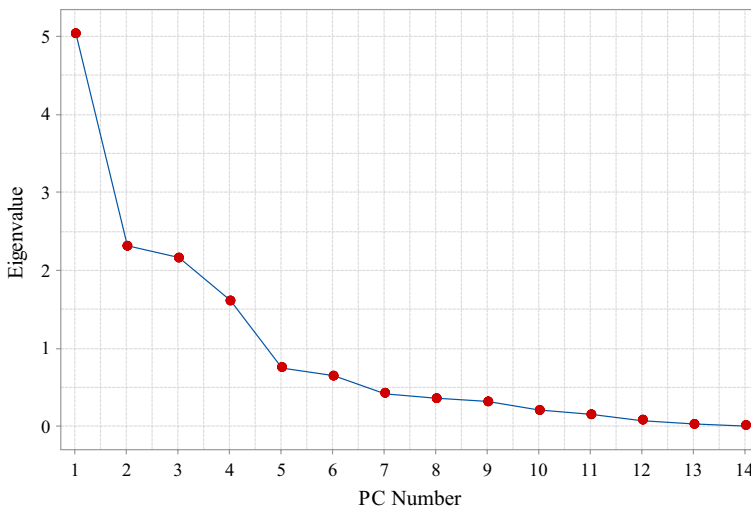
$$z_t^{\text{PCA}} = 0.3595y_{t1} + 0.1648y_{t2} + 0.1538y_{t3} + 0.1145y_{t4} \quad \forall t \in \{1, 2, \dots, 15\} \quad (37)$$

Table 3 Correlation coefficient matrix

	NOP	NOA	NOI	NOL	WP	IV	ECO	ESO	ENO	EN	PR	EV	ROI	VA
NOP	1													
NOA	-0.187	1												
NOI	0.362	-0.219	1											
NOL	-0.160	0.505	-0.152	1										
WP	-0.127	0.165	-0.449	0.051	1									
IV	0.356	-0.265	0.828	-0.064	-0.538	1								
ECO	0.490	-0.394	-0.065	-0.064	-0.294	0.066	1							
ESO	0.331	-0.095	-0.384	0.130	0.167	-0.263	0.447	1						
ENO	-0.645	0.343	-0.465	0.396	0.399	-0.477	-0.292	-0.123	1					
EN	0.581	0.167	0.119	-0.008	0.124	0.116	0.220	0.266	-0.289	1				
PR	-0.033	0.053	0.187	0.331	-0.439	0.008	0.137	0.290	-0.130	-0.273	1			
EV	-0.582	0.357	-0.528	0.169	0.666	-0.535	-0.423	-0.266	0.695	-0.292	-0.482	1		
ROI	-0.156	-0.394	-0.146	-0.514	0.417	-0.257	-0.045	-0.086	0.156	-0.273	-0.380	0.456	1	
VA	-0.585	0.213	-0.481	0.130	0.663	-0.563	-0.333	-0.333	0.648	-0.243	-0.470	0.940	0.414	1

Table 4 Eigenvalue and weight of PCs

PC	Eigenvalue	Weight	Cumulative weight
PC ₁	5.034	0.3595	0.3595
PC ₂	2.308	0.1648	0.5244
PC ₃	2.154	0.1538	0.6782
PC ₄	1.604	0.1145	0.7927
PC ₅	0.747	0.0533	0.8461
PC ₆	0.639	0.0456	0.8917
PC ₇	0.416	0.0297	0.9214
PC ₈	0.351	0.0250	0.9464
PC ₉	0.310	0.0222	0.9686
PC ₁₀	0.197	0.0141	0.9827
PC ₁₁	0.146	0.0104	0.9931
PC ₁₂	0.072	0.0051	0.9982
PC ₁₃	0.023	0.0016	0.9999
PC ₁₄	0.002	0.0001	1.0000

**Fig. 2** Scree plot of eigenvalues

4.4 Results of NT

Table 8 computes the distance matrix. The distance vector is the minimum distance of annual records given in the last row of the table. Figure 5 illustrates the line chart of the distance vector. Since all of the values of this vector are within the bounds, the homogeneity is confirmed.

Table 5 Coefficients of PCs

Factor	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11	PC12	PC13	PC14
NOP	-0.308	0.109	-0.297	-0.280	-0.029	0.177	0.403	0.145	-0.173	0.570	0.140	-0.272	-0.221	-0.116
NOA	0.158	-0.428	0.043	-0.370	0.236	-0.232	0.401	-0.117	-0.455	-0.149	-0.248	0.152	0.141	-0.188
NOI	-0.288	0.111	0.385	-0.188	0.006	0.407	0.200	-0.286	0.153	-0.052	-0.232	-0.263	0.492	0.205
NOL	0.081	-0.540	0.008	-0.092	-0.383	0.411	0.035	0.241	0.029	-0.103	0.503	0.124	0.195	0.016
WP	0.306	0.126	-0.275	-0.229	0.208	0.481	-0.032	0.132	0.230	0.148	-0.354	0.514	-0.010	0.072
IV	-0.308	0.091	0.335	-0.193	-0.279	0.273	-0.274	0.117	-0.337	-0.237	-0.249	0.136	-0.444	-0.243
ECO	-0.206	0.047	-0.378	0.227	-0.635	-0.153	0.336	-0.100	0.024	-0.138	-0.332	0.244	0.095	0.113
ESO	-0.087	-0.157	-0.557	0.191	0.171	0.285	-0.335	0.020	-0.296	-0.238	-0.203	-0.408	0.194	-0.091
ENO	0.350	-0.177	0.054	0.068	-0.303	0.119	-0.247	-0.638	-0.160	0.466	-0.085	-0.061	-0.084	-0.063
EN	-0.144	-0.011	-0.330	-0.566	-0.006	-0.105	-0.139	-0.451	0.309	-0.333	0.228	-0.031	-0.229	0.067
PR	-0.166	-0.396	0.061	0.430	0.285	0.251	0.330	-0.216	0.269	-0.107	-0.076	-0.003	-0.487	-0.017
EV	0.425	0.062	0.034	-0.094	-0.122	0.069	0.172	0.143	-0.193	-0.194	-0.086	-0.348	-0.328	0.652
ROI	0.181	0.497	-0.056	0.198	0.083	0.276	0.283	-0.298	-0.336	-0.279	0.416	0.173	0.033	-0.176
VA	0.409	0.099	0.030	-0.075	-0.219	0.027	0.191	0.130	0.376	-0.173	-0.167	-0.394	-0.031	-0.600

