Lecture Notes in Civil Engineering

Lelitha Devi
Gowri Asaithambi
Shriniwas Arkatkar
Ashish Verma *Editors*

Proceedings of the Sixth International Conference of Transportation Research Group of India

CTRG 2021 Volume 2

Springer

Lecture Notes in Civil Engineering

Volume 272

Lecture Notes in Civil Engineering (LNCE) publishes the latest developments in Civil Engineering—quickly, informally and in top quality. Though original research reported in proceedings and post-proceedings represents the core of LNCE, edited volumes of exceptionally high quality and interest may also be considered for publication. Volumes published in LNCE embrace all aspects and subfields of, as well as new challenges in, Civil Engineering. Topics in the series include:

- Construction and Structural Mechanics
- Building Materials
- Concrete, Steel and Timber Structures
- Geotechnical Engineering
- Earthquake Engineering
- Coastal Engineering
- Ocean and Offshore Engineering; Ships and Floating Structures
- Hydraulics, Hydrology and Water Resources Engineering
- Environmental Engineering and Sustainability
- Structural Health and Monitoring
- Surveying and Geographical Information Systems
- Indoor Environments
- Transportation and Traffic
- Risk Analysis
- Safety and Security

To submit a proposal or request further information, please contact the appropriate Springer Editor:

– Pierpaolo Riva at pierpaolo.riva@springer.com (Europe and Americas);
– Swati Meherishi at swati.meherishi@springer.com (Asia - except China, and Australia, New Zealand);
– Wayne Hu at wayne.hu@springer.com (China).

All books in the series now indexed by Scopus and EI Compendex database!

Lelitha Devi · Gowri Asaithambi ·
Shriniwas Arkatkar · Ashish Verma
Editors

Proceedings of the Sixth International Conference of Transportation Research Group of India

CTRG 2021 Volume 2

Springer

Editors
Lelitha Devi
Department of Civil Engineering
Indian Institute of Technology Madras
Chennai, India

Gowri Asaithambi
Department of Civil and Environmental Engineering
Indian Institute of Technology Tirupati
Tirupati, India

Shriniwas Arkatkar
Department of Civil Engineering
Sardar Vallabhbhai National Institute of Technology
Surat, India

Ashish Verma
Department of Civil Engineering
Indian Institute of Science
Bangalore, Karnataka, India

ISSN 2366-2557 ISSN 2366-2565 (electronic)
Lecture Notes in Civil Engineering
ISBN 978-981-19-3493-3 ISBN 978-981-19-3494-0 (eBook)
https://doi.org/10.1007/978-981-19-3494-0

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

Preface

We are pleased to present Volume 2 of the Proceedings of the 6th Conference of Transportation Research Group of India (CTRG), which was held at Trichy, India, during December 14 to December 17, 2021. The conference was organized by the Transportation Research Group of India (TRG) and NIT, Trichy, and was supported by IISc, Bangalore; IIT Madras; IIT Palakkad; and NATPAC. The ATPIO, WCTRS (SIG C3, D3, and H2), and IATBR were also co-sponsors (non-financial) of the 6th CTRG.

TRG (https://www.trgindia.org/) is a not-for-profit registered society under Karnataka Societies Registration Act, 1960. It is a national society with board members from various top institutions like IISc, IITs, IIMs, NITs, and CRRI. The society got registered on May 28, 2011, in Bangalore, India, and has already completed 10 years of successful existence. CTRG is the flagship conference of TRG and is held once in two years in different parts of India (1st CTRG at Bangalore in December 2011, 2nd CTRG at Agra in December 2013, 3rd CTRG at Kolkata in December 2015, 4th CTRG at Mumbai in December 2017, 5th CTRG at Bhopal in December 2019, and 6th CTRG at Trichy in December 2021). CTRG follows a strict and journal-style two-stage double-blind peer-review policy for reviewing and selecting the papers. It has been formally supported in the past by various national and leading international institutions/bodies of transportation research like IISc, Bangalore; IIT Kanpur; IIT Kharagpur; IIT Guwahati; IIT Madras; IIT Palakkad; IIT Bombay; SVNIT Surat; MANIT Bhopal; NIT, Trichy; CSIR-CRRI, TRB, T&DI-ASCE, EASTS, ATPIO, WCTRS, and IATBR. The conference typically has about 200 finally accepted paper presentations from 12 to 15 countries and is attended by about 350–400 delegates.

This Volume 2 contains 20 peer-reviewed papers in three sections covering broad topics: traffic flow theory, operations, and facilities; travel behavior and transport demand; and other transportation modes (including NMT) and pedestrian. Through a rigorous peer-review process, editors have ensured high quality for all papers that are included in this volume. We strongly believe that this proceeding volume will serve as a very useful reference for the transportation community in India as well as elsewhere in the world.

Bangalore, India — Ashish Verma
Tirupati, India — Gowri Asaithambi
Chennai, India — Lelitha Devi
Surat, India — Shriniwas Arkatkar

About TRG and CTRG

Transportation Research Group of India (TRG) is a not-for-profit registered society with the mission to aid India's overall growth through focused transportation research, education, and policies in the country. It was formally registered on May 28, 2011, and has completed 10 years of its journey this year. The following are the Vision and Objectives of TRG

Vision

- To provide a unique forum within India for the interchange of ideas among transportation researchers, educators, managers, and policymakers from India and all over the world, with the intention of covering all modes and sectors of transport (road, rail, air, and water; public and private; motorized and non-motorized) as well as all levels (urban, regional, inter-city, and rural transport), and for both passenger and freight movement, in India, and at the same time, to also address the

transportation-related issues of safety, efficiency, economic, and social development, local and global environmental impact, energy, land use, equity and access for the widest range of travelers with special needs, etc.
- To serve as a platform to guide and focus transportation research, education, and policies in India toward satisfying the country's needs and to assist in its overall growth.

Objectives

- To conduct a regular peer-reviewed conference in India so as to provide a dedicated platform for the exchange of ideas and knowledge among transportation researchers, educators, managers, and policymakers from India and all over the world, from a perspective which is multimodal, multi-disciplinary, multi-level, and multi-sectoral, but with an India-centric focus. Initially, this conference will be held every two years; however, the frequency may change as per the decision of the society from time to time.
- To publish a peer-reviewed journal of good international standard that considers and recognizes quality research work done for Indian conditions, but which also encourages quality research focused on other developing and developed countries that can potentially provide useful learning lessons to address Indian issues.
- To conduct other activities such as seminars, training and research programs, meetings, and discussions, as decided by the society from time to time, toward fulfilling the mission and vision of the society.
- To identify pertinent issues of national importance, related to transportation research, education, and policy through various activities of the society, and promote transportation researchers, educators, managers, and policymakers in an appropriate manner to address the same.
- To collaborate with other international societies and organizations like WCTRS, ASCE, TRB, etc., in a manner that works toward fulfilling the mission and vision of the society.

The Conference of Transportation Research Group of India (CTRG) is the premier event of TRG. It is held every two years and traditionally moves around India. In the past, CTRG has been organized in Bangalore (December 2011), Agra (December 2013), Kolkata (December 2015), Mumbai (December 2017), Bhopal (December 2019), Trichy (December 2021), and Surat (upcoming in December 2023 jointly with SVNIT Surat). CTRG has been getting a wide scale recognition from reputed Indian and international institutions/organizations like IIT Kanpur, IIT Kharagpur, IIT Guwahati, IIT Bombay (Mumbai), SVNIT Surat, MANIT Bhopal, NIT Trichy, TRB, WCTRS, CSIR-CRRI, ATPIO, T&DI-ASCE, EASTS, to name a few. CTRG is a large conference typically attended by around 400–500 participants, usually from 12–15 countries, with about 200 double-blind peer-reviewed technical papers being presented. The conference provides a wide range of executive courses, tutorials, workshops, technical tours, keynote sessions, and special sessions.

Transportation in Developing Economies (TiDE) is the official journal of TRG and is published by Springer. TiDE was formally launched in 2014 and has so far published eight volumes.

Prof. Akhilesh Kumar Maurya
Indian Institute of Technology Guwahati
Current President, TRG

List of Reviewers

Ammu Gopalakrishnan, School of planning and architecture, Delhi, India
Ankit Gupta, Shiv Nadar University Dadri, India
Niraj Sharma, CRRI, Delhi, India
Padmarekha, SRM, Chennai, India
Saurabh Dandapat, PwC India, India
Yogeshwar Navandar, NIT Calicut, India
Abdul Pinjari, IISc Bangalore, India
Abdullah Ahmad, NIT Srinager, India
Aboelkasim Diab, Aswan University, Egypt
Aditya Kumar Das, S'O'A University Bhubaneswar, India
Aditya Medury, IIT Kanpur, India
Agnivesh Pani, IIT (BHU) Varanasi, India
Akhilesh Chepuri, GITAM School of Technology Hyderabad, India
Akhilesh Kumar Maurya, IIT Guwahati, India
Ambika Behl, CRRI, New Delhi, India
Ambika Kuity, NIT Silchar, India
Amit Agarwal, IIT Roorkee, India
Amit Kumar Yadav, Central University of Jharkhand, India
Aneena Mohan, IIT Bombay, India
Aninda Bijoy Paul, SVNIT, Surat, India
Anjan Kumar Siddagangaiah, IIT Guwahati, India

Ankit Kathuria, IIT Jammu, India
Ankit Kumar Yadav, IIT Bombay, India
Ankit Yadav, IIT Bombay, India
Anna Charly, IIT Bombay, India
Anuj Budhkar, IIEST, Shibpur, India
Anush Chandrappa, IIT Bhubaneswar, India
Anusha S. P., College of Engineering Trivandrum, India
Apparao Gandi, NICMAR Hyderabad, India
Aravind Krishna Swamy, IIT Delhi, India
Arkopal Goswami, IIT Kharagpur, India
Arpan Mehar, NIT Warangal, India
Arpita Saha, VNIT, India
Arpita Suchismita, IIT Bombay, India
Arun Rajkumar, IIT Madras, India
Asha Latha, College of Engineering Trivandrum, India
Ashish Dhamaniya, NIT Surat, India
Ashish Verma, IISc Banglore, India
Ashok Julaganti, NIT Warangal, India
Ashutosh Arun, Queensland University of Technology, Brisbane, Australia
Avijit Maji, IIT Bombay, India
Avinash Chaudhari, Government Engineering College Daman, India
Ayyanna Habal, PDEU Gandhinagar, India
B. Raghuram Kadali, NIT Warangal, India
B. K. Bhavathrathan, IIT Palakkad, India
Babak Mehran, University of Manitoba, Winnipeg, Canada
Bachu Anil Kumar, IIT Patna, India
Balaji Ponnu, The Ohio State University, USA
Bandhan Majumdar, BITS Pilani Hyderabad, India
Bharat Pathivada, IIT Bombai, India
Bharat Rajan, IIT Bombai, India
Bhargava Chilukuri, IIT Madras, India
Bhaswati Bora, IIT Kanpur, India
Bhupendra Singh, IIT Jalandhar, India
Binanda Khungur Narzary, Tezpur University Assam, India
Bishwajit Bhattacharjee, IIT Delhi, India
Bivina G. R., MANIT Bhopal, India
Burhan Showkat, NIT Surathkal, India
Darshana Othayoth, NIT Trichy, India
Debapratim Pandit, IIT Kharagpur, India
Debashish Roy, NIT Sikkim, India
Debasis Basu, IIT Bhubaneswar, India
Deepa L., IISC Banglore, India
Digvijay Pawar, IIT Hyderabad, India
Dipanjan Mukherjee, IIT Kharagpur, India
Emanuele Renzi, Ministry of Sustainable Infrastructure and Mobility, Italy

Fabio Croccolo, Ministry of Sustainable Infrastructure and Mobility, Italy
Gaurang Joshi, SVNIT Surat, India
Gavadakatla Vamsikrishna, IIT Bombay, India
Gopal Patil, IIT Bombay, India
Gottumukkala Bharath, CSIR-CRRI New Delhi, India
Gourab Sil, IIT Indore, India
Goutham Sarang, VIT Chennai, India
Gowri Asaithambi, IIT Tirupati, India
Gunasekaran Karuppanan, Anna University, Chennai, India
Hareshkumar Golakiya, Govt. Engineering College Surat, India
Hari Krishna Gaddam, National Rail and Transportation Institute, India
Harikrishna Madhavan, NIT Calicut, India
Harish Puppala, BML Munjal University Haryana, India
Harsha Vajjarapu, iCED CAG Jaipur, India
Hemanth Kumar, IIT Bombay, India
Hyuk-Jae Roh, University of Regina, Canada
Indrajit Ghosh, IIT Roorkee, India
Ipsita Banerjee, University of California, Berkeley
Jaikishan Damani, IIT Bombay, India
Jakkula Nataraju, CSIR-CRRI Delhi, India
Jaydip Goyani, SVNIT Surat, India
Jiten Shah, IITRAM Ahmedabad, India
Jithin Raj, IIT Bombay, India
Kalaanidhi Sivagnanasundaram, Technion-Isreal Institute of Technology, Isreal
Kinjal Bhattacharyya, LICIT-ECO7, Université Gustave Eiffel, France
Kirti Mahajan, IIT Bombay, India
Kranthi Kuna, IIT Kharagpur, India
Kshama Puntambekar, SPA, Bhopal, India
Lekhaz Devulapalli, Technion-Isreal Institute of Technology, Isreal
M. N. Sharath, IIT Bombay, India
M. Sivakumar, NIT Calicut, India
M. R. Nivitha, PSG, Coimbatore, India
Madhu Errampalli, CRRI, New Delhi, India
Madhumita Paul, IIT Kharagpur, India
Madhuri Kashyap N. R., IIT Thirupathi, India
Malavika Jayakumar, IIT Bombay, India
Mallikarjuna Chunchu, IIT Guwahati, India
Manoj Malayath, IIT Delhi, India
Manoranjan Parida, IIT Roorkee, India
Marisamynathan S., NIT Trichy, India
Minal Minal, CRRI, Delhi, India
Mithun Mohan, NITK, Surathkal, India
Mohan Rao, CRRI, Delhi, India
Monalisa Patra, IIT Bombay, India
Mounisai Siddartha Middela, IIT Madras, India

Mukti Advani, CSIR-CRRI, India
Munavar Fairooz Cheranchery, TKM College of Engineering, India
Muthulingam Subramaniyan, IIT Ropar, India
Naga Siva Pavani Peraka, IIT Tirupati, India
Nagendra Velaga, IIT Bombay, India
Nandan Dawda, SVNIT Surat, India
Narayana Raju, TU Delft, Netherlands
Naveen Kumar Chikkakrishna, VNRVJIET Hyderabad, India
Nikhil Bugalia, IIT Madras, India
Nikhil Saboo, IIT Roorkee, India
Ninad Gore, Ryerson University, Toronto, Canada
Nipjyoti Bharadwaj, IIT Guwahati, India
Nishant Pawar, IIT Bombay, India
Nivedya M. K., Worcester Polytechnic Institute, USA
Nur Izzi Md. Yusoff, Universiti Kebangsaan, Malaysia
Padma Seetharaman, CRRI, Delhi, India
Pallav Kumar, MIT Muzaffarpur, India
Partha Pratim Dey, IIT Bhubaneswar, India
Parthan K., BMS, Bangalore, India
Phani Kumar Chintakayala, University of Leeds, England
Phani Kumar Patnala, University of Manitoba, Canada
Pooja Raj, Gopalan College of Engineering and Management, India
Prabin Ashish, IIT Kanpur, India
Pranamesh Chakraborty, IIT Kanpur, India
Prasanta Sahu, BITS Pilani, Hyderabad Campus, India
Prateek Bansal, National University of Singapore, Singapore
Pravin Telang, IIT Bombay, India
Pritikana Das, MANIT Bhopal, India
Priyansh Singh, IIT Palakkad, India
Pushpa Choudhary, IIT Roorkee, India
R. Sivanandan, IIT Madras, India
Rahul T. M., IIT Ropar, India
Rajan Choudhary, IIT Guwahati, India
Rajat Rastogi, IIT Roorkee, India
Rajeev Mishra, Delhi Technological University, India
Rajiv Kumar, NIT Jalandhar, India
Ramachandra Rao K., IIT Delhi, India
Ramesh Anbanandam, IIT Roorkee, India
Ramya Sri Mullapudi, IIT Hyderabad, India
Ranja Bandyopadhyaya, NIT Patna, India
Ranju Mohan, IIT Jodhpur, India
Raunak Mishra, University of North Carolina, USA
Ravi Sekhar Chalumuri, CRRI, New Delhi, India
Ravindra Kumar, CRRI, New Delhi, India
Reema Bera, IIT Kharagpur, India

Remya K. P., IIT Bombay, India
Rupali Zope, COEP, Pune, India
Rushikesh Amrutsamanvar, TU Dresden, Germany
S. P. Atul Narayan, IIT Madras, India
S. Vasantha Kumar, VIT, Vellore, India
Sadguna Nuli, Vardhaman College of Engineering, Hyderabad, India
Saladi Sv Subbarao, Mahindra University Hyderabad, India
Sandeepan Roy, IIT Bombay, India
Sanhita Das, IIT Roorkee, India
Sanjay Dave, MSU Baroda, India
Sarah Mariam, BITS Pilani, India
Sarfaraz Ahmed, NUST, Pakistan
Shahana A., IIT Bombay, India
Shankar Sabavath, NIT Warangal, India
Shashi Bhushan Girimath, IIT Bombay, India
Shobhit Saxena, IISc Banglore, India
Shriniwas Arkatkar, SVNIT Surat, India
Shubhajit Sadhukhan, IIT Roorkee, India
Siddharth S. M. P., IIT Bombay, India
Siddhartha Rokade, MANIT, Bhopal, India
Siksha Swaroopa Kar, CSIR-CRRI, India
Sita Rami Reddy, VNIT, Nagpur, India
Smruti Mohapatra, ISM Dhanbad, India
Sonu Mathew, University of North Carolina at Charlotte, USA
Sridhar Raju, BITS Pilani, Hyderabad Campus, India
Srinath Mahesh, IIT Madras, India
Srinivas Geedipally, Texas A&M Transportation Institute, USA
Subhojit Roy, IIT Bhubaneswar, India
Sudhakar Reddy Kusam, IIT Kharagpur, India
Sudhir Varma, IIT Patna, India
Sunitha K. Nayar, IIT Palakkad, India
Sunitha V., NIT Trichy, India
Suresha S. N., NIT Suratkal, India
Sushma Mb, IIT Bombay, India
Suvin P. Venthuruthiyil, IIT Guwahati, India
Swati Maitra, IIT Kharagpur, India
Tarun Rambha, IISc Bangalore, India
Tushar Choudhari, IIT Bombay, India
Udit Jain, VNIT Nagpur, India
Uma M. Arepalli, SRM University AP, Andhra Pradesh, India
Umesh Sahoo, IIT Bhubaneswar, India
Vasudevan N., SVNIT, Surat, India
Vedagiri Perumal, IIT Bombay, India
Veena Venudharan, IIT Palakkad, India
Velmurugan S., CRRI, India

Venkaiah Chowdary, NIT Warangal, India
Venkatesan Kanagaraj, IIT Kanpur, India
Vinamra Mishra, IIT Bombay, India
Vinayaka Ram V., BITS Pilani-Hyderabad, India
Vinayaraj V. S., IIT Bombay, India
Vishwas Sawant, IIT Roorkee, India
Vivek R. Das, Ramaiah Institute of Technology, India
Yawar Ali, SVNIT, Surat, India
Yogesh Shah, IITRAM Ahmedabad, India

Contents

About the Editors

Dr. Lelitha Devi is a Professor in the Transportation Division of the Department of Civil Engineering at Indian Institute of Technology (IIT) Madras and holds a Ph.D. from Texas A&M University, USA. Her teaching and research interests are in the general area of transportation systems with emphasis on traffic flow modelling, traffic operations, and Intelligent Transportation systems. She has published around 90 referreed journal papers and more than 100 conference papers. She serves in various editorial boards and is a member of various national and international societies. She has been involved in 15 research projects and has graduated more than 10 MS/Ph.D. scholars.

Dr. Gowri Asaithambi is currently working as an Assistant Professor in the Department of Civil and Environmental Engineering at Indian Institute of Technology (IIT) Tirupati, India. She received her Ph.D. from IIT Madras in 2011. Her main research interests are traffic flow modelling and simulation, traffic operations and management, pedestrian safety, and intelligent transportation systems.

Dr. Shriniwas Arkatkar is currently working as an Associate Professor in the Civil Engineering Department at SVNIT Surat. Recently, he is appointed as Adjunct Professor at Department of Civil Engineering, Ryerson University, Canada. Prior to joining SVNIT Surat, he has worked in the Department of Civil Engineering at BITS Pilani, India. He pursued his Ph.D. research in the Transportation Engineering Division, Department of Civil Engineering, IIT Madras. His research interests include traffic flow modelling and simulation, traffic operation and management, data collection using new technologies, intelligent transportation systems, transportation systems planning, design and operation, public transportation and sustainable transportation, road safety and simulation. He has published more than 200 research papers in international/national journals and conference proceedings. He has guided eight doctoral students and more than forty masters' students. He serves in various editorial boards and works as a member of various national and international societies.

Ashish Verma is the Convenor of "IISc Sustainable Transportation Lab. (IST Lab.)". He is Ph.D. from Indian Institute of Technology (IIT) Bombay and is currently serving as Associate Professor of Transportation Systems Engineering at the Department of Civil Engineering, Centre for infrastructure, Sustainable Transportation, and Urban Planning (CiSTUP), and Robert Bosch Centre for Cyber Physical Systems (RBCCPS) at Indian Institute of Science (IISc), Bangalore, India. Further, he was a Visiting Professor at ITMO University, Russia and a Visiting Fellow at Queensland University of Technology (QUT), Brisbane, Australia. Before joining IISc, he has served in IIT Guwahati, and Mumbai Metropolitan Region Development Authority (MMRDA).

Traffic Flow Theory, Operations, and Facilities

Machine Learning-Based Gap Acceptance Model for Uncontrolled Intersections Under Mixed Traffic Conditions

A. R. Arathi, M. Harikrishna, and Mithun Mohan

Abstract Uncontrolled intersections are the most common type of intersections in a transportation network. The study modeled minor road driver's decision of accepting or rejecting a gap at four-legged uncontrolled intersections having similar geometric characteristics using Artificial Neural Network (ANN) model, Logistic Regression (LR) model, and Support Vector Machine (SVM) model. The results reveal that the performance of LR and SVM models are somewhat similar, while the performance of ANN model exceeds the performance of both LR and SVM models with a correct prediction of about 96.2%. Also, the higher values of the goodness of fit measures like F1 score and R^2 value together with a lower value of MSE show that ANN model is better in distinguishing between the classes. The variable gap duration has a major influence on model prediction comparing to other variables. The effect of the critical gap, occupancy time, conflicting volume, and vehicle type are also found remarkable.

Keywords Gap acceptance model · Mixed traffic conditions · Uncontrolled intersections · Artificial neural networks (ANN)

A. R. Arathi (✉) · M. Harikrishna
Department of Civil Engineering, National Institute of Technology Calicut, Calicut, Kerala, India
e-mail: arathirajkumar@gmail.com

M. Harikrishna
e-mail: harikrishna@nitc.ac.in

M. Mohan
Department of Civil Engineering, National Institute of Technology Karnataka, Surathkal, Karnataka, India
e-mail: mithun@nitk.edu.in

L. Devi et al. (eds.), *Proceedings of the Sixth International Conference of Transportation Research Group of India*, Lecture Notes in Civil Engineering 272,
https://doi.org/10.1007/978-981-19-3494-0_1

1 Introduction

Uncontrolled intersections are complex at-grade intersections formed either when one major and one minor road intersects or when two major roads/minor roads intersect each other. Analysis of at-grade intersections is important since such intersections involve crossing and turning movements, which will create conflicts between the vehicles as well as between vehicles and pedestrians. Gap acceptance is the widely adopted theoretical basis for the analysis of uncontrolled intersection.

HCM [1] defined gap acceptance as "the process by which a minor street vehicle accepts an available gap to maneuver." Based upon the relative priority of the movements, the driver of a minor road vehicle on approaching the intersection needs to determine whether the gap offered on the major road is sufficient enough to permit entry and when to enter. A gap (measured in seconds) is the time interval between two successive vehicles concerning the reference line. The critical gap is an important parameter that describes gap acceptance, and it is difficult to estimate since it varies among drivers as well as among intersections. A critical gap is defined as "the minimum time in seconds, between successive major-stream vehicles, in which a minor street vehicle can make a maneuver" [1]. It cannot be measured directly from the field.

The popular critical gap estimation methods are Siegloch, Greenshields, Acceptance Curve, Lag, Harders, Ashworth, Raff, Logit, Probit, Hewitt, MLM, PEM, Clearing time, and Occupancy time methods [2–5]. It was found that most of the methods were inefficient to provide actual critical gap value under heterogeneous traffic conditions [4]. VISSIM simulation quantified the accuracy of these methods and found that MLM, PEM, Raff, and Acceptance curve methods provide consistent results [5]. The occupancy time method (OTM) for critical gap estimation considers the actual driver behavior at unsignalized intersections in developing countries and provides reasonably good results when applied to the data collected at one intersection in the US also [6]. It was found that the critical gap estimation becomes realistic by the consideration of clearing time and aggressive behavior of the drivers [7]. The critical gap obtained for right-turning movements was found to be smaller than that obtained for through movements since more than 60% of right-turning from minor stream vehicles forced opposing vehicles to slow down [4]. The critical gap was found to be highest for right-turn (RT) movement from the minor street [8]. Critical gaps of cars for various movements in India are found to be lower than that in the USA, as well as HCM values [9]. Critical gap values have implications on safety at intersections, and two-wheelers having lesser critical gap than that of cars were found to be involved more in road crashes [9]. With an increase in the proportion of large-size vehicles in the conflicting stream, critical gaps of two-wheelers and cars increased linearly [8]. The forceful entries result in higher critical gap values [10]. Among various methods applied for estimating critical gaps, the one suggested by the Indo-HCM was found to give precise values of the critical gap under mixed traffic conditions [11]. In most countries, the gap acceptance modeling is being a common area of interest to research.

Various researchers used different techniques/approaches to model the gap acceptance behavior, such as traffic signal analogy-based gap acceptance model [12, 13], microscopic decision model using risk-reward loop process [14], behavioral model using ANFIS (Adaptive Neuro-Fuzzy Interface System) [15], Binary Logit Model (BLM) [7, 16–26], binary probit model [27], alternative unbiased model [28], random utility model [29], fuzzy logic model [29], Support Vector Machines (SVM) [19, 21, 25, 30], Random Forests (RF) [30], Decision Trees (DT) [30], and generalized gap acceptance model [31]. Among these methods, the deterministic method of gap acceptance modeling was easy to apply [16]. The exponential model of rejected proportion was established as more practical than the linear model [32]. Both the logit and fuzzy models were good for representing gap acceptance behavior [29]. It was observed that SVM outperforms BLM, and thus, SVM can be used to classify and predict accepted gaps and rejected gaps [21]. The performance of SVM and DT models were found similar while the performance of the RF model exceeds the SVM and DT [30]. BLM and SVM were relatively close in their estimation of the spatial critical gap values [25]. The gap acceptance values for truck traffic were found higher than that for passenger traffic [33]. The increase in service delay was found to provoke the drivers to accept smaller gaps [34]. Drivers accept shorter gaps when turning across traffic than merging with traffic [35]. It was observed that the conflicting vehicle type affects the gap acceptance behavior [36]. As the number of rejected gaps and the mean time interval of the rejected gaps increase, the probability of accepting a given gap also increases [23]. The size and type of vehicle were found to affect the decision of the drivers to accept a presented gap [24]. It was also observed that the age of the driver was an effective parameter in gap acceptance decision than of the conflicting vehicle type [15]. Middle-aged and old-aged drivers accepted comparatively larger gaps than younger ones [17, 37]. When the driver was young, and the conflicting vehicle was a two-wheeler (or both), the probability of accepting a gap was higher [38]. Older female drivers were found to be a more vulnerable group due to their poor driving ability [37]. Researchers had established drivers to be less sensitive to the approach speed of the vehicle and more sensitive to the distance or position of oncoming vehicles [37]. The accepted temporal gaps were independent of approach speed, while accepted spatial gaps were dependent on approach speed [19]. In developing countries like India, the drivers were aggressive and accepted smaller gaps [20, 39]. The two-wheelers were more aggressive than three-wheelers for most of the observed major-stream vehicular combinations [10].

It can be summarized that different researchers developed different approaches for gap acceptance modeling and critical gap estimation. In the case of gap acceptance modeling, different approaches such as deterministic approach (Raff, Siegloch, Greenshields methods), probabilistic approach (Logit model, Probit model, Neural Networks), stochastic approach (Bayesian, Bootstrap approaches), and fuzzy logic approaches were developed by the researchers. Some of these approaches were simple to use, whereas some involve complex computation tasks. From the review of literature, the variables that may influence driver's gap acceptance decision in developing countries can be summarized as gap duration, critical gap, temporal/spatial gaps, number of rejections, occupancy, traffic volume, intersection type, time of day, size and type of minor road vehicle, type of movement, waiting time, aggressive behavior/forced entry, distractions on driving, driver's socio-economic characteristics like age and gender, size and type of major road/conflicting vehicle, conflicting volume, major road vehicle combination, major road traffic speed, and distance or position of the oncoming vehicle. The gap acceptance studies at the uncontrolled intersection are necessary for developing countries like India. In comparison with developed countries, the lane discipline, movement priorities, right of way allocation, fixation of stop, or yield signs are not strictly obeyed in developing countries like India.

Uninterrupted flow condition on the major street is violated in India because the vehicles on the major street are slowing down or forced to slow down by vehicles on the minor street [8]. Lack of proper control at uncontrolled intersections makes traffic behavior complex and difficult to predict. This is especially so in heterogeneous traffic conditions, as prevalent in developing countries like India. The drivers with varied characteristics and aggressive behavior, along with violation of priority rules, compound the issues at uncontrolled intersections. The risk-taking maneuver and aggressive behavior of Indian drivers were found as the major causes of road crashes at uncontrolled intersections [9]. Hence, the traffic situation at uncontrolled intersections is often chaotic and unpredictable making maneuvering unsafe. Drivers approaching the intersection need to evaluate whether the gaps available are sufficient to execute maneuvers safely and when to enter the intersection. The inappropriate traffic control and misjudgments of the drivers may lead to the poor operational performance of the intersection and may lead to road crashes at such intersections. The operational analysis of uncontrolled intersections, therefore, becomes necessary, for which gap acceptance-based analysis is mainly adopted. There are several kinds of research dealing with gap acceptance characteristics at uncontrolled intersections since it affects performance measures like capacity, delay, and safety. Even though there were different approaches to model gap acceptance behavior, the simplest method with the least error and no compromise on accuracy is still a research question in front of each researcher. Artificial Neural Network (ANN) is an alternative to conventional methods like regression analysis. Unlike the traditional statistical methods of analysis, the machine learning-based method like ANN does not impose any restrictions on the input variables and residual distributions. In the area of gap acceptance studies, the ANN-based modeling approach is rare. The main objective of this study is to develop a machine learning-based gap acceptance model for four-legged uncontrolled intersections having ideal geometric conditions under mixed traffic conditions in India and to evaluate how well the ANN model makes predictions compared to other popular models such as Logistic Regression (LR) model and Support Vector Machine (SVM) model.

Fig. 1 Manassery intersection (site-1)

2 Study Site Selection and Data Collection

The study sites were selected considering the roadway geometry, traffic movements, and availability of a suitable location for mounting the video camera. The important considerations for the selection of study sites were (i) approach legs of the intersection meet at right angles, (ii) intersection is at level grade, (iii) similar geometry, and (iv) free from pedestrians, presence of bus-stops and presence of speed breakers. Two intersections, in the state of Kerala, India, fulfilling the above criteria having similar geometric characteristics, were selected for the study purpose. Site-1 was Manassery Junction, situated in the district of Kozhikode and site-2 was Kacheripady Junction in the district of Malappuram. Both intersections were four-legged uncontrolled intersections having two-lane undivided major and minor roads. The approach width of major and minor roads at site-1 were 14.9 m and 10 m, and at site-2 were 14.3 m and 12.9 m. Both intersections were having comparatively similar geometric characteristics. The traffic data was collected by conducting a videographic survey at the selected intersections. The traffic flow was recorded for 1 hour during the morning peak. The traffic composition on the roads was highly heterogeneous comprising vehicles of widely varying static and dynamic characteristics. The snapshot of the traffic flow at these intersections is shown in Figs. 1 and 2.

3 Data Extraction

The traffic data collected using the videographic method were transferred to a computer, and the required data were extracted manually from the video. The format of the data extraction sheet is shown in Fig. 3. The time at which the vehicles occupy the required positions on the intersection was noted in the Excel sheet simultaneously by playing the video repeatedly on the computer. The following data were noted down from the recorded video:

Fig. 2 Kacheripady intersection (site-2)

Minor Road						Major Road					
Vehicle type	Movement (LT/T/RT)	Time when vehicle stop at the entry line (t_1)	Time when vehicle start from the entry line (t_2)	Time when vehicle cross the exit line (t_3)	Occupancy time (t_{ot}=t_3-t_2)	Leading vehicle type	Rear wheel at reference line (t_4)	Following vehicle type	Front wheel at reference line (t_5)	Gap duration (t_g=t_5-t_4)	Accepted / Rejected gap (A/R)

Fig. 3 Format of data extraction sheet

- Time when the minor road vehicle stops at the entry line (t_1)
- Time when the minor road vehicle starts to move from the entry line (t_2)
- Time when the minor road vehicle crosses the exit line (t_3)
- Time when the rear wheel of the leading vehicle on the major road comes at reference line (t_4)
- Time when the front wheel of the following vehicle on a major road comes at reference line (t_5)

The gaps presented to the driver while entering the intersection and the driver's decision about accepting or rejecting that particular gap (*A* or *R*) were also noted. Occupancy time (t_{ot}) and gap duration (t_g) were calculated from the extracted data. The time difference between t_3 and t_2 was calculated as the occupancy time ($t_{ot} = t_3$-t_2) of each vehicle, and the time difference between two vehicles on the major road ($t_g = t_5$-t_4) was calculated as the gap duration.

3.1 Traffic Volume Count and Vehicle Proportion

Classified movement-wise traffic volume counts of vehicles passing through the selected intersections were retrieved from the recorded video. As per vehicle classification given in Indo-HCM [40], the vehicles were categorized into Two-Wheelers (TW), Auto-rickshaws (Auto), Standard/Small Cars (SC), Big Cars (BC), Light Commercial Vehicles (LCV), Two/Three-Axle Trucks (TAT), Multi-axle Trucks (MAT), and Buses (B). Total traffic volumes of 2900 veh/h and 3581 veh/h were observed during morning peak hour at site-1 and site-2. The extracted vehicle counts in each vehicle category were converted into passenger car units (PCU) as per the PCU values provided in Indo-HCM [40]. The vehicle proportions at both sites are represented in Figs. 4 and 5.

Comparing to other modes of vehicles, the proportion of TW is higher (56–57%) than SC (15–18%), Auto (15–17%), and LCV (3–5%). The buses constitute only 3% out of total vehicles. It indicates higher private mode usages compared to public transit. Proportion of heavy vehicle such as TAT is only about 2%, which is found very less compared to other modes.