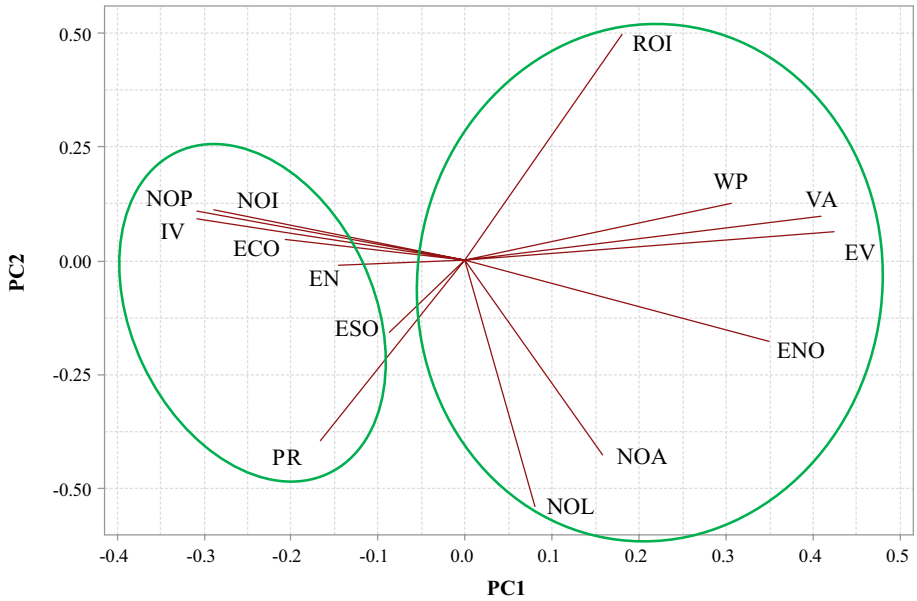


Fig. 3 2D loading plot

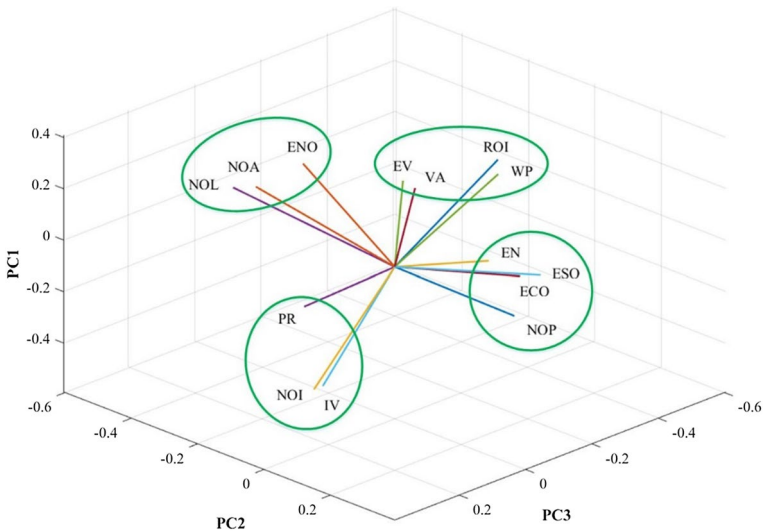


Fig. 4 3D loading plot

Positive and negative ideal values are extracted, and subsequently, the Euclidean distance from these values and positive and negative growth levels are calculated for each annual record. Table 9 reports the results of the annual ranking from the NT. Eventually, Table 10 gives the utility and importance degree of factors for NT.

Table 6 Performance values, scores, and annual ranking by PCA

Annual record	Performance value				z_t^{PCA}	R_t^{PCA}
	y_{t1}	y_{t2}	y_{t3}	y_{t4}		
AR ₁	3.699	1.185	- 1.287	- 2.367	14.787	2
AR ₂	1.793	- 1.336	0.093	0.208	6.477	6
AR ₃	1.755	0.194	0.643	- 0.716	9.518	5
AR ₄	2.216	0.610	0.492	0.677	14.703	3
AR ₅	1.455	0.840	1.016	1.209	13.387	4
AR ₆	3.090	0.435	1.229	1.162	21.067	1
AR ₇	0.800	- 3.377	- 0.624	0.447	- 4.393	9
AR ₈	- 0.453	- 0.259	- 0.788	1.690	- 1.866	7
AR ₉	- 0.957	0.168	- 0.458	- 1.767	- 8.247	11
AR ₁₀	- 1.994	0.227	- 1.800	1.808	- 10.492	12
AR ₁₁	- 1.128	- 0.656	- 1.579	- 1.251	- 12.597	13
AR ₁₂	- 2.829	1.753	2.600	- 0.845	- 5.951	10
AR ₁₃	- 2.904	- 2.818	1.631	- 0.987	- 19.192	15
AR ₁₄	- 2.054	1.219	1.447	0.546	- 3.536	8
AR ₁₅	- 2.488	1.815	- 2.613	0.186	- 13.666	14

Table 7 Utilities and importance degree of factors by PCA

Factor	A_v^{PCA}	U_v^{PCA}	Importance degree
NOP	- 9.338	- 1.000	14
NOA	0.974	0.013	6
NOI	- 5.415	- 0.615	10
NOL	- 1.022	- 0.183	7
WP	6.569	0.563	5
IV	- 6.262	- 0.698	12
ECO	- 6.138	- 0.686	11
ESO	- 5.138	- 0.587	9
ENO	8.356	0.738	3
EN	- 6.712	- 0.742	13
PR	- 4.935	- 0.568	8
EV	11.021	1.000	1
ROI	7.489	0.653	4
VA	10.852	0.983	2

Table 8 Distance matrix

	AR1	AR2	AR3	AR4	AR5	AR6	AR7	AR8	AR9	AR10	AR11	AR12	AR13	AR14	AR15
AR1	–														
AR2	4.98	–													
AR3	4.44	2.65	–												
AR4	4.51	3.45	2.79	–											
AR5	5.38	3.64	2.88	3.01	–										
AR6	4.98	3.70	3.33	2.13	2.93	–									
AR7	6.45	3.64	4.97	4.93	5.35	5.37	–								
AR8	6.33	3.85	4.63	4.57	3.57	4.93	4.39	–							
AR9	5.64	4.49	3.99	5.01	4.64	5.76	5.23	4.40	–						
AR10	7.35	5.23	5.56	5.39	5.05	6.27	5.44	3.40	4.46	–					
AR11	6.21	4.66	4.34	5.41	5.16	6.10	4.90	3.89	3.19	4.62	–				
AR12	7.98	6.58	5.93	6.35	5.73	6.92	7.27	5.67	4.89	5.88	5.91	–			
AR13	8.59	5.98	6.14	6.71	6.58	7.42	5.52	5.90	5.19	5.75	5.27	5.55	–		
AR14	7.31	5.24	4.65	4.88	4.24	5.62	6.21	3.87	3.97	4.22	4.62	3.19	4.43	–	
AR15	7.18	6.41	6.00	5.88	6.08	7.18	6.81	4.81	4.71	3.20	4.30	5.77	6.60	4.68	–
q_i	4.44	2.65	2.65	2.13	2.88	2.13	3.64	3.40	3.19	3.20	3.19	3.19	4.43	3.19	3.20

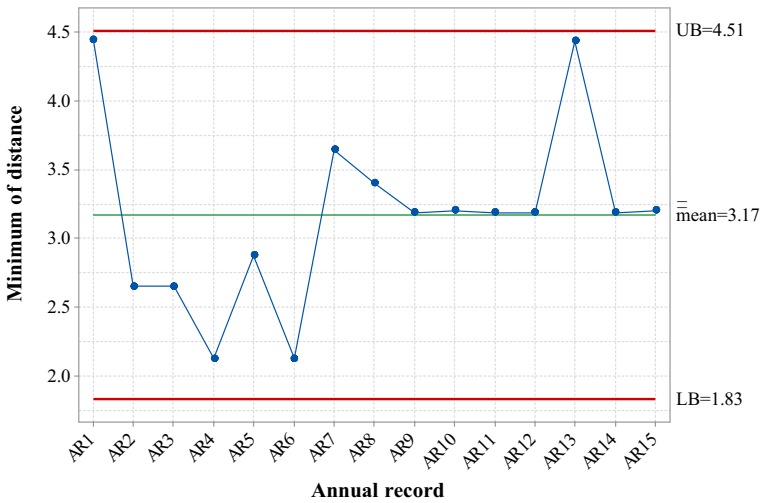


Fig. 5 Line chart of distance vector

Table 9 Results of annual ranking from the NT

Annual record	c_t^{PIV}	c_t^{NIV}	f_t^{PIV}	f_t^{NIV}	z_t^{NT}	R_t^{NT}
AR ₁	7.736	6.695	0.7991	0.8124	0.5041	4
AR ₂	8.494	6.905	0.8774	0.8378	0.4885	6
AR ₃	8.859	7.001	0.9151	0.8495	0.4814	7
AR ₄	7.633	6.882	0.7884	0.8351	0.5144	3
AR ₅	7.401	7.507	0.7645	0.9109	0.5437	2
AR ₆	7.539	8.595	0.7787	1.0430	0.5725	1
AR ₇	8.721	6.359	0.9009	0.7717	0.4614	8
AR ₈	8.405	5.856	0.8681	0.7106	0.4501	9
AR ₉	8.559	5.616	0.8842	0.6815	0.4353	12
AR ₁₀	8.961	5.686	0.9257	0.6899	0.4270	13
AR ₁₁	8.773	5.985	0.9062	0.7263	0.4449	10
AR ₁₂	8.920	5.975	0.9214	0.7251	0.4404	11
AR ₁₃	9.448	5.128	0.9760	0.6223	0.3894	15
AR ₁₄	8.204	6.952	0.8475	0.8436	0.4989	5
AR ₁₅	9.028	5.654	0.9325	0.6861	0.4239	14

5 Verification and validation

The verification and validation of results are performed in different aspects:

1. Categorizing factors in different groups,
2. Investigating the annual SP,
3. Assessing the importance degree of factors,
4. Comparing the results of developed NT with traditional NT.