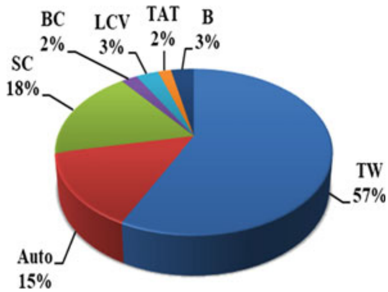


Fig. 4 Vehicle composition at site-1

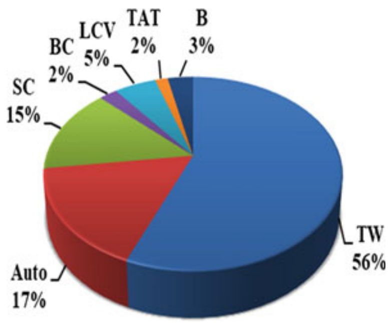


Fig. 5 Vehicle composition at site-2

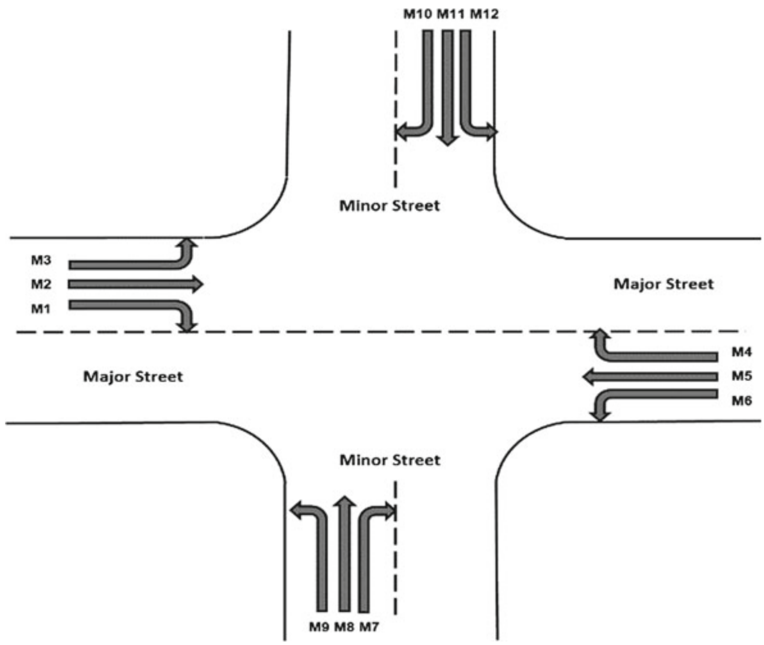


Fig. 6 Vehicular movements at a typical four-legged uncontrolled intersection

3.2 Conflicting Traffic Volume

The conflicting traffic volume was calculated based on the extracted traffic volumes as per the equation provided in Indo-HCM [40]. Vehicular movements at a typical four-legged uncontrolled intersection are shown in Fig. 6. The conflicting traffic volume corresponding to each movement (M_x) from minor roads was calculated at both intersections. The right-turn movements from minor roads (M7 and M10) were subjected to a conflicting volume of 1203 PCU/h and 1203 PCU/h at site-1 and 1625 PCU/h and 1625 PCU/h at site-2. The through movements from minor roads (M8 and M11) were subjected to a conflicting volume of 1634 PCU/h and 1385 PCU/h at site-1 and 1921 PCU/h and 1747 PCU/h at site-2.

3.3 Average Approach Speed

The approach speed was extracted by noting the time taken by a major road vehicle to pass each of the entry line and exit line marked on the site with a known distance. The ratio of distance and time to cross the distance gives the speed of the corresponding vehicle. Then, the average approach speeds were calculated for each vehicle category at both sites.

3.4 Critical Gap

In the case of gap data set, the number of rejected gaps (1183) was found to be higher than the number of accepted gaps (779). The critical gap (t_{cg}) is determined based on the occupancy time method. The graph corresponds to the occupancy time method shows time (in seconds) on the X-axis and cumulative frequency distribution of occupancy time (F_{ot}) and 1 minus cumulative frequency distribution of accepted gap (1-F_a) on the Y-axis. The intersection point of the cumulative frequency distribution of the (1-F_a) curve and occupancy time (F_{ot}) curve was determined, and that intersection point was then extended to X-axis, which gives the critical gap (t_{cg}) value. Likewise, the critical gap values corresponding to each vehicle type executing a particular movement were determined for both sites and they are summarized in Table 1.

From the tabled values, a general trend as obtained in previous researches [5, 6, 8, 9] can be observed such as the value of the critical gap is increasing with the size of the vehicle except for two movements (exception cases may be due to lesser volume of vehicle executing that particular movement or due to aggressive behavior of the drivers). The critical gap obtained for right-turning movements (M7, M10) is found to be smaller than that obtained for through movements (M8, M11), similar to the results obtained in the previous researches [4, 10].

4 Gap Acceptance Modeling

The problem of predicting the driver's decision to accept or reject an available gap is a type of classification problem. This study developed a machine learning-based gap acceptance model using ANN with the help of the python programming language. Factors influencing the drivers' decision were taken as independent variables, and the driver's decision, such as accepting the gap, was considered as the dependent variable. The categorical variables (nominal variables) were converted into ordinal variables with the help of a hot encoding process. The total data set was divided into two parts such that 80% of data is for training and 20% is for testing purposes. A four-fold cross-validation process was carried out in training data. That is, out of

Table 1 Critical gap values at both sites

Vehicle type	Critical gap values							
	Movements at site-1				Movements at site-2			
	M7	M8	M10	M11	M7	M8	M10	M11
TW	4.8	7.0	5.0	7.5	3.8	6.5	4.0	7.9
Auto	7.5	8.0	6.5	9.0	5.5	7.0	5.8	8.0
SC	7.5	9.0	9.1	6.5	6.0	8.0	6.5	9.5
LCV	–	–	–	–	–	7.0	7.0	10.5

80% training data set, a split was made as fourfolds. In onefold, 20% data was used for cross-validation and the remaining threefolds for training 60% of data. Then, this cross-validation was performed repeatedly in these fourfolds with replacement. Three models were developed, such as Artificial Neural Network (ANN) model, Logistic Regression model (LR model), and Support Vector Machine model (SVM model). Different trial runs were conducted until the model performs well and gives better goodness of fit measures. Finally, the performance of the developed ANN model was evaluated by comparing its accuracy and other goodness of fit measures with corresponding LR and SVM models.

4.1 Theoretical Considerations

Logistic regression (LR) is a widely adopted supervised learning classification algorithm which provides probabilistic interpretation of outcome variable. The advantages of LR algorithm are as follows: easy to implement, highly interpretable, very efficient to train, provides feature importance along with its direction of association, less training time, training requires less computational power, performs well when data has linearly separable features, easy to regularize, produce discrete outcome, unlikely to overfit, easy to extend into multiple classes, and used as a benchmark model. The disadvantages are as follows: assumption of linearly in logit can rarely hold, cannot solve non-linear problems, tough to obtain complex relationships, cannot be used if the number of observations are lesser than the number of features, requires less or no multi-collinearity, can only predict categorical variable since outcome variable bound to discrete number set, creates linear boundaries, model may overfit on high dimensional data set, sensitive to outliers, require large data set, and not immune to missing data.

Support Vector Machine (SVM) is one of the most popular supervised learning algorithms which can be used to solve both classification and regression problems. For classification, SVM finds suitable hyperplane that has maximum margin in N-dimensional space and that hyperplane act as decision boundaries to classify the data points below and above it. The advantages of SVM are as follows: can handle non-linear data efficiently, can solve complex relationship, can be used when there is no idea on data or when data have unknown distribution, less risk of overfitting, able to handle high dimensional data, has convex optimization nature, better computational complexity, outliers have less influence, applicable to semi-supervised learning models, and can be used when total number of samples lesser than the number of dimensions. The disadvantages are as follows: difficult to understand and interpret the final model, lack of transparency for results, not suitable for large data set, works poorly for overlapping classes (when data set has more noise), may underperform if the number of features exceed the number of training samples, long training time, no probabilistic explanation for classification since it directly gives the resultant classes, difficult to choose a good kernel function, requires feature scaling, and requires high memory.

Artificial Neural Network (ANN) is another popular supervised learning algorithm which is the computing systems inspired by the structure and function of human brain. The advantages of ANN are as follows: There is no need to decide what kind of model will fit, can easily handle nominal and ordinal scale variables either as dependent or independent variables, can handle non-linearities easily without beforehand knowing exactly which type of non-linearity exists, can model both non-linear and complex relationships, can infer unseen relationships on unseen data, predicts the results from a learned experience, can found hidden intrinsic relations, can handle multi-class problems, less chances for overfitting, can produce a number of outputs, can train the model in one go, and have fixed model size. The disadvantage of ANN is that the hidden layer is somewhat called a black box, where no one knows what is going inside that one. Other disadvantages are as follows: ANN covers only local minima, may overfit if training goes for too long time, and might start to consider the noise as part of pattern.

4.2 Data Pre-Processing

The data set consists of the parameters like minor road vehicle type, movement (M), inter-arrival time (IAT), occupancy time (OT), gap offering major road following vehicle type, gap duration (GD), accepted/rejected gap (A/R), critical gap (CG), conflicting volume (CV), and average approach speed of major road vehicles (AAS). The driver's decision of accepting the gap was considered as the dependent variable. The correlation coefficient was determined between these parameters. Independent variables were finalized based on the correlation coefficient and its significance. Data pre-processing processes such as removing outliers and removing multi-collinearity were carried out in the gap data set.

4.3 Gap Acceptance Model

The gap data set consists of eight independent variables like minor road vehicle type (considered as four different variables: TW, Auto, SC, LCV), occupancy time, gap duration, critical gap, conflicting volume, and one dependent variable such as accepted gap. The minor road vehicle type was the only categorical variable (with four categories) in the selected data set. The results obtained from the ANN, LR, and SVM models are describing as follows.

ANN—Gap Acceptance Model. A gap acceptance model was developed using the MLP classifier modeling method with 'Logistic' as activation function and 'Adam' as the solver. Three hidden layers were used. The different trial runs were carried out by changing the number of hidden layers and the number of neurons repeatedly to achieve good results. The neural network architecture which provides reasonably good result is given in Fig. 7. Figures 8, 9, and 10 represent the confusion matrix, receiver operating characteristic (ROC) curve, and model loss curve obtained for the ANN model based on selected neural architecture.

Fig. 7 Neural network architecture

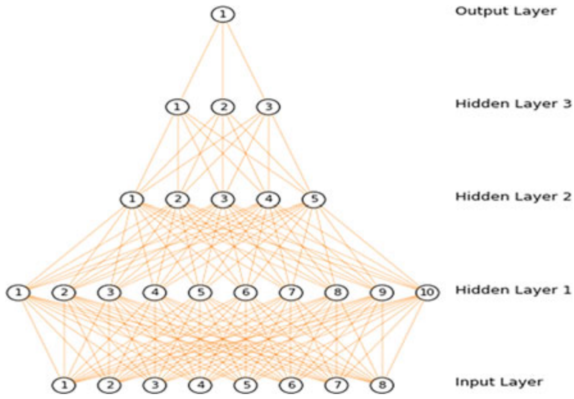


Fig. 8 Confusion matrix

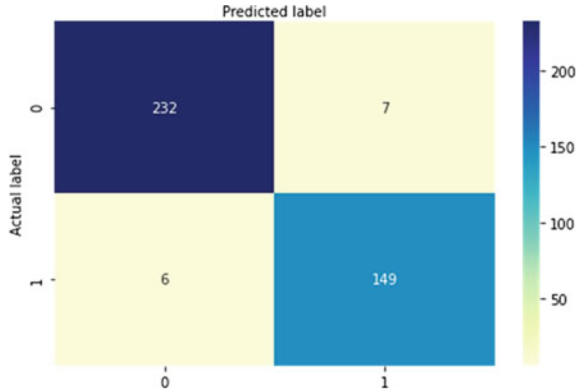
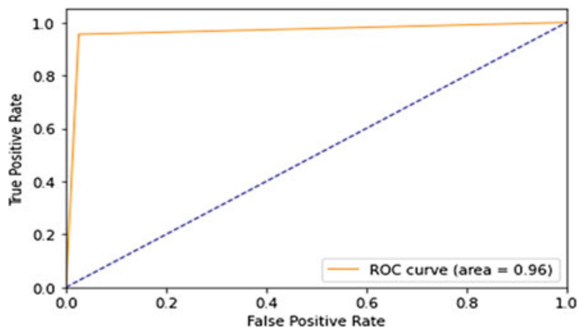


Fig. 9 ROC curve



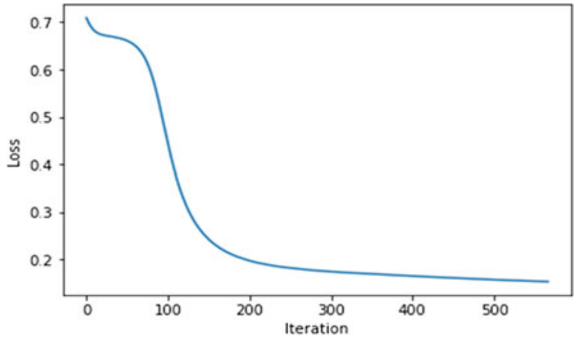


Fig. 10 Loss curve

The ANN model gives 92% accuracy with a standard deviation of 0.01 during validation and gives 96.2% accuracy during testing. The obtained values for Mean Absolute Error (MAE), Mean Squared Error (MSE), ROC-Area Under the Curve (AUC), F1 score, and R^2 value are 2.617, 0.038, 0.96, 0.955, and 0.843. The least loss value interpreted from the loss curve (Fig. 10) is 0.156.

LR—Gap Acceptance Model. A gap acceptance model was developed using the logistic regression (LR) modeling technique. The LR model gives 95% accuracy with a standard deviation of 0.01 during the validation stage and results in 95.1% accuracy during the testing stage. The obtained values for MAE, MSE, ROC-AUC, F1 score, and R^2 value are 4.553, 0.041, 0.96, 0.951, and 0.833. The model coefficients of the LR model are given in Table 2, where β_1, β_2, β_3, β_4, β_5, β_6, β_7, and β_8 are the coefficients corresponding to occupancy time (OT), gap duration (GD), critical gap (CG), conflicting volume (CV), Minor road vehicle type-Auto, Minor road vehicle type-LCV, Minor road vehicle type-SC, Minor road vehicle type-TW, and C is the constant.

The negative valued coefficients indicate that the occupancy time, critical gap, and conflicting volume are inversely proportional to the gap acceptance decision, while the gap duration is directly proportional to the gap acceptance decision. As

Table 2 Logistic regression model coefficients

Model coefficients	Coefficient values
β_1 (OT)	−0.588
β_2 (GD)	6.585
β_3 (CG)	−0.658
β_4 (CV)	−0.301
β_5 (minor road vehicle type—Auto)	0.264
β_6 (minor road vehicle type—LCV)	−0.176
β_7 (minor road vehicle type—SC)	−0.102
β_8 (minor road vehicle type—TW)	−0.120
Constant	0.375

Table 3 Summary of gap acceptance model results

Results	ANN model	LR model	SVM model
Accuracy (validation)	0.92	0.95	0.95
	With a SD of 0.01	With a SD of 0.01	With a SD of 0.01
Accuracy (testing)	0.962	0.951	0.959
ROC—AUC	0.96	0.96	0.96
MAE	2.617	4.553	4.553
MSE	0.038	0.041	0.041
F1 score	0.955	0.951	0.951
R^2 value	0.843	0.833	0.833
Model loss	0.156	_	_

gap duration increases, the driver tends to accept the gaps, whereas the increase in occupancy time/critical gap/conflicting volume provokes the driver to reject those gaps. Hence, the proposed LR model provides a logical relationship between the dependent and independent variables. The variable gap duration is found to have a major influence on the model prediction comparing to other variables. Other than gap duration, the effect of the critical gap, occupancy time, conflicting volume, and vehicle type are also found remarkable.

SVM—Gap Acceptance Model. A support vector machine (SVM)-based gap acceptance model was also developed in this research. The SVM model gives 95% accuracy with a standard deviation of 0.01 during the validation stage and results in 95.9% accuracy during the testing stage. The obtained values for MAE, MSE, ROC-AUC, F1 score, and R^2 value are 4.553, 0.041, 0.96, 0.951, and 0.833.

Findings and Inferences. The results obtained from the three gap acceptance models, such as the ANN model, LR model, and SVM model, are summarized in Table 3.

From the table, it can be observed that the three models are performing well. Comparing the results obtained from these three classification models, the ANN model can be selected as a reasonably good predicting model which predicts the results with higher values for the accuracy (in the validation and testing), ROC-AUC, F1 score, and R^2 value. Therefore, the ANN model outperforms the LR model and the SVM model. LR and SVM models give approximately the same results. ANN model produced 96.2% correct predictions with an MSE of 0.038. F1 and R^2 values that correspond to the ANN model are 0.955 and 0.843, which is near 1 indicates a good model. Overall, the developed machine learning-based ANN model is better in distinguishing between the classes (accept or reject).

5 Conclusions

This study provides a machine learning-based gap acceptance model to predict the decision of drivers of minor road vehicles to accept/reject the particular gap presented to them at four-legged uncontrolled intersections under mixed traffic conditions in Indian context. The three different models, such as ANN, LR, and SVM models, were developed with the aid of the python programming language. Then, the correctness and goodness of fit measures of the developed ANN-gap acceptance model were compared with the results of both LR and SVM models. It revealed that the performance of LR and SVM models were somewhat similar, while the performance of the ANN model exceeds the performance of both LR and SVM models. Comparing to LR and SVM models, the ANN model performs better and gives reasonably good results. ANN model is better in distinguishing between the classes (accept or reject). It gives 96.2% correct predictions for gap data. As stated in previous research [7, 9, 24], the proposed LR model also implies that the variable gap duration has a major influence on the model prediction compared to other variables. Other variables like the critical gap, occupancy time, and conflicting volume also significantly affect gap acceptance decisions. Also, similar to the result from the previous study [24], it was found that the vehicle type also produces some effects on the decision of the driver to accept the gap.

The proposed machine learning-based gap acceptance model can be used for predicting the driver's decision to accept or reject the available gap presented to them at four-legged uncontrolled intersections in Indian context in an effective manner, since ANN-based gap acceptance modeling is rare in previous researches. In the real scenario, the gap and lag are different and need to be studied separately. So, a similar procedure can be adopted to develop a lag acceptance model in further studies. In this modeling, the forced entry/aggressive behavior and other driver behavior-related variables, presence of pedestrians, geometric factors, and time of day were not considered. These variables can be included in future studies. Also, this study can be extended to three-legged intersections with the same procedure. Here, the gap data parameters extracted from two intersections were only used for modeling purposes. Hence, the data set is somewhat small. ANN can provide good and much better results for a comparatively large data set. So, with an increase in data points, the efficiency of the model can be improved.

References

1. Transportation Research Board (TRB) (2000) Highway capacity manual (HCM). National Research Council, Washington, D.C.
2. Golias JC, Kanellaidis GC (1990) Estimation of driver behavior model parameters. J Transp Eng 116(2):153–166
3. Gattis JL, Low ST (1999) Gap acceptance at a typical stop-controlled intersections. J Transp Eng 125(3):201–207

4. Amin HJ, Maurya AK (2015) A review of critical gap estimation approaches at uncontrolled intersection in case of heterogeneous traffic conditions. J Transp Lit 9(3):5–9
5. Mohan M, Chandra S (2016) Review and assessment of techniques for estimating critical gap at two-way stop-controlled intersections. Eur Transp Transp Europei 61:1–18
6. Mohan M, Chandra S (2018) Critical gap estimation at two-way stop-controlled intersections based on occupancy time data. Transp A: Transp Sci 14(4):316–329
7. Dutta M, Ahmed MA (2018) Gap acceptance behavior of drivers at uncontrolled T-intersections under mixed traffic conditions. J Mod Transp 26(2):119–132
8. Mohan M, Chandra S (2020) Influence of major stream composition on critical gap at two-way stop-controlled intersections–A case study. Transp Lett 12(1):1–8
9. Chandra S, Mohan M (2018) Analysis of driver behaviour at unsignalized intersections. J Indian Roads Congr 79(2):5–10
10. Advani M, Gupta NJ, Velmurugan S, Madhu E, Chandra S (2020) Defining and analyzing forceful gap behavior at unsignalised intersections. Transp Res Rec
11. Deepthi MS, Ramesh A (2021) A model for estimation of critical gap and its distribution behaviour at un signalised intersections. Adv Civ Eng 83:537–549
12. Akçelik R (1994) Gap acceptance modelling by traffic signal analogy. Traffic Eng Control 35(9):498–506
13. Akçelik R (2007) A review of gap-acceptance capacity models. In: The 29th conference of Australian institutes of transport research (CAITR), University of South Australia, Adelaide, Australia
14. Pollatschek MA, Polus A, Livneh M (2002) A decision model for gap acceptance and capacity at intersections. Transp Res Part B: Methodological 36(7):649–663
15. Sangole JP, Patil GR, Patare PS (2011) Modelling gap acceptance behavior of two-wheelers at uncontrolled intersection using neuro-fuzzy. Procedia Soc Behav Sci 20:927–941
16. Cassidy MJ, Madanat SM, Wang MH, Yang F (1995) Unsignalized intersection capacity and level of service: Revisiting critical gap. Transp Res Rec 1484:16–23
17. Dissanayake S, Lu JJ, Yi P (2002) Driver age differences in day and night gap acceptance capabilities. IATSS Res 26(1):71–79
18. Devarasetty PC, Zhang Y, Fitzpatrick K (2012) Differentiating between left-turn gap and lag acceptance at unsignalized intersections as a function of the site characteristics. J Transp Eng 138(5):580–588
19. Patil GR, Pawar DS (2014) Temporal and spatial gap acceptance for minor road at uncontrolled intersections in India. Transp Res Rec 2461:129–136
20. Patil GR, Sangole JP (2015) Gap acceptance behavior of right-turning vehicles at T-intersections—A case study. J Indian Roads Congr 76(1):44–54
21. Pawar DS, Patil GR, Chandrasekharan A, Upadhyaya S (2015) Classification of gaps at uncontrolled intersections and midblock crossings using support vector machines. Transp Res Rec 2515:26–33
22. Pawar DS, Patil GR (2017) Minor-street vehicle dilemma while maneuvering at unsignalized intersections. J Transp Eng Part A: Syst 143(8):1–11
23. Zhou H, Ivan JN, Gårder PE, Ravishanker N (2017) Gap acceptance for left turns from the major road at unsignalized intersections. Transport 32(3):252–261
24. Mohan M, Kezhuvathodi V (2019) Simulation of gap acceptance at three-legged unsignalized intersections under heterogeneous traffic conditions. Adv Transp Stud 49
25. Pawar DS, Patil GR (2019) Analyzing variations in spatial critical gaps at two-way stop controlled intersections using parametric and non-parametric techniques. J Traffic Transp Eng (English Edition), 1–10
26. Ma W, Zhao J (2019) Investigating gap acceptance behavior at two-way stop-controlled intersections in China. Transp Lett 1–11
27. Hamed MM, Easa SM, Batayneh RR (1997) Disaggregate gap-acceptance model for unsignalized T-intersections. J Transp Eng 123(1):36–42
28. McGowen P, Stanley L (2012) Alternative methodology for determining gap acceptance for two-way stop-controlled intersections. J Transp Eng 138(5):495–501

29. Riccardo R, Massimiliano G, Gregorio G, Claudio M (2012) Comparative analysis of random utility models and fuzzy logic models for representing gap-acceptance behavior using data from driving simulator experiments. Procedia Soc Behav Sci 54:834–844
30. Nagalla R, Pothuganti P, Pawar DS (2017) Analyzing gap acceptance behavior at unsignalized intersections using support vector machines, decision tree and random forests. Procedia Comput Sci 109C:474–481
31. Abhishek, Boon MAA, Mandjes M (2019) Generalized gap acceptance models for unsignalized intersections. Math Methods Oper Res 89(3):385–409
32. Guo R, Lin B (2011) Gap acceptance at priority-controlled intersections. J Transp Eng 137(4):269–276
33. Fitzpatrick K (1991) Gaps accepted at stop-controlled intersections. Transp Res Rec 1303:103–112
34. Lu JJ, Lall BK (1995) Empirical analysis of traffic characteristics at two-way stop-controlled intersections in Alaska. Transp Res Rec 1495:49–56
35. Beanland V, Lenné MG, Candappa N, Corben B (2013) Gap acceptance at stop-controlled T-intersections in a simulated rural environment. Transp Res F: Traffic Psychol Behav 20:80–89
36. Sangole JP, Patil GR (2014) Adaptive neuro-fuzzy interface system for gap acceptance behavior of right-turning vehicles at partially controlled T-intersections. J Mod Transp 22(4):235–243
37. Yan X, Radwan E, Guo D (2007) Effects of major-road vehicle speed and driver age and gender on left-turn gap acceptance. Accid Anal Prev 39(4):843–852
38. Patil GR, Sangole JP (2016) Behavior of two-wheelers at limited priority uncontrolled T-intersections. IATSS Res 40(1):7–18
39. Pawar DS, Patil GR (2018) Response of major road drivers to aggressive maneuvering of the minor road drivers at unsignalized intersections: A driving simulator study. Transp Res F: Traffic Psychol Behav 52:164–175
40. Indo-HCM (2017) Indian Highway Capacity Manual. CSIR-Central Road Research Institute, New Delhi

Modelling Decision to Overtake on Two-Lane Two-Way Highway

Jose Jais and M. Harikrishna

Abstract One of the most complex and unavoidable manoeuvres on undivided two-lane two-way roads is overtaking, where the vehicles use the opposing lane to overtake the slow-moving vehicles, in the presence of vehicles coming from opposite direction. Overtaking manoeuvre is complicated in mixed traffic conditions, where the vehicles have different static and dynamic characteristics. In this study, a binary logistic regression model was developed to study the decision to overtake on the roads. Different variables that affect the overtaking behaviour of vehicles were considered in this study. The results show that speed of overtaking, overtaken and opposing vehicle, space headway before overtaking and lateral adjustment by overtaken vehicle during overtaking manoeuvre are significant at 95% confidence level. The results that obtained from this paper are useful to predict the overtaking behaviour of vehicles in mixed traffic conditions, and it can also be used as an input in a traffic simulation model.

Keywords Overtaking · Two-lane two-way highways · Logistic regression model

1 Introduction

India, a country having mixed traffic conditions on its roads, is estimated to have 132,500 km of National Highways in a road network of 5.4 million km and a major part of it are two-lane two-way type [1]. In a mixed traffic stream, vehicles do not follow lane discipline, and overtaking manoeuvers are completed by using the lane occupied by the opposing stream of vehicles. A considerable percentage of the fatal road traffic crashes on two-lane two-way highways could be directly linked to overtaking manoeuvers [2]. Hence, overtaking manoeuvers have a significant impact

J. Jais (✉) · M. Harikrishna
National Institute of Technology Calicut, Calicut, Kerala, India
e-mail: jaisaryankala55@gmail.com

M. Harikrishna
e-mail: harikrishna@nitc.ac.in

L. Devi et al. (eds.), *Proceedings of the Sixth International Conference of Transportation Research Group of India*, Lecture Notes in Civil Engineering 272,
https://doi.org/10.1007/978-981-19-3494-0_2

on the traffic flow performance and safety of two-lane two-way highways. In order to develop mitigation strategies to augment the road safety scenario on Indian roads, a comprehensive understanding about the overtaking manoeuvers is necessary. This study aims to bridge the gap in the understanding on the overtaking behaviour of different categories of vehicles on two-lane two-way highways.

There were different studies carried out on overtaking manoeuvers on two-lane roads to improve and to understand the driver's behaviour before, during and after an overtaking manoeuver [3–6]. Rahul et al. [7] studied the overtaking behaviour of vehicles on Indian highways under vehicle heterogeneity. Data collection was done using LiDAR and video cameras from an instrumented vehicle. LiDAR was used to measure the distance between the vehicles. In this study, a new variable called 'excess distance' was introduced to understand overtaking behaviour. The results showed that on divided four lane roads, the side of overtaking, and on undivided two-lane roads, the type of overtaken vehicle plays an important role. And also, if overtaking vehicle undergoes overtaking at higher speed, then it was seen that overtaking vehicle maintained greater excess distances with the overtaken vehicle. This study also shows that overtaking decision of vehicles on undivided roads is affected by the presence of vehicles coming from opposite direction and the type of overtaken vehicles. It was observed that the decision to overtake was made before an appropriate gap in the opposing traffic stream is available [6]. Farah et al. [3] developed a passing decision acceptance model using the data collected on two-lane highways, using an interactive driving simulator. This model mainly took into account the impact of the road geometry and driver's characteristics. It mainly focussed on the subject vehicle's decision as to pass the leading vehicle or not. The results indicated that the variables like size of the available passing gap, the speed of the subject vehicle, the vehicle in front and the opposing vehicle, gap between the front vehicle, the subject vehicle and the type of the front vehicle were statistically significant at 95% confidence level [3]. Mocsari [6] carried out a study on overtaking behaviour of vehicles on two-lane rural roads in Hungary. The data required for the study was collected with the help of an instrumented vehicle. The results indicated that the space headway before overtaking was not influenced by the speed of the vehicle to be overtaken. Asaithambi and Shrarani [8] carried out a similar study on a two-lane two-way National Highway in India. The results of this study showed that the number of overtaking manoeuvres decreases with increase in flow in opposite direction. However, the influence of space headway between vehicles was not accounted for. Mizanur and Nakamura [9] developed a model in terms of total traffic volume for studying the overtaking manoeuvres of vehicles. They studied the effects of non-motorized rickshaws on passing manoeuvres.

Compared to the studies done on homogeneous traffic conditions, there are only limited studies which focus on the overtaking behaviour of vehicles on mixed traffic condition. Most of the researchers were focussed on studying the overtaking behaviour of vehicle by considering either overtaking or overtaken vehicles. In this study, the decision to overtake is modelled considering various aspects pertaining to overtaking, overtaken and opposing vehicles. The work aims to bridge the gap in the knowledge base relating to overtaking manoeuvres in mixed traffic conditions.

The subsequent sections deal with detailed methodology and data collection effort for the study, preliminary analysis of data, model development and the conclusions derived from the study.

2 Methodology and Data Collection

2.1 Methodology

From the studies reviewed, it is identified that the factors which affect the overtaking behaviour of vehicles are speed of overtaking vehicle, speed of overtaken and the speed of opposing vehicles, time taken for overtaking, length of overtaking, space headway before overtaking, space headway after overtaking, type of vehicles involved in overtaking manoeuvre, width of carriage way and shoulder, and lateral adjustment by overtaken vehicle during overtaking. The data required for this study is mainly classified into two: traffic data and geometric data. Geometric data was collected using rodometer and traffic data using video cameras.

In order to collect the traffic data relevant for the study, a straight section of 300 m length in the NH 66 Bypass, a two-lane two-way highway, at Pantheerankavu, Calicut, Kerala, was identified. As a microscopic analysis of traffic operations on considerable length of road is required, a uniform straight road section of considerable length on which traffic operations is visible without any interferences was selected. Figure 1 shows the video snapshot of selected stretch of study area and Fig. 2 shows the map view of the same.

The relevant data was manually extracted from the collected videos. Preliminary analysis of extracted data was then carried out. The detailed analysis of extracted data was done using R software. The data was used for developing logistic regression model using R software. Logistic regression model (LRM) was developed using two-third of the collected data in order to identify the pertinent variables influencing the decision to overtake on two-lane two-way highways. The developed model was also validated using one-third of the data.

Fig. 1 Video snapshots of selected stretch of study area

Fig. 2 Map view of selected stretch of study area

2.2 Data Collection

Traffic data was collected in the selected study stretch for 6 h from10:00 am to 4:00 pm using videographic method. Two cameras used for data collection were fixed on the top of building near to highway in such a way that each camera captures a maximum coverage of 150 m. Care was taken to ensure that no tilting of cameras occurred. Flow data was collected using TIRTL, which is a portable infrared traffic logger. TIRTL is a multi-purpose traffic sensor, in which infrared light cones are sent from transmitter and is received by a receiver which is situated on the opposite side of the road, perpendicular to the flow of traffic. It was fixed on the selected stretch of study area with transmitter on one side of carriageway and receiver on the other side to collect the flow data. Geometric characteristics of the selected road stretch were obtained using rodometer which measures distances on the road. The carriageway width and shoulder width on both the sides of road were found for every 10 m interval, using the rodometer.

The average, maximum and minimum values of shoulder width and carriageway width towards Pantheerankavu and towards Calicut are given in Table 1. According to IRC, for two-lane two-way highway, the minimum width of carriageway is 7 m. Therefore, the selected study stretch conforms to the standards.

Data extraction from the collected videos was done manually. If a vehicle overtakes another vehicle at high speed from a long distance, it is referred to as flying type overtaking. If a vehicle overtakes a vehicle in such a way that, the overtaking vehicle follows the overtaken vehicle at a speed similar to overtaken vehicle for some time and on observing a suitable gap if the overtaking manoeuver is completed, it is called

Table 1 Geometric details of study section

	Average width (m)	Maximum width (m)	Minimum width (m)
Carriageway width	7.35	7.7	7
Shoulder width towards Pantheerankavu	1.13	1.6	0.8
Shoulder width towards Calicut	1.23	1.6	0.9

as accelerative overtaking. Flying and accelerative type overtaking are considered in this study because, these are the major type of overtaking manoeuvers observed in the selected study stretch. The details extracted from the videos are (1) type of overtaking, overtaken and opposing vehicle; (2) speed of overtaking, overtaken and opposing vehicle; (3) total time required for overtaking; (4) length of overtaking; (5) space and time headway between overtaking and overtaken vehicle before and after overtaking; (6) gap between overtaken vehicle and opposing vehicle when overtaking vehicle starts overtaking; (7) type of overtaking (accelerative or flying overtaking); (8) direction in which overtaking happened; (9) lateral adjustment by overtaken vehicle during overtaking. The speed of vehicle is obtained by dividing the distance covered by vehicle to the time taken for travelling that distance. Length of overtaking is obtained by measuring the distance from the recorded video and then converting it to actual scale. Types of vehicle were noted by visual examination. The collected data is first used for preliminary data analysis, which is given in the next section.

3 Preliminary Data Analysis

Preliminary data analysis of data helps the analyst to get an overall idea about the various overtaking characteristics of vehicles involved in overtaking manoeuvers. The subsequent sections detail the preliminary data analysis carried out with the collected data.

3.1 Successful and Aborted Overtaking

A total of 1615 overtaking manoeuvers were observed in the study stretch. If the overtaking process is completed, then the overtaking is called as successful overtaking. Aborted overtaking occurs when an overtaking vehicle is forced to terminate the overtaking operation due to some unexpected reasons. These reasons may be due to insufficient speed of overtaking vehicle, insufficient gap in front of overtaking vehicle or the overtaken vehicle does not undergo lateral adjustment during overtaking, etc. 1329 (82.29%) were successful overtaking and 286 (17.71%) were aborted overtaking.

3.2 Types of Overtaking

Different types of overtaking were observed in the study area in which most of them are accelerative and flying overtaking. Out of 1329 total successful overtaking observed, 888 (67%) was accelerative type overtaking and 441 (33%) was flying type overtaking. Under low traffic volume condition, it was seen that the number of flying type overtaking is more. On the other hand, when the traffic flow increases, the number of accelerative types overtaking increases.

3.3 Type of Vehicles Involved

Different types of vehicles were found to involve in the overtaking manoeuver. The types of vehicles identified are two-wheelers (2 W), cars, heavy commercial vehicles (HCV), light commercial vehicles (LCV), buses, three-wheeled passenger auto (3P), three-wheeled commercial auto (3A) and four-wheeled passengers auto (4P). Table 2 shows the relative proportion of different types of vehicles as overtaking, overtaken and opposing vehicle. Overtaking vehicles may have a different response towards different types of overtaken vehicles. For example, the way in which a car overtakes a heavy vehicle may be differ from the way a car overtakes a two-wheeler. Therefore, the type of vehicles is an important variable that should consider in the modelling of overtaking decision of vehicles.

According to motor vehicles statistical year book, India [10], cars and two-wheelers population are more compared to other category of vehicles. In the selected stretch also, the cars and two-wheelers proportion are higher in all category as overtaking, overtaken and opposing vehicles. In recent years, the crashes by two-wheelers and cars are observed to be more, in which the majority are caused during overtaking road accidents in India, MoRTH [11]. Since a greater number of cars and two-wheelers were observed in study stretch, a study on overtaking behaviour is relevant. A car overtaking a slow-moving car with another car as opposing vehicle was mostly observed in the field, followed by a car overtaking a three-wheeled passenger auto with another car as opposing vehicle.