Table 10 Utilities and importance degree of factors by NT

Factor	A_v^{NT}	U_v^{NT}	Importance degree
NOP	-0.519	-1.000	14
NOA	-0.046	-0.091	6
NOI	-0.285	-0.551	10
NOL	-0.133	-0.258	7
WP	0.276	0.529	5
IV	-0.322	-0.621	11
ECO	-0.348	-0.672	12
ESO	-0.280	-0.541	9
ENO	0.397	0.761	3
EN	-0.390	-0.753	13
PR	-0.199	-0.386	8
EV	0.502	0.963	2
ROI	0.365	0.699	4
VA	0.521	1.000	1

Table 11 Cluster analysis results

Two partitions		Four partitions	
Cluster no	Factors	Cluster no	Factors
Cluster 1	NOP, EN, ECO, ESO, NOI, IV, PR	Cluster 1	NOP, EN, ECO, ESO
Cluster 2	NOA, NOL, ENO, WP, EV, VA, ROI	Cluster 2	NOA, NOL, ENO
		Cluster 3	NOI, IV, PR
		Cluster 4	WP, EV, VA, ROI

5.1 Cluster analysis

Based on PCA results, the factors were categorized into two and four partitions. In support of this, cluster analysis is applied. Table 11 shows the results of cluster analysis for both two and four partitions. The nearest neighbor method and the Euclidean distance have been utilized in cluster analysis. Figure 6 shows the dendrogram plot of this analysis. In this figure, line (a) and line (b) identify two and four partitions in clustering, respectively. Two partitions of cluster analysis are the same as the two extracted groups of Fig. 3. In addition, four partitions of cluster analysis are the same as the four determined groups of Fig. 4.

5.2 Comparison of results

It was shown that the first four PCs explain about 80 percent of data variability in PCA. Here, the effect of considering more PCs on the results is investigated. To do so, 90%, 99%,

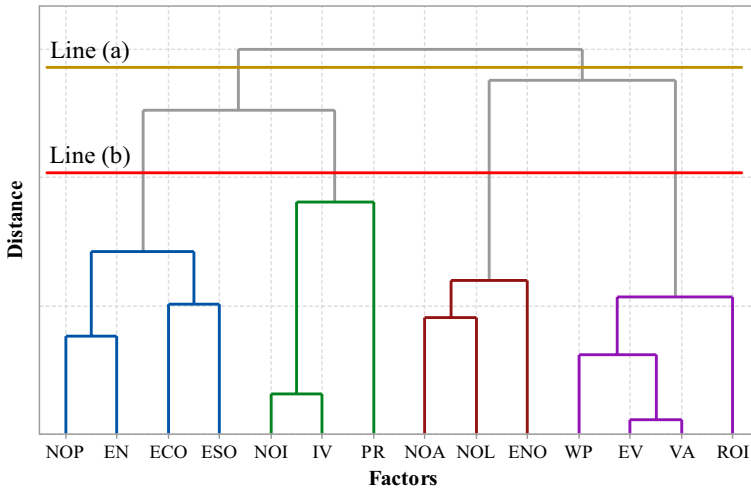


Fig. 6 Dendrogram plot of cluster analysis

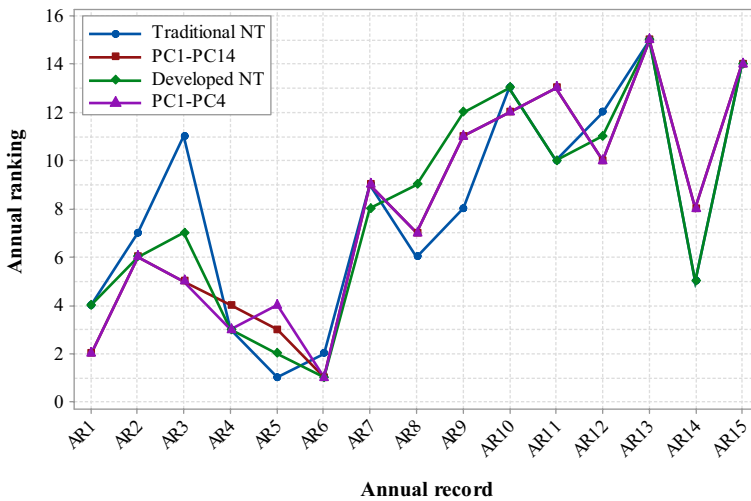


Fig. 7 Comparison of annual ranking by different analyses

and 100% of data variability, which forms PC1–PC6, PC1–PC10, and PC1–PC14, respectively, are examined. Therefore, the annual ranking and the importance degree of factors by PCA are studied with different variabilities. On the other hand, the NT was performed in two modes: developed and traditional. Tables 12 and 13 give the annual ranking and importance degree of factors with different analyses. These findings express the results by PCA for 90%, 99%, and 100% variabilities are the same. Therefore, one of them is taken for pairwise comparison and compatibility tests. Figures 7 and 8 schematically compare the annual ranking and importance degree of factors by different analyses.

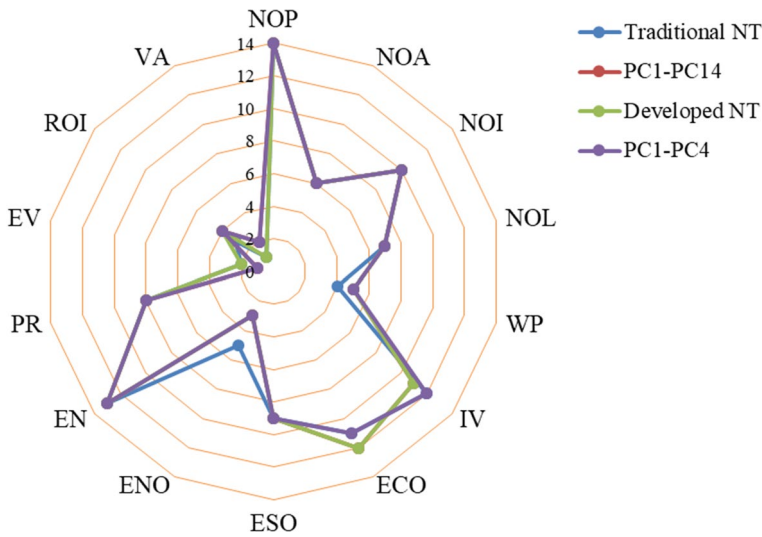


Fig. 8 Comparison of importance degree of factors by different analyses

To statistically compare the results of different analyses, differences between the results are examined. In this paper, the level of significance is considered 95%. Table 14 contains the nonparametric stepwise Wilcoxon signed-rank test for annual ranking results by PCA with PC1-PC4 and developed NT. The Wilcoxon statistic is less than the critical value. Therefore, statistically, the null hypothesis is accepted.

Table 15 summarizes the results of all pairwise comparison tests by Wilcoxon signed-rank and Spearman's rank correlation. It is concluded that there is a direct relationship between the results of all pairwise analyses.

6 Discussion and managerial insights

Two quantitative methodologies are presented, for which the description of steps of data preparation, computation, and inference are provided. Comparable results on the importance degree obtained from the two methodologies increase the confidence in the results of the importance and rankings. Hypothesis testing procedures conduct the comparison of the results of the two methods. After dimensionality reduction to only the important ones, a company's performance is studied based on the clustered factors and ranks of the annual records. Concerning three introduced criteria to reduce the dimension of PCA, instead of 14 factors, the first four PCs can be used as new variables. The results of the Wilcoxon signed-rank test showed no statistically significant difference between the results of PCA with the first 4 PCs and those with more PCs. Spearman's results also showed a high correlation between the results. Therefore, the introduced criteria can reduce the computational complexity and ensure that dimensional reduction is desirable.

2D loading plot and 3D loading plot of PCA schematically categorized the factors in two and four different groups. Cluster analysis for two and four partitions showed precisely the same results. The first partition includes NOP, ECO, ESO, and EN, and the second one comprises NOA, NOL, and ENO, which are related to social

and environmental dimensions of sustainability. Reducing workplace pollution can affect personnel performance. Workplace-related factors not only endanger the health of operators but also avoid them being active for several months. Therefore, the SP of the organization can be improved via environmental aspects and providing facilities for personnel. The third partition contains NOI, IV, and PR factors related to the social and economic dimensions of sustainability. NOI causes short-term interruption for personnel and production fluctuations. Thus, a short-term plan for personnel rest seems necessary for appropriate production. In addition, IV and PR are both related to production. Therefore, balancing imports and adapting with PR reduces costs such as inventory and improves the SP of the organization. Finally, the fourth partition encompasses WP, EV, ROI, and VA factors primarily related to the economic dimension of sustainability. All these factors are in the direction of the organization's outcomes and can guarantee the competitiveness and resiliency of the organization in the future. Increasing exports and reducing WP will improve VA and ROI. Therefore, improving these factors should be one of the priorities of the organization.

The annual ranking specified poor SP of organization in recent years. The management should make more efforts to improve it and also learn from lessons from partitioning the factors. EV and VA were identified as the most influential factors in the SP. Both of these factors are related to the economic dimension. Data on these factors indicated interconnected causality. The following remarks can help management to improve SP.

- Proper marketing expedites the organization to improve EV.
- Having a plan to reduce WP and using modern technology cause to increase PR.
- Focusing on EV and IV along with NOP lead to balancing work-related NOI and NOA.
- Paying attention to PR along with WP accelerates the VA and ROI.

7 Conclusion

This paper dealt with sustainability performance (SP) assessment via a multivariate quantitative approach. To do so, a five-phase methodology was proposed. Related factors were identified by considering sustainability dimensions. For monetary factors, a unique procedure was proposed to transform them. Principal component analysis (PCA) was developed so that it can simultaneously measure the annual SP and the importance degree of the factors. The utility function was defined to recognize the influential factors. Then, the traditional numerical taxonomy (NT) was expanded. The developed NT considers the distances of both positive and negative ideal values by integrating them for ranking. Moreover, it can rank annual records besides identifying the importance degree of factors by proposing a suitable utility function.

An upstream oil and gas company was investigated as a case study. The results derived from the two developed NT and PCA analyses were statistically compared. These analyses were made from four different aspects: (1) the impact of dimension reduction, (2) clustering of factors in different partitions, (3) the assessment of annual SP, and (4) the evaluation of the importance degree of factors.

Cluster analysis showed that the environmental factors affect only social factors related to the personnel. In addition to the environmental dimension, social-based factors influence the economic dimension, especially in production and imports. Also, economic factors have a high interaction in the financial output of the organization and create value-added. The correlation

of developed NT with PCA is higher than that of traditional NT and PCA. Regarding the case of the oil and gas company, it can be concluded that the company SP in recent years has deteriorated. Exports value and value-added were identified as the most influential factors.

The paper's contributions are twofold: First, the methodology can be easily scaled to larger data sets which may consider a more significant number of performance factors than those considered in this paper. Second, it can be extended to different industries. The approach of this paper can be applied to other industry organizations to improve their SP. The impact of other factors can be investigated as other topics for future researches. Furthermore, the approach of this paper is extended simultaneous recognition of the importance of variables and alternatives. The knowledge in this paper will be helpful for further academic investigations into this topic and immediate industrial practice.

Appendix

See Tables 12, 13, 14 and 15.