Table 2 Types of vehicles involved

		Overtaking vehicle	Overtaken vehicle	Opposing vehicle
CAR		1006 (62%)	420 (26%)	738 (46%)
HCV		66 (4%)	174 (11%)	143 (9%)
LCV		118 (7%)	94 (6%)	117 (7%)
2 W		322 (20%)	600 (37%)	430 (27%)
3A		9 (1%)	89 (5%)	48 (3%)
3P		37 (2%)	171 (11%)	99 (6%)
BUS		33 (2%)	36 (2%)	19 (1%)
4P		24 (2%)	31 (2%)	21 (1%)

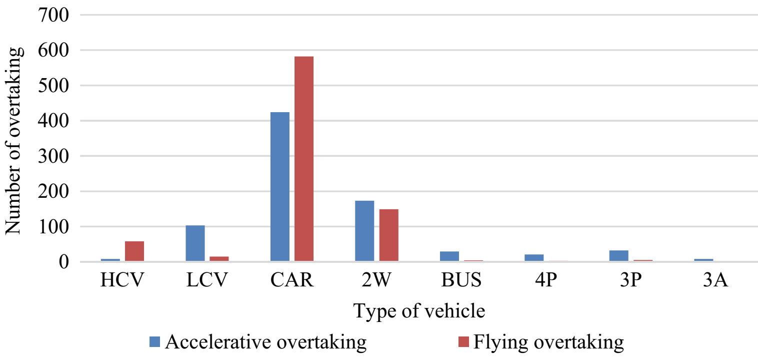


Fig. 3 Number of overtaking versus type of overtaking for different categories of vehicles

3.4 Number of Overtaking Versus Type of Overtaking

Figure 3 shows the relation between the number of overtaking manoeuvers and the type of overtaking vehicle. It was seen that 87.8% of HCV and 58% of cars are involved in flying overtaking, compared to other categories of vehicles. This may be due to higher operating speeds of cars and aggressive behaviour of heavy vehicles. 88% of buses, 86.49% of 3P, 87.3% of LCV were found to be involved in more accelerative overtaking manoeuvers, compared to other categories of vehicles. This may be attributed to their lower operating characteristics and larger size of vehicles. Moreover, two-wheelers were found to be involved in both types of overtaking manoeuvers in equal numbers, possible due to their higher manoeuvrability and smaller size.

4 Logistic Regression Modelling

The relation between the decision to overtake and identified influencing variables was explored using binary logistic regression. The decision to overtake was considered as the dependent variable and traffic-related variables were considered as independent variables. Spearman's correlation test was used to identify salient influencing variables. Out of different variables listed earlier, 10 variables were found to be highly correlated (correlation greater than 0.6) with dependent variable. Multi-collinearity was also accounted for while selecting the independent variables. The binary dependent variable 'Y' was coded as '1' for successful overtaking and '0' for aborted overtaking. Equation 1 is used for estimating the parameters of logistic regression model.

$$\begin{aligned} &\mathrm{Log}\big(p/1-p\big) = \beta_o + \beta_i X_i \quad \text{and,} \\ &p(\text{successful overtaking}) = \exp(\beta_o + \beta_i X_i) \end{aligned} \tag{1}$$

where 'p' is the probability of successful overtaking, β_0 is the intercept and β_i is the vector of coefficients of X_i which is the vector of explanatory variables (input variables). One of the important statistical measures associated with logistic regression is the odds ratio. If Y is an event with 1 indicating that the event has occurred and 0 indicating that it has not, then,

$$\text{odds}(\mathrm{Y} = 1) = \frac{\Pr(\mathrm{Y} = 1)}{\Pr(\mathrm{Y} = 0)} = \mathrm{e}^{\beta} \tag{2}$$

where β is the coefficient corresponding to variable Y. This implies that the expectation of the probability of the event to occur is e^{β} times that of the probability of the event not occurring. In the developed model, the base category is taken as two-wheelers. The model variables and coefficients are shown in Table 3.

Higher absolute value of coefficient 'B' indicates that higher increase per unit change of the corresponding predictor variable. Negative value of coefficients indicates decrease of the odds upon increase of corresponding predictor variable.

The effect of type of overtaking vehicle shows that, compared to two-wheelers, if overtaking vehicle is a car, then chances of successful overtaking is more and least, as size of overtaking vehicles increases. This is because the cars possess higher manoeuvrability, as compared to heavy vehicles. It was seen that when the overtaken

Table 3 Results of logistic regression model

Variables		B	S.E	Z value	Pr(>|z|)	e^{β}
Overtaking vehicle category	Car	1.789	0.284	6.264	0.035	5.924
	LCV	−2.772	0.886	−3.241	0.098	0.056
	HCV	−2.887	0.557	−5.363	0.687	0.051
Overtaken vehicle category	Car	−1.570	0.337	−4.955	0.002	0.188
	LCV	0.854	0.178	4.854	0.458	2.373
	HCV	0.943	0.237	3.557	0.002	2.323
Opposing vehicle category	Car	−0.967	0.487	−1.986	0.476	0.380
	LCV	−0.654	0.241	−2.714	0.086	0.519
	HCV	−1.271	0.807	-1.575	0.004	0.281
Speed of overtaking vehicle		0.445	0.113	3.938	0.000	1.561
Speed of overtaken vehicle		−0.542	0.499	1.086	0.000	0.582
Speed of opposing vehicle		−0.285	0.484	0.589	0.000	0.752
Space headway before overtaking		−0.191	0.548	−0.348	0.000	0.826
Gap between overtaking and opposing vehicle during overtaking		0.081	0.112	0.723	0.018	1.084
Time headway before overtaking		−0.489	0.305	−1.275	0.001	0.678
Lateral adjustment by overtaken vehicle		1.899	0.575	3.476	0.000	7.382
Constant		3.690	0.937	4.045	0.008	44.256

vehicle is a car, then the chance of successful overtaking is less. If the overtaken vehicle is HCV or LCV, then the chance of successful overtaking is more. This is because the overtaking vehicles has less tendency to follow an HCV or LCV due to their large size and visibility reasons so overtaking vehicle always try to overtake them. From the model, it is observed that the effect of type of opposing vehicle on making successful overtaking is negatively related, as compared to two-wheelers. Moreover, the chances to make a successful overtaking decrease when heavy vehicles are encountered from the opposite direction. Speed of overtaken and opposing vehicles, space and time headway before overtaking are negatively related to dependent variable, which implies that the chance of successful overtaking decreases with an increase in their values. The speed of overtaking vehicle is found to have a positive coefficient value. The reason could be attributed to the fact that, when the speed of overtaken vehicle is less compared to overtaking vehicle, the driver of overtaking vehicle will try to complete the overtaking manoeuvre as soon as possible. Space headway and time headway before overtaking are negatively related to the dependent variable. This is because of the fact that if the headway distance is less, then the distance between overtaking and overtaken vehicle is less and overtaking vehicle can easily undergo overtaking manoeuvre. If the gap between overtaking and opposing vehicle is more, then the overtaking vehicle can easily perform the overtaking manoeuvre. Hence, the variable has a positive coefficient. During overtaking manoeuvre, if the overtaken vehicle undergoes lateral adjustment, then the chances of successful overtaking is more. Hence, the particular variable has a positive coefficient value.

In logistic regression modelling, e^{β} (odd ratio) is the factor by which the odds change when 'X' increases by one unit. The coefficient 0.445 implies that a one unit change in speed of overtaking vehicle results in a 1.561 unit change in making a successful overtaking. In other words, the probability of making a successful overtaking is 1.561 times the probability of event not occurring. If space and time headway before overtaking change by one unit, the resultant change in making successful overtaking is 0.826 and 0.678, respectively. Similarly, if the gap between overtaking and opposing vehicle during overtaking increases by one unit, the resultant increase in making successful overtaking is 1.084 times.

The developed logistic regression model is validated with one-third of the data (538 observations). Caret package is used using R-programming language to check the accuracy of the model. A confusion matrix, shown in Table 4, is made to visualize the performance of the model. The accuracy score of developed logistic regression model is obtained as 84.39%.

Table 4 Confusion matrix

		Predicted value	
		1	0
Actual values	1	342	59
	o	25	112

The Cox and snell, Nagelkerke and McFadden pseudo R square values are 0.686, 0.737 and 0.589, respectively. This suggests that the model explains around 73.7 to 58.9% of the variation in the outcome.

5 Conclusions

This paper presents the methodology to model the decision to overtake by vehicles on two-lane two-way highways based on type and speed of overtaking, overtaken and opposing vehicle, gap between overtaking and opposing vehicle during overtaking, space and time headway between overtaking and overtaken vehicle before overtaking and lateral adjustment by overtaken vehicle. Logistic regression method of modelling is used to predict whether the overtaking is successful or aborted. R-programming language was used for developing the model. Developed logistic regression model is validated with one-third of the data. The model results indicate that these variables play a significant role in the decision to overtake by the drivers. The chances of successful overtaking are found to increase with increase in speed of overtaking vehicle and decreases with increase in speed of overtaken and opposing vehicles. The model results show that space and time headways between overtaking and overtaken vehicles affect the decision to overtake. The accuracy score of developed logistic regression model is obtained as 84.39%. In this study, unlike from other works [3, 6, 8], lateral adjustment by overtaken vehicle during overtaking was taken as one variable and the result shows that the variable is significant.

In future, simulation model of traffic pattern including overtaking behaviour of vehicles can also be developed. This study can be extended to divided highways. While developing the model, the presence of median can be chosen as a variable and also driver's characteristics can be included.

References

1. Ministry of Road Transport & Highways, Government of India, New Delhi (2019–2020), Annual report: 2019–2020
2. Gebru MK (2017) Road traffic accident: Human security perspective. Road Traffic Accid Hum Secur Perspect 8(2):15–24
3. Farah H, Bekhor S, Polus A, Toledo T (2009) A passing gap acceptance model for two-lane rural highways. Transportmetrica 5(3):159–172
4. Figueira AC, Larocca APC (2020) Analysis of the factors influencing overtaking in two-lane highways: A driving simulator study. Transp Res Part F 69:38–48
5. Hegeman G, Hoogendoorn S, Brookhuis A (2005) Observations overtaking manoeuvres on bi-directional road. Advanced OR and AIin Transportation, Publishing House of Poznam University of Technology, Poznan, p 505
6. Mocsari T (2009) Analysis of the overtaking behaviour of motor vehicle drivers. Acta Tech J 2(1)
7. Rahul S, Vasudevan V, Dutta B (2020) Analysis of overtaking patterns of Indian drivers with data collected using a LiDAR. Transp Res part F: Traffic Psychol Behav 74:139–150 (2020)

8. Asaithambi G, Shrarani G (2017) Overtaking behaviour of vehicles on undivided roads in non-lane based mixed traffic conditions. J Traffic Transp Eng 4(3):252–261
9. Mizanur RM, Nakamura F (2005) A study on passing-overtakingd level of service of heterogeneous traffic flow. J Eastern Asia Soc Transp 6:1471–1483
10. Ministry of Statistics and Programme Implementation (2018) Motor vehicles-statistical year book, India
11. Road Accidents in India (2019) Government of India, Ministry of road transport and highways, Transport Research Wing, New Delhi

Influence of Skew Angle on the Gap Parameters at Uncontrolled Intersections

Unnikrishnan Sreevishnu, A. R. Arathi, and M. Harikrishna

Abstract Occupancy time and critical gap are the key parameters used in the determination of the capacity of uncontrolled intersections. Indo-HCM (CSIR-Central Road Research Institute (2017) Unsignalised intersections, Indo-HCM, Chapter 8) suggests the method of occupancy time to determine the capacity of uncontrolled intersections but is limited to right-angled intersections. Geometric factors influence the occupancy time of movements that occur at the intersections. This study considers skew angle as the influencing factor. Three sites are considered in this study with varying skew angles. The occupancy times and critical gaps for various movements for different vehicle classes have been recorded. Linear regression models are developed for the three types of non-priority movements at the intersection. It is found that with increased distance to be traversed for a particular movement, the occupancy time and critical gap also increases. The occupancy time values also vary according to vehicle classes with the larger vehicles tending to have larger values.

Keywords Occupancy time · Skew angle · Critical gap

1 Introduction

Uncontrolled intersections are intersections where no form of space or time control exists. Such intersections in countries like India are seen to have complicated vehicle interactions due to the mixed traffic condition and aggressive driver behavior. Hence, vehicles entering the intersections are given no indication of whether the movement they intend to make is possible at that point of time. It is up to the driver of the vehicle to determine whether the gaps available in the conflicting streams are suitable to be

U. Sreevishnu (✉) · A. R. Arathi · M. Harikrishna
National Institute of Technology Calicut, Calicut, Kerala, India
e-mail: usreevishnu@gmail.com

M. Harikrishna
e-mail: harikrishna@nitc.ac.in

L. Devi et al. (eds.), *Proceedings of the Sixth International Conference of Transportation Research Group of India*, Lecture Notes in Civil Engineering 272,
https://doi.org/10.1007/978-981-19-3494-0_3

acceptable or not. Considering the complexity of the situation, accurate estimation of the capacity of such intersections is also a challenge.

One of the parameters required for estimating the capacity of uncontrolled intersections is the occupancy time, which is used for the estimation of the critical gap [1]. Critical gap refers to the minimum gap in the priority stream, which is acceptable to a driver executing a non-priority movement [1]. Occupancy time refers to the time that a vehicle incurs in completely clearing the conflict area of the intersection. It is measured as the time elapsed between the arrival of subject vehicle at the edge of the conflict area and its complete exit from the intersection conflict area gap [1]. Occupancy time varies for each different vehicle class and for each type of movement occurring at the intersection. Hence, it is indicative of the clearing behavior of various vehicles at the intersection.

One of the earlier research works related to the method of occupancy time was by Chandra and Mohan [2]. They recommended the use of occupancy time for both heterogeneous conditions as in India, as well as, for homogenous traffic conditions. They observed that the critical gaps obtained by this method depict driver behavior better than previously existing methods. Moreover, they discerned that the critical gap was found to be lower in value for Indian drivers, which indicate their aggressive nature. They also found that motorized two-wheelers have a lower critical gap value due to their small size and easy handling [3]. Further, they observed that the capacity computed was found to vary from field capacity, and this difference was used to develop an exponential correlation factor [4]. Ashalatha and Chandra [5] proposed the method of clearing time to determine the critical gap for vehicle categories executing the various movements at the intersection. It considers both the accepted gap and clearing time distribution to calculate critical gap. Asaithampi and Anuroop [6] performed an analysis of occupancy time of vehicles at urban unsignalized intersections in non-lane-based mixed traffic conditions. The influence of conflicting traffic on occupation time at an uncontrolled and a semi-controlled intersection was studied by them, and occupancy time was taken as the sum of service delay and clearing time. Mathematical models were formulated to relate the occupancy time to volume of conflicting traffic. In their study, occupancy time was found to increase when the volume of conflicting traffic was higher. The most recent work on the determination of capacity of unsignalized intersections was enunciated in the Indo-HCM 2017. However, the methodology suggested can be used only for intersections that fulfill certain criteria such as:

- The number of intersecting approaches should be either 3 or 4
- The angle of intersection should be 90° on a three or four-legged intersection with a deviation of ±10°
- The intersecting arms should have either a 2 or 4 lane-divided major road
- Negligible presence of non-motorized traffic, on-street parking, hawkers, or any other land use activities within 75 m from the center of the intersection
- There should be no gradient present on the intersecting approaches
- Safe stopping sight distance is available
- No speed breakers on any approach within 75 m from the center of intersection.

As intersection legs meet at angles other than right angle, on account of the site restriction, it is necessary to study the influence of intersection geometry in the gap parameters at uncontrolled intersections. Previous research works are silent on the influence of geometric factors on the gap parameters at uncontrolled intersections. This study aims to bridge the gap in the knowledge base by considering the influence of skew angle on the gap parameters at uncontrolled intersections. For this study, data pertaining to occupancy times and accepted gaps was collected from three study sites. All three study sites selected had varying skew angles. Once data were collected, the occupancy time for each non-priority movement for various vehicle classes was computed in the data analysis stage. Critical gap values for all non-priority movements for vehicle classes were also computed using the method of occupancy time. Using the datasets from the three sites, linear regression models for occupancy time and critical gap were formed for each type of non-priority movement.

In the subsequent sections, the methodology adopted for the study, study site selection, and data collection details are given. The details of the data extracted from the videos and their relevance are explained. The details of the analysis with the extracted data and model development are explained. Finally, the conclusions derived from the study and scope for future work are also given.

2 Methodology and Data Collection

In this section, the details of the methodology adopted for the study and that of data collection and data extraction are included.

2.1 Methodology Adopted for the Study

As it is required to analyze the influence of skew angle on the critical gap of non-priority movements at uncontrolled intersections, a videographic survey was conducted at three identified uncontrolled intersections. The data were collected for one hour each in the FN and AN session of a typical working day. Due care was taken to select a day in which the traffic movements are not influenced by any extraordinary events such as festivals. From the videos, using the KINOVEA software, details such as major stream traffic volume, time at which minor road vehicle stops at the entry line, time at which minor road vehicle starts to move, and time at which minor road vehicle crosses the exit line are extracted. The entry time and exit time of major road vehicles in the intersection area are also noted. From the collected details, variables such as accepted gap, rejected gap, occupancy time, and delay were derived. The occupancy time for each non-priority movement for various vehicle classes was computed in the data analysis stage. Critical gap values for all non-priority movements for vehicle classes were also computed using the method of occupancy time. Using the datasets from the three sites, linear regression models for occupancy time were formed for each type of non-priority movement. The relationship between occupancy time and critical gap with the skew angle was explored, and models were developed to predict the occupancy time of non-priority movements based on the skew angle and vehicle type.

2.2 Study Intersection Details

The first site considered for the study is Kodakara Junction, in the district of Thrissur, Kerala. It is a four-legged skewed intersection with the minor roads skewed at an angle of 30° toward the right. Both the major and minor roads are 2-lane 2-way undivided roads with road widths of 6.5 m and 5.5 m, respectively. The traffic composition consists of 63% two-wheelers, 14% auto-rickshaws, 19% small cars, and 4% light commercial vehicles with the rest being vehicle classes not considered for this study.

Site 2 identified for the study is Polytechnic Junction, Thriprayar, in the district of Thrissur, Kerala. It is a four-legged skewed intersection with the minor roads skewed at an angle of 14° toward the left. Both the major and minor roads are 2-lane 2-way undivided roads with road widths of 7.5 m and 5.5 m, respectively. The traffic composition consists of 57% two-wheelers, 13% auto-rickshaws, 19% small cars, and 4% light commercial vehicles with the rest being vehicle classes not considered for this study.

Site 3 considered for the study is Market Junction, Irinjalakuda, in the district of Thrissur, Kerala. It is a four-legged skewed intersection with the minor roads skewed at an angle of 18° toward the left. Both the major and minor roads are 2-lane 2-way undivided roads with road widths of 7.5 m and 5.5 m, respectively. The traffic composition consists of 50% two-wheelers, 14% auto-rickshaws, 20% small cars, and 5% light commercial vehicles with the rest being vehicle classes not considered for this study.

2.3 Data Collection

Videographic survey was used to collect the traffic data at the selected intersections. The data were collected at the sites for morning and evening peak periods from 9.30 AM to 10.30 AM and from 4 to 5 PM. The peak periods were identified after a reconnaissance visit to the study sites. The geometric details such as carriageway width were collected using measuring tape.

Data extraction was done using KINOVEA software. This software helps to play the collected video at slower speeds so that accurate data can be extracted. The various movements that occur at an uncontrolled intersection are depicted in Fig. 1. Based on the level of priority, the movements are classified into four classes.

Priority rankings for movements are assigned as suggested in Indo-HCM [1]. Movements 2, 3, 5, and 6 are given the highest priority with rank 1. They are followed

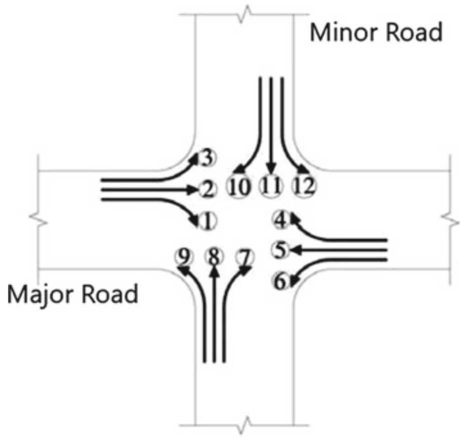


Fig. 1 Various movements occurring at a four-legged intersection. (*Source* Indo-HCM [1])

by movements 1 and 4 with a priority rank of 2 and further by movements 7 and 10 with a priority rank of 3. Movements 8 and 11 are given the least priority. Movements 9 and 12 are left turning movements from minor road which are omitted in the rankings as they are said to provide no interference to any traffic movements. The data extracted from the videos include major stream traffic volume, time at which minor road vehicle stops at the entry line, time at which minor road vehicle starts to move, and time at which minor road vehicle crosses the exit line. It also includes entry time and exit time of major road vehicles. Using the extracted data, values of accepted gap, rejected gap, occupancy time and delay were estimated. The classes of vehicles that were used in the study were auto-rickshaws (3 or 4 wheeled) which were denoted as AUTO, small cars denoted as SC, two-wheelers denoted as TW, and light commercial vehicles denoted as LCV.

3 Data Analysis

The occupancy time for the various study sites was classified on the basis of vehicle class and turning movement. The median values of the occupancy times were considered in this study.

Skew angle is the angle by which the angle of the intersection deviates from 90°. For this study, skew angle is taken with respect to a rectangular coordinate system. Any skew angles measured toward the right are taken as positive and those to the left are taken as negative as depicted in Fig. 2.

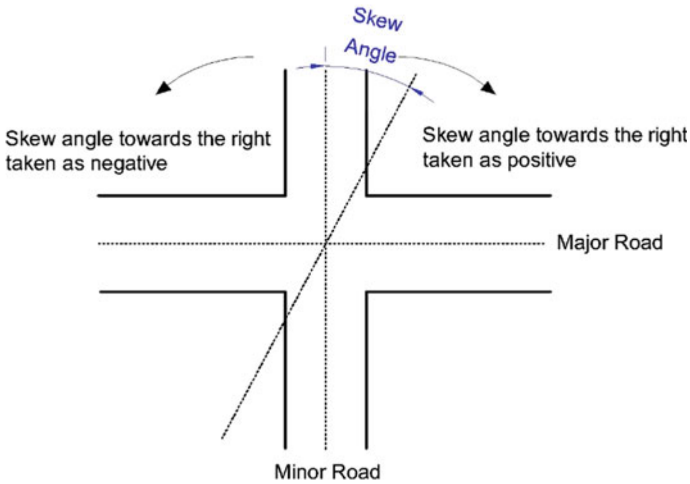


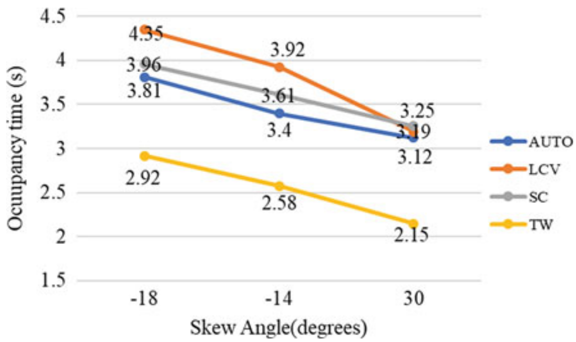
Fig. 2 Sign convention followed for measuring skew angle

3.1 Variation of Occupancy Time and Critical Gap with Skew Angle

In this section, the variation of occupancy time with skew angle is explored. Figure 3 shows the variation of occupancy time for right-turn movements from major road with the skew angle of the intersection. Figures 3 and 4 show the variation of occupancy time for right turns from major roads and minor roads, respectively. Occupancy time for right turns from major road decreases as the skew angle increases because the distance to be covered across the intersection decreases. The opposite is true for right turns from minor roads as occupancy time increases with increasing skew angle due to the increasing distance to be traversed across the intersection.

Figure 5 depicts the variation of occupancy time for through movement on minor roads. In this case, sign conventions are ignored because the distance to be traversed

Fig. 3 Variation of occupancy time for right-turn movements from major road with respect to skew angle



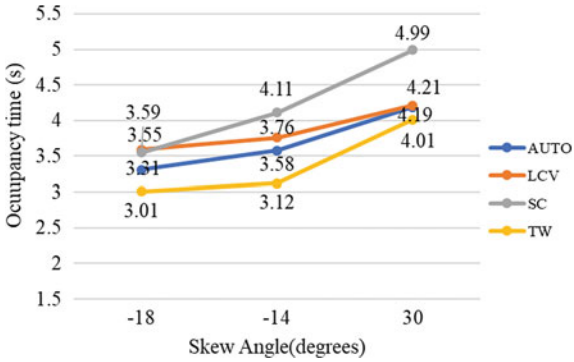


Fig. 4 Variation of occupancy time for right turn from minor movements with respect to skew angle

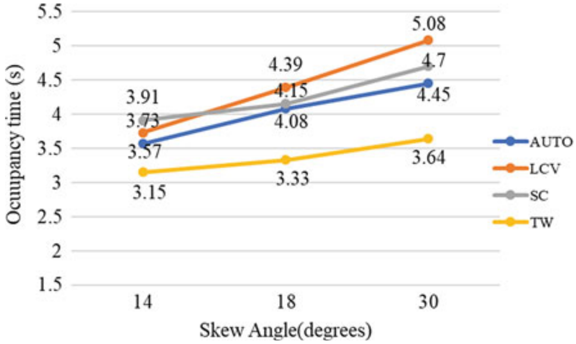


Fig. 5 Variation of occupancy time for through on minor road movements with respect to deviation angle

across the intersection increases with an increase in skew angle irrespective of direction. Hence, it can be inferred that occupancy times also increase when the skew angle increases. It can also be inferred from Figs. 3, 4, and 5 that two-wheelers have a significantly lower occupancy time compared to other modes due to the ease of maneuverability of the vehicle.

Figures 6 and 7 show the variation of critical gaps for right turns from major roads and minor roads, respectively. Critical gaps for right turns from major road

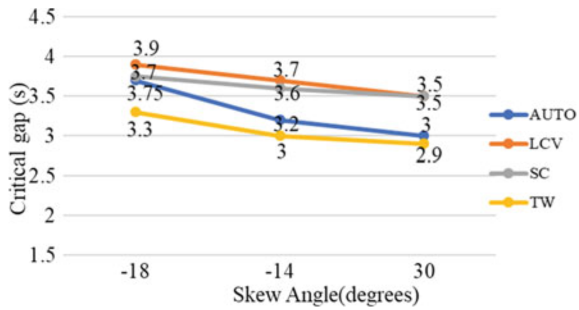


Fig. 6 Variation of critical gap for right turn from major road movements with respect to skew angle

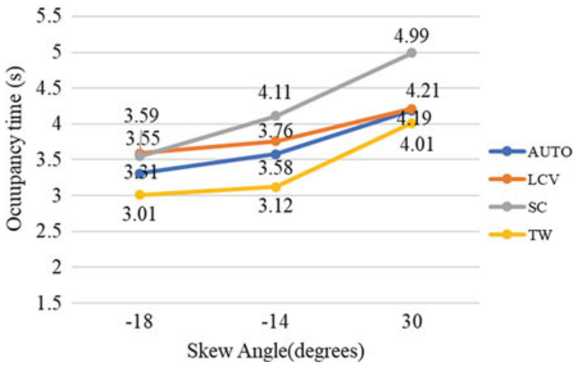


Fig. 7 Variation of critical gap for right turn from minor road movements with respect to skew angle

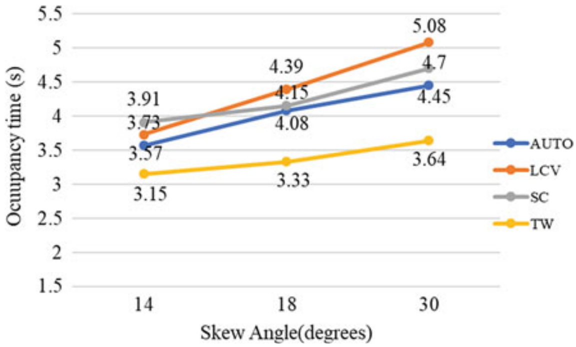


Fig. 8 Variation of critical gap for through on minor movements with respect to skew angle

and minor road show the same trend as observed with occupancy times. Critical gaps are lower for smaller vehicles such as two-wheelers as they are willing to accept smaller gaps in the traffic stream. Figure 8 shows the variation of critical gap for through movements on minor roads. It can be inferred from Figs. 6, 7, and 8 that vehicles accept larger gaps in the conflicting stream when the distance to be covered across the intersection is larger. This may be because larger distances force drivers to be more careful about the gap they are accepting in the conflicting streams.

4 Modeling of Occupancy Time

Linear regression models are developed to predict the occupancy times of different vehicle movements at the study intersections, considering the vehicle composition and skew angle as independent variables. The following sections detail out the models developed.

4.1 Right-Turn Movements from Major Road

The linear regression model for predicting the occupancy time of vehicles taking a right turn from the major road is given in Eq. (1).

$$\begin{aligned}\mathrm{OT} = 2.536 - 0.0143 \times (\text{Skew angle}) + 0.89 \\ \times X_{\mathrm{Auto}} + 1.27 \times X_{\mathrm{LCV}} + 1.06 \times X_{\mathrm{SC}}\end{aligned} \tag{1}$$

where

OT is occupancy time in seconds

Skew angle is the angle of the intersection in degrees

$X_{\mathrm{Auto}} = 1, X_{\mathrm{LCV}} = 0, X_{\mathrm{SC}} = 0$ if the vehicle considered is an auto-rickshaw

$X_{\mathrm{Auto}} = 1, X_{\mathrm{LCV}} = 0, X_{\mathrm{SC}} = 0$ if the vehicle considered is an auto-rickshaw

$XX_{\mathrm{Auto}} = 0, X_{\mathrm{LCV}} = 1, {}_{\mathrm{sc}} = 0$ if the vehicle considered is a light commercial vehicle

$X_{\mathrm{Auto}} = 0, X_{\mathrm{LCV}} = 0, X_{\mathrm{sc}} = 1$ if the vehicle considered is a small car

$X_{\mathrm{Auto}} = 0, X_{\mathrm{LCV}} = 0, X_{\mathrm{sc}} = 0$ if the vehicle considered is a two-wheeler.

The model has an adjusted R-squared value of 0.88. The p-values of all the coefficients are less than 0.05, implying that coefficients in the model are statistically different from zero, at 95% confidence level. It is inferred from the model that occupancy time decreases when the skew angle is toward the right (positive) and increases when it is toward the left (negative). This is in line with the previous observation that occupancy time for this movement decreases due to the decreased distance to be traversed across the intersection. Occupancy times increase with the size of the vehicle which is depicted by the coefficients of X_{Auto}, X_{LCV}, and $X_{\mathrm{SC.}}$

4.2 Right-Turn Movements from Minor Road

The linear regression model for predicting the occupancy time of vehicles that take a right turn from the minor road is given below.

$$\begin{aligned}\mathrm{OT} = 3.522 + 0.0178 \times (\text{Skew angle}) + 0.243 \\ \times X_{\mathrm{Auto}} + 0.283 \times X_{\mathrm{LCV}} + 0.707 \times X_{\mathrm{SC}}\end{aligned} \tag{2}$$

where

OT is occupancy time in seconds

Skew angle is the angle of the intersection in degrees

$X_{\mathrm{Auto}} = 1, X_{\mathrm{LCV}} = 0, X_{\mathrm{TW}} = 0$ if the vehicle considered is an auto-rickshaw

$X_{\mathrm{Auto}} = 0, X_{\mathrm{LCV}} = 1, X_{\mathrm{TW}} = 0$ if the vehicle considered is a light commercial vehicle

$X_{\text{Auto}} = 0$, $X_{\text{LCV}} = 0$, $X_{\text{TW}} = 1$ if the vehicle considered is a two-wheeler
$X_{\text{Auto}} = 0$, $X_{\text{LCV}} = 0$, $X_{\text{TW}} = 0$ if the vehicle considered is a small car.

The model has an adjusted *R*-squared value of 0.78. The *P*-values of the coefficients, except that for X_{LCV} and X_{Auto}, are less than 0.05, implying that coefficients in the model are statistically different from zero, at 95% confidence level. It is found that the p-value of X_{LCV} and X_{Auto} is not below 0.05, due to the smaller number of observations in the dataset. From the model, it can be inferred that occupancy time increases when the skew angle is toward the right (positive) and decreases when it is toward the left (negative). This is the trend observed earlier that occupancy times increase with increase in skew angle as the distance to be traversed increases. Occupancy times decrease when the size of the vehicle under consideration is smaller. Two-wheelers show the greater decrease in occupancy time for the same skew angle as depicted by the coefficient of X_{TW}.

4.3 Through Movements from Minor Roads

The linear regression model for predicting the occupancy time of through movements from the minor road is given in Eq. (3). In this model, the sign of skew angle is ignored as mentioned earlier.

$$\begin{aligned} \text{OT} = 2.53 + 0.040 \times (\text{Skew Angle}) + 0.66 \\ \times X_{\text{Auto}} + 1.03 \times X_{\text{LCV}} + 0.86 \times X_{\text{SC}} \end{aligned} \tag{3}$$

where
OT is occupancy time in seconds
Skew angle is the angle of the intersection in degrees (ignoring sign conventions)
$X_{\text{Auto}} = 1$, $X_{\text{LCV}} = 0$, $X_{\text{SC}} = 0$ if the vehicle considered is an auto-rickshaw
$X_{\text{Auto}} = 0$, ${}_{\text{LCV}} = 1$, $X_{\text{sc}} = 0$ if the vehicle considered is a light commercial vehicle
$X_{\text{Auto}} = 0$, $X_{\text{LCV}} = 0$, $X_{\text{sc}} = 1$ if the vehicle considered is a small car
$X_{\text{Auto}} = 0$, $X_{\text{LCV}} = 0$, $X_{\text{sc}} = 0$ if the vehicle considered is a two-wheeler.

The model has an adjusted R-squared value of 0.63. The *p*-values of the coefficients are less than 0.05, except that for X_{Auto}. This implies that all the coefficients, except that for X_{AUTO}, are statistically significant at 95% confidence level. From the model, it can be inferred that occupancy time increases with an increase in skew angle irrespective of direction. Hence, in this case, the lowest occupancy times will be recorded when the skew angle is 0. Any deviation from this values implies that the distance to be traversed across the intersection increases, and hence, vehicles will take a longer time to clear the intersection. Occupancy times increase when the size of the vehicle under consideration increases. Two-wheelers show the greater decrease in occupancy time for the same skew angle as depicted by the coefficient of X_{TW}.

5 Modeling of Critical Gap

5.1 Right-Turn Movements from Major Roads

The linear regression model for occupancy time of right turn from the major road is as follows.

$$\begin{aligned} \mathrm{CG} = 3.0622 - 0.0068 \times (\text{Skew Angle}) + 0.233 \\ \times X_{\text{Auto}} + 0.633 \times X_{\text{LCV}} + 0.55 \times X_{\text{SC}} \end{aligned} \tag{4}$$

where
CG is critical gap in seconds
Skew angle is the angle of the intersection in degrees
$X_{\text{Auto}} = 1, X_{\text{LCV}} = 0, X_{\text{SC}} = 0$ if the vehicle considered is an auto-rickshaw
$X_{\text{Auto}} = 0, X_{\text{LCV}} = 1, X_{\text{sc}} = 0$ if the vehicle considered is a light commercial vehicle
$X_{\text{Auto}} = 0, X_{\text{LCV}} = 0, X_{\text{sc}} = 1$ if the vehicle considered is a small car
$X_{\text{Auto}} = 0, X_{\text{LCV}} = 0, X_{\text{sc}} = 0$ if the vehicle considered is a two-wheeler.

The model has an adjusted *R*-squared value of 0.74. The *P*-values of all the variables are given in the table below. All the variables except that of X_{Auto} are found to be lower than 0.05. The higher *P*-value for auto-rickshaws may be due to the smaller sample size that was obtained for auto-rickshaws in this study. From the model, it can be inferred that critical gap decreases when the skew angle is toward the right (positive) and increases when it is toward the left (negative). Since the distance to be traversed across the intersection decreases with increasing skew angle, drivers will be able to accept smaller gaps to move across the intersection, and hence, the trend of decreasing critical gaps with increasing skew angles is observed. Occupancy times increase with the size of the vehicle which is depicted by the coefficients of X_{Auto}, X_{LCV}, and X_{SC}.

5.2 Right Turn from Minor Roads

The linear regression model for occupancy time of right turns from the minor road is as follows.

$$\begin{aligned} \mathrm{CG} = 3.04 + 0.01573 \times (\text{Skew Angle}) + 0.95 \\ \times X_{\text{Auto}} + 0.9 \times X_{\text{LCV}} + 0.8 \times X_{\text{SC}} \end{aligned} \tag{5}$$

where

CG is critical gap in seconds

Skew angle is the angle of the intersection in degrees

$X_{\text{Auto}} = 1$, $X_{\text{LCV}} = 0$, $X_{\text{TW}} = 0$ if the vehicle considered is an auto-rickshaw

$X_{\text{Auto}} = 0$, $X_{\text{LCV}} = 1$, $X_{\text{TW}} = 0$ if the vehicle considered is a light commercial vehicle

$X_{\text{Auto}} = 0$, $X_{\text{LCV}} = 0$, $X_{\text{TW}} = 1$ if the vehicle considered is a two-wheeler

$X_{\text{Auto}} = 0$, $X_{\text{LCV}} = 0$, $X_{\text{TW}} = 0$ if the vehicle considered is a small car.

The model has an adjusted *R*-squared value of 0.96. The *P*-values of all the variables are given in the table below. Since the *p*-values of all variables are below 0.05, all the variables are treated as significant. From the model, it can be inferred that critical gap increases when the skew angle is toward the right (positive) and decreases when it is toward the left (negative). Critical gap timings increase with increase in skew angle because the distance to be traversed across the intersection increases in this case. Hence, the gaps that vehicles are ready to accept would be larger. Critical gap timings decrease when the size of the vehicle under consideration is smaller. Two-wheelers show the greater decrease in occupancy time for the same skew angle as depicted by the coefficient of $X_{\text{TW.}}$

5.3 Through Movements on Minor Roads

The linear regression model for occupancy time of through movement from the minor road is as follows. In this model, the sign convention relating to skew angle is ignored as mentioned in previous sections.

$$\begin{aligned} \text{CG} = 2.3112 + 0.0591 \times (\text{Skew angle}) + 0.7 \\ \times X_{\text{Auto}} + 0.6 \times X_{\text{LCV}} + 1.117 \times X_{\text{SC}} \end{aligned} \tag{6}$$

where

CG is critical gap in seconds

Skew angle is the angle of the intersection in degrees (ignoring sign convention)

$X_{\text{Auto}} = 1$, $X_{\text{LCV}} = 0$, $X_{\text{SC}} = 0$ if the vehicle considered is an auto-rickshaw

$X_{\text{Auto}} = 0$, $X_{\text{LCV}} = 1$, $X_{\text{sc}} = 0$ if the vehicle considered is a light commercial vehicle

$X_{\text{Auto}} = 0$, $X_{\text{LCV}} = 0$, $X_{\text{sc}} = 1$ if the vehicle considered is a small car

$X_{\text{Auto}} = 0$, $X_{\text{LCV}} = 0$, $X_{\text{sc}} = 0$ if the vehicle considered is a two-wheeler.

The model has an adjusted R-squared value of 0.72. The *P*-values of all the variables are given in the table below. The *P*-value of all the variables is below 0.05 except for XLCV, which implies that they are statistically different from zero at 95% level of significance level. The distance to be traversed increases with increased skew angle. Hence, this leads to vehicles accepting larger gaps because the distance to be covered is larger. Similar to the previous cases, critical gap is dependent on the size of the vehicle with LCVs having the largest coefficient values and hence generating the larger values of critical gaps.

6 Conclusions

For this study, three intersections with varying skew angle were selected, and data relating to gap parameters for various vehicles were collected. The collected data were analyzed, and models that predict the critical gap and occupancy time of the different vehicle classes for intersections with varying skew angles were formulated. The following conclusions were drawn from the study:

1. Distance to be traversed across the intersection decreases with increase in skew angle in the case of right turn from major road movements and decreases with increase in skew angle in the case of right turn from minor road movements.
2. For through movements on minor roads, the distance to be traversed across the intersection increase with respect to skew irrespective of direction of skew.
3. Occupancy time increases when the distance to be traversed across the intersection increases, as expected. This is explained simply by understanding that vehicles will take longer times to cover larger distances.
4. Critical gap increases when the distance to be traversed across the intersection increases. When the distance to be crossed is larger, drivers tend to show a more careful behavior and only accept larger gaps that they feel would be sufficient to cover the increased distance.
5. Critical gap for right turn movements from major road decreases with increase in skew angle. Critical gap for right-turn movements from minor road and through movements on minor road increases with increase in skew angle.
6. Both critical gap and occupancy times are dependent on the size of the vehicle. Smaller vehicles are able to cross the intersection more quickly and will be able to accept smaller gaps in the conflicting stream.

To further explore this area or to obtain more clarity in this area, other skewed intersections can be considered in future works to obtain a more refined model. Intersections with varying road widths and other geometric factors can be considered.

References

1. CSIR-Central Road Research Institute (2017) Unsignalised intersections, Indo-HCM, Chapter 8
2. Chandra S, Mohan M, Gates TJ (2014) Estimation of critical gap using intersection occupancy time. In: 19th international conference of Hong Kong society for transportation studies
3. Chandra S, Mohan M (2018) Analysis of driver behavior at unsignalized intersections. J Indian Road Congr 79–2(675)
4. Chandra S, Mohan M (2020) Capacity estimation of unsignalized intersections under heterogeneous traffic conditions. Can J Civ Eng 47(6):651–662
5. Ashalatha R, Chandra S (2011) Critical gap through clearing behavior of drivers at unsignalized intersections. KSCE J Civ Eng 15(8):1427–1434
6. Asaithambi G, Anuroop C (2016) Analysis of occupation time of vehicles at urban unsignalised intersections in non-lane-based mixed traffic conditions. J Mod Transp

Multi-Branch Traffic Flow Prediction Based on Temporal Speed

Nisha and Kranti Kumar

Abstract Efficient traffic management is a major issue for developing countries. Traffic flow prediction is an important problem in intelligent transportation system (ITS). Various studies have been reported in the literature for traffic flow prediction in which combined models have been proposed only using traffic flow data. It is evident from the traffic flow theory that speed and flow are inter-related. Therefore, considering speed to predict flow in a model can help in improving the performance of a prediction method. Keeping this in mind, we propose a traffic flow prediction method consisting of two branches. First branch predicts traffic flow using past flow data through long short-term memory (LSTM) neural network. Second branch predicts volume using Gaussian process regression (GPR) based on temporal speed. Finally, prediction from both the branches was combined through weighted average. The mean squared error (MSE), root mean squared error (RMSE), and Pearson's correlation coefficient (r) were used to evaluate the effectiveness of the proposed model. Based on these measures, it was found that results of our proposed model are promising.

Keywords Multi-branch · GPR · Traffic flow · LSTM

1 Introduction

Transportation systems are among the resources that are vital to the economic and social welfare as well as the safety of a nation. The versatility and effectiveness of the transportation system contribute significantly to the vitality and competitiveness of the economy of a nation. In recent years, there has been a continuous rise in traffic demand on the roads resulting in congestion which causes the idling of engines, leads to more energy consumption, and causes pollution [1]. Air pollution adversely affecting human health has become a prominent problem nowadays due to the high density of various type of vehicles on road [2]. It has been noticed that vehicle

Nisha · K. Kumar (✉)
School of Liberal Studies, Dr. B. R. Ambedkar University Delhi, Delhi 110006, India
e-mail: kranti@aud.ac.in

L. Devi et al. (eds.), *Proceedings of the Sixth International Conference of Transportation Research Group of India*, Lecture Notes in Civil Engineering 272,
https://doi.org/10.1007/978-981-19-3494-0_4

emissions led to high pollution levels especially during peak time [3]. As far as India is concerned, being the second most populated country is facing extreme traffic conditions such as congestion, noise and air pollution, and traffic fatalities on roads. Indian government has spent a huge amount in the urban infrastructure section such as developing public transport systems like bus rapid transit and making metro services available for more new places. But still, the use of personal vehicles has increased in past years [4].

Infrastructure development, implementation of traffic control guidance, congestion pricing, and utilization of advanced technologies coming under ITS are some important steps to alleviate traffic problems. Amid various options, sustainable solutions like ITS are mainly chosen than capital-intensive development plans. Intelligent transport system makes use of modern information and advanced technologies to solve various transportation problems arising in transportation system like level of congestion, traffic fatalities, air and noise pollution, etc. ITS has emerged as an effective way to improve strategies for traffic congestion and to manage traffic demand on roads. Traffic state prediction which uses information collected from various devices is a significant issue in the framework of ITS. It helps in traffic congestion prevention in time and accomplishes intelligent traffic reminder and evacuation management. Short-term traffic forecasting directly predicts traffic conditions expected at a future time and provides continuous short-term traffic information feedback. Several modeling attempts have been made to deal with the problem of short-term forecasting in last decade. Traffic flow is highly non-linear and stochastic in nature which makes its prediction difficult. Due to this, machine learning-based methods have been applied extensively to predict traffic flow because of their capability to approximate any function regardless of its degree of non-linearity. Recently, the literature has shown that combined methods are most effective than a single prediction method for short-term forecasting especially combining traditional methods with machine learning methods as they combine the capability of different methods [5, 6].

Although several studies have been reported in the literature to deal with traffic flow prediction problems, still some gaps exist.

(1) Research on combining different methods that helps in studying different aspects of traffic flow prediction is relatively limited.
(2) Most existing combined methods make use of only historical flow to predict traffic flow, and it is evident from traffic flow theory that speed and flow are inter-related and hence considering historical temporal speed in these methods can have an influence on traffic flow prediction in different ways.
(3) Traffic flow is highly non-linear and stochastic in nature and hence leads to uncertainties in prediction. Therefore, providing an interval of prediction rather than a single value is more realistic and beneficial.

The main objectives of this paper are to alleviate abovementioned research gaps. In this paper, we propose (1) a multi-branch method which uses historical temporal speed and flow information for traffic flow forecasting. As historical speed and flow influence flow prediction, the method proposed in this paper uses separate branches to predict flow, one using historical flow and other using temporal speed with the help of GPR. (2) The choice of a regression tool such as GPR enables to directly capture the complex relationships between variables and uncertainty of the model. It gives an interval of prediction rather than a single value which is more realistic in this scenario. (3) This paper also presents a novel way of merging the prediction from two branches using their correlation coefficient as weights.

The rest of this paper is organized as follows: In section two, the studies found in literature dealing with traffic flow and speed prediction are reviewed. Section three describes the methodology of the multi-branch method in detail. Data description and experiment results are presented in section four. Finally, a conclusion is drawn based on the results in section five.

2 Related Work

In this section, a literature review of methods applied to predict traffic flow is presented. Mainly, these methods are classified into following categories: statistical, machine learning (ML) based, and hybrid methods.

Statistical methods are widely applied to traffic prediction due to their good theoretical interpretation ability and visible computational structure. Some of the common statistical methods include linear regression [7], autoregressive integrated moving average (ARIMA) [8–10], Kalman filter [11], and generalized autoregressive conditional heteroscedasticity (GARCH) [12]. ARIMA is one of the most widely utilized statistical methods for predicting traffic flow. Many experiments proved that ARIMA and its variants have been very useful in stochastic traffic prediction. These methods are helpful when traffic varies temporarily or regularly but fails to capture high complex traffic variations.

ML-based methods have received great attention in traffic flow forecasting due to their strong ability to capture the non-linear and stochastic relationships of traffic flow as well as adaptability. To overcome the problem associated with statistical methods, numerous studies have used ML-based methods such as k-nearest neighborhoods (KNN), kernel based, and artificial neural network (ANN). Among them, ANN are most widely used. Chang et al. [13] applied KNN to develop a dynamic multi-variate traffic flow prediction model whereas Tan et al. [14] showed that KNN is superior to traditional statistical methods. Kernel-based methods such as support vector machines (SVM) and the Gaussian processes (GP) model have gained special attention in short-term traffic flow forecasting and other problem of time series analysis because of their outstanding capability to generalize and perform non-linear approximation. GP has been shown to be a powerful tool in exploring implicit relationships and addressing difficult non-linear regression problems in the field of traffic volume prediction [15–17]. Comparative studies have demonstrated GP superiority over other methods on short-term traffic prediction [18]. Theja and Vanajakskshi [19] used SVM to predict traffic volume under heterogeneous traffic conditions and compared the result with multi-layer feed forward neural network. Wang and shi [20] highlighted that selection of appropriate kernel in SVM is an important and challenging task. Smith and Demetsky (1994), Ledoux (1997), and Dougherty and cobbett (1997) provided some early applications of ANN in traffic flow forecasting [21–23]. Afterward, several researchers employed ANN in their studies such as RBFNN [24], GRNN [25], FNN [26], and WNN [27]. Recently, with advancement of technology and availability of large amount of data, deep learning methods include LSTM, gated recurrent unit (GRU), and convolutional neural network (CNN) and among them, LSTM is mostly employed to study temporal characteristics of the time series due to its excellent predictive power and ability to capture long-term dependency. Tian et al. [28] provided early application of LSTM to predict traffic flow. Fu et al. [29] introduced GRU and compared it with LSTM and showed LSTM-outperformed GRU. However, these models are not much effective in studying spatial dependencies. To overcome this weakness, researchers implemented CNN-based models [30].

Hybrid methods combine different methods and hence the capabilities of these methods to improve prediction accuracy. A combined model is shown to be more accurate than a single predicting short-term traffic forecasting [5]. Mainly statistical methods are combined with machine learning-based methods for better results. Zheng et al. [31] combined the probability theory with ANN, and Yangchong et al. [32] combined statistical methods with SVM to improve the results. A hybrid LSTM model was developed to study traffic conditions' impact on the predictions and showed that the prediction error for road section and intersection were much closer proving the efficiency of model under different traffic conditions [33]. LSTM is well known for its ability to capture the temporal characteristic of the variables but cannot capture spatial features of the variables, and hence to learn spatial–temporal of the variables, different hybrid methods were suggested such as incorporating sparse encoder, auto-encoder, and RBFNN to capture spatial features with LSTM to predict traffic flow [34–36]. Liu et al. [6] integrated ARIMA and LSTM into a combined method to capture linear and non-linear features in time series to predict traffic flow. The predicted results from both methods were merged through dynamical weightage to obtain final traffic flow in the short term. As studies involving combined methods are interesting as well as developing, this paper presents a combined method called multi-branch method to help in enhancing traffic flow forecasting strategies.

3 Methodology

In this section, first concepts of methods used to develop the proposed methodology are discussed in detail. Finally, the proposed structure of combined method is introduced.

3.1 Long Short-Term Memory Neural Network

Long short-term memory (LSTM) neural network proposed by Hochreiter and Schmidhuber in 1997 is a special form of recurrent neural networks (RNN) which is capable of capturing non-linear behaviors of time series problems. It was developed to solve the problem of vanishing gradient in RNN during training as it incorporates gating functions into their state dynamics.

LSTM model consists of three layers: input layer, recurrent layer, and output layer. The recurrent layer's basic unit is called memory block which consists of four main cells: forget gate (f_t), input gate (i_t), output gate (o_t), and memory cell to read, update, and delete the operations of a cell. The features of the input are reflected by state of memory cell as shown in Fig. 1.

Let x, h and c represent the input vectors, hypothesis, and cell state, respectively.

$$i_t = \sigma\left(W_i.\left[h_{t-1}, x_t\right]\right) + b_x \tag{1}$$

$$f_t = \sigma\left(W_f.\left[h_{t-1}, x_t\right]\right) + b_f \tag{2}$$

$$o_t = \sigma\left(W_o.\left[h_{t-1}, x_t\right]\right) + b_o \tag{3}$$

$$c_t = f_t * c_{t-1} + i_t * \tan\, h\left(W_c.\left[h_{t-1}, x_t\right]\right) + b_c \tag{4}$$

$$h_t = o_t * \tanh(c_t) \tag{5}$$

where $*$ refers to Hadamard products; W is weight matrix; and σ is the sigmoid function. In this paper, branch 1 used LSTM neural network to predict traffic flow which is then merged with traffic flow from branch 2.

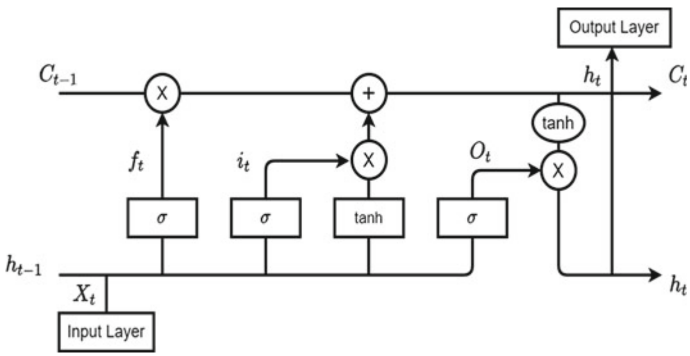


Fig. 1 LSTM neural network structure

3.2 Gaussian Process Regression

Gaussian process regression (GPR) is a kernel-based approach, getting high attention in the area of ML. Gaussian process (GP) models are useful in exploring the implicit relations between set of variables based on training dataset and are based on Bayesian framework which allow them to address difficult non-linear regression and probabilities interpretation of model outputs.

GPs unlike other Bayesian models elicit posterior predictive distribution for new test inputs. A GP is a collection of random variables $f(X)$ that are indexed by a finite $X = [x_1, x_2, \ldots\ldots, x_n]^T$ which have joint Gaussian distribution with mean function $u(X)$ and covariance function $K\left(X, X'\right)$ such that

$$\begin{pmatrix} f(x_1) \\ \vdots \\ f(x_n) \end{pmatrix} \sim N\left(\begin{bmatrix} u(x_1) \\ \vdots \\ u(x_n) \end{bmatrix}, \begin{bmatrix} K(x_1, x_1) & \cdots & K(x_1, x_n) \\ \vdots & \ddots & \vdots \\ K(x_n, x_1) & \cdots & K(x_n, x_n) \end{bmatrix}\right) \tag{6}$$

This can also be written as:

$$f(.) = GP(u(.), K(., .)) \tag{7}$$

where the prior GP is fully specified by a mean function u and covariance function K.

Consider the training data $\{(x_i, y_i)\}$ for $i = 1, 2, \ldots, n$ from dataset D. GPR assume that response variable y_i satisfies

$$y_i = f(x_i) + e_i, \quad e_i \overset{iid}{\sim} N\left(0, \sigma^2\right) \tag{8}$$

Let $\left\{\left(\{x_i\}^*, \{y_i\}^*\right)\right\}$ for $i = 1, 2, \ldots, m$ be the test data drawn from D.

The objective is to determine predictive distribution of $Y*$ which will be multivariate normal, i.e

$$Y^*|X^*, X \sim N\left(u^*, \sum\nolimits^*\right) \tag{9}$$

with, $u^* = K(X^*, X)\left(K(X, X) + \sigma^2 I\right)^{-1} Y$

$$\sum\nolimits^* = \mathrm{K}\left(\mathrm{X}^*, \mathrm{X}^*\right) + \sigma^2\mathrm{I} - \mathrm{K}\left(\mathrm{X}^*, \mathrm{X}\right)\left(\mathrm{K}(\mathrm{X}, \mathrm{X}) + \sigma^2\mathrm{I}\right)^{-1}\mathrm{K}\left(\mathrm{X}, \mathrm{X}^*\right)$$

where $K(X, X)$ is the $n \times n$ matrix with $Kij = K\left(x_i, x_j\right)$, $K(X^*, X^*)$ is the $m \times m$ matrix with $K_{jj} = K\left(x^j, x^j\right)$, $Y^* = \left[y_1^*, \ldots., y_m^*\right]$ and $X^* = \left[x_1^*, \ldots., x_m^*\right]$.

As traffic flow and speed is expected to have some complex relationship, it would be hard to model this relation using simple parametric models such as linear or polynomial functions. Thus, in this paper, GPR is used to study relation between them as GPR has the ability to fit arbitrary-shaped curves due to its nonparametric nature. GPR enables to directly capture the complex relationships between variables, uncertainty of the model, and an interval of prediction rather than a single value which is more realistic in this scenario.

3.3 Multi-Branch Method

Considering the fact that speed and flow are inter-related and have influence on flow prediction in a different way, the proposed method consists of two branches. First branch predicts traffic flow using past traffic flow data through LSTM neural network. Second branch predicts flow using GPR based on temporal speed. Finally, prediction from both the branches are merged through weighted average (WA). Structure of multi-branch method is illustrated in the Fig. 2.

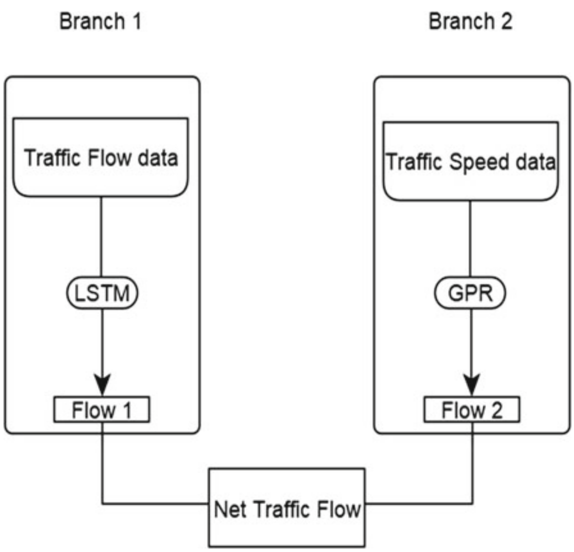


Fig. 2 Structure of the proposed multibranch method

3.4 Weighted Average

The predicted net flow (f) from both the branches is merged using below mentioned weighting:

$$f = w_1 \times f_1 + w_2 \times f_2 \tag{10}$$

where f_1 and f_2 are predicted flow from branch 1 and branch 2, respectively, w is the weight corresponding to each branch such that

$$w_1 = \frac{r_1}{r_1 + r_2} \tag{11}$$

$$w_2 = \frac{r_2}{r_1 + r_2} \tag{12}$$

where w_1 and w_2 are weights associated with results from branch 1 and branch 2, and r_1 and r_2 are correlation coefficient corresponding to predicted flow of branch 1 and branch 2, respectively.

4 Experiments and Results

4.1 Data Description

In this paper, an arterial route called Inner Ring Road, in South Extension II (28∘34′ 7.41″, 77∘13′ 3.38″) in Delhi city, India, was considered as the study stretch (see Fig. 3). It is a three-lane route and traffic run in both the direction from ISBT Azadpur to Naraina and vice versa. South Extension is known to be one of Delhi's most high-end and premier markets. It attracts significant crowd from across various parts of Delhi NCR and hence frequent traffic throughout the day. Traffic data were collected from a foot-over bridge using Handy-cam from 8 to 11 am and 4 pm to 7 pm consecutively for five days in the month of February 2020.

The traffic data were aggregated every 5 min, and 675 datasets were obtained in total for a period of five days. Seven vehicle classes were considered, namely Car/Jeep/Van, 3-Wheeler, 2-Wheeler, Light commercial vehicle, Truck, Bus, and Bicycle/Rickshaw to account for vehicle heterogeneity. Traffic volume for each class of vehicles was manually extracted at 5-min intervals using the collected video data whereas class-wise speed was calculated in both directions by operating on time difference values (T1-T2) and trap length (15 m), i.e., two points were marked at a distance of 15 m and the average time taken by vehicle of each class to travel between these points was measured by time displayed on the video screen with an accuracy of 0.01 s. The observed class-wise traffic volume was converted to equivalent passenger

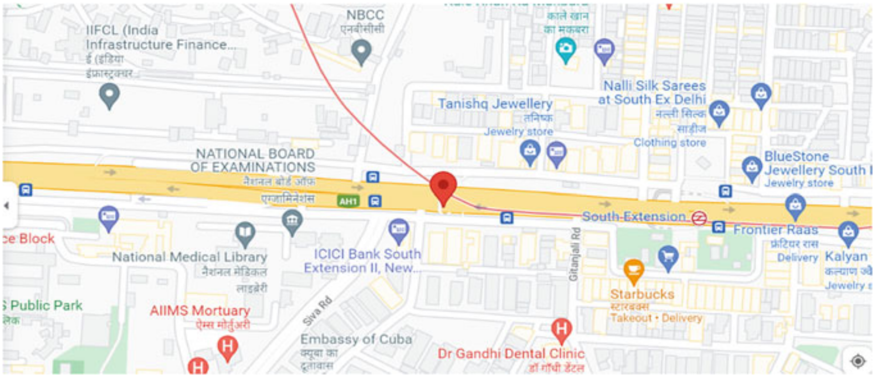


Fig. 3 Location of study stretch on Google Maps

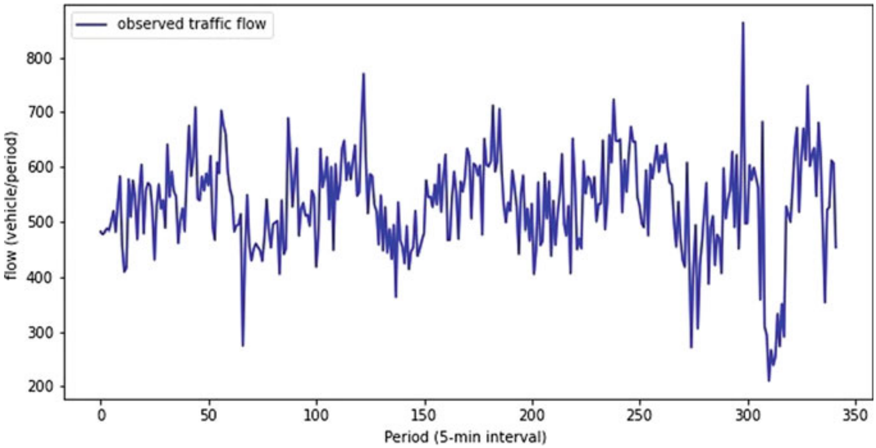


Fig. 4 Variation in traffic flow over time from Naraina to Azadpur

car units (PCU) using Indian Highway Capacity Manual (Indo-HCM) [37]. Both side traffic was considered in this study. The variation in collected traffic flow data over time for one side is shown in Fig. 4.

4.2 Evaluation Metrics

The proposed method was evaluated on the basis of three measures of effectiveness: MSE, RMSE, and Pearson's correlation coefficient (r) described as:

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2 \tag{13}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2} \tag{14}$$

$$r\left(y_i, \hat{y}_i\right) = \frac{n\sum_{i=1}^{n} y_i \times \hat{y}_i - \sum_{i=1}^{n} y_i \times \sum_{i=1}^{n} \hat{y}_i}{\left[n\sum_{i=1}^{n} y_i^2 - \left(\sum_{i=1}^{n} y_i\right)^2\right]\left[n\sum_{i=1}^{n} \widehat{y_i}^2 - \left(\sum_{i=1}^{n} \widehat{y_i}\right)^2\right]} \tag{15}$$

where y_i is the predicted flow and $\hat{y}_i$ is the actual flow.

4.3 Experiment Design and Results

The first four days of collected data from both sides were taken for training, last day's each hour's first 30 min data were taken for validation, and last 30 min data were taken for testing purpose. Since the current flow depends on the flow at previous time steps, for branch 1, historical flow with different window sizes (previous time steps) was taken as input. For branch 1, different architectures with one, two, and three LSTM layers and hidden neurons (HN) were compared to determine the one with highest prediction accuracy as shown in the Table 1.

From Table 1, architectures with two LSTM layers were found more efficient and thus analyzed further to find the best architecture based on prediction accuracy named as Train 1, Train 2, and Train 3 as shown in the Table 2.

A LSTM neural network consisting of 128 neurons in first layer and 64 neurons in second layer with 250 epochs was selected (see Tables 2 and 3). Dropout layer with dropout probability 0.2 was used while training. The whole network was trained

Table 1 Details of prediction results on different number of LSTM layers

Number of LSTM layers [Hidden neurons]	MSE	RMSE	r
1 [64]	1.031	1.015	0.474
1 [128]	1.449	1.203	0.407
1 [256]	1.425	1.193	0.377
2 [64, 32]	1.095	1.046	0.592
2 [128, 64]	**0.827**	**0.909**	**0.657**
2 [256, 128]	1.046	1.022	0.530
3 [128, 64, 32]	1.147	1.070	0.480
3 [256, 128, 64]	1.397	1.819	0.364

Bold represents the best LSTM architecture with two layers based on results

Table 2 Branch 1 network with two LSTM layers and different hidden neurons (HN)

Trial No	HN in LSTM layer 1	HN in LSTM layer 2
Train 1	64	32
Train2	128	64
Train 3	256	128

Table 3 Details of experiment results with different architectures and epochs

Epoch		Training			Testing		
		MSE	RMSE	*r*	MSE	RMSE	*r*
200	Train 1	0.094	0.307	0.946	1.031	1.015	0.601
	Train 2	0.228	0.477	0.854	1.310	1.145	0.476
	Train 3	0.105	0.324	0.936	1.047	1.023	0.542
250	Train 1	0.058	0.241	0.965	1.095	1.046	0.592
	Train 2	**0.013**	**0.112**	**0.992**	**0.827**	**0.910**	**0.657**
	Train 3	0.009	0.096	0.995	1.046	1.337	0.530
300	Train 1	0.031	0.175	0.983	1.020	1.010	0.595
	Train 2	0.010	0.099	0.995	0.911	0.954	0.619
	Train 3	0.006	0.078	0.996	0.959	0.979	0.587

Bold represents the best results with Train 2 and 250 epochs

with learning rate of 0.001. Adam optimizer was used to fine tune the network. Effect of different activation functions (tanh, ReLU, sigmoid) was also studied on selected LSTM architecture, and sigmoid function was used as activation function in dense layer based on prediction accuracy (see Table 4).

We used traffic flow at previous t (window) time steps as an input for the LSTM neural network to predict $(t + 1)$th time step flow. Window size was used as a hyper-parameter and window 10 was selected as input window for branch 1 as prediction accuracy tended to be flat after 10 window size as shown in the Table 5.

For branch 2, historical temporal speed with different window size was used as input feature to predict traffic flow; however, window 5 was chosen based on validation metric. GPR with different kernels was experimented with different parameters to obtain traffic flow. Product kernel which is a multiplication of Constant kernel with Dot-Product kernel was chosen and tuned during model selection based on validation accuracy. Constant kernel was used as a part of product kernel to scale down

Table 4 Effect of activation function on prediction results

Activation function	MSE	RMSE	*r*
tanh	1.034	1.016	0.519
ReLU	1.378	1.173	0.361
Sigmoid	**0.827**	**0.910**	**0.657**

Bold represents the best activation function based on results

Table 5 Branch 1 prediction result comparison for different input time windows

Window	MSE	RMSE	r
11	0.993	0.996	0.562
10	**0.872**	**0.933**	**0.657**
9	0.927	0.962	0.654
8	0.954	0.974	0.583
7	1.295	1.137	0.429

Bold represents the best input window based on results

the magnitude of the other kernel and is given by:

$$K\left(X, X'\right) = \text{constant value} \tag{16}$$

Dot-product kernel is parameterized by a parameter σ_0 which controls the inhomogeneity of the kernel and is defined as:

$$K\left(X, X'\right) = \sigma_0^2 + X.X' \tag{17}$$

After specifying the kernel functions and their hyperparameters, other parameters of the GP model were specified such as alpha that is the variance of the $i.i.d.$ (Independent and identically distributed) noise on the labels and normalize_y refers to the constant mean function which is 0 if sets to False and the mean of training data if sets to True. All these parameters were tuned on training data, and the best GP model was chosen based on validation accuracy. GP model returned not only the predicted flow but also the standard deviation for each prediction. This provided an interval for prediction rather than a single value. The 95% confidence interval can also be calculated which is 1.96 time the standard deviation for the predicted flow. All these experiments were carried out with the help of scikit-learn library in Python.

Next, the branch 1, branch 2, multi-branch equal weightage (EW) method that combine both the branches in equal ratio (average of both the branches), and multi-branch weighted average (WA) method results are compared as shown in Table 6.

Table 6 clearly shows that the multi-branch (WA) method is effective and better than individual branches. Multi-branch (EW) method has also shown its effectiveness proving that combined methods are more effective that individual methods. This also

Table 6 Details of comparison of proposed model with different methods

Methods	MSE	RMSE	r
Branch 1 (LSTM)	0.827	0.909	0.657
Branch 2 (GPR)	1.015	1.00	0.579
Multi-branch (EW)	0.803	0.896	0.702
Multi-branch (WA)	**0.797**	**0.892**	**0.703**

Bold represents the results of multi branch method

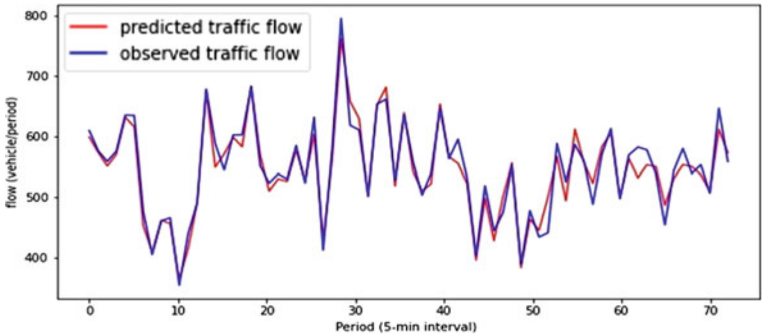


Fig. 5 Predicted results of the multi-branch method (WA)

proved that use of temporal speed in flow prediction can improve the performance of a method. Though branch 2 produced greater error than branch 1 but combining their results using EW and WA improved the results significantly. Figure 5 shows the comparison between the observed traffic flow and the predicted one using multi-branch method (WA).

5 Conclusion

This paper presents a combined method called multi-branch method consisting of two branches, where one branch used LSTM neural network and other applied GPR to predict traffic flow. Two branches complement each other, and model traffic flow in a different manner which provides promising results even on lesser amount of data. GPR was applied to obtain flow from speed which is a powerful tool to explore implicit relationship between variables. Traffic speed in a separate branch was used to obtain traffic flow keeping in mind that they both are interdependent and hence have influence on traffic flow prediction. The proposed method used a weighted average using correlation coefficient of prediction of the two branches which proved to be better than a simple equal weightage. Prediction results proved that the multi-branch (WA) method performed better than the LSTM NN, GPR, and multi-branch (EW) method. Study results also show that the proposed method performed satisfactory even with less amount of data under heterogeneous traffic conditions. Future study will concentrate on the collection of more datasets for longer time period and extended number of days. With the availability of more data, robustness of the proposed method can be tested and method can be generalized for diverse traffic conditions.

Acknowledgements "This study was supported by the University Grants Commission (UGC), New Delhi, India, through the start-up grant research project ***Modelling and simulation of vehicular***

traffic flow problems through the grant No. F.30-403/2017(BSR), which is thankfully acknowledged. Authors also convey their gratitude to anonymous reviewers for their valuable suggestions which significantly improved this article."

Conflict of Interest "Authors declare that they have no conflict of interest."