Table 12 Annual ranking by different analyses

Annual record	PCA				NT	
	$PC_1 - PC_4$	$PC_1 - PC_6$	$PC_1 - PC_{10}$	$PC_1 - PC_{14}$	Developed NT	Traditional NT
AR ₁	2	2	2	2	4	4
AR ₂	6	6	6	6	6	7
AR ₃	5	5	5	5	7	11
AR ₄	3	4	4	4	3	3
AR ₅	4	3	3	3	2	1
AR ₆	1	1	1	1	1	2
AR ₇	9	9	9	9	8	9
AR ₈	7	7	7	7	9	6
AR ₉	11	11	11	11	12	8
AR ₁₀	12	12	12	12	13	13
AR ₁₁	13	13	13	13	10	10
AR ₁₂	10	10	10	10	11	12
AR ₁₃	15	15	15	15	15	15
AR ₁₄	8	8	8	8	5	5
AR ₁₅	14	14	14	14	14	14

Table 13 Importance degree of factors by different analyses

Factor	PCA				NT	
	PC ₁ – PC ₄	PC ₁ – PC ₆	PC ₁ – PC ₁₀	PC ₁ – PC ₁₄	Developed NT	Traditional NT
NOP	14	14	14	14	14	14
NOA	6	6	6	6	6	6
NOI	10	10	10	10	10	10
NOL	7	7	7	7	7	7
WP	5	5	5	5	5	4
IV	12	11	11	11	11	12
ECO	11	12	12	12	12	11
ESO	9	9	9	9	9	9
ENO	3	3	3	3	3	5
EN	13	13	13	13	13	13
PR	8	8	8	8	8	8
EV	1	1	1	1	2	2
ROI	4	4	4	4	4	3
VA	2	2	2	2	1	1

Table 14 Wilcoxon signed-rank test between results of the annual ranking of PCA and NT

Record no	R_t^{PCA}	R_t^{NT}	$ R_t^{PCA} - R_t^{NT} $	$\text{sgn}(R_t^{PCA} - R_t^{NT})$	R_t^W	$\text{sgn}(R_t^{PCA} - R_t^{NT}) \times R_t^W$
AR1	2	4	2	-1	6.5	-6.5
AR2	6	6	0	0		
AR3	5	7	2	-1	6.5	-6.5
AR4	3	3	0	0		
AR5	4	2	2	1	6.5	6.5
AR6	1	1	0	0		
AR7	9	8	1	1	2.5	2.5
AR8	7	9	2	-1	6.5	-6.5
AR9	11	12	1	-1	2.5	-2.5
AR10	12	13	1	-1	2.5	-2.5
AR11	13	10	3	1	9.5	9.5
AR12	10	11	1	-1	2.5	-2.5
AR13	15	15	0	0		
AR14	8	5	3	1	9.5	9.5
AR15	14	14	0	0		
Wilcoxon test statistic:						1
The critical value of Wilcoxon test in 95% level of significance:						8
p-value:						0.959

Table 15 Statistical comparison

Comparison	First analysis	Second analysis	Wilcoxon p -value	Spearman's rank correlation
Annual ranking	$PC_1 - PC_4$	$PC_1 - PC_{14}$	1.000	0.966
	$PC_1 - PC_4$	Developed NT	0.959	0.932
	$PC_1 - PC_4$	Traditional NT	0.753	0.859
	$PC_1 - PC_{14}$	Developed NT	0.892	0.936
	$PC_1 - PC_{14}$	Traditional NT	0.751	0.857
	Developed NT	Traditional NT	0.829	0.918
Importance degree	$PC_1 - PC_4$	$PC_1 - PC_{14}$	1.000	0.996
	$PC_1 - PC_4$	Developed NT	1.000	0.991
	$PC_1 - PC_4$	Traditional NT	1.000	0.982
	$PC_1 - PC_{14}$	Developed NT	1.000	0.996
	$PC_1 - PC_{14}$	Traditional NT	1.000	0.978
	Developed NT	Traditional NT	1.000	0.982





References

- Ahi, P., Jaber, M. Y., & Searcy, C. (2016). A comprehensive multidimensional framework for assessing the performance of sustainable supply chains. *Applied Mathematical Modelling*, 40(23–24), 10153–10166.
- Ahmad, M., Chandio, A. A., Solangi, Y. A., Shah, S. A. A., Shahzad, F., Rehman, A., & Jabeen, G. (2021). Dynamic interactive links among sustainable energy investment, air pollution, and sustainable development in regional China. *Environmental Science and Pollution Research*, 28(2), 1502–1518.
- Ahmad, M., Jan, I., Jabeen, G., & Alvarado, R. (2021). Does energy-industry investment drive economic performance in regional China: Implications for sustainable development. *Sustainable Production and Consumption*, 27, 176–192.
- Ahmad, M., Muslija, A., & Satrovic, E. (2021). Does economic prosperity lead to environmental sustainability in developing economies? Environmental Kuznets curve theory. *Environmental Science and Pollution Research*, 28(18), 22588–22601.
- Alnajem, M., Mostafa, M. M., & ElMelegy, A. R. (2021). Mapping the first decade of circular economy research: A bibliometric network analysis. *Journal of Industrial and Production Engineering*, 38(1), 29–50.
- Al-Tuwaijri, S. A., Christensen, T. E., & Hughes Ii, K. (2004). The relations among environmental disclosure, environmental performance, and economic performance: A simultaneous equations approach. *Accounting, Organizations and Society*, 29(5–6), 447–471.
- Alves, J. A., Silva, L. T., & Remoaldo, P. C. C. (2015). The influence of low-frequency noise pollution on the quality of life and place in sustainable cities: A case study from Northern Portugal. *Sustainability*, 7(10), 13920–13946.
- Anser, M. K., Usman, M., Godil, D. I., Shabbir, M. S., Tabash, M. I., Ahmad, M., Zamir, A., & Lopez, L. B. (2021). Does air pollution affect clean production of sustainable environmental agenda through low carbon energy financing? Evidence from ASEAN countries. *Energy and Environment*.
- Arikan, E., Fichtinger, J., & Ries, J. M. (2014). Impact of transportation lead-time variability on the economic and environmental performance of inventory systems. *International Journal of Production Economics*, 157, 279–288.
- Azadeh, A., Ghaderi, S., & Nasrollahi, M. (2011). Location optimization of wind plants in Iran by an integrated hierarchical data envelopment analysis. *Renewable Energy*, 36(5), 1621–1631.
- Azadeh, A., Ghaderi, S., Omrani, H., & Eivazy, H. (2009). An integrated DEA–COLS–SFA algorithm for optimization and policy making of electricity distribution units. *Energy Policy*, 37(7), 2605–2618.
- Azadeh, A., Salmanzadeh-Meydani, N., & Motevali-Haghighi, S. (2017). Performance optimization of an aluminum factory in economic crisis by integrated resilience engineering and mathematical programming. *Safety Science*, 91, 335–350.
- Battisti, M., Beynon, M., Pickernell, D., & Deakins, D. (2019). Surviving or thriving: The role of learning for the resilient performance of small firms. *Journal of Business Research*, 100, 38–50.