References

1. Mohan RA (2012) Measuring urban traffic congestion-a review. Int J Traffic Transp Eng 2(4):286–305
2. Patankar AM, Trivedi PL (2011) Monetary burden of health impacts of air pollution in Mumbai, India: implications for public health policy health policy. Publ Health 125(3):157–164
3. Choudhary A, Gokhale S (2016) Urban real-world driving traffic emissions during interruption and congestion. Transp Res Part D Transp Environ 43:59–70
4. Sharma B, Kumar S, Tiwari P, Yadav P, Nezhurina MI (2018) ANN based short-term traffic flow forecasting in undivided two-lane highway. J Big Data 5(1):1–16
5. Tselentis DI, Vlahogianni EI, Karlaftis MG (2015) Improving short-term traffic forecasts: to combine models or not to combine? IET Intel Transp Syst 9(2):193–201
6. Liu B, Tang X, Cheng J, Shi P (2020) Traffic flow combination forecasting method based on improved LSTM and ARIMA. Int J Embed Syst 12(1):22–30
7. Sun H, Liu HX, Xiao H, He RR, Ran B (2003) Use of local linear regression model for short-term traffic forecasting. Transp Res Rec 1836(1):143–150
8. Chandra SR, Al-Deek H (2009) Predictions of freeway traffic speeds and volumes using vector autoregressive models. J Intell Transp Syst 13(2):53–72
9. Williams BM (2001) Multivariate vehicular traffic flow prediction: evaluation of ARIMAX modeling. Transp Res Rec 1776(1):194–200
10. Kumar SV, Vanajakshi L (2015) Short-term traffic flow prediction using seasonal ARIMA model with limited input data. Eur Transp Res Rev 7(3):1–9
11. Xie Y, Zhang Y, Ye Z (2007) Short-term traffic volume forecasting using Kalman filter with discrete wavelet decomposition. Comput-Aided Civ Infrastruct Eng 22(5):326–334
12. Wang QP, Shu Q, Huang HG (2016) Study on GARCH effect in traffic flow data and prediction model. Comput Simul 33(2):194–197
13. Zhang L, Liu Q, Yang W, Wei N, Dong D (2013) An improved k-nearest neighbor model for short-term traffic flow prediction. Procedia Soc Behav Sci 96:653–662
14. Chang H, Lee Y, Yoon B, Baek S (2012) Dynamic near-term traffic flow prediction: system-oriented approach based on past experiences. IET Intel Transp Syst 6(3):292–305
15. Tan MC, Wong SC, Xu JM, Guan ZR, Zhang P (2009) An aggregation approach to short-term traffic flow prediction. IEEE Trans Intell Transp Syst 10(1):60–69
16. Zhao J, Sun S (2016) High-order Gaussian process dynamical models for traffic flow prediction. IEEE Trans Intell Transp Syst 17(7):2014–2019
17. Wang P, Kim Y, Vaci L, Yang H, Mihaylova L (2018) Short-term traffic prediction with vicinity Gaussian process in the presence of missing data. In: Sensor data fusion: trends, solutions, applications (SDF), IEEE, pp 1–6
18. KI Williams C (2006) Gaussian processes for machine learning. Taylor & Francis Group
19. Theja PVV, Vanajakshi L (2010) Short term prediction of traffic parameters using support vector machines technique. In: 3rd international conference on emerging trends in engineering and technology. IEEE, pp 70–75
20. Wang J, Shi Q (2013) Short-term traffic speed forecasting hybrid model based on chaos–wavelet analysis-support vector machine theory. Transp Res Part C: Emerg Technol 27:219–232

21. Smith BL, Demetsky MJ (1994) Short-term traffic flow prediction models-a comparison of neural network and nonparametric regression approaches. In: Proceedings of IEEE international conference on systems, man and cybernetics, vol. 2. IEEE, pp 1706–1709
22. Ledoux C (1997) An urban traffic flow model integrating neural networks. Transp Res Part C Emerg Technol 5(5):287–300
23. Dougherty MS, Cobbett MR (1997) Short-term inter-urban traffic forecasts using neural networks. Int J Forecast 13(1):21–31
24. Park B, Messer CJ, Urbanik T (1998) Short-term freeway traffic volume forecasting using radial basis function neural network. Transp Res Rec 1651(1):39–47
25. Leng Z, Gao J, Qin Y, Liu X, Yin J (2013) Short-term forecasting model of traffic flow based on GRNN. In: 25th chinese control and decision conference (CCDC). IEEE, pp 3816–3820
26. Yin H, Wong S, Xu J, Wong CK (2002) Urban traffic flow prediction using a fuzzy-neural approach. Transp Res Part C Emerg Technol 10(2):85–98
27. Guorong G, Yanping L (2010) Traffic flow forecasting based on PCA and wavelet neural network. In: International conference of information science and management engineering, vol. 1. IEEE, pp 158–161
28. Tian Y, Pan L (2015) Predicting short-term traffic flow by long short-term memory recurrent neural network. In: IEEE international conference on smart city/SocialCom/SustainCom (SmartCity). IEEE, pp 153–158
29. Fu R, Zhang Z, Li L (2016) Using LSTM and GRU neural network methods for traffic flow prediction. In: 31st youth academic annual conference of Chinese association of automation (YAC). IEEE, pp 324–328
30. Zhang W, Yu Y, Qi Y, Shu F, Wang Y (2019) Short-term traffic flow prediction based on spatio-temporal analysis and CNN deep learning. Transportmetrica A Transp Sci 15(2):1688–1711
31. Zheng W, Lee DH, Shi Q (2006) Short-term freeway traffic flow prediction: Bayesian combined neural network approach. J Transp Eng 132(2):114–212
32. Yanchong C, Darong H, Ling Z (2016) A short-term traffic flow prediction method based on wavelet analysis and neural network. In: Chinese control and decision conference (CCDC). IEEE, pp 7030–1034
33. Xiao Y, Yin Y (2019) Hybrid LSTM neural network for short-term traffic flow prediction. Information 10(3):105
34. Lin F, Xu Y, Yang Y, Ma H (2019) A spatial-temporal hybrid model for short-term traffic prediction. Mathematical Problems in Engineering
35. Wei W, Wu H, Ma H (2019) An autoencoder and LSTM-based traffic flow prediction method. Sensors 19(13):2946
36. Chen YC, Li DC (2021) Selection of key features for PM2. 5 prediction using a wavelet model and RBF-LSTM. Appl Intell 51(4):2534–2555
37. Chandra S, Gangopadhyay S, Velmurugan S, Ravinder K (2017) Indian highway capacity manual (Indo-HCM). Council of Scientific and Industrial Research-Central Road Research Institute (CSIR-CRRI). http://www.crridom.gov.in/sites/default/files/Indo-HCM%20Snippets.pdf

Following Behavior of Motorized Two-Wheelers in Mixed Traffic Conditions

Jaikishan Damani and Perumal Vedagiri

Abstract The traffic scenario found in most developing countries including India is mixed traffic, wherein a significant proportion is composed of motorized two-wheelers (MTWs). The irregular driving patterns of MTWs increase the complexity of driver behavior modeling as compared to homogeneous car-based traffic. In this direction, this study aims at investigating the single leader following behavior of MTWs in mixed traffic conditions based on field data. Multiple linear regression (MLR) models were developed for quantifying the following behavior, and support vector machines (SVMs) were used to investigate the discrepancies and highlight variations with longitudinal and lateral parameters. The findings highlight clear differences in the following behavior between cases with different leader vehicle classes and an aggressive driving behavior by MTW drivers. Results from this study can be useful for developing accurate microsimulation models, vehicle-to-vehicle (V2V) communication systems, etc.

Keywords Mixed traffic · Motorized two-wheelers · Vehicle following · Driving behavior

1 Introduction

Mixed traffic is widely prevalent in developing countries. The main features in mixed traffic include heterogeneity in vehicle types and lane indiscipline. Heterogeneity is the presence of a variety of modes of transport which exhibit varying static and dynamic characteristics. The modes of transport referred here may be motorized (passenger cars, auto-rickshaws, etc.) or non-motorized (cycles, bullock carts, etc.). This heterogeneity forms one of the major reasons of lane indiscipline. This may

J. Damani (✉) · P. Vedagiri
Department of Civil Engineering, Indian Institute of Technology (IIT), Bombay, Powai 400076, India
e-mail: jaikishan.damani1@gmail.com

P. Vedagiri
e-mail: vedagiri@civil.iitb.ac.in

L. Devi et al. (eds.), *Proceedings of the Sixth International Conference of Transportation Research Group of India*, Lecture Notes in Civil Engineering 272,
https://doi.org/10.1007/978-981-19-3494-0_5

be explained by the fact that every class of transport has its own dimensions, operating speeds, and maneuverability, which highly affect their specific driving behaviors, which is in contrast to the broadly passenger car-based homogeneous traffic conditions as experienced in some of the developed parts of the world.

Motorized two-wheelers (MTWs) account for almost 73.5% of the total registered vehicles in India [1]. For some regions such as Indonesia and Taiwan, the percentage share of MTWs in total registered vehicles is up to 95% [2]. This is not surprising since MTWs are cheaper to buy, operate and maintain, and are easier to maneuver through traffic and park while not in use [3, 4]. However, earlier literature has highlighted the various problems that are encountered by transportation engineers due to this unique attribute. A significant proportion of MTWs introduces a complexity in understanding the microscopic parameters that determine the driving behavior. A higher proportion of MTWs is also associated with higher fatality rates [5]. It is therefore imperative to investigate such kind of vehicle interactions in order to establish more suitable traffic flow theories and enhance safety [6].

In this regard, several studies have tried to model the car-following behavior in the past several decades [7–9]. Car following essentially describes how the follower vehicle reacts to the leader vehicle. One of the methods of describing a car-following framework is the safe distance model, which assumes that the follower will maintain a perceived safe distance with the leader vehicle in order to avoid a rear-end collision [9]. Ye and Zhang carried out a vehicle-type specific analysis of time headways for a freeway and highlighted the importance of the type of vehicle involved in the scenario [10]. However, it considered only cars and trucks. Gunay et al. attempted to modify car following for weak lane-disciplined scenario [11]; however, MTWs were again not considered for this study. On the other hand, another study conducted in Taiwan developed a safe distance longitudinal model for motorcycles [12]. However, all vehicle types found in mixed traffic conditions were not considered as leader vehicles. Another study tried to model the car-following behavior for heterogeneous traffic conditions based on the data collected on an inter-urban express highway in Mumbai, India [13]. However, these studies considered only cars and trucks for their analysis. Although MTWs were considered for analysis in a study by Durrani et al. [14], a following model could not be developed due to a very small proportion in the study area. Another study presented a data-driven framework for development of a car-following model, but MTWs were not considered a part of analysis. Recently, Das et al. [15] attempted to investigate the time headway with a special focus on the staggered car-following conditions. But, it considered MTW followers with respect to only passenger car leaders. Additionally, Das and Maurya modeled the time headway for heterogeneous conditions in their recent study, which did not consider the various types of leader vehicles such as auto-rickshaw and heavy vehicles [16].

It is therefore established that although several studies have attempted to model car-following conditions in the past, the studies focusing on the specific behavior found in the mixed traffic conditions (characterized by heterogeneity and lane indiscipline) are sparse. Moreover, MTWs are also characterized by staggered following in mixed traffic conditions due to higher maneuverability. These patterns have been mentioned to have a significant influence on the following behavior, and therefore, these should be given special importance while modeling driving behavior of MTWs. However, no such comprehensive study exists for such mixed traffic conditions which considers the driving behavior of an MTW follower with respect to various conditions generally found in mixed traffic conditions, such as different classes of leader vehicles and staggered driving behavior.

This study therefore attempts to model the single leader following behavior of MTWs in mixed traffic conditions with respect to various categories of leader vehicles commonly found in such conditions. Specifically, the objective of this study is to model the single leader following behavior of MTWs in mixed traffic conditions. The pairs of vehicles where the follower was an MTW, and the leader was either of the considered classes of vehicles were considered for the analysis. Multiple linear regression (MLR) models were developed to quantitatively investigate the differences, which were followed by development of support vector machine (SVM) models, which emphasized on the differences between various classes of leader vehicles. Observations and related conclusions have been listed toward the end of the paper.

2 Methodology

2.1 Study Framework

To fulfill the aim of the present study, real-world data were used from Mumbai (India). Site investigation visits were conducted and suitable locations were shortlisted for data collection. Video data were collected from two locations with the help of high-definition video recorder cameras. Frames were extracted from the video data, which resulted in two images per second of video data. The on-screen coordinates were extracted manually from the images and transformed into real-world coordinates using a methodology developed by Fung et al. [17]. Stitching of the coordinates gave a high-definition trajectory of vehicles across the desired length. Moreover, parameters such as speed, acceleration, distance headway (DH), and centerline separation (CS) were determined from the available parameters. To smoothen the noise for every frame, a two-point moving average was used. The framework adopted for the current study is shown in Fig. 1.

2.2 Data Collection and Extraction

Data Collection

Since the purpose of the current study was modeling of the following behavior of MTWs, video data collection was found to be suitable form of data collection to

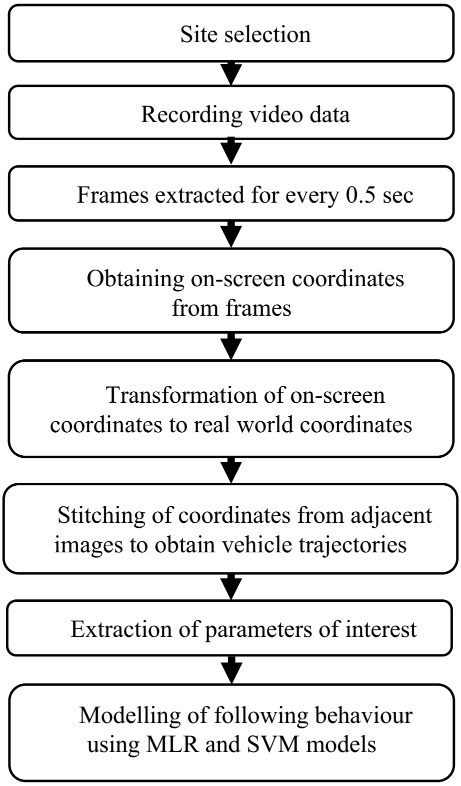


Fig. 1 Methodology framework for the current study

investigate real-world driving behavior. For this purpose, two locations were chosen in Mumbai (India), which are described below.

(1) Location 1 is the Western Express Highway near NESCO in Mumbai. This location is an inter-urban expressway which has almost five lanes in the direction of study apart from service lanes. The data were recorded at this location by installing video camcorders above a foot-over bridge (FOB) during afternoon. Top view of the location is shown in Fig. 2a.
(2) Location 2 is the Mrinal Tai Gore Flyover in Goregaon (Mumbai). This is a two-lane unidirectional flyover. At the study section, the flyover has negligible gradient that can be considered as insignificant in affecting driving behavior. Video data were recorded by installing a video camera above a high-rise residential complex adjacent to the facility. Top view of the location is shown in Fig. 2b.

Video data of around 3 h were recorded at both the locations during afternoon period to ensure good visibility. Both the locations had a very good road surface to ensure no hindrance to the driving behavior. Four points, that represented vertices of a rectangle, were marked on the road for calibration purposes.

(a) (b)

Fig. 2 Top view of the locations of data collection

Data Extraction

After completion of data collection process, the video data were converted into images for further extraction and analysis. Two frames were extracted for every second of video data. This is based on the generally adopted value of 0.5 s, as the reaction time of MTW drivers in earlier literature [18, 19]. Sample screenshots for Location 1 (Western Express Highway) and Location 2 (Mrinal Tai Gore flyover) are shown in Fig. 3a and b, respectively. The positions of vehicles were clicked on the screen which gave the on-screen coordinates of the vehicles. Once the on-screen coordinates were extracted, they were transformed into real-world coordinates using the framework developed by Fung et al. [17]. This manual process was adopted because, although very time-consuming, manual methods are generally more accurate and have almost no false identification of vehicle classes, which otherwise reduces the efficiency for automatic data extraction software.

(a) (b)

Fig. 3 Sample screenshots of the collected video data

2.3 Variables Extracted from Video Data

Video data can be used to extract various parameters of interest. Since the purpose of the current study was modeling of the following behavior of MTWs, the video cameras were placed at high vantage points which could give a good view of sufficient length of the road section. The focus was on spatial parameters which mainly deal with the position of a vehicle on road at a specific instant. The difference in the longitudinal and lateral coordinates of the vehicles gave their relative positions. The parameters that indicate the relative longitudinal and lateral distances between a pair of vehicles adopted in this study are shown in Fig. 4.

Distance headway (DH) is defined as the longitudinal distance between front of leader vehicle to front of follower vehicle. Centerline separation (CS) is the lateral center-to-center distance between the leader and follower vehicle. Moreover, time headway (TH) is a temporal parameter that measures the time taken for the leader vehicle to reach the position of the follower vehicle. TH is therefore calculated by dividing the DH by the speed of the follower MTW. Speed and acceleration of vehicles were calculated on the basis of differences in their positions in adjacent frames. For these calculations, two-point moving average was employed to avoid noise in data. A typical trajectory (not to scale) obtained by the exercise has been visually presented in Fig. 5. The values corresponding to the closest interaction between the pair of vehicles were considered for further analysis.

Other parameters extracted from the video data include type of leader and follower vehicle, helmet use, and apparent age and gender. It is important to mention that four classes of vehicles were considered for the analysis, which generally represent the largest proportion of traffic found in Indian conditions. The four classes of vehicles considered for this study are as follows: motorized two-wheelers (MTWs), passenger cars, heavy vehicles (including trucks, buses, etc.), and auto-rickshaws (motorized three-wheelers).

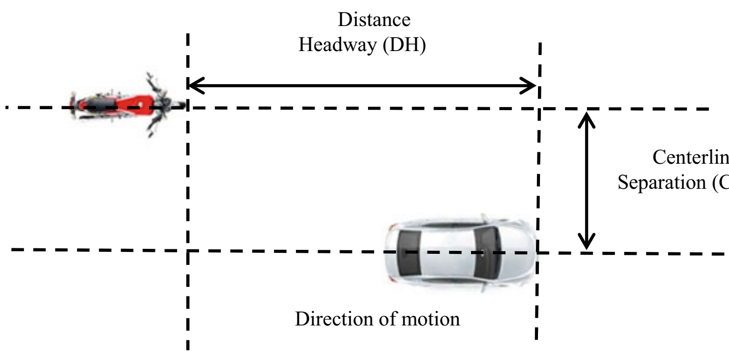


Fig. 4 Theoretical representation of parameters considered for a single leader staggered following

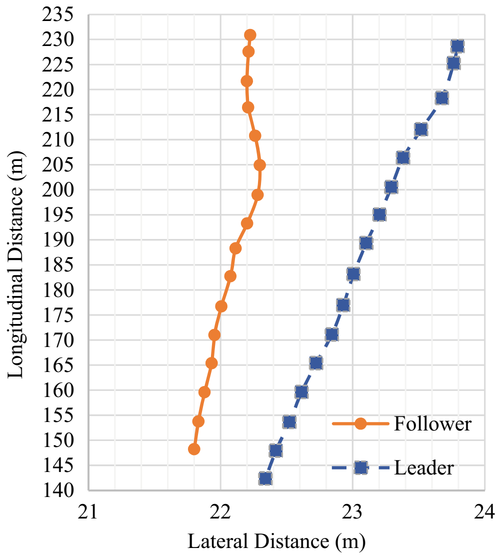


Fig. 5 Obtained trajectory from a single leader following case

3 Preliminary Analysis

3.1 Data Filtering

The data extraction process generated almost 3502 data points corresponding to about 415 pairs of vehicles. However, it is important to make sure that there is a certain interaction between the leader and follower vehicle. For the current study, the objective of interest is vehicle following, and therefore, spatial proximity was therefore important, in absence of which, the pair of vehicles might be considered as independent. To achieve this, the following filters were applied to the dataset, which are generally in line with previous research [20, 21].

(1) The cases with distance headway (less than 30) were considered for further analysis. This is important since 30 m has been found out as the maximum range of effect of leader vehicle on the follower vehicle.

(2) Pairs of vehicles with centerline separation (CS) of only less than 3 m were considered for further analysis. Although there is a general tendency to drive with certain off-centeredness in such conditions, 3 m has been considered as the maximum zone of lateral influence. A value greater than 3 m might indicate no effective interaction between the pair of vehicles.

After the application of the said filters, a dataset of only about 397 pairs of vehicles was used for further analysis.

3.2 Descriptive Statistics

It was evident from the data that a vast majority of MTW riders were males, wore helmets, and are riding solo. The apparent age of most of the riders is below 40, and most of them are riding a motorcycle. A summary of such data has been presented in Table 1.

Around 75% of the total vehicles was either MTWs or passenger cars, as shown in Fig. 6. These two vehicle classes were also found to be the fastest among all, while expectedly, the heavy vehicles were the slowest. Box plots representing speed profile of all considered classes of vehicles are shown in Fig. 7.

Scatter plots representing distance headway (DH) and centerline separation (CS) between various classes of leader vehicles and follower MTWs were drawn as shown in Fig. 8, and a clear visual difference was found between most of the leader vehicle classes. The dots show the relative positions between the leader vehicle and follower

Table 1 Descriptive statistics of observed data

Sr. No	Variable of interest	Proportion of observed data (%)	
1	Apparent gender	Male (96.8)	Female (3.2)
2	Apparent age	Younger (below 35, 67)	Elder (above 35, 33)
3	Helmet usage	Wearing helmet (92.67)	Not wearing helmet (7.33)
4	Type of follower MTW	Motorcycle (64.73)	Scooter (35.27)
5	Presence of pillion	No pillion (79.8)	One pillion (20.2)

Fig. 6 Observed proportion of classes of vehicles

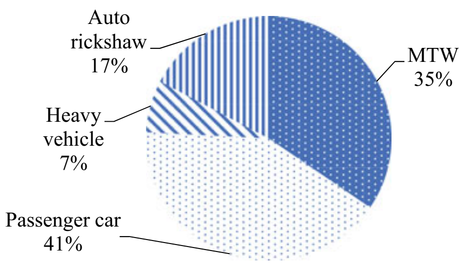
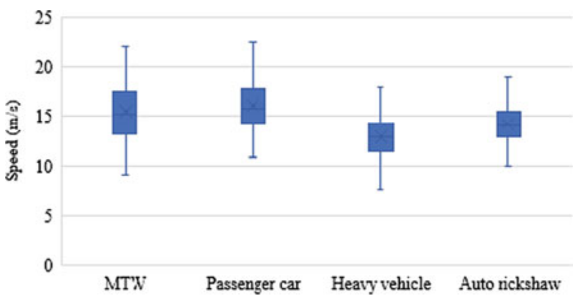


Fig. 7 Speed profiles of the observed vehicle classes



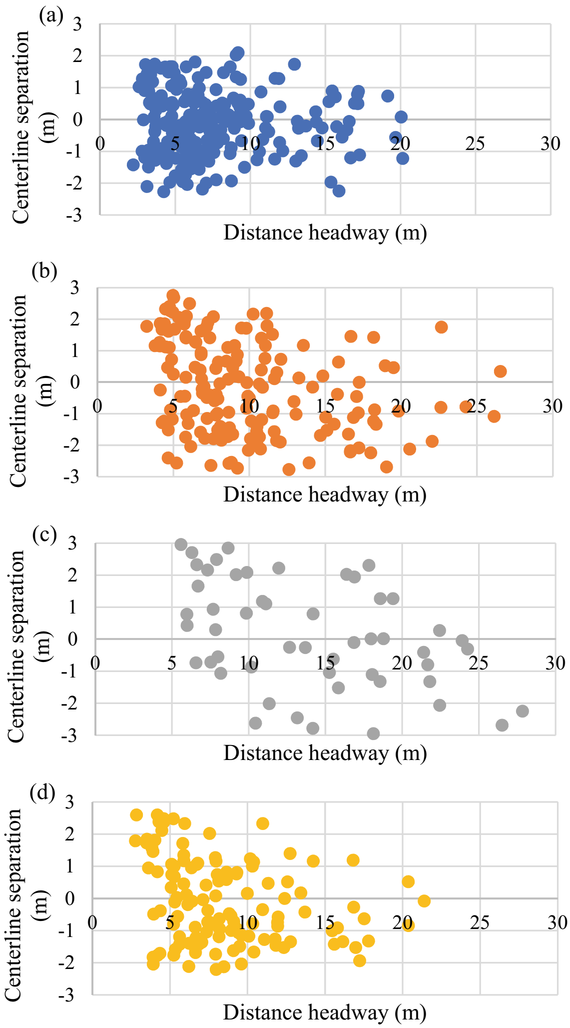


Fig. 8 Scatter plots between distance headway (DH) and centerline separation (CS) for leader vehicle as **a** MTW, **b** passenger car, **c** heavy vehicle, **d** auto-rickshaw

MTW at the closest interaction, and it can be clearly seen that the dots are much more concentrated near the origin for MTW leaders, as opposed to much scattered dots for heavy-vehicle leaders. Visual inspection also suggests that MTW followers are more willing to accept shorter distance headways (DH) with an increase in centerline separations (CSs). This is indicative of the off-centered style of following adopted by a majority of MTW riders in mixed traffic conditions.

3.3 Preliminary Verification of Differences Between Groups

The data consisted of all combined cases of following, which included all categories of leader-vehicle classes being followed by MTWs. Moreover, the MTWs usually have a tendency to follow the leader vehicle with a certain off-centeredness. That is, rather than following the leader vehicle with a centerline separation (CS) of close to zero, they usually tend to follow either from left or right, with a certain non-negligible value of CS.

To accurately model the following behavior, it therefore becomes important to determine if there is a significant difference between the following behavior with respect to class of leader vehicle and also the direction of following (left or right). Statistical tests were conducted for the same. To verify if there was a statistically significant difference in the following behavior according to various leader vehicle classes, Kruskal Wallis test is conducted. Kruskal Wallis is a more general form of ANOVA, and therefore, it does not require all the assumptions of ANOVA to hold true. The test is performed separately on both the considered dependent variables, LG and TH, separately for both the locations. The asymptotic significances therefore derived are as shown in Table 2. Since the all the values are less than 0.05, it means that there is indeed a statistically significant difference between following behavior based on the class of leader vehicle. This is consistent with other studies that have highlighted the importance of considering the class of lead vehicle due to different driving behaviors by the followers [15, 22].

Similar tests were conducted to confirm if there was a statistically significant difference in following behavior between MTW followers following the leader vehicle from left or right direction. Since there were only two available options (left or right), Mann Whitney U test and Kolmogorov–Smirnov test were conducted. The asymptotic significance values thus obtained are shown in Table 3. Since the values are greater than 0.05, it can be concluded that there does not appear to be a

Table 2 Results from statistical tests on the response variables based on class of leader vehicle

Sr. No	Test	Asymptotic significance values	
		Distance headway (DH)	Time headway (TH)
1	Kruskal Wallis test	0.000	0.000
2	Median test	0.000	0.000
3	Jonckheere–Terpstra test	0.000	0.000

Table 3 Results from statistical tests on the response variables based on direction of following

Sr. No	Test	Asymptotic significance values	
		Distance headway (DH)	Time headway (TH)
1	Mann Whitney U test	0.081	0.089
2	Kolmogorov–Smirnov test	0.085	0.053

statistically significant difference between following behavior based on the direction of following (left or right).

The MTW followers tend to treat various classes of leader vehicles differently with respect to following behavior. This necessitates the inclusion of a variable which represents the type of leader vehicle. However, they do not tend to drive differently in conditions where they follow a vehicle from left or right. Therefore, the absolute values of centerline separation (CS) can be considered for further analysis.

3.4 Distribution of Distance Headway and Time Headway for Different Classes of Leader Vehicles

Percentile plots were developed for distance headway (DH) as well as time headway (TH) as shown in Fig. 9a and b. It is evident that the minimum as well as the median values of both the parameters are the least for MTW leaders, followed by auto-rickshaw leaders, passenger car leaders, and finally heavy-vehicle leaders. The plots also indicate that the median values for distance headway (DH) are about 8.21 m, 10.82 m, 18.13 m, and 9.32 m, respectively, for MTW, passenger car, heavy vehicle, and auto-rickshaw as the leader vehicle, respectively. Corresponding values for time headway (TH) are about 0.61 s, 0.79 s, 1.36 s, and 0.77 s, respectively.

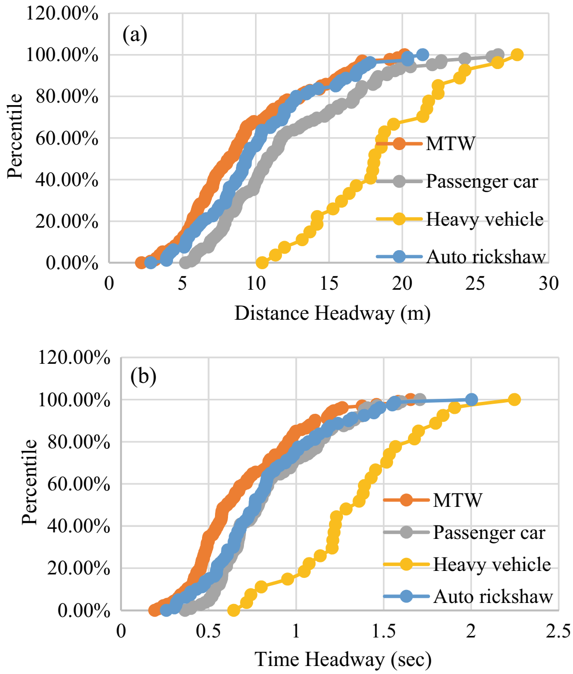


Fig. 9 Percentile graphs for **a** Distance headway (DH) and **b** Time headway (TH)

It is evident that these values are much lesser than the results from other studies [15, 23, 24], which can be largely attributed to an increased driving confidence by MTWs (due to the tendency to drive staggered to the leader vehicle) combined with the risky driving attitudes of MTWs. This is further corroborated by the fact that almost 85% of the TH values with MTW leaders are found to be lesser than 1 s, while the percentage being almost 72 and 75% for passenger car leaders and auto-rickshaw leaders, respectively.

4 Model Development for Following Behavior

4.1 Choice of Variables for Modeling

To model the following behavior of MTWs, multiple linear regression (MLR) technique was adopted. Two sets of equations were developed, with either of distance headway (DH) and time headway (TH) as a dependent variable in each set. In all cases, the dependent variable was a continuous variable. The independent variables used for the analysis are listed in Table 4.

The parameter estimates for the MLR model with distance headway (DH) as a dependent variable are shown in Table 5, where the statistically significant parameters have been presented in bold. The discrete variable representing the type of leader

Table 4 Details of explanatory variables included in the analysis

Sr. No	Variable	Type of variable	Description
1	CS (Centerline separation)	Continuous	Lateral distance between the middle axes of both the vehicles (m)
2	Type of leader vehicle	Discrete	This shows the class of influential vehicle, i.e., whether it is MTW (0), Passenger car (1), heavy vehicle (2), or auto-rickshaw (3)
3	Apparent gender	Binary	Apparent gender; male (0) or female (1)
4	Apparent age	Binary	Apparent age based on video data, drivers were classified as below 35 (younger, 0) or above 40 (older, 1)
5	Helmet usage	Binary	Whether the MTW rider has a helmet (0) or not (1)
6	Type of follower MTW	Binary	Whether the follower MTW is a motorcycle (0) or a scooter (1)
7	Presence of pillion	Binary	Whether there is a pillion (1) or not (0)
8	Speed of follower MTW	Continuous	Speed of following MTW (km/h)

Table 5 Parameter estimates with distance headway (DH) as a response variable

Sr. No	Variable	Coefficient	Std. error	t-stat	Significance
1	(Constant)	1.952	1.849	1.056	0.293
2	**Car leader**	**1.192**	**0.658**	**1.810**	**0.072**
3	**Heavy leader**	**8.374**	**0.948**	**8.833**	**0.000**
4	**Auto-leader**	**2.016**	**0.669**	**3.014**	**0.003**
5	Gender	−1.111	1.477	−0.752	0.453
6	Age	0.652	0.516	1.264	0.208
7	Helmet	0.858	1.185	0.725	0.470
8	MTW type	0.246	0.533	0.462	0.645
9	Pillions	−0.226	0.610	−0.371	0.711
10	**Centerline separation (m)**	**−1.875**	**0.387**	**−4.842**	**0.000**
11	**Average speed (m/s)**	**0.256**	**0.140**	**1.832**	**0.069**
12	Average acceleration (m/s^2)	0.069	0.709	0.097	0.923
13	Average relative speed (m/s)	−0.315	0.239	−1.319	0.189

Bold indicates the statistically significant parameters

Table 6 Model summary with distance headway (DH) as a response variable

Parameter	Value
Adjusted R square	0.350
Std. error of estimate	3.371
F-stat	9.043
Significance	< 0.001

vehicle was replaced by three dummy variables namely "car leader," "heavy leader," and "auto-leader" which represented a passenger car, a heavy vehicle, and an auto-rickshaw, respectively, as a leader. The model summary has been presented in Table 6. It is evident from the analysis that apart from category of leader vehicle, average speed and centerline separation are statistically significant variables affecting the following behavior of MTWs. The values indicate that in comparison to MTW leaders, the follower MTWs tend to maintain about 1.2 m, 2.0 m, and 8.4 m higher gaps with passenger cars, auto-rickshaws, and heavy-vehicle leaders, respectively. Centerline separation (CS) is a very sensitive parameter since every 1 m increase in CS tends to decrease the distance headway (DH) by about 1.88 m. Moreover, a speed increase by 1 m/s increases the DH by about 0.25 m.

Similar models were developed for time headway (TH) as a dependent variable, as shown in Tables 7 and 8. Similar to the previous case, the same set of variables has been found to be statistically significant although with a better fit. The values indicate that the difference in time headway (TH) of cases with MTW leaders and heavy-vehicle leaders is very prominent, at about 1.1 s. Corresponding values for car leaders and auto-rickshaw leaders are 0.25 s and 0.23 s, respectively. Similar to the

Table 7 Parameter estimates with time headway (TH) as a response variable

Sr. No	Variable	Coefficient	Std. error	t-stat	Significance
1	**(Constant)**	**1.007**	**0.151**	**6.677**	**0.000**
2	**Car leader**	**0.253**	**0.054**	**4.718**	**0.000**
3	**Heavy leader**	**1.088**	**0.077**	**14.076**	**0.000**
4	**Auto-leader**	**0.234**	**0.055**	**4.284**	**0.000**
5	Apparent gender	−0.064	0.120	−0.527	0.599
6	Apparent age	0.028	0.042	0.671	0.503
7	Helmet	0.064	0.097	0.665	0.507
8	MTW type	0.020	0.044	0.457	0.648
9	Pillions	-0.027	0.050	-0.537	0.592
10	**Centerline separation (m)**	**−0.150**	**0.032**	**−4.740**	**0.000**
11	**Average speed (m/s)**	**−0.033**	**0.011**	**−2.922**	**0.004**
12	Average acceleration (m/s^2)	−0.002	0.058	−0.027	0.978
13	Average relative speed (m/s)	−0.025	0.019	−1.275	0.204

Bold indicates the statistically significant parameters

Table 8 Model summary with time headway (TH) as a response variable

Parameter	Value
Adjusted R square	0.539
Std. error of estimate	0.275
F-stat	18.443
Significance	<0.001

previous case, an increase in centerline separation (CS) reduced the time headway (TH), by about 0.15 s. However, an increase in average speed by 1 m/s also reduced TH by about 0.03 s.

It can therefore be established that the MTW followers tend to maintain highest spatial (DH), as well as temporal headway (TH) when the leader vehicle is a heavy vehicle. On the other hand, the headways tend to be the minimum when the leader is an MTW itself. Centerline separation (CS) was found to be statistically significant for both the models, which suggests that driving with an off-centeredness affects the following behavior of MTW riders, both spatially as well as temporally. This has necessary implications in countries like India where such staggered following is a common practice. Other parameters which were not found to be significant included apparent age, apparent gender, helmet usage, type of follower MTW, presence of a pillion, average acceleration, and average relative speed. However, previous literature has established the role of demographics in affecting the riding behavior in general [25, 26]. This might be attributed to the divergence between the actual and apparent observation for both age and gender.

4.2 Developing Support Vector Machines (SVMs)

Advanced data visualization techniques can be used to complement statistically developed models. While the models (such as MLR in this paper) can be used to quantify certain parameters, use of sophisticated techniques such as support vector machines (SVMs) can help in visualizing the existing data. SVMs are machine learning-based classification tools, which have been widely used in qualitative analysis and representation of categorized datasets. The provision of a "soft margin" enables the SVM to analyze linearly inseparable data. For this study, an error-tolerant soft margin was developed as shown in Fig. 10a and b for distance headway (DH) and time headway (TH), respectively. The soft boundary could not be sufficiently developed for cases where leader vehicle is MTW itself, which may be because of insufficient representation or dispersity in data points. The general trend observed in both the figures is similar with respect to leader-vehicle class. Heavy vehicles as leaders usually demand the highest gaps, temporally as well as spatially.

On the basis of the developed SVM models, the minimum observed distance headway (DH) between MTW followers and heavy-vehicle leaders is about 15 m. On the other hand, MTWs usually tend to maintain not more than 25 m with the leader vehicle when it is any other class of vehicle. Moreover, it is evident that the MTWs tend to maintain a time headway (TH) of not more than 1.5 s with any other class of vehicles except for heavy-vehicle leaders.

5 Discussion

5.1 Conclusion

The study aimed at developing following models for MTWs in mixed traffic conditions. This was done with the help of field data obtained from videos recorded in Mumbai, India. Special focus was given on the class of leader vehicle, which proved to have significant impact on following behavior of MTWs. SVM models were developed which helped visualize the data better and classify the following behavior on the basis of class of leader vehicle. Some observations and conclusions that can be drawn on the basis of current study are discussed in this section.

Statistical tests which were conducted to verify the effect of class of leader vehicle on the following behavior of MTWs showed that the class of leader vehicle had a statistically significant impact on the following behavior of MTWs in such conditions. This may be due to the fact that MTWs try to perceive the damage caused by rear-end crashes with different classes of vehicles differently. However, the direction of following did not have a statistically significant effect on the following behavior. This may be explained by the highly unordered traffic regime of mixed traffic in general and MTWs in particular. The observed distance headways (DHs) and time headways (THs) are in general lesser than other similar studies [15, 23], which highlights the

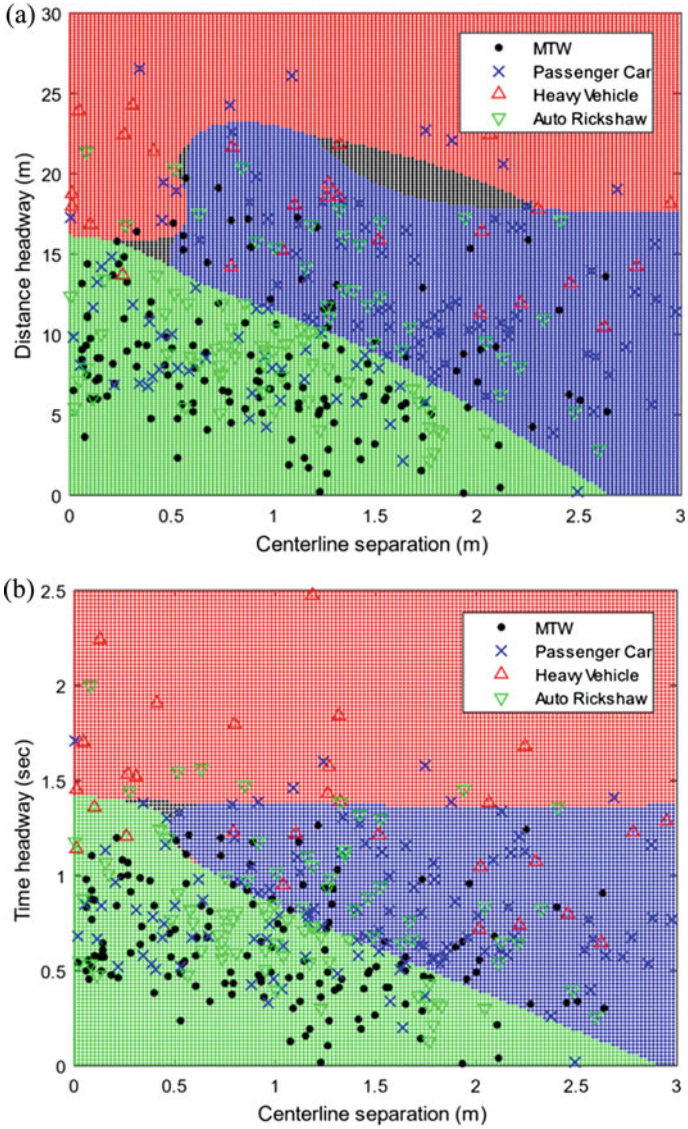


Fig. 10 Support vector machine models developed for **a** Distance headway (DH) and **b** Time headway (TH)

higher spatial and temporal proximity maintained by MTWs with leader vehicles during following. Earlier studies have not fully captured the various combinations of following which might be possible in a typical mixed traffic condition. However, it is to be noted that there was found to be a clear difference among the different classes of leader vehicles. For example, median values of distance headway (DH) between a leader MTW and follower MTW are just about 8.21 m. However, corresponding values for leader as passenger car, heavy vehicle, and auto-rickshaw are 10.82 m, 18.13 m, and 9.32 m, respectively. This emphasizes the fact that MTW followers tend to maintain minimum spatial and temporal headways with MTW leaders, and highest headways with heavy-vehicle leaders. This may be due to higher level of perceived damage in a rear-end crash with a heavy-vehicle leader. The distribution of the DH and TH can be implemented in the development of microsimulation software which can closely replicate the following behavior of MTWs. Moreover, threshold values can be fixed to implement in the in-vehicle warning systems and development of similar driving assistance systems. Another factor that might have cause the excessive difference in headway values with heavy leader vehicles in comparison with other classes of leader vehicles is the higher width of heavy vehicles, which might make the swerving maneuver in case of an emergency more difficult, which therefore necessitates maintaining higher headways with heavy-vehicle leaders.

An important observation is that almost 85% of the follower MTWs maintained a time headway (TH) of less than 1 s when the leader was an MTW. The corresponding values for passenger car leaders and auto-rickshaw leaders were 72% and 75%, respectively. This indicates the excessively small TH values adopted by MTW in mixed traffic conditions, which are much lesser than earlier literature [27, 28]. Earlier studies have either not considered the staggered driving behavior or various classes of leader vehicles. The staggered driving pattern is generally adopted by MTWs in such conditions, wherein instead of exactly following the leader vehicle with a centerline separation (CS) of almost zero, they maintain some off-centeredness, increases their field of view, and perceived safety. This behavior encourages the follower MTWs to adopt very small headways with the leader vehicles.

MLR technique was used to model the following behavior. The fit of the model with time headway (TH) as a response variable was found to be better than that of the model with distance headway (DH) as a response variable. In both the models, the type of leader vehicle, centerline separation (CS), and average speeds were found to be the significant parameters that influenced the following behavior. The parameter estimates corroborate the intuitively highest headways with heavy-vehicle leaders and lowest headways with MTW leaders. Both the distance headway (DH) as well as time headway (TH) decrease with an increase in centerline separation (CS). However, although an increase in average speed increases the distance headway (DH), it results in reduction in time headway (TH). This indicates that the increase in DH is offset by the increase in speeds which causes TH to reduce. Apparent gender and apparent age were not found to be significant in affecting the following behavior. However, earlier literature points to risky driving behavior of males and younger drivers [29, 30]. This discrepancy may be attributed to the inability to determine age and gender accurately from the video data. Other factors such as helmet usage, type of follower MTW (motorcycle or scooter), presence of pillion, average acceleration, and average relative speed were found to be insignificant in affecting the following behavior.

The SVM models also indicate the highest spatial (DH) as well as temporal (TH) gaps maintained by MTW followers when the leader is a heavy vehicle. The minimum general DH and TH values were found to be 15 m and 1.5 s, respectively. Most of the DH and TH values for leader vehicles other than heavy vehicles are less than 25 m and 1.5 s, respectively. Although a soft boundary could not be developed for cases where the leader is an MTW itself, which may be due to dispersity in data, the general trends observed in the SVM models further corroborate the observations from the MLR models. These outcomes can be applied in collision detection systems as well as microsimulation models focusing on replicating mixed traffic conditions.

Specialized microscopic models are important for correct replication of the real-world traffic found in specific conditions such as mixed traffic found in India. One of the essential components of such simulation models is the car-following models which can account for staggered following, tailgating, etc. Moreover, such investigations are important to accurately model traffic flow theories, intelligent transportation systems (ITSs) applications (such as in-vehicle warning systems), and safety studies. It is therefore important to establish following behavior adopted by MTWs in such specific conditions. The major application of the study is the development of the following behavior model that can help the traffic engineers to understand the following behavior more accurately. The quantitative results can be directly applied into models such as simulation tools and traffic flow models, which are planned to be developed for similar traffic conditions.

5.2 Limitations and Future Scope

The study attempted to formulate the following behavior of MTWs in mixed traffic conditions with the help of MLR models. The study has some limitations which are discussed below. These may form the basis for further research in this subject.

(1) Only, four vehicle classes were considered. Although these classes account for an overwhelming majority of the traffic composition in general in such conditions, further studies may be required for more comprehensive models that have varying proportions of traffic. Sensitivity analysis may be carried out on such data to develop even more accurate systems based on the analysis. Moreover, only, a single leader was considered for all cases, which might not always be the case as an MTW might follow multiple leader vehicle simultaneously.

(2) Apparent age and apparent gender were not found to significantly affect the following behavior. This may be due to the difference between apparent and actual demographics, which may be attributed to difficulty in correct identification by the authors since the video was taken from a high vantage point.

(3) Certain variables such as helmet usage and presence of a pillion were found to be insignificant in affecting the following behavior. However, most of the data points corresponded to solo riders with helmets. Therefore, there were few samples which had a pillion or who were not wearing helmets, which might cause insufficiency in sample size. Therefore, it is recommended that more

data, preferably from different regions, can be collected to establish the effect of helmet usage, presence of pillion, etc., on the following behavior.

(4) Although the manual method of data extraction is a very tedious process, it has been adopted owing to higher accuracy. However, better data-driven extraction techniques based on machine learning (ML) frameworks may be incorporated in the future, which may result in better accuracy than the currently available image processing-based techniques. These sophisticated tools may help in extracting larger datasets, which can be used to develop site-specific models.

References

1. MORTH (2017) Road Accidents in India—2017
2. Kumar S, Sharma B, Nezhurina MI (2018) Responsible factors of powered two wheeler accidents: a review. In: Proceedings of the 2018 3rd IEEE international conference on research in intelligent and computing in engineering, RICE 2018, pp 1–6
3. Vlahogianni EI, Yannis G, Golias JC (2012) Overview of critical risk factors in Power-Two-Wheeler safety. Accid Anal Prev 49:12–22
4. Damani J, Vedagiri P (2021) Safety of motorised two wheelers in mixed traffic conditions: Literature review of risk factors. J Traffic Trans Eng (Engl Ed) 8(1):35–56
5. Bhalla K, Ezzati M, Mahal A, Salomon J, Reich M (2007) A risk-based method for modeling traffic fatalities. Risk Anal 27(1):125–136
6. Kanagaraj V, Asaithambi G, Toledo T, Lee TC (2015) Trajectory data and flow characteristics of mixed traffic. Transp Res Rec 2491:1–11
7. Das S, Maurya AK (2018) Multivariate analysis of microscopic traffic variables using copulas in staggered car-following conditions. Transportmetrica Trans Sci 14(10):829–854
8. Chandler RE, Herman R, Montroll EW (1958) Traffic dynamics: studies in car following. Oper Res 6(2):165–184
9. Asaithambi G, Kanagaraj V, Srinivasan KK, Sivanandan R (2016) Study of traffic flow characteristics using different vehicle-following models under mixed traffic conditions. Trans Lett 10(2):92–103
10. Ye F, Zhang Y (2009) Vehicle type-specific headway analysis using freeway traffic data. Transp Res Rec 2124:222–230
11. Gunay B (2007) Car following theory with lateral discomfort. Transp Res Part B Methodological 41(7):722–735
12. Cho HJ, Wu YT (2004) Modeling and simulation of multiphase traffic flow. AIP Conf Proc 963(2):1036–1039
13. Ravishankar KVR, Mathew TV (2011) Vehicle-type dependent car-following model for heterogeneous traffic conditions. J Transp Eng 137(11):775–781
14. Durrani U, Lee C, Maoh H (2016) Calibrating the Wiedemann's vehicle-following model using mixed vehicle-pair interactions. Transp Res Part C: Emerg Technol 67:227–242
15. Das S, Maurya AK, Budhkar AK (2017) Determinants of time headway in staggered car-following conditions. Transp Lett 11(8):447–457
16. Das S, Maurya AK (2020) Bivariate modeling of time headways in mixed traffic streams: a copula approach. Transp Lett 12(2):138–148
17. Fung GSK (2003) Camera calibration from road lane markings. Opt Eng 42(10):2967
18. Minh CC, Sano K, Matsumoto S (2006) Models of motorcycle deceleration behavior at signalized intersections. Infrastruct Planning Rev 927–936

19. Long NX (2012) A concept of safety space for describing non-lane-based movements of motorcycles, (Sept)
20. Das S, Maurya AK, Budhkar AK (2019) Dynamic data collection of staggered-following behavior in non-lane-based traffic streams. SN Appl Sci 1(6)
21. Asaithambi G, Joseph J (2018) Modeling duration of lateral shifts in mixed traffic conditions. J Transp Eng Part A: Syst 144(9):04018055
22. Brackstone M, Waterson B, McDonald M (2009) Determinants of following headway in congested traffic. Trans Res F: Traffic Psychol Behav 12(2):131–142
23. Duan J, Li Z, Salvendy G (2013) Risk illusions in car following: is a smaller headway always perceived as more dangerous? Saf Sci 53:25–33
24. Das S, Budhkar A, Maurya AK, Maji A (2019) Multivariate analysis on dynamic car-following data of non-lane-based traffic environments. Transp Developing Econ 5(2)
25. Wong JT, Chung YS, Huang SH (2010) Determinants behind young motorcyclists' risky riding behavior. Accid Anal Prev 42(1):275–281
26. Pawar NM, Khanuja RK, Choudhary P, Velaga NR (2020) Modelling braking behaviour and accident probability of drivers under increasing time pressure conditions. Accid Anal Prev 136(Aug 2019):105401
27. Siebert FW, Oehl M, Pfister HR (2014) The influence of time headway on subjective driver states in adaptive cruise control. Transp Res Part F: Traffic Psychol Behav 25(PART A):65–73
28. Taieb-Maimon M, Shinar D (2001) Minimum and comfortable driving headways: reality versus perception. Hum Factors 43(1):159–172
29. Chang HL, Yeh TH (2007) Motorcyclist accident involvement by age, gender, and risky behaviors in Taipei, Taiwan. Transp Res Part F: Traffic Psychol Behav 10(2):109–122
30. Chung YS, Wong JT (2012) Beyond general behavioral theories: structural discrepancy in young motorcyclist's risky driving behavior and its policy implications. Accid Anal Prev 49:165–176

Spatio-Temporal Traffic Characteristics Analysis on Multi-lane Highways Under Varying Traffic Flow and Density Levels

Sandeep Singh, Vidya Rajesh, and S. Moses Santhakumar

Abstract The study of traffic flow characteristics is essential for designing highway systems for which different traffic influencing variables are of primary importance. The study aims to assess the impact of a range of traffic flow and density levels on the spatio-temporal (lateral placement and time headway) traffic characteristics of highways. A substantial amount of traffic data like speed, time headway (TH), and lateral placement (LP) of vehicles was collected continuously for 12 h on six Indian highway sections. The DPCU for different vehicles with different flow and density rates were calculated. The speed and TH descriptive statistics and probability distribution functions at varying flow and density levels were also analyzed. To determine the most appropriate probability distribution function, the goodness of fit test was used. Furthermore, the impact of traffic speed, flow, and density on the LP of vehicles under different traffic conditions was assessed using descriptive statistics. It was observed that the traffic flow and density levels had an unusual effect on the DPCUs of vehicles. The mean traffic speed under high flow and density decreased by 33.3 and 18.3%, while the mean TH under high flow and density decreased by 28.4 and 29.9%. Also, the results indicated that traffic speed and TH data exhibit a different distribution of probability functions depending on the traffic flow and density. The descriptive statistics on the LP of vehicles show the existence of a significant difference of the same concerning the different ranges of traffic speed, flow, and density.

Keywords Speed · Flow · Density · Time headway · Lateral placement · Mixed traffic

S. Singh (✉) · S. M. Santhakumar
Department of Civil Engineering, National Institute of Technology Tiruchirappalli, Tiruchirappalli, Tamil Nadu, India
e-mail: sandeepsingh.nitt@gmail.com

S. M. Santhakumar
e-mail: moses@nitt.edu

V. Rajesh
School of Civil Engineering, SASTRA Deemed University, Thanjavur, Tamil Nadu, India
e-mail: vidya@civil.sastra.ac.in

L. Devi et al. (eds.), *Proceedings of the Sixth International Conference of Transportation Research Group of India*, Lecture Notes in Civil Engineering 272,
https://doi.org/10.1007/978-981-19-3494-0_6

1 Introduction

Traffic data is essential for solving complex traffic engineering problems, especially under mixed traffic conditions [1]. Under ubiquitous heterogeneous traffic situations, different vehicles have varying maneuverability characteristics. These characteristics allow them to operate at different speed levels by maintaining different time headways and occupying any lateral space available in the roadway for a given traffic condition. These parameters are used to evaluate the changes in driver behavior on highways. Because speed, flow, and density are the necessary measures essential to understand the interaction among them collectively rather than independently, specifically on multi-lane rural highways.

Time headway (TH) is the elapsed time between the fronts of the previous vehicle to the front of the present vehicle crossing the IR sensor beams. The lateral placement (LP) is the distance between the center of the vehicle and the edge of the roadway on the curbside when the vehicle is moving. The poor lane discipline and heterogeneous traffic nature in Indian traffic conditions make it necessary to analyze the TH (temporal parameter) and LP (spatial parameter) of vehicles under mixed traffic conditions to understand traffic behavior better. Further, accurate acquisition of vehicular traffic information in real-time is necessary [2].

The majority of research conducted in developing nations with mixed traffic is confined to examining the traffic efficiency using the temporal and longitudinal parameters like speed and TH. Very few studies have been made to explore the spatial and lateral characteristics like LP of vehicles in the past. Speed and TH are critical performance measures to be examined under different traffic conditions such as flow and density. Furthermore, there is no doubt that it is worth researching the sensitivity of flow and density on dynamic passenger car units (DPCU) to assess traffic characteristics. Also, the LP of vehicles is one of the most vital operational characteristics of the highways, which needs to be extensively analyzed at a different speed, flow, and density levels.

For a clear understanding of vehicle dynamics and the development of traffic flow models, more observational studies using huge traffic data are needed. Besides, precise modeling with comprehensive research analysis representing the real-field conditions is essential to regulate, manage, and control traffic. Hence this research is undertaken to assess the impact of traffic flow and density on vehicle speed and TH. Also, the DPCU factors for different vehicle types with varying traffic flow and density were evaluated. The analyses of the probability distributions on the speed and TH at different flow rates and density levels were conducted to determine the best-fitted distribution. Besides, this study aims to examine the LP characteristics of vehicles at the varying speed, flow, and density of the traffic stream using the (Infra-red) IR sensor-based traffic data obtained from Indian highways.

2 Literature Review

Zhang et al. [3] obtained headway data from urban areas and analyzed the probability distributions. Similarly, Jang et al. [4] collected headway data on Korean multi-lane highways carrying homogenous traffic using loop detectors and researched theoretical progress models by categorizing the traffic flow into five different rates. The study found that Gamma distribution fits well at light to medium flows, and the Pearson VI distribution fits well at all flow rates.

Gunay and Erdemir [5] conducted a lateral analysis of longitudinal TH between the interacting vehicles. The authors observed that most of the vehicles chose either to pass by or lag instead of making a side-by-side movement. Araghi et al. [6] presented how to use GPS loggers to capture the travel time data. Bhaskar and Chung [7] provided a basic understanding of the use of Bluetooth scanner data. Brennan et al. [8] studied the efficiency of data collected from Bluetooth using the lateral distance of vehicles. Li et al. [9] used Bluetooth sensors to calculate the travel time of vehicles.

Goodall [10] analyzed the nature of Wi-Fi re-identification technology and suggested using Wi-Fi sensors for low-speed and low-volume traffic. Gore et al. [2] investigated the performance of Wi-Fi sensors in vertical and horizontal positions for evaluating the efficiency of traffic data under Indian traffic conditions. The study revealed that the sensor location significantly influences the stream speed, time headway, and time-to-detection. Jang [11] used laser-sensor data to investigate the time headway characteristics on an interrupted traffic stream at various levels of traffic flow to model many stochastic distributions. Mahapatra and Maurya [12] used V-Box equipment and analyzed the lateral and longitudinal vehicle behavior of the mixed traffic on the highways in India. The study explored the effect of vehicle speed in the longitudinal direction on the yaw rate of the vehicles and studied the relationships with the longitudinal speed of different vehicle types. All these studies have attempted to analyze traffic characteristics using different technologies.

Another piece of a study by Isaac [13] reported that speed has a linear relation with lateral clearance at different speed levels. However, a later study by Pal and Chunchu [14] concluded that lateral clearance/gaps are a direct consequence of several other aspects other than the speed of the passing/overtaking vehicle, and it depends upon the vehicle type. Furthermore, increasing the degree of the model improves the fitting of the regression curve showing higher sensitivity to speed. Budhkar and Maurya [15] determined the lateral clearance for vehicles during overtaking in mixed traffic using an instrumented vehicle with ultrasonic sensors and GPS devices with cameras. The analysis findings indicated that the lateral clearance preserved by similar vehicle pairs is less than dissimilar vehicle pairs. Mallikarjuna et al. [16] elaborated image processing-based vehicle detection techniques using computer software for computing the lateral interaction between vehicles. However, this technique lack accuracy in detection due to heterogeneous traffic condition on Indian roads.

In India, due to the presence of abreast driving behavior and mixed traffic driving conditions, vehicles interact longitudinally and laterally with other vehicles [17].

Modeling traffic characteristics under non-lane-based and mixed traffic is a daunting task. Considerable research efforts have been directed toward analyzing traffic characteristics on different types of roadway facilities using various technologies. A little amount of research has been concentrated on analyzing traffic parameters such as speed, flow, density, TH, and LP of vehicles on multi-lane rural highways using IR sensor-based technologies. Also, it is understandable from past research works that the use of advanced data collection techniques for acquiring mixed traffic characteristics is limited, mainly in developing countries like India.

3 Study Framework

The field data on traffic characteristics was collected using the IR sensor-based device. Initially, the data were classified at flow increments of 500 Veh/h, ranging from 0–500 up to 1500–2000 Veh/h and at a density rate of 25 Veh/km increments, ranging from 25–50 up to 100–125 Veh/km. The DPCU was estimated using the speed-area ratio method for each vehicle category at varying traffic flow and density levels. The descriptive statistics were carried out, and several speed and TH distribution functions concerning the traffic flow and density rates were determined. The goodness of fit of the probability distributions was measured using the Kolmogorov–Smirnov (K–S) statistical test at a 5% level of significance for finding the best-fit distribution.

Besides, the association of LP of vehicles with respect to the speed, flow, and density was analyzed to evaluate their influence. The traffic speeds are classified into four groups with the size of the bin as 20 km/h, ranging from 20–40 up to 80–100 km/h. Similarly, the flow and density were also classified into four groups with the size of the bin as 500 Veh/h and 25 Veh/km, respectively. The range of flow adopted was from 0–500 up to 1500–2000 Veh/h, and the range of density adopted was from 25–50 to 100–125 Veh/km. The impact of increasing magnitude of speed, flow, and density on the vehicular LP was analyzed.

4 Data Collection

4.1 Study Section Details

The national highways (NH) in the southern part of India were chosen as the study sections based on the criteria that the test sections did not have any pedestrian crosswalks, median openings for making U-turning movement and were free from any other such side frictions. The data was collected for 12 h by setting up the IR sensors device on the four-lane divided highway sections. Extraction of the speed and volume data at every 15-min interval was carried because "the aggregation of vehicle count provides the realistic estimation of hourly traffic volume in heterogeneous traffic

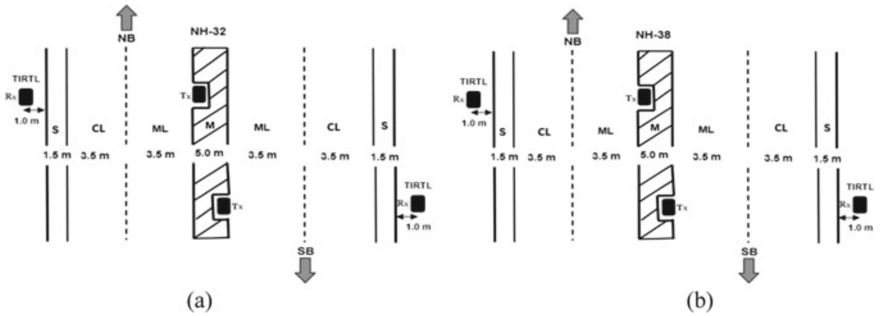
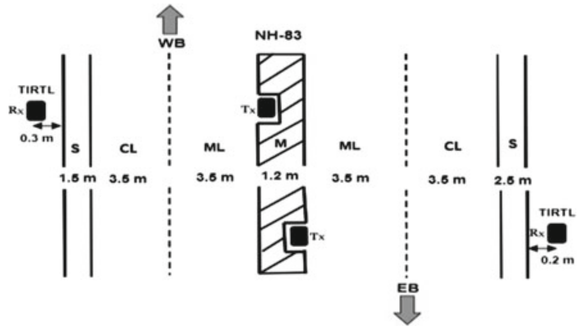


Fig. 1 Geometric features **a** NH-32. **b** NH-38

Fig. 2 Geometric features (NH-83)



flow conditions" [18]. Figures 1a, b and 2 depict the geometric features of NH-32, and NH-38, and NH-83, respectively.

4.2 Traffic Data Acquisition Using IR Sensor Device and Its Working Principle

The conventional traffic data collection methods, such as the video-graphic, pneumatic tubes, inductive loop detectors, etc., possess many limitations in their use by demanding various prerequisites. Also, these techniques are invasive and do not adapt to the mixed traffic scenarios peculiar to India. This difficulty can be overcome by using Intelligent Transportation Systems (ITS)-based new technological devices, such as the IR sensor devices. The IR sensor device, Transportable Infra-Red Traffic Logger (TIRTL), works on IR sensor-based technology to record traffic-related parameters such as speed, TH, LP, spacing, clearance, gap, CVC, volume, vehicle dimensions.

Both transmitter (TX) and receiver (RX), as shown in Fig. 3, are aligned so that they are less than 150 mm above the road surface, ensuring that the IR beams do

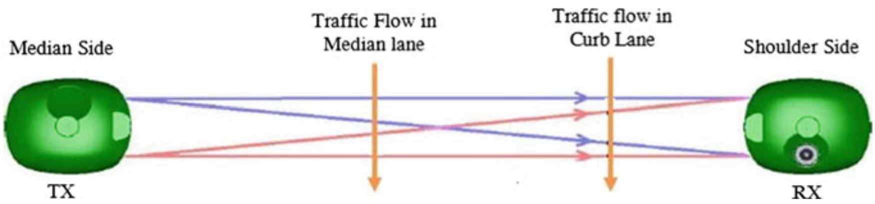


Fig. 3 IR lights transmission for traffic detection

not clip the vehicle bodies. The 12.0 V batteries are the power source for TX and RX. The TX is the IR beam source for traffic detection. On the other hand, the RX detects the disturbance caused by the passing vehicle wheels. The RX is connected to a laptop device through an RS232 serial port to access the system interface and store the data file in .csv format.

For any detection, the four beam events generate eight timestamps, and the vehicle travel direction is determined by the order of occurrence of beam events. The use of the "free-flowing" condition option optimizes vehicle detection for different traffic conditions. The make and break beam events are measured as time intervals to determine vehicle positions. Hence the LP of the vehicles is recorded. The field snapshot of the IR sensor device setup is shown in Fig. 4.

Accuracy of the IR sensor device. The traffic during the data collection was videotaped using a video camera for 1 h. It covered a 50 m trap length to capture the actual traffic movement. The speed, vehicle classification, TH, and LP of vehicles were extracted manually from the recorded videos and compared with the IR sensor device recorded data. An accuracy match of 96% for vehicle speed and 94% for CVC was obtained. The TH of vehicles was manually extracted at microscopic levels from the recorded

Fig. 4 Field snapshot of IR sensor device set up across the highway segment

Table 1 Average dimensions of the vehicle

Class of vehicle	Length (m)	Width (m)	Area (m^2)
2-wheeler (2W)	1.65	0.76	1.25
3-wheeler (3W)	2.74	1.25	3.43
Small car (SC)	3.84	1.45	5.57
Big car (BC)	4.60	1.80	8.28
Light commercial vehicle (LCV)	5.62	1.52	8.54
Medium commercial vehicle (MCV)	8.28	2.20	18.22
Heavy commercial vehicle (HCV)	9.10	2.30	20.93
Multi-axle vehicle (MAV)	12.62	2.42	30.54

videos, and the match percentage with the IR sensor device recorded TH was found to be 96%. As far as the LP of the vehicles is concerned, the collected IR sensor device data revealed a 95% accuracy match with the videotaped data. Other studies achieved up to an accuracy of 99% for the vehicle's speed [19] and 94–97% for classified vehicle count [20].

4.3 Vehicular Dimensions

The vehicles detected by the IR sensor are compared with the pre-loaded vehicle type scheme (up to 21 standard classes specific to Indian road conditions) for describing the vehicle class. The recorded classification is displayed in the user interface once all the axles of the vehicles cross the IR detection zone. Table 1 shows the average physical dimensions of the vehicle classes.

4.4 Traffic Composition at the Study Sections

The mainstream traffic is composed of eight different vehicle categories. The large-sized new generation cars of more than 5.57 m^2 area were categorized as BC, and the others were categorized as SC. LCV included light and small commercial vehicles. MCV included the two-axle buses/trucks, while the HCV included three-axle buses/trucks. The vehicles with four-axles, five-axles, and six-axles are categorized as one vehicle type as MAV. Figure 5 illustrates the vehicle composition of different vehicle classes.

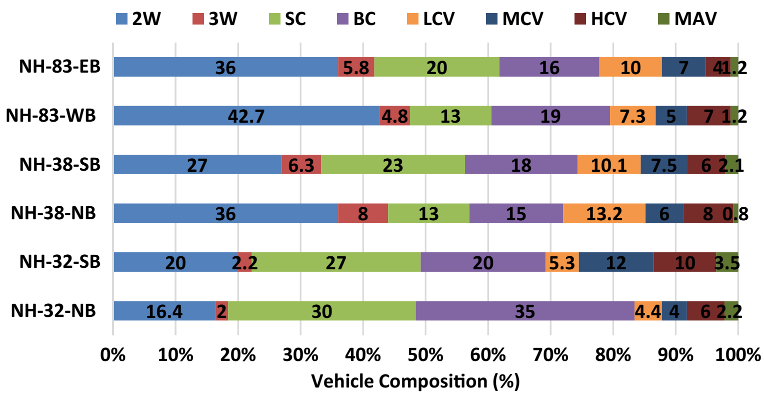


Fig. 5 Traffic composition for the subject study sections

4.5 *Traffic Flow Variation at the Study Sections*

The traffic flow variation on an hourly basis at the highway sections was examined to measure the peak and off-peak hours of vehicle traffic. It can be observed from Fig. 6 that the recorded volume of traffic is nearly the same at all sites except in the NH-32-NB section, which has the highest volume of traffic almost every hour. The hourly variation of traffic at all the subject study sections is depicted in Fig. 6.

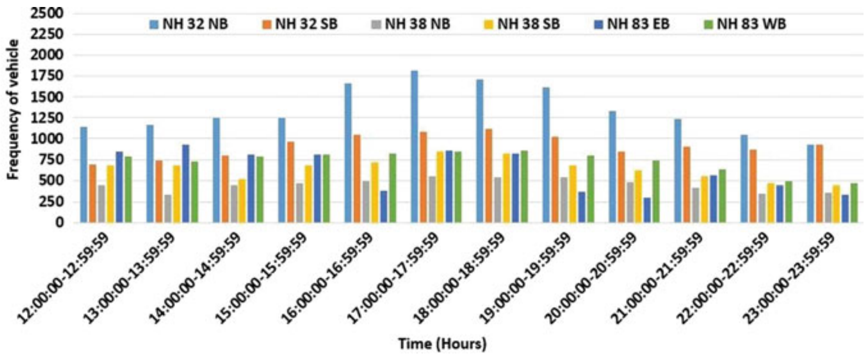


Fig. 6 Traffic variation at the subject study sections

5 Dynamic Passenger Car Unit (DPCU) Determination

The PCU values are typically assigned to homogenize the traffic [21]. The conventional PCE values suggested by HCMs cannot be used at different traffic flow conditions [22]. Additionally, the Indian Roads Congress-IRC:64 [23], the standard code provides single and static PCU values for the different ranges of traffic flow and density. Hence, the DPCU factors have been evaluated by the speed-area ratio model suggested Indo-HCM: 2017 [24], which is shown in Eq. (1).

$$\mathrm{DPCU}_i = (V_c/V_i)/(A_c/A_i) \tag{1}$$

The DPCU values estimated for different traffic flow and density levels, as presented in Tables 2 and 3, respectively.

Tables 2 and 3 indicate that the DPCU of 2Ws, 3Ws, and BI decrease with the traffic flow and density increase. Conversely, the DPCU increases for BC, LCV, MCV, HCV, and MAV when the traffic flow and density increase. This was due to

Table 2 DPCU at different traffic flow levels

Traffic flow level (Veh/h)					
Class of vehicle	0–500	500–1200	1000–1500	1500–2000	Average DPCU
2W	0.35	0.33	0.28	0.25	0.30
3W	1.18	1.12	1.05	1.00	1.09
SC	1.00	1.00	1.00	1.00	1.00
BC	1.48	1.55	1.59	1.63	1.56
LCV	2.75	3.08	3.36	3.50	3.17
MCV	3.51	3.82	4.01	4.46	3.95
HCV	3.65	3.91	4.22	4.76	4.14
MAV	6.11	6.30	6.65	6.91	6.49

Table 3 DPCU at different traffic density levels

Traffic density level (Veh/km)					
Class of vehicle	25–50	50–75	75–100	100–125	Average DPCU
2W	0.38	0.35	0.29	0.25	0.32
3W	1.15	1.13	1.12	1.05	1.11
SC	1.00	1.00	1.00	1.00	1.00
BC	1.44	1.52	1.54	1.59	1.52
LCV	2.47	2.82	3.57	3.49	3.09
MCV	3.28	3.77	4.18	4.44	3.92
HCV	3.78	3.99	4.47	4.61	4.21
MAV	6.2	6.38	6.59	6.75	6.48

Table 4 DPCU values for the different types of vehicles as per Indo-HCM [24]

Class of vehicle	Indo-HCM [24] DPCU value	
	Range	Median
2W	0.3–0.5	0.4
3W	1.1–1.3	1.2
SC	1.00	1.00
BC	1.4–1.5	1.45
LCV	2.7–3.3	3.1
MCV	3.5–4.6	4.4
HCV	3.5–4.6	4.4
MAV	6.3–7.0	6.6

the inferior operation of these vehicles compared to an SC, 3W, and 2W, which have higher speeds. In both cases, the average DPCU for each type of vehicle was found to be logical and reasonably acceptable compared to the Indo-HCM [24] DPCU values, as presented in Table 4.

6 Analysis of Speed Data

The use of speed data without considering traffic composition, traffic flow, and density may yield skewed and inconsistent outcomes [25]. Further vehicle speed is affected by overall density, which includes individual vehicle densities [26] that are required for the development of macroscopic traffic models, microscopic traffic characteristics evaluation, capacity determination, and level of service analysis. Concerning this aspect, the study considers conducting the descriptive statistical analysis and finding the best-fitted distribution function for the traffic speed at varying flow and density levels.

6.1 Descriptive Statistics of Speed

The descriptive statistic of speed provides an understanding of the vehicle's quality of service. Table 5 summarizes the descriptive statistics on the speed at different traffic flow and density levels.

From Table 5, it can be observed that the vehicles have different minimum speed, maximum speed, mean speed, range value, mode value, standard deviation (SD), coefficient of variation (CV), standard error (SE), skewness, kurtosis, and percentile speeds over the varying range of traffic flow and density. The mean value of traffic speed under high traffic flow decreased from 57 to 38 km/h, which represents a reduction of 33.3%. The mean value of traffic speed under high traffic density decreased

Table 5 Descriptive statistics for speed data

Traffic characteristics	Flow range (Veh/h)				Density range (Veh/km)			
Statistical parameters	0–500	500–1000	1000–1500	1500–2000	25–50	50–75	75–100	100–125
Minimum	35	30	25	23	35	30	29	28
Maximum	81	78	72	73	82	76	72	70
Range	46	48	47	50	47	46	43	42
Mean	57	44	40	38	60	56	52	49
Mode	50	49	45	32	55	52	50	48
SD	12.4	12.6	13.1	14.4	13.6	14.5	15.4	16.1
CV	0.28	0.30	0.38	0.46	0.23	0.39	0.48	0.56
SE	0.20	0.16	0.14	0.21	0.27	0.24	0.21	0.14
Skewness	4.28	5.19	6.02	6.91	3.57	3.86	4.25	4.53
Kurtosis	3.14	3.25	4.65	4.97	4.58	4.95	5.89	6.41
15th percentile	40	36	28	26	38	32	30	27
50th percentile	56	42	38	36	58	54	51	47
85th percentile	75	73	68	67	78	74	70	67

from 60 to 49 km/h, which represents a reduction of 18.3%. It is observed that the speed gets affected because of the interaction of different classes of vehicles using the same road space. The higher the traffic flow and density, the lower is the speed value.