- Brandenburg, M. (2016). Supply chain efficiency, value creation and the economic crisis—An empirical assessment of the European automotive industry 2002–2010. *International Journal of Production Economics*, 171, 321–335.
- Chen, W., Kucukyazici, B., & Saenz, M. J. (2019). On the joint dynamics of the economic and environmental performances for collective take-back systems. *International Journal of Production Economics*, 218, 228–244.
- D’Inverno, G., Carosi, L., & Romano, G. (2020). Environmental sustainability and service quality beyond economic and financial indicators: A performance evaluation of Italian water utilities. *Socio-Economic Planning Sciences*, 100852.
- Dal Pozzo, A., Guglielmi, D., Antonioni, G., & Tugnoli, A. (2018). Environmental and economic performance assessment of alternative acid gas removal technologies for waste-to-energy plants. *Sustainable Production and Consumption*, 16, 202–215.
- D’Ambra, L., Crisci, A., Meccariello, G., Della Ragione, L., & Palma, R. (2020). Evaluation of the social and economic impact of carbon dioxide (CO₂) emissions on sustainable mobility using a trend odds model. *Socio-Economic Planning Sciences*, 100817.
- Das, D. (2018). Sustainable supply chain management in Indian organisations: An empirical investigation. *International Journal of Production Research*, 56(17), 5776–5794.
- Dev, N. K., Shankar, R., & Swami, S. (2020). Diffusion of green products in industry 4.0: Reverse logistics issues during design of inventory and production planning system. *International Journal of Production Economics*, 223, 107519.
- Ding, H., Liu, Q., & Zheng, L. (2016). Assessing the economic performance of an environmental sustainable supply chain in reducing environmental externalities. *European Journal of Operational Research*, 255(2), 463–480.
- Elfarouk, O., Wong, K. Y., & Wong, W. P. (2021). Multi-objective optimization for multi-echelon, multi-product, stochastic sustainable closed-loop supply chain. *Journal of Industrial and Production Engineering*, 1–19.
- Forteza, F. J., Carretero-Gomez, J. M., & Sese, A. (2017). Occupational risks, accidents on sites and economic performance of construction firms. *Safety Science*, 94, 61–76.
- Ganda, F., & Milondzo, K. S. (2018). The impact of carbon emissions on corporate financial performance: Evidence from the South African firms. *Sustainability*, 10(7), 2398.
- Goyal, S., Garg, D., & Luthra, S. (2021). Sustainable production and consumption: Analysing barriers and opportunities for maintaining green tomorrow by using fuzzy-AHP-fuzzy-TOPSIS hybrid framework. *Environment, Development and Sustainability*, 1–47.
- Hammad, A. W., Akbarnezhad, A., & Rey, D. (2017). Sustainable urban facility location: Minimising noise pollution and network congestion. *Transportation Research Part E: Logistics and Transportation Review*, 107, 38–59.
- Hayami, H., Nakamura, M., & Nakamura, A. O. (2015). Economic performance and supply chains: The impact of upstream firms’ waste output on downstream firms’ performance in Japan. *International Journal of Production Economics*, 160, 47–65.
- Huo, B., Gu, M., & Wang, Z. (2018). Supply chain flexibility concepts, dimensions and outcomes: An organisational capability perspective. *International Journal of Production Research*, 56(17), 5883–5903.
- Hussain, A., Oad, A., Ahmad, M., Irfan, M., & Saqib, F. (2021). Do financial development and economic openness matter for economic progress in an emerging country? Seeking a sustainable development path. *Journal of Risk and Financial Management*, 14(6), 237.
- Khan, S. A. R., & Qianli, D. (2017). Does national scale economic and environmental indicators spur logistics performance? Evidence from UK. *Environmental Science and Pollution Research*, 24(34), 26692–26705.
- Król-Badziak, A., Pishgar-Komleh, S. H., Rozakis, S., & Książak, J. (2021). Environmental and socio-economic performance of different tillage systems in maize grain production: Application of Life Cycle Assessment and Multi-Criteria Decision Making. *Journal of Cleaner Production*, 278, 123792.
- Ledesma, R. D., Valero-Mora, P., & Macbeth, G. (2015). The scree test and the number of factors: a dynamic graphics approach. *The Spanish journal of psychology*, 18.
- Ling-zhi, R., Xin-gang, Z., Yu-zhuo, Z., & Yan-bin, L. (2018). The economic performance of concentrated solar power industry in China. *Journal of Cleaner Production*, 205, 799–813.
- Nishitani, K. (2011). An empirical analysis of the effects on firms’ economic performance of implementing environmental management systems. *Environmental and Resource Economics*, 48(4), 569–586.
- Nishitani, K., & Kokubu, K. (2020). Can firms enhance economic performance by contributing to sustainable consumption and production? Analyzing the patterns of influence of environmental performance in Japanese manufacturing firms. *Sustainable Production and Consumption*, 21, 156–169.