6.2 Probability Distributions for Speed Data

Conventionally, normal distribution defines the vehicle speed under homogeneous traffic conditions, but the distributions differ considerably under heterogeneous traffic conditions [22]. To assess the best-fitted distribution, the Kolmogorov–Smirnov (K–S) test was used with a 5% significance level. The distribution functions for speed are represented in Table 6.

The speed at a low traffic flow of 0–500 Veh/h exhibits Log-normal distribution, whereas, at high traffic flow of 1500–2000 Veh/h, the speed follows the Erlang distribution. Further, Table 6 shows that at a low traffic density of 25–50 Veh/km, the vehicle speed follows the Log-logistic distribution, whereas at a high traffic density of 100–125 Veh/km, the speed follows the Inverse-Gaussian distribution. The distribution functions varied for the vehicle speeds at several traffic flow rates and density levels because of variation in vehicles' static and dynamic characteristics

Table 6 Probability distribution functions for speed data

Traffic flow (Veh/h)	Type of distribution	K–S value	*p*-value	Best-fitted distribution	Traffic density (Veh/km)	Type of distribution	K–S value	*p*-value	Best-fitted distribution
0–500	Log-normal	0.032	0.941	Log-normal	25–50	Log-logistic	0.224	0.959	Log-logistic
	Beta	0.042	0.934			Beta	0.334	0.942	
	Log-logistic	0.054	0.920			Burr	0.568	0.912	
500–1200	Log-logistic	0.124	0.965	Log-logistic	50–75	Burr	0.347	0.954	Burr
	Gamma	0.322	0.953			Pearson 5	0.468	0.941	
	Weibull	0.392	0.941			Weibull	0.689	0.902	
1000–1500	Beta	0.155	0.941	Beta	75–100	Weibull	0.754	0.933	Weibull
	Weibull	0.184	0.883			Exponential	0.896	0.915	
	Log-Pearson III	0.210	0.875			Inverse-Gaussian	0.922	0.902	
1500–2000	Erlang	0.147	0.988	Erlang	100–125	Inverse-Gaussian	0.154	0.927	Inverse-Gaussian
	Log-logistic	0.171	0.972			Beta	0.247	0.910	
	Burr	0.189	0.966			Erlang	0.368	0.894	

and abreast driving behavior under prevailing traffic conditions. This has confirmed that speeds do not follow a particular distribution under different traffic conditions.

7 Analysis of Time Headway Data

7.1 Descriptive Statistics for TH Data

The TH data is the basis for building and analyzing the microscopic traffic simulation models. Table 7 summarizes the descriptive statistics on TH at different traffic flows and densities.

From Table 7, it can be observed that the TH varies with changes in traffic flow and density. Hence proving that TH is not a constant value. In general, the TH of vehicles typically decreased as the traffic flow and density values increased. From Table 7, it can be observed that because of the difference in the dynamic features of the vehicles, the TH has a different minimum speed, maximum speed, mean speed, range value, mode value, SD, CV, SE, skewness, kurtosis, and 15th, 50th, and 85th percentile speeds over the varying traffic flows and densities. The mean value of TH under high traffic flow decreased from 8.1 to 5.8 s, which represents a reduction of

Table 7 Descriptive statistics for TH data

Traffic characteristics	Flow range (Veh/h)				Density range (Veh/km)			
Statistical parameters	0–500	500–1000	1000–1500	1500–2000	25–50	50–75	75–100	100–125
Minimum	5.7	4.0	3.5	2.9	5.8	5.5	4.6	3.9
Maximum	16.4	14.5	13.2	12.8	12.6	10.2	8.6	7.1
Range	10.7	10.5	9.7	9.9	6.8	4.7	4.0	3.2
Mean	8.1	7.5	6.6	5.8	6.7	6.1	5.6	4.7
Mode	7.5	6.9	6.2	5.6	6.2	5.8	5.4	4.2
SD	7.21	8.48	9.02	9.77	5.68	6.52	7.41	8.47
CV	1.12	2.04	2.96	3.84	1.25	2.11	2.45	2.97
SE	0.11	0.11	0.10	0.15	0.11	0.11	0.10	0.13
Skewness	4.35	5.19	5.62	5.94	3.47	3.89	4.12	4.55
Kurtosis	2.20	3.14	3.78	4.12	2.78	2.99	3.47	3.79
15th percentile	6.2	4.8	3.9	3.6	6.0	5.8	4.9	4.1
50th percentile	7.8	7.2	6.3	5.7	6.5	6.2	5.7	4.8
85th percentile	12.2	10.5	9.8	8.6	9.4	8.5	7.7	5.9

28.4%. Meanwhile, the mean value of TH under high traffic density decreased from 6.7 to 4.7 s, which represents a reduction of 29.9%. It can be inferred that the TH is significantly influenced by the variation in the flow and density from a low level to a high level.

7.2 Probability Distributions for TH Data

The TH probability distributions were identified from the vehicle arrival pattern recorded in the field data. The K–S test results showed various probability distribution functions as the best-fitted distribution for TH at different traffic flow and density levels, which is represented in Table 8. It shows that the TH of vehicles follows the Exponential distribution when the traffic flow is as low as 0–500 Veh/h, whereas the TH of vehicles follows the Gumbel-max distribution when the traffic flow is as high as 1500–2000 Veh/h.

During high traffic flows, the repeated platoon formations on the highway by slow-moving vehicles increase the composition of following vehicles, thereby forming shorter headways [27]. This eventually identifies the form of distribution for this degree of traffic flow. Further, it can be observed that the TH of vehicles follows the Weibull distribution when the density is as low as 25–50 Veh/km, whereas the TH of vehicles follows the Log-logistic distribution when the density is as high as 100–125 Veh/km. Finally, it can be seen from the non-parametric test results of the distribution that the TH varies for every density level based on the traffic conditions.

8 Analysis of Lateral Placement of Vehicles

Lateral movements of vehicles in the traffic streams have a major effect on the traffic flow [28]. The vehicle's lateral interaction is influenced by the vehicle type, the speed of the vehicle, the behavior of the driver [29], and area-occupancy [30]. In this study, the LP of a vehicle is the distance between the center of the vehicle and the edge of the roadway on the curbside. The edge of the curbside of the roadway was chosen as a common reference point for the measurement of the LP of vehicles.

8.1 Frequency Distribution Analysis

The lane position of the vehicles on the carriageway is influenced by traffic characteristics [31]. The analysis of the frequency of LP of vehicles shows that many vehicles travel on the center of the road (between 2 and 7 m of the carriageway) as the vehicles try to maintain a safer longitudinal and lateral distance with the other

Table 8 Probability distribution functions for TH data

Traffic flow (Veh/h)	Type of distribution	K–S value	*p*-value	Best-fitted distribution	Traffic density (Veh/km)	Type of distribution	K–S value	*p*-value	Best-fitted distribution
0–500	Exponential	0.478	0.991	Exponential	25–50	Weibull	0.114	0.992	Weibull
	Log-normal	0.322	0.984			Erlang	0.176	0.974	
	Log-Pearson III	0.305	0.981			Burr	0.198	0.961	
500–1000	Gamma	0.216	0.964	Gamma	50–75	Beta	0.151	0.984	Beta
	Erlang	0.255	0.948			Erlang	0.186	0.977	
	Log-normal	0.277	0.934			Log-logistic	0.194	0.971	
1000–1500	Log-Pearson III	0.147	0.909	Log-Pearson III	75–100	Pearson 5	1.887	0.987	Pearson 5
	Log-logistic	0.161	0.891			Erlang	1.964	0.942	
	Pearson 5	0.188	0.872			Burr	1.992	0.968	
1500–2000	Gumbel-Max	0.115	0.974	Gumbel-Max	100–125	Log-logistic	1.511	0.991	Log-logistic
	Beta	0.126	0.955			Pearson 5	1.644	0.984	
	Weibull	0.145	0.950			Gamma	1.667	0.979	

interacting vehicles [32]. The frequency distribution of the LP of vehicles is shown in Fig. 7.

The analysis of the frequency of speed of vehicles shows that most of the vehicles travel, maintaining a moderate speed (30–40 km/h) to high speeds (60–80 km/h) and very high speeds (90–110 km/h) because of their maneuverability at nearly free-flow conditions. The frequency distribution for the speed of vehicles is shown in Fig. 8.

The speed and the LP of the vehicles have been found to follow the increasing nature of the relationship. Figure 9 indicates that the speed of vehicles increases as the vehicles shift toward the center of the highway sections. This may be reasoned to be because of the high speed vehicles wanted to have ease of maneuverability and lesser interaction with other vehicles.

As seen in Fig. 9, it can also be said in another way that the LP of vehicles increases as the speed of vehicles increases. Additionally, it can be noted that the observed LP maintained by the vehicles depends on the speed of the vehicles. Therefore, variation in traffic speed causes variation in the LP of vehicles and vice-versa.

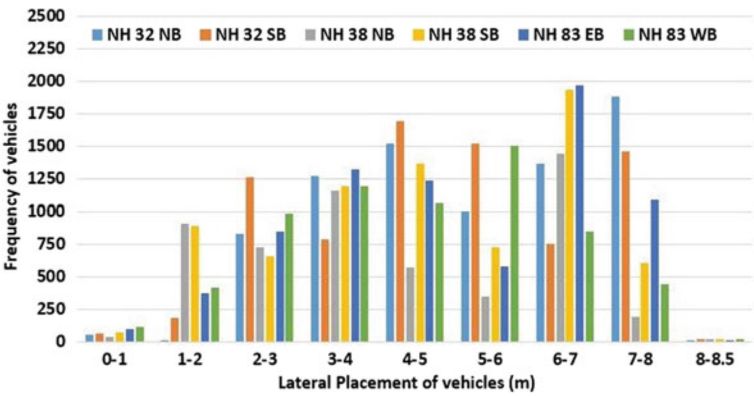


Fig. 7 Frequency distribution of the LP of vehicles

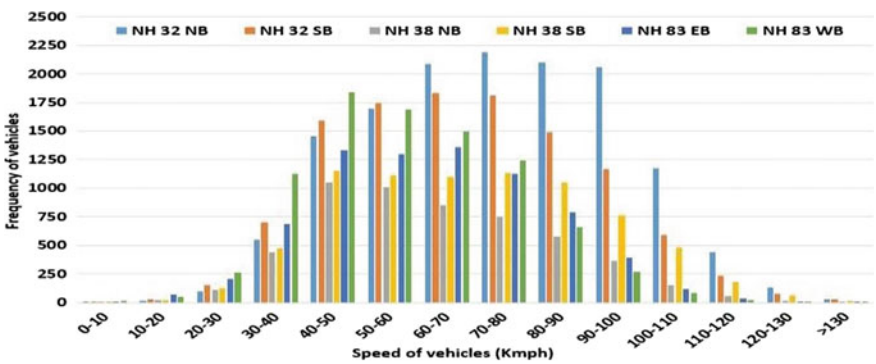


Fig. 8 Frequency distribution of the speed of vehicles

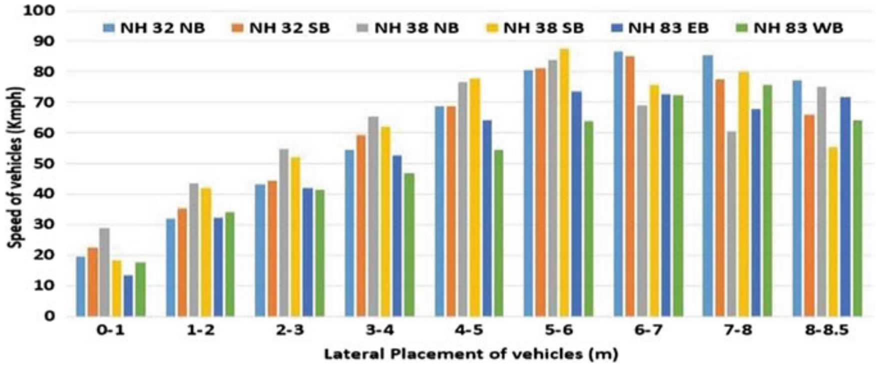


Fig. 9 Variation of speed with respect to the LP of vehicles

8.2 Descriptive Statistical Analysis of LP of Vehicles

The observed vehicle speeds were classified at an increment of 20 km/h to determine the frequency distribution of the LP of vehicles. Speeds were assumed as 20–40, 40–60, 60–80, and 80–100 km/h. Similarly, the observed traffic flows were classified at 500 Veh/h increments as 0–500, 500–1000, 1200–1500, and 1500–2000 Veh/h. Also, the traffic density levels were classified at an increment of 25 Veh/km, ranging from 25–50 to 100–125 Veh/km. To comprehend the effect of speed, flow, and density on the LP of vehicles, descriptive statistics have been carried out, which are provided in Table 9.

From Table 9, it was interpreted that the mean vehicle LP at the lowest speed level is 2.2 m, and at the highest speed level is 4.6 m. Similarly, the mean vehicle LP at the lowest flow level is 1.2 m, and at the highest flow level is 4.2 m, while the mean LP of vehicles at the lowest density level is 1.4 m and at the highest density level is 4.6 m. The LP of vehicles was skewed toward the curbside of the road at low traffic speed, flow, and density range. As the speed, flow, and density of traffic increase, the LP of vehicles shifts from the curbside of the road to the median side. Because under free-flowing traffic conditions, the speed, flow, and density increase, and the lane occupancy of the vehicles in the median lane increases. The slow-moving vehicles alone are forced to occupy the curbside of the roadway. It can, therefore, be said that as the mean speed, flow, and density increase, the LP of vehicles increases.

From Table 9, it was interpreted that at the lowest speed range of 20–40 km/h, the minimum and maximum vehicle LP lies between 1.5 m and 2.7 m, respectively. While at the highest speed range of 80–100 km/h, the minimum and maximum vehicle LP lies between 2.7 m and 5.4 m, respectively. Similarly, at the lowest and highest flow ranges, the minimum and maximum LP of vehicles are 0.5 m and 4.8 m, respectively, while at the lowest and highest density ranges, the minimum and maximum LP of vehicles are 0.8 m and 5.1 m, respectively. Table 9 shows that as the speed, flow, and density range increases, the mean LP of the vehicles also increases. As the

Table 9 Descriptive statistics for LP of vehicles

Traffic characteristics	Speed range (km/h)				Flow range (Veh/h)				Density range (Veh/km)			
Statistical parameters	20–40	40–60	60–80	80–100	0–500	500–1000	1000–1500	1500–2000	25–50	50–75	75–100	100–125
Minimum	1.5	1.8	2.5	2.7	0.5	1.7	2.5	2.8	0.8	1.8	2.8	3.2
Maximum	2.7	3.9	5.0	5.4	1.7	3.1	4.2	4.8	1.6	3.2	4.6	5.1
Range	1.2	2.1	2.5	2.7	1.2	1.4	1.7	2.0	0.8	1.4	1.8	1.9
Mean	2.2	3.4	4.2	4.6	1.2	2.5	3.4	4.2	1.4	2.6	3.8	4.6
Mode	2.0	2.2	3.2	4.0	1.0	2.1	2.3	2.4	1.1	2.2	2.6	3.3
SD	1.01	2.18	3.56	3.87	0.82	1.57	2.42	4.45	1.86	2.84	3.57	3.80
CV	0.46	0.64	0.85	0.84	0.68	0.63	0.71	1.06	1.33	1.89	2.94	3.83
SE	0.02	0.03	0.04	0.05	0.01	0.02	0.03	0.07	0.04	0.05	0.05	0.06
Skewness	−0.24	0.49	0.72	0.95	0.48	0.69	1.62	1.94	−0.88	−1.58	1.73	2.34
Kurtosis	−1.15	1.24	1.65	1.76	1.10	−1.46	1.75	2.17	1.62	−1.90	2.43	3.11
15th percentile	1.6	2.0	2.6	3.2	0.7	1.9	2.2	3.0	0.9	2.1	3.0	3.4
50th percentile	1.9	3.3	3.8	4.1	1.1	1.8	2.8	3.9	1.2	2.4	3.5	4.2
85th percentile	2.5	3.7	4.7	5.1	1.3	2.8	3.8	4.4	1.5	2.8	4.4	4.7

highway's speed, flow, and density increase, the vehicles keep a more considerable distance from the edge of the road to avoid road crashes.

The SD and CV of the LP of vehicles were less for reduced speed, flow, and density ranges, while the SD and CV of the LP of vehicles were high for increased speed, flow, and density ranges. The SE, skewness, and kurtosis are seen to increase as traffic speed, flow, and density increase. Therefore, it can be said that stream speed, flow, and density influence the LP of vehicles. Moreover, the observed variation in the 15th, 50th, and 85th percentile for the LP of vehicles shows an increment with the increase in speed, flow, and density. Hence for safety reasons, vehicles maintain a more considerable distance from the interacting vehicles. This can be attributed to the drivers' propensity to travel at the optimal road speed for safety purposes, retaining relatively high speed limits.

9 Results and Discussion

The analyzed statistical data indicates that the distribution of the likelihood, followed by vehicle speed and TH, varies with changes in traffic flow and density rates. The study indicated distinct and significant variations, even with respect to the LP of vehicles. These results support the argument that the speed and TH analysis should consider traffic characteristics separately at the different flow and density levels. Therefore, it is important to explicitly consider the variations in the speed and TH distribution functions across the different flow and density levels when the microscopic simulation models are developed for analysis. More specifically, in calibration and validation of the generation of vehicle car-following models and lane changing algorithms. The LP of the vehicles increased with the increase in stream speed, flow, and density. It was found that at different traffic rates, the LP of the vehicles changes according to the traffic condition. Finally, the movement of the vehicle by maintaining a steady speed and desired TH, and optimal LP, is observed to be solely reliant on the traffic conditions.

The present study contributes to the existing literature by analyzing the traffic parameters under different mixed traffic conditions. The authors foresee that the researched traffic data from the study will aid in developing microscopic simulation models for various traffic conditions and can be used to calibrate and validate the build models. The results from this research provide significant insights into the extensive analysis of traffic variables for developing realistic simulation tools and human-like self-driving vehicles under varying traffic conditions on non-lane-based and heterogeneous highways.

Further, the study has explored and provided the working and technological usefulness of IR sensor devices for collecting traffic data under the different traffic conditions prevailing in India. Previous research used different methods for gathering traffic data; however, as per the authors' knowledge, no study has reported using IR sensor devices to capture the spatio-temporal characteristics of the traffic effectively and efficiently under mixed traffic environments. Additionally, the present findings can assist in managing and controlling traffic to improve the efficiency of the highway traffic system, particularly in developing countries such as India, where ITS-based solutions are currently being looked upon.

10 Conclusion

This study uses a huge amount of traffic data collected to estimate the DPCU values for vehicle types at different traffic flow and density levels. This research has studied the variations in the speed and TH distribution functions at various traffic flow and density rates and has analyzed them. The study also seeks to identify and evaluate the impact of traffic parameters, such as speed, flow, and density, on the LP of vehicles. The distribution of speed and TH computed for different traffic flow and density levels shows a difference in the probability distribution patterns. Some noteworthy variations with regard to the speed and TH distributions were also observed. Furthermore, the LP of vehicles was also found to be significantly affected by variations in speed, flow, and density. This study could benefit traffic engineers and highway researchers to operate, control, and manage the NH in the future paradigm. The different traffic characteristics regarding similar and dissimilar specific-vehicle-type leader–follower pairs were not considered in this study, which could be examined extensively in future research.

References

1. Singh S, Moses Santhakumar S (2021) Empirical analysis of impact of multi-class commercial vehicles on multi-lane highway traffic characteristics under mixed traffic conditions. Int J Transp Sci Technol. https://doi.org/10.1016/j.ijtst.2021.07.005
2. Gore N, Arkatkar S, Joshi G et al (2019) Exploring credentials of Wi-Fi sensors as a complementary transport data: an Indian experience. IET Intell Transp Syst 13(12):1860–1869
3. Zhang G, Wang Y, Wei H et al (2007) Examining headway distribution models with urban freeway loop event data. Transp Res Rec 1999:141–149
4. Jang J, Kim B, Choi N et al (2011) Analysis of time headway distribution on Korean multi-lane highway using loop event data. J East Asian Soc Transp Stud 9:1447–1457
5. Gunay B, Erdemir E (2011) Lateral analysis of longitudinal headways in traffic flow. Int J Eng Appl Sci 3(2):90–100
6. Araghi BN, Olesen JM, Krishnan R et al (2013) Reliability of Bluetooth technology for travel time estimation. J Intell Transp Syst Technol Plann Oper 19(3):240–255
7. Bhaskar A, Chung E (2013) Fundamental understanding on the use of Bluetooth scanner as a complementary transport data. Transp Res Part C Emerg Technol 37:42–72
8. Brennan TM, Ernst JM, Day CM et al (2010) Influence of vertical sensor placement on data collection efficiency from Bluetooth MAC address collection devices. J Transp Eng Part A Syst 136(12):1104–1110
9. Li Q, Chen L, Huang Y et al (2017) Estimation of travel time based on Bluetooth MAC address identification. In: Information technology and intelligent transportation systems. Advances in intelligent systems and computing, vol 454, pp 645–651. https://doi.org/10.1007/978-3-319-38789-5_71

10. Goodall NJ (2016) Fundamental characteristics of Wi-Fi and wireless local area network re-identification for transportation. IET Intell Transp Syst 11(1):37–43. https://doi.org/10.1049/iet-its.2016.0087
11. Jang J (2012) Analysis of time headway distribution on suburban arterial. KSCE J Civ Eng 16:644
12. Mahapatra G, Maurya AK (2013) Study of vehicles lateral movement in non-lane discipline traffic stream on a straight road. Procedia Soc Behav Sci 104:352–359
13. Isaac KP (1995) Studies on mixed traffic flow characteristics under varying composition. Ph.D. thesis, Bangalore University, India
14. Pal D, Chunchu M (2019) Effect of vehicular lateral gap models on heterogeneous traffic stream behaviour. Transp Lett 12(6):399–407
15. Budhkar AK, Maurya AK (2017) Characteristics of lateral vehicular interactions in heterogeneous traffic with weak lane discipline. J Mod Transp 25(2):74–89. https://doi.org/10.1007/s40534-017-0130-1
16. Mallikarjuna C, Tharun B, Pal D (2013) Analysis of the lateral gap maintaining behavior of vehicles in heterogeneous traffic stream. In: 2nd conference of the transportation research group, Agra
17. Mahapatra G, Maurya AK (2015) Study on lateral placement and speed of vehicles under mixed traffic condition. In: Proceeding of conference of Eastern Asia society for transportation studies, Cebu City
18. Khan T, Mohapatra SS (2020) Lateral placement characteristics of U-turning vehicles: a statistical investigation. Transp Dev Econ 6(3)
19. An assessment of the uncertainty achieved by the CEOS TIRTL calibrator (2004) National Measurement Institute, Clayton
20. Kotzenmacher J, Erik M, Bingwen H (2005) Evaluation of portable non-intrusive traffic detection system. Technical report, MN-RC-2005-37. Minnesota Department of Transportation, Research Services Section, pp 29–34. https://www.pooledfund.org/documents/TPF-5_073/final_report.pdf
21. Singh S, Rajesh V, Moses Santhakumar S (2022) Effect of Mixed Traffic Platooning by Commercial Vehicle Types on Traffic Flow Characteristics of Highways. Periodica Polytechnica Transportation Engineering. https://doi.org/10.3311/PPtr.18200
22. Roh CG, Park BJ, Kim J (2017) Impact of heavy vehicles on highway traffic flows: case study in the Seoul metropolitan area. J Transp Eng Part A Syst 143(9):05017008
23. IRC-64 (1990) Indian roads congress: guidelines for capacity of urban roads in plain areas. Indian Roads Congress, New Delhi
24. Indo-HCM (2017) Indian highway capacity manual. Central Road Research Institute (CRRI), New Delhi
25. Singh S, Moses Santhakumar S (2022) Modeling traffic parameters accounting for platoon characteristics on multilane highways. Transp Dev Eco. https://doi.org/10.1007/s40890-022-00166-3
26. Patkar M, Dhamaniya A (2020) Influence of nonmotorized vehicles on speed characteristics and capacity of mixed motorized traffic of urban arterial midblock sections. J Transp Eng Part A Syst 146(4):04020013. https://doi.org/10.1061/JTEPBS.0000325
27. Singh S, Moses Santhakumar S (2022) Assessing the impacts of heavy vehicles on traffic characteristics of highways under mixed traffic platooning conditions. Eur Transp\Trasp Eur. https://doi.org/10.48295/ET.2022.86.3
28. Singh S, Vidya R, Panda RK, Moses Santhakumar S (2021) A study on characteristics of vehicular lateral position on rural highways. In: The international conference on community based research and innovations in civil engineering (CBRICE-2021), Manipal University, Jaipur. IOP Conf Ser Earth Environ Sci 796(1):012046. IOP Publishing. https://doi.org/10.1088/1755-1315/796/1/012046
29. Kotagi PB, Raj P, Asaithambi G (2019) Modeling lateral placement and movement of vehicles on urban undivided roads in mixed traffic: a case study of India. J Transp Eng Part A Syst

30. Singh S, Vidya R, Shukla BK, Moses Santhakumar S (2021) Analysis of traffic flow characteristics based on area-occupancy concept on urban arterial roads under heterogeneous traffic scenario—a case study of Tiruchirappalli city. In: Mehta YA et al (eds) Advances in water resources and transportation engineering. Lecture notes in civil engineering, vol 149. Springer Nature, Singapore, pp 69–84. https://doi.org/10.1007/978-981-16-1303-6_6
31. Singh S, Moses Santhakumar S (2021) Evaluation of lane-based traffic characteristics of highways under mixed traffic conditions by different methods. J Inst Eng (India) Ser A 102(3):719–735. https://doi.org/10.1007/s40030-021-00549-6
32. Singh S, Moses Santhakumar S (2021) Platoon-based impact assessment of heavy-duty vehicles on traffic stream characteristics of highway lanes under mixed traffic environment. 20, 29–45. Int J Intell Transp Syst Res. https://doi.org/10.1007/s13177-021-00268-z

Development of Car-Following Models with Multiple Leader Vehicles for Mixed Traffic Conditions in Urban Areas

Madhu Errampalli, Himanshu Verma, and Nisha Radhakrishnan

Abstract For quick and efficient implementation of transport policies, the assessment of traffic conditions with high precision techniques namely, microscopic simulation models is important. However, the models that are developed for homogeneous traffic conditions would not yield realistic estimations for Indian conditions which are highly heterogeneous and lane indiscipline. Further car-following model being core part of simulation considers single leader vehicle which is not the case especially in mixed traffic condition. In the present study, an attempt has been done to consider all the surrounding vehicles as influencing variables in developing car-following model apart from driver's behaviour. For this, vehicle trajectory data consist of time wise positions, speed and acceleration has been extracted from the video graphic data collected. Multiple linear regression models are developed for three vehicles types, i.e. cars, two-wheelers and three-wheelers and subsequently compared with General Motor car-following using RSME and MAPE values. It was found that developed model is able to reduce these values by 44% and 26%, respectively, for cars similar observation are observed for two-wheeler and three-wheeler which demonstrate the suitability for the mixed traffic condition. Further the developed model is implemented in a microscopic traffic simulation to study Impact of traffic composition of heavy vehicles on road capacity and free speed.

Keywords Heterogeneous traffic · Microscopic characteristics · Simulation · Car-following · Traffic composition

M. Errampalli (✉)
TPE Division, CSIR-Central Road Research Institute (CRRI), Mathura Road, New Delhi 110025, India
e-mail: madhu.crri@nic.in

H. Verma · N. Radhakrishnan
Department of Civil Engineering, National Institute of Technology (NIT), Tiruchirappalli, Tamil Nadu 620015, India

L. Devi et al. (eds.), *Proceedings of the Sixth International Conference of Transportation Research Group of India*, Lecture Notes in Civil Engineering 272,
https://doi.org/10.1007/978-981-19-3494-0_7

1 Introduction

Due to rapid development in there is an increase in urban population, but most of the urban areas are not well equipped infrastructure wise to manage this increase this has led to problem of traffic congestion, the solution of this problem is tricky because increasing the traffic infrastructure rapidly may increase the number of new vehicles on road which will increase the congestion and travel time. Introduction of public transport is another solution but it has its own challenges of last mile connectivity which make it less attractive for commuters. So, to tackle this problem there is a need to formulate new traffic management policies. But, implementing these policies is more difficult, as it needs assessment of these policies before implementing them.

In general, the impacts of transportation policies are assessed by estimating vehicular behaviour under given conditions. The standard assessment approaches may wrongly estimate vehicular behaviour due to the limitations in traffic models. To solve these challenges, effective technique that estimate the true behaviour of the driver are necessary. The traffic simulation methodology is commonly used to study vehicular movements on a road network. Microscopic traffic simulation is recognised as effective tools for assisting with transportation feasibility studies [1]. The microscopic simulation models give realistic result because they take into account of individual vehicle behaviour, which. Car-following model is critical and has a significant impact on the accuracy of simulation model predictions. Though microscopic simulation seems to be the most appropriate for evaluating the policies, its accuracy and validity are mostly determined by the accuracy of the underlying microscopic models in the simulation model.

In India people driving behaviour is bit aggressive most of people does not follow lane discipline, and travel very closely to the leader vehicle which is unsafe and may lead to rare end collision [1]. The mixed traffic condition on Indian roads has made it difficult for microscopic simulation to predict the results accurately because most of the models are developed for homogenous traffic condition with lane discipline. In light of the foregoing, a car-following model is proposed in this study that takes into account mixed traffic situations, variable driver behaviour and lack of lane discipline.

1.1 Car-Following Model

Car-following model is mainly used to describe the movement of following vehicle (FV) based on leader vehicle (LV) characteristics and behaviour. This model represents interaction between two consecutive vehicles in a reasonably congested traffic stream. The car-following model formulated by General Motors (GM) laboratories is considered to be the standard car-following model [1], in which, acceleration is considered as a function of relative speed of FV with LV and sensitivity. The equation is given below:

$$a_{n+1}(t+\Delta t) = \left\{\frac{\alpha(l,m) * (V_{n+1}(t))^m}{(X_n(t) - X_{n+1}(t))^l}\right\}\left[V_n(t) - V_{n+1}(t)\right] \quad (1)$$

where: $a_{n+1}(t + \Delta t)$ = acceleration of follower vehicle $(n + 1)$ at a time interval $(t + \Delta t)$

$X_n(t)$ = position of leader vehicle (n) at previous time interval (t)
$X_{n+1}(t)$ = position of follower vehicle $(n + 1)$ at previous time interval (t)
$V_n(t)$ = speed of leader vehicle (n) at previous time interval (t)
$V_{n+1}(t)$ = speed of follower vehicle (n) at previous time interval (t)
α = sensitivity coefficient
m = speed exponent
l = distance headway exponent.

GM model is widely used and considered to be base model. Subsequently researchers namely, Gipps, Hidas, etc. found some limitations and accordingly modified the formulations. Chakraborty and Kikuchi further evaluated GM model with respect to some important car-following behaviour highlighting human behaviour in modelling approach [2]. For this purpose, fuzzy logic reasoning has been considered in modelling car-following behaviour incorporating human behaviour along with safety. Errampalli examined the variables influencing car-following behavouir especially considering road side friction, deviation from desire speed etc. with fuzzy inference rules [3]. Arasan and Koshy [4] and Metkari, et. al. [5] emphasized the lane discipline behavior under mixed traffic conditions in car-following model. Different researchers focused on the evaluation of different car following models and found their suitability for mixed traffic conditions [6]. Some researchers studied different combination of leader-follower vehicle types in mixed traffic conditions and also highlighted staggered following [7]. Kashyap, et. al. [8] identified two new variables, oblique spacing (R) and the angle between the leader and the follower (θ), influence car-following behaviour [8]. To address weak lane discipline in mixed traffic, an integrated methodology was proposed for different behaviours such as car-following, lane-changing or free flow [9].

Further, all these prevailing models have primary assumption of having one leader vehicle at front only influencing car-following behaviour which may not be the case especially in heterogeneous and no-lane discipline condition. In this paper an attempt is being made to develop a car-following model for such heterogeneous traffic conditions.

1.2 Objective

Considering the need to explain the car-following under mixed traffic condition as discussed in previous section the objective of the study has been formulated as given below:

- To identify the influencing parameters for car-following
- To develop car-following model with multiple leader vehicles.

In addition to above, the developed car-following model is used to study the effect of composition of commercial vehicles on road capacity, free speed and density.

2 Methodology

The methodology to achieve the objectives as mentioned in previous section is discussed in the form of flow chart as shown in Fig. 1.

Firstly, an urban road section is considered for this study and videography data of 4 hours is collected from the study area. Since, the offline image processing-based data collection system is more suitable for mixed traffic conditions [10] and calibration and validation of car following models require detailed vehicle trajectory data [11], videography data was collected. From the video, leader and subject vehicle trajectories for every 0.5 s are extracted with the help of IITB traffic data extractor

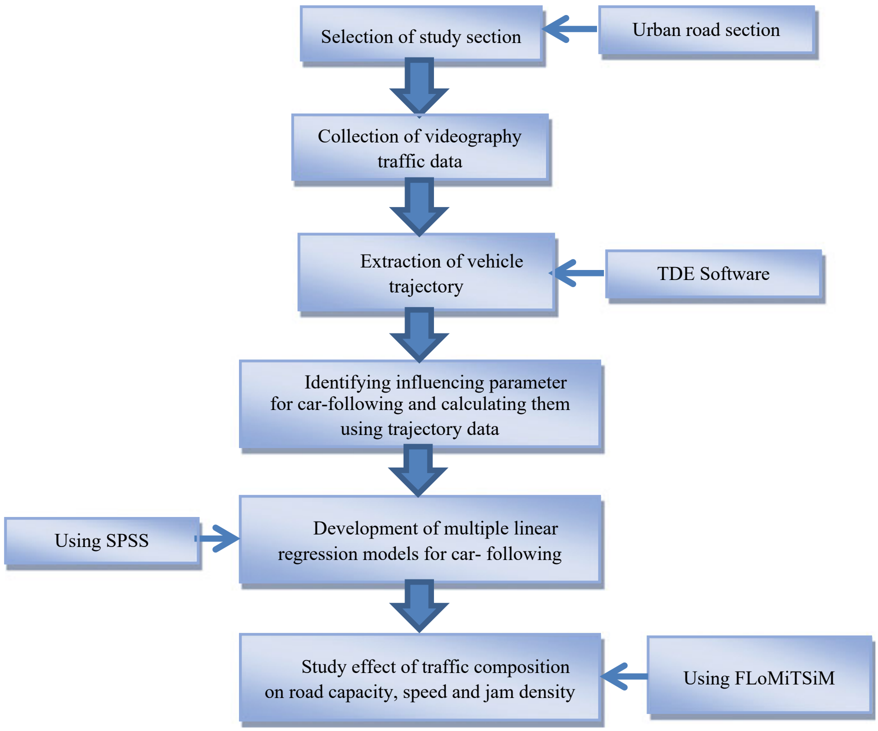


Fig. 1 Methodology adopted

software [12], using the trajectory data different parameters such as speed, acceleration, longitudinal gap, lateral gap, lateral position of vehicles, associated to leader and subject vehicles are calculated and relations are being established between parameters of leader and subject vehicle in order to study the car-following behaviour of the subject vehicle. Multiple linear regression models for car-following are developed using the selected influencing parameters in the software, SPSS. The developed car-following models are implemented in a fuzzy logic based microscopic traffic simulation model (FLoMiTSiM) which was developed by Errampalli [3] to study effect of traffic composition on road capacity, free speed and density.