- Pal, R., Torstensson, H., & Mattila, H. (2014). Antecedents of organizational resilience in economic crises—an empirical study of Swedish textile and clothing SMEs. *International Journal of Production Economics*, *147*, 410–428.
- Pashapour, S., Bozorgi-Amiri, A., Azadeh, A., Ghaderi, S. F., & Keramati, A. (2019). Performance optimization of organizations considering economic resilience factors under uncertainty: A case study of a petrochemical plant. *Journal of Cleaner Production*, *231*, 1526–1541.
- Rahimi, M., Ghezavati, V., & Asadi, F. (2019). A stochastic risk-averse sustainable supply chain network design problem with quantity discount considering multiple sources of uncertainty. *Computers and Industrial Engineering*, *130*, 430–449.
- Sabbaghnia, A., Heydari, J., & Salmanzadeh-Meydani, N. (2019). Optimization and coordination of a sustainable supply chain with emission reduction considerations. In *2019 15th Iran International Industrial Engineering Conference (IIIEC)*.
- Sabogal-De La Pava, M. L., Vidal-Holguín, C. J., Manotas-Duque, D. F., & Bravo-Bastidas, J. J. (2021). Sustainable supply chain design considering indicators of value creation. *Computers and Industrial Engineering*, 107294.
- Salmanzadeh-Meydani, N., & Fatemi Ghomi, S. (2019). The causal relationship among electricity consumption, economic growth and capital stock in Iran. *Journal of Policy Modeling*, *41*(6), 1230–1256.
- Schwab, L., Gold, S., & Reiner, G. (2019). Exploring financial sustainability of SMEs during periods of production growth: A simulation study. *International Journal of Production Economics*, *212*, 8–18.
- Shan, S., Ahmad, M., Tan, Z., Adebayo, T. S., Li, R. Y. M., & Kirikkaleli, D. (2021). The role of energy prices and non-linear fiscal decentralization in limiting carbon emissions: Tracking environmental sustainability. *Energy*, *234*, 121243.
- Shirali, G. A., Mohammadfam, I., Ebrahimipour, V., & Safety, S. (2013). A new method for quantitative assessment of resilience engineering by PCA and NT approach: A case study in a process industry. *Reliability Engineering and System Safety*, *119*, 88–94.
- Shirali, G. A., Shekari, M., & Angali, K. (2016). Quantitative assessment of resilience safety culture using principal components analysis and numerical taxonomy: A case study in a petrochemical plant. *Journal of Loss Prevention in the Process Industries*, *40*, 277–284.
- Singh, A., & Misra, S. C. (2021). Safety performance & evaluation framework in Indian construction industry. *Safety science*, *134*, 105023.
- Solangi, Y. A., Longsheng, C., Shah, S. A. A., Alsanad, A., Ahmad, M., Akbar, M. A., Ahmad, M., Akbar, M. A., Gumaei, A., & Ali, S. (2020). Analyzing renewable energy sources of a developing country for sustainable development: an integrated fuzzy based-decision methodology. *Processes*, *8*(7), 825.
- Taherdangkoo, M., Ghasemi, K., & Beikpour, M. (2017). The role of sustainability environment in export marketing strategy and performance: A literature review. *Environment, Development and Sustainability*, *19*(5), 1601–1629.
- Toutouchian, S., Abbaspour, M., Dana, T., & Abedi, Z. (2018). Design of a safety cost estimation parametric model in oil and gas engineering, procurement and construction contracts. *Safety Science*, *106*, 35–46.
- Tsao, Y.-C., Thanh, V.-V., Lu, J.-C., & Yu, V. (2018). Designing sustainable supply chain networks under uncertain environments: Fuzzy multi-objective programming. *Journal of Cleaner Production*, *174*, 1550–1565.
- Tseng, M.-L., Tran, T. P. T., Ha, H. M., Bui, T.-D., & Lim, M. K. (2021). Sustainable industrial and operation engineering trends and challenges Toward Industry 4.0: A data driven analysis. *Journal of Industrial and Production Engineering*, 1–18.
- Wilcoxon, F., Katti, S., & Wilcoxon, R. A. (1970). Critical values and probability levels for the Wilcoxon rank sum test and the Wilcoxon signed rank test. *Selected Tables in Mathematical Statistics*, *1*, 171–259.
- Zhong, Y., & Wu, P. (2015). Economic sustainability, environmental sustainability and constructability indicators related to concrete-and steel-projects. *Journal of Cleaner Production*, *108*, 748–756.

Authors and Affiliations

Navid Salmanzadeh-Meydani^{1,2}  · S. M. T. Fatemi Ghomi¹  ·
Seyedhamidreza Shahabi Haghighi¹  · Kannan Govindan² 

Navid Salmanzadeh-Meydani
salmanzadeh@aut.ac.ir

Seyedhamidreza Shahabi Haghighi
shahabi@aut.ac.ir

Kannan Govindan
kgov@iti.sdu.dk

¹ Department of Industrial Engineering, Amirkabir University of Technology, 424 Hafez Avenue, 1591634311 Tehran, Iran

² Centre for Sustainable Supply Chain Engineering, Department of Technology and Innovation, Danish Institute for Advanced Study, University of Southern Denmark, 5230 Odense M, Denmark