3 Data Collection

3.1 Site Selection

A study section of NH-2 of Delhi-Mathura highway near CRRI in New Delhi which is six lane divided carriageway was selected for the purpose of data collection to develop car-following model. The videography data was collected by placing a camera over nearby flyover. The dimensions of the study section considered with length as 60 m and width as 10.5 m, as shown in Fig. 2.

Fig. 2 Study section of 60 m length and 10.5 m width near CRRI, New Delhi

3.2 Data Collection and Extraction

Videography data collection was done for 4 h from the selected site, i.e. from 7:30 to 11:30 a.m. in order to get both the morning peak and off-peak traffic condition with good mixed traffic condition. The trajectory data was extracted from the video with the help of IITB traffic data extractor [12] software by manually tracking the subject and all influencing surrounding vehicles considered as multiple leader. The other parameters such as speed, acceleration, longitudinal gap, lateral gap, lateral position and deviation angle of vehicles are also extracted with the help of trajectory data using python programing. For, identifying multiple leader vehicles around the subject vehicles the area around the subject vehicles was divided into four zones as shown in Fig. 3 for each zone one most influencing vehicle is selected as a leader vehicle except in back zone. In the present study, it is assumed that the vehicles in the back zone will not influence subject vehicle in terms of its car-following behaviour.

The influencing vehicles as leader vehicle in each zone is selected on the basis of gap from subject vehicle and also Passenger Car Unit (PCU) value of influencing vehicle. Smaller gap or nearest vehicle is generally considered as leader vehicle and in case two vehicles are at the same distance (a sensitivity of 0.5 m is considered) higher PCU value vehicle is considered as leader vehicle. Considering these points, the subject vehicle data along with influencing vehicles as leader vehicle have been created from extracted to develop multiple linear regression models to estimate speed of subject vehicle (following vehicle) as depended variable and data of leader vehicles as independent variables. The extracted data mainly include: positions of subject vehicle as following vehicle and all influencing vehicles as leader vehicles in each time frame. From this, speed of all the vehicles, gap or distance from subject vehicles, angle from subject vehicle are calculated.

4 Car-Following Behaviour Analysis

The data from the video is extracted to analysis the car-following behaviour of subject vehicles by creating scatter plots to find relation between parameters of subject vehicles and leaders vehicles as shown in Fig. 4. By studying these, five parameters

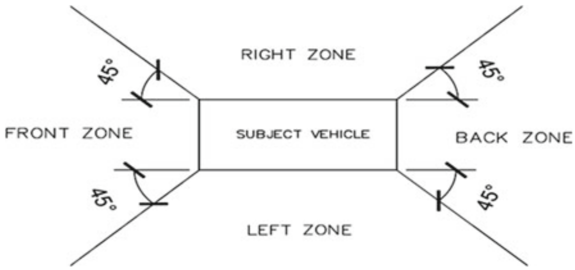


Fig. 3 Four zones considered surrounding to subject vehicle

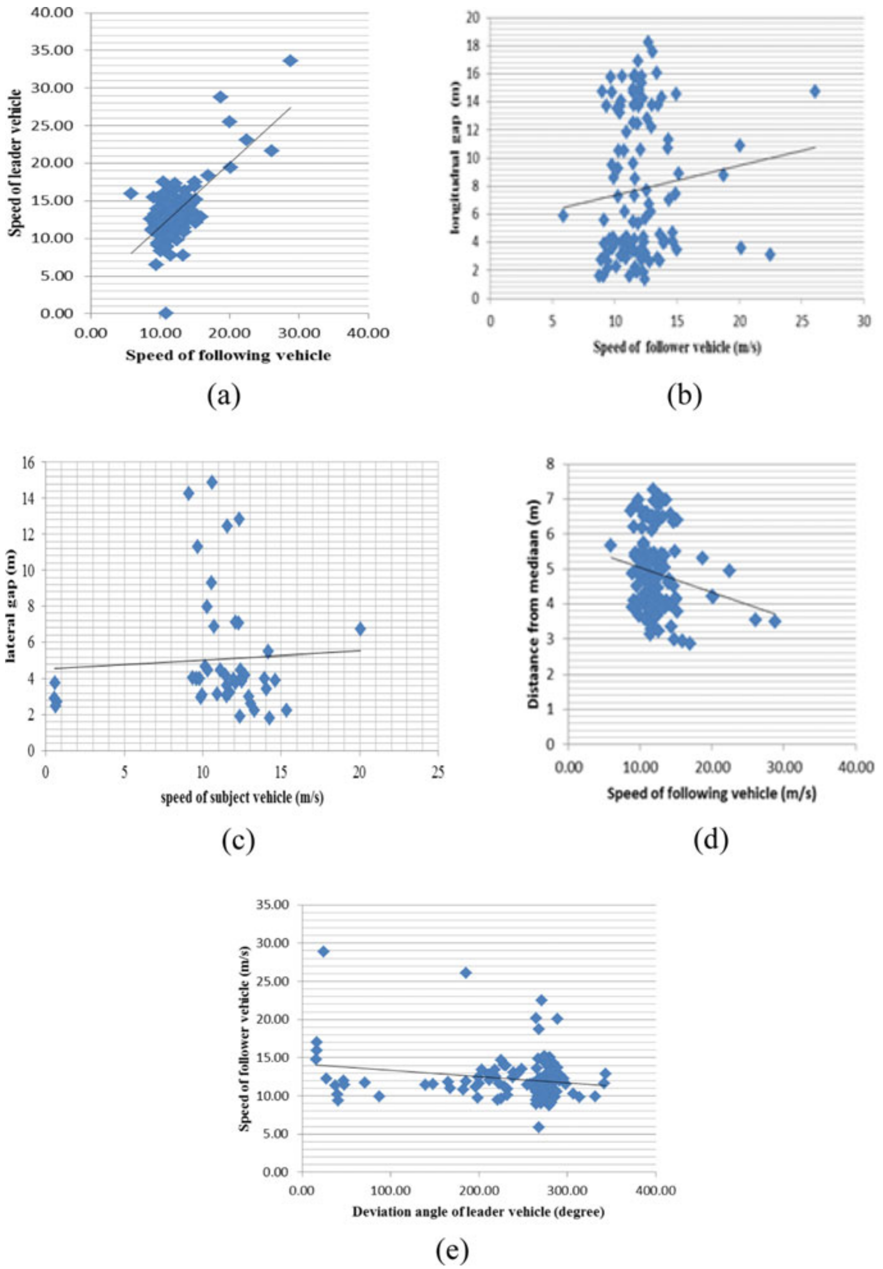


Fig. 4 Variation in **a** speed of leader vehicle, **b** longitudinal gap, **c** distance from median, **d** lateral gap and **e** deviation angle with respect to speed of follower vehicle

of leader vehicle in each zone are taken as influencing parameters: speed, gap from subject vehicle, lateral position from median, vehicle type (PCU value) and deviation angle with respect to subject vehicle. Apart from that lateral position of subject vehicle is also considered as another independent variable which became a total of 16 variables that are being used for model development.

The selected influencing parameters in each zone are: speed of leader, longitudinal gap, lateral position from median, vehicle type (PCU values), Deviation angle of leader and lateral position of subject vehicles, hence total 16 variables are being used for modelling.

5 Development of Car-Following Model

For modelling the car-following behaviour, multiple linear regression method is used to develop models for three types of subject vehicles (car, two-wheeler and three-wheeler). A dataset of 200 observations is considered for developing each model. For, developing these models SPSS is being used. The developed models are shown in Table 1 with 16 independent variables. From Table 1, it can be seen that R^2 values

Table 1 Developed car-following models for car, two-wheeler and three-wheeler

Variables in regression equation	Car	Two-wheeler	Three-wheeler
Constant	−1.38	–	–
Lateral position of following vehicle	−0.82	0.382	1.64
Gap of front zone leader vehicle	0.22	0.039	0.23
Deviation angle of front zone leader vehicle	−0.003	0.014	−0.01
Lateral position of front zone leader vehicle	0.59	0.447	−1.15
Speed of front zone leader vehicle	0.89	0.573	0.41
Gap of left zone leader vehicle	−0.45	−2.395	−0.46
Deviation angle of left zone leader vehicle	0.008	–	−0.01
Lateral position of left zone leader vehicle	0.012	7.73	−1.55
Speed of left zone leader vehicle	−0.32	0.483	0.66
Gap of right zone leader vehicle	0.55	−2.323	0.88
Deviation angle of right leader vehicle	0.02	0.01	0.01
Lateral position of front zone leader vehicle	–	1.73	
Speed of right zone leader vehicle	0.13	−1.06	0.69
Type of front zone leader vehicle (PCU value)	1.82	10.67	2.33
Type of left zone leader vehicle (PCU value)	7.21	−15.781	9.05
Type of right zone leader vehicle (PCU value)	1.18	4.883	−8.40
R^2	0.92	0.98	0.96

Table 2 RMS and MAP error from developed and GM models

Models	RMS error[a]	MAP error[a]
Developed model (car)	2.04	1.04
GM model (car)	3.70	1.41
Developed model (two-wheeler)	1.80	0.11
GM model (two-wheeler)	5.13	0.25
Developed model (three-wheeler)	2.43	0.18
GM model (three-wheeler)	2.70	0.22

[a] *RMS* Root Mean Square; *MAP* Mean Absolute Percentage

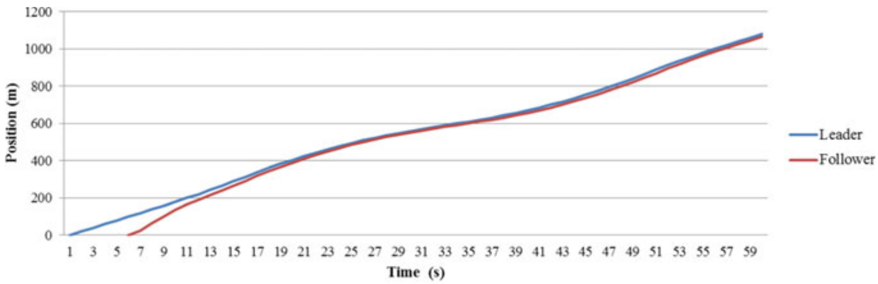
of all the models are higher which shows that the developed models are statically significant.

5.1 Validation of Models

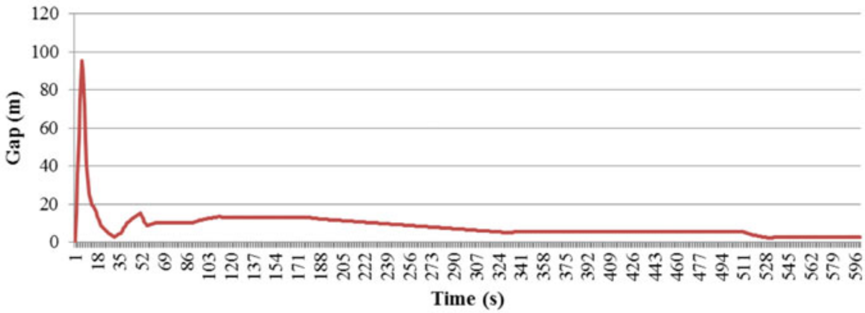
In order to validate the developed models, comparison of error from developed mode and GM model is made and presented in Table 2. There is a reduction of about 45% in RMS error and about 26% in MAP error by developed model for car compared to GM model. In case of two-wheeler model, there is a reduction of about 65% and 54% in RMS and MAP error values, respectively. For three-wheeler model, reduction of about 10% and 18% found in RMS and MAP error, respectively. The developed models are more suitable to explain vehicle behaviour in mixed traffic conditions as error values are less.

6 Evaluating Car-Following Behaviour

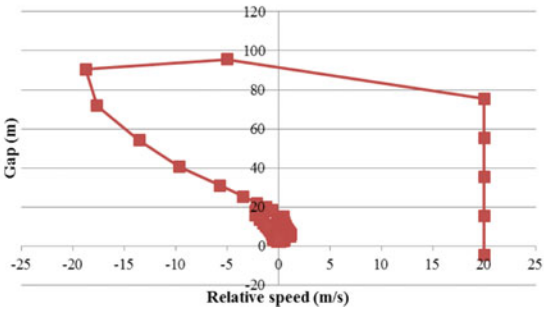
For evaluating car-following behaviour by the developed models, a simulation of follower and leader vehicle for 10 min duration has been carried out. A leader vehicle (car) has been placed in the front zone of the subject vehicle and speed and positions of the leader is updated manually and the speed of following vehicle has been calculated with the developed model as shown in Figs. 5, 6 and 7 for car, two-wheeler and three-wheeler, respectively.



(a) Position of leader and follower vehicle



(b) Variation of gap

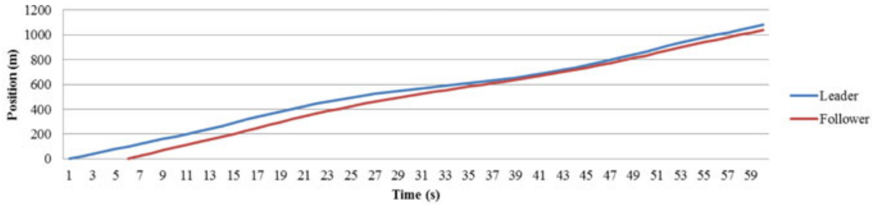


(c) Relative speed and gap

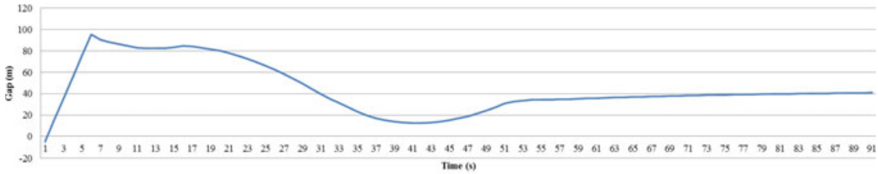
Fig. 5 Car-following behaviour estimated for car

7 Impact of Heavy Vehicle Traffic on Traffic Flow Parameter

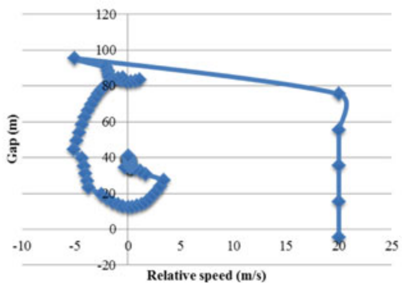
Using the developed model, it is proposed to assess the impact of heavy vehicle traffic composition on traffic flow parameters. For this purpose, a single four lane divided carriageway has been considered with 600 m length and different traffic composition



(a) Position of leader and follower vehicle



(b) Variation of gap

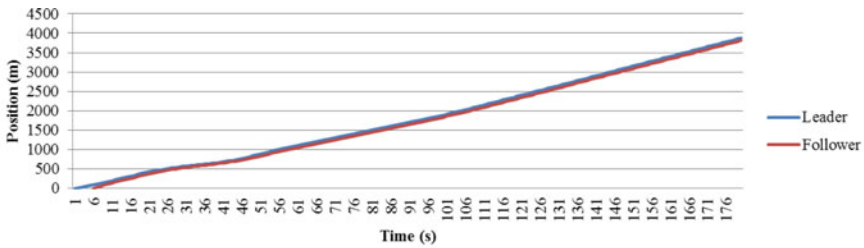


(c) Relative speed and gap

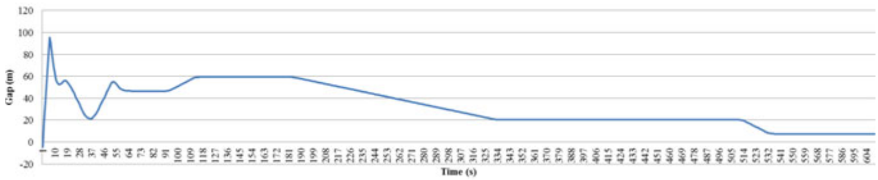
Fig. 6 Car-following behaviour estimated for two-wheeler

of heavy vehicles is considered [3]. Traffic composition is changed from 0 to 16% and effect on parameter such as speed, road capacity and jam density are studied using FLoMiTSiM which is developed by Errampalli [3]. It can simulate vehicles and passengers on urban road network involving a set of nodes and links is used. The simulation is done for four lane divided carriage way for three composition of traffic as shown in Table 3 for a traffic flow of 3600 veh/h.

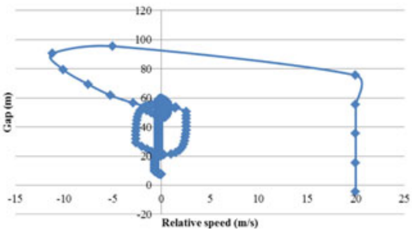
After running the simulation programme for the above traffic compositions, the results have been estimated. From the simulation results, speed-flow-density relation are made and using the green-shield model the speed of vehicle stream, capacity and jam density are calculated the result from the analysis are shown in Fig. 8. From Fig. 8, it can be observed that road capacity is going to impact with the increase in heavy vehicle composition and free speeds of cars are going to reduce. Further, it can be seen that density at capacity and jam density is also going to impact with the increase in heavy vehicle composition. This could be mainly due to higher PCU values for heavy vehicles [3].



(a) Position of leader and follower vehicle



(b) Variation of gap



(c) Relative speed and gap

Fig. 7 Car-following behaviour estimated for three-wheeler

8 Conclusion

A car-following model has been developed considering all the surrounding vehicles as influencing variables in developing apart from driver behaviour for mixed traffic conditions. In this study, there are some specific and some general conclusion that can be made which are as following:

- It is observed from the comparison of GM and developed models that for mixed traffic flow that the performance of GM model is not good because in GM model the speed of subject vehicles only depend on the position and speed of leader present in the front zone, where as in mixed traffic follow condition the subject vehicle can have multiple leader which effect the car-following behaviour of the subject vehicle. In mixed traffic flow the vehicles running of either side and at back of the subject vehicle are also responsible for changing driving behaviour of the driver.

Table 3 Different commercial vehicle composition considered for simulation

Vehicle type	No heavy vehicle (%)	Normal composition (%)	Higher proportion (%)
Small car/standard car (CS)	37	35	33
Big car/SUV (CB)	17	15	13
Bus	0	2	4
Two-axle heavy commercial vehicle/truck (HCV)	0	2	4
Light commercial vehicle (LCV)	0	2	4
Auto rickshaw (auto)	12	10	8
Two-wheeler (TW)	34	32	30
Tractor/trailer (TT)	0	0	0
Multi-axle heavy commercial vehicle (MCV)	0	2	4

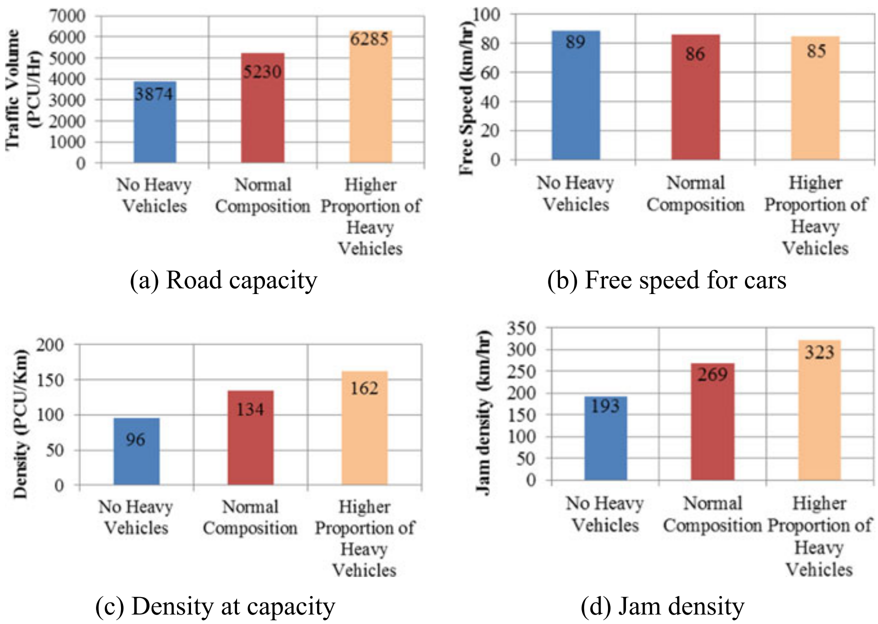


Fig. 8 Impact of traffic composition on road capacity, free speed of cars, density at capacity and jam density

- As the speed of leader vehicle decreases the gap between leader and follower also decreases to the minimum following gap and the subject vehicle starts to follow the leader vehicle at that minimum longitudinal gap at 2.8 m for cars with leader speed of 18 km/h.
- The position-time plot of leader and follower vehicle runs parallel to each other which depicts following behaviour of cars and as the time passes the subject vehicles covers the longitudinal gap and starts to follow the leader at minimum possible gap at that speed.
- As the speed of leader vehicle becomes constant the gap between leader and follower vehicle is also becoming constant with minimum value of gap for that speed.
- As the composition of heavy vehicle increases from zero to 16% there is an increase in capacity of road, density at capacity and Jam density but there was decline in free speed of the vehicles this is due to high PCU values of the heavy vehicle and their slow speeds.
- It is observed that for 1% rise in heavy vehicle there is an increase of 151 PCU/h. in road capacity, there was reduction of 0.25 km/h. in speed of cars and an increase of 8 PCU/km in jam density.

In this study, models are developed for cars, two-wheeler and three-wheeler so, there is a scope to develop models for other vehicle types such as bus, HCV and MCV. The limitation of the study is that the models developed under this study are developed on the data based on only one city and only for mid-section of the road this makes these models difficult to other cities which have different traffic composition as many cities may have domination of other type of vehicles on road traffic. The models need modification when applied for intersection. Models also not take into account of the vehicle present in the back zone of the subject vehicle which can influence the car-following behaviour of the vehicle, the development model uses the IITB traffic data extractor software in which it is very difficult to extract the trajectory due to manual tracking, this may introduce some error in data.

Acknowledgements The authors acknowledge the Council of Scientific and Industrial Research (CSIR) India who sponsored this research under Fast-Track Translational (FTT) Project titled "Microscopic Traffic Network Simulation Model for Mixed Traffic Conditions (MiTraNS)" MLP-0628. The Director CRRI is also acknowledged for allowing to publish the paper from the project. The team members of this project are also highly acknowledged for their contribution in data collection and data extraction.

References

1. Errampalli M, Mallela SS, Chandra S (2020) Calibration of car-following model for Indian traffic conditions. Transp Res Procedia 48:829–839
2. Chakraborty P, Kikuchi S (1999) Evaluation of the general motors based car-following models and a proposed fuzzy inference model. J Transp Res C 7:209–235

3. Errampalli M (2008) Fuzzy logic based microscopic traffic simulation model for transport policy evaluation. PhD thesis, Gifu University, Gifu, pp 3–26, 76–82
4. Arasan VT, Koshy RZ (2005) Methodology for highly heterogeneous traffic flow. J Transp Eng ASCE 131:544–551
5. Metkari M, Budhkar A, Maurya AK (2013) Development of simulation model for heterogeneous traffic with no lane discipline. Procedia Soc Behav Sci 104:360–369
6. Kanagaraj V, Asaithambi G, Kumar CHN, Srinivas KK (2013) Evaluation of different vehicle following models under mixed traffic conditions. Procedia Soc Behav Sci 104:390–401
7. Asaithambi G, Basheer S (2017) Analysis and modeling of vehicle following behavior in mixed traffic conditions. Transp Res Procedia 25:5094–5103
8. Kashyap NR, Madhuri C, Bhargava SR, Kartik KA, Asaithambi G (2020) Analysis of vehicle-following behavior in mixed traffic conditions using vehicle trajectory data. Transp Res Rec J Transp Res Board (TRB) 2674:842–855
9. Papathanasopoulou V, Antoniou C (2018) Flexible car-following models for mixed traffic and weak lane—discipline conditions. Eur Transp Res Rev 10:62
10. Mallikarjuna C, Phanindra A, Rao KR (2009) Traffic data collection under mixed traffic conditions using video image processing. J Transp Eng ASCE 135:174–182
11. Kanagaraj V, Asaithambi G, Toledo T, Lee T-C (2015) Trajectory data and flow characteristics of mixed traffic. Transp Res Rec J Transp Res Board (TRB) 2491:1–11
12. TDE (2013) Traffic data extractor user manual. IIT Bombay

Travel Behavior and Transport Demand

Travel Behavior of Access-Egress Mode Users at Rapid Transit Stations-Case Study Kolkata, India

Payel Roy and Sudip Kumar Roy

Abstract In a multimodal public transportation system, feeder service is often more disappointing to commuters than rapid transit service. This paper aims to investigate commuters' needs from access-egress modes by analyzing their travel behavior using a stated preference survey. A travel chart is used to assess modal preferences of access-egress modes where names of travel modes are obscured. The characteristics of different alternative modes during peak hours are represented by relevant images to avoid bias on any particular mode and to get authentic as well as consistent responses. Regular rapid transit users at three metro stations in Kolkata, India stated their preferences for access-egress trips using this travel chart. According to this study, majority of commuters are willing to pay more to reduce their in-vehicle travel time by accepting longer waiting time. They are also willing to pay more for comfortable daily trips when travel time is fixed. Better convenience and lesser number of transfers get more priority than lower travel cost and shorter travel time. However, travel cost remains as a secondary parameter upto a certain cost increment. This limit varies depending on the user's perceptions and socio-economic characteristics.

Keywords Access-egress mode · Travel behavior · Stated preference survey · Rapid transit

1 Introduction

Presently, majority of commuters in metropolitan cities have to avail a multimodal public transportation system. In such cases, sometimes the service of feeder system

P. Roy (✉)
Department of Civil Engineering, Indian Institute of Technology, Madras, Chennai, Tamil Nadu, India
e-mail: payel.roy.jgec@gmail.com

S. K. Roy
Department of Civil Engineering, Indian Institute of Engineering Science and Technology, Shibpur, Howrah, West Bengal, India

L. Devi et al. (eds.), *Proceedings of the Sixth International Conference of Transportation Research Group of India*, Lecture Notes in Civil Engineering 272,
https://doi.org/10.1007/978-981-19-3494-0_8

may not be as much satisfactory as of rapid transit due to improper planning of feeder service. These usual problems are forcing citizens to use private transport instead of public transport. The consequence of that is making the situation even worse by inventing issues like severe traffic congestion, uncountable traffic accidents, intolerable air, and noise pollution. Commuters' satisfaction is definitely a measure of effectiveness of public transport services. The aim of this paper is to search the needs of users from public transportation in terms of travel mode attributes.

The specific objectives of this paper include:

- To investigate users' preference on access or egress modes at rapid transit stations by stated preference survey
- To identify users' willingness to pay and ability to pay.

To search for the most accountable factor for increasing ridership a nationwide survey of transit agencies in the USA between 1995 and 1999 was conducted by Hess et al. [8]. Their report revealed that the most frequently cited factors were service improvement, i.e., more reliable, safer, cleaner service with attractive stops, and also an acceptable fare structure which attracted new riders and stretched the ridership growth. For a sustainable public transportation system, a balance between its customer responsiveness and cost-effectiveness is to be ensured. A case study of sub-regional governance of bus services in Los Angeles Country was conducted by Chen [4]. Users have some expectations or needs from public transportation system in terms of its service aspects; that can be identified by evaluating how much money they are willing to pay to improve the service quality. Those qualities for which commuters are ready to pay are also to be identified. In 2008, Eboli and Mazzulla [6] examined commuter's willingness to pay to improve the service quality of bus service in Cosenza, Italy by considering qualitative service aspects, namely reliability, frequency, proximity to bus stop, bus stop facilities, bus overcrowding, ticket fare, cleanliness, helpfulness of personnel, and information at bus stops. In 2017, Bachok and Ponrahono [2] stated that most of the commuters, who were willing to pay increased fare for an improved bus service in Malaysia, choose mostly cleanliness and comfort following by safety or security, punctuality, waiting facilities, and infrastructure as bus improvement priority areas. In 2017, Li et al. [13] considered five attributes, namely speed, connecting distance, transfer time, degree of crowding, and reliability for analyzing willingness to pay of commuters to get more improved service quality in Beijing, China. In 2017, Ingvardson and Nielsen [9] revealed that travel time reduction in bus rapid transit can increase attractiveness to commuters significantly. In 2005, Advani and Tiwari [1] presented an evaluation of Delhi metro, India concerning its travel time, capacity and accessibility from commuters' perspective. In 2015, Goel and Tiwari [7] conducted a detailed study on access-egress service of Delhi metro, India. In 2016, Chandra et al. [3] evaluated the modal share for work trips only using multinomial logit model along a busy street corridor in Kolkata, India. Three types of surface transportation mode including transit, paratransit, and personal vehicles were considered in this study. The model was developed based on both the quantitative criteria such as travel time, travel cost, waiting time, and the qualitative criteria such as comfort, reliability, dust, and noise. Above literature states that users' satisfaction has a great influence on ridership and users can pay more for better service. Moreover, there would be a trade-off between service quality and user cost. The present study wants to capture this trade-off, i.e., users are choosing which parameter over which.

2 Methodology

2.1 Adopted Methodology

To reach rapid transit station, commuters often avail a mode. This mode is known as access mode. After getting off from rapid transit, commuters get dispersed to avail their suitable travel mode to reach their destination; that is egress mode. The characteristics of interface between rapid transit and surface transportation are to be studied through reconnaissance survey.

Based on those characteristics, a travel chart is to be prepared. The names of travel modes are to be suppressed there, and the characteristics of different types of existing modes are to be represented by relevant images. Using this chart, regular rapid transit users have to state their preferences of modes as access or egress according to their suitability for daily trips for a certain distance (e.g. upto 10 km travel). From these data, preferences perceived by individual of access/egress mode of rapid transit along with travel attributes of public transport modes would be evaluated. In travel chart, relevant images are to be used to avoid biasness to any particular mode, to get more authentic and consistent results and also to reduce the monotony of questionnaire.

Then, to clarify the reason for choosing any mode, some stated preference questions are to be added in questionnaire. Every question would represent a pair of situations with the help of suitable images, and commuters have to choose one from them as per their suitability. In most of these cases, travel cost would be the objective function. These data would represent commuters' willingness to pay. These questions also check the reliability (if same answer is obtained in the same context) and validity (if the intention of asking question and users' response match accurately).

Commuters' socio-economic and trip characteristics are also to be noted by revealed preference survey to identify the reason of variation in responses and to reveal whether the respondents are true representatives of this study or not. The following flowchart states the adopted methodology of this paper briefly (Fig. 1).

2.2 Case Study Region: Kolkata

Kolkata is the capital of Indian state West Bengal. Versatility can be observed in surface public transport modes in this city such as fixed route transit like suburban rail and ancient tram, air conditioned and non-air conditioned buses, paratransit like

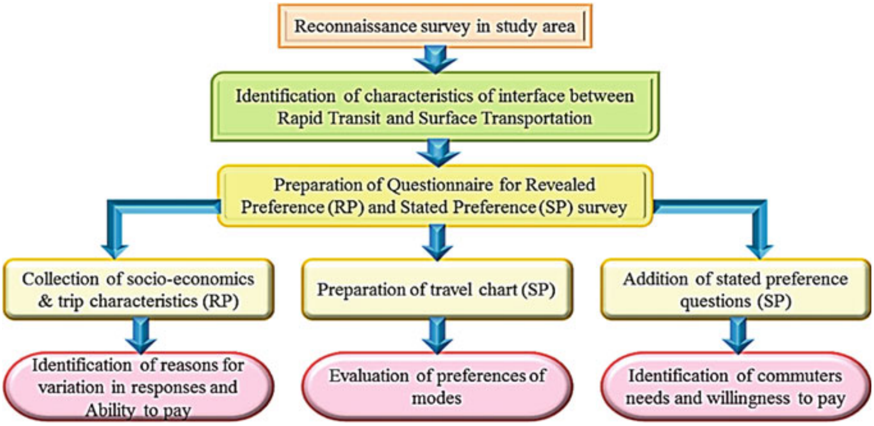


Fig. 1 Adopted methodology

auto-rickshaw and Carpool, demands responsive transit like Meter taxi and App-cab. 687,918 motorized vehicles are registered (including two and three wheelers) on roads in Kolkata [5]. In this city, approximately 54% of all trips are made through public transportation and that is the highest in India [10].

The rapid transit system, Kolkata Metro, has made the public transportation system in Kolkata strong. Line 1 (North–South Metro) Metro Network runs between Noapara and Kabi Subhas over 24 stations and Line 2 (East–West Metro) is partially operational between Saltlake Sector V and Saltlake Stadium, serving six stations, and two more networks are under construction and two are in proposal to connect surrounding suburban areas with this city [11, 12, 14].

Study Area. Three existing stations of Line 1 in Kolkata Metro Network, namely Dumdum, Mahanayak Uttam Kumar and Rabindra Sadan have been selected as survey area; one was toward north, one was toward south, and rest one was exactly middle of the corridor. All of them are one of the busiest stations and have enough versatility in present interface between rapid transit and surface transportation (Fig. 2).

3 Data Collection and Analysis

3.1 Design of Questionnaire

The questionnaire has been designed to achieve the aim of this paper. There were two parts in that designed questionnaire. In the first part, commuters' trip characteristics such as frequency of trip, purpose of journey, starting time of journey, trip

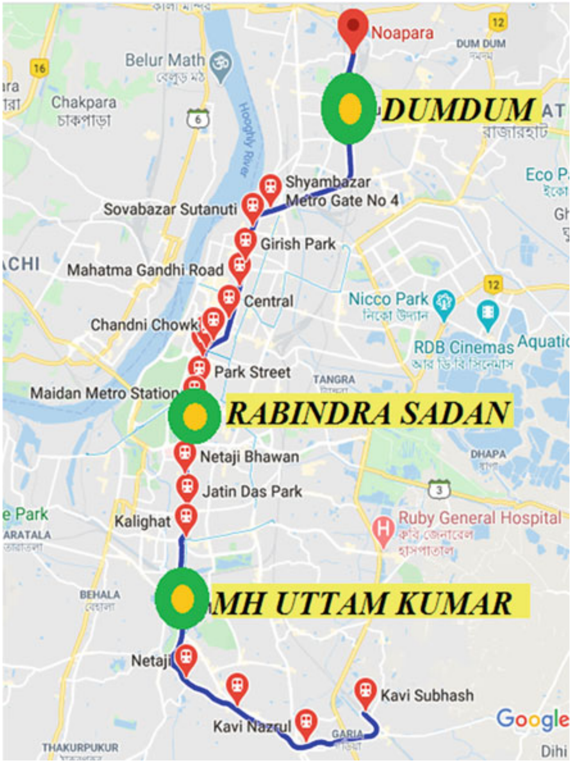


Fig. 2 Study area. *Source* https://www.google.co.in/maps

length, and socio-economic characteristics such as age group, gender, highest education, occupation, no. of cars and motorbikes in household, and monthly income and monthly expenditure on public transportation were added. In some previous study, it was observed that respondents showed reluctance in revealing their exact age and income; hence, these questions are modified by making classification groups for age and monthly income.

In the second part, a travel chart (Table 1) was attached including all the existing surface transport modes but the names of modes were suppressed. The subjective characteristics for each mode such as length of waiting queue and crowding level were perceived by author during reconnaissance survey. Based on that, the characteristics of different existing modes are represented by relevant images (Sect. 2.1). Respondents were asked their preferences among those modes as access or egress mode for upto 10 km according to suitability for their daily trips. To get more clarity and validity in the reason for choosing any mode, some stated preference questions were kept in this questionnaire with relevant images (Sect. 4.2).

In Table 1, App-cab is denoted by Mode A, Meter taxi by Mode B, Suburban rail by Mode C, Non-AC bus by Mode D, Carpool by Mode E, Auto-rickshaw by Mode F, and AC bus by Mode G.

Table 1 Travel chart in designed questionnaire

Mode	Waiting time during peak hours	Scheduled time-table	Seat availability during peak hours	Comfort	Travel-time delay and average speed during peak hours	Fare Rs./10 km
A	*Too short* 	*NO*	*Seat is assured and very comfortable* 	*AC*	*Delay ~ Low* *Speed ~ 35 km/h* 	*₹180*
B	*Short* 	*NO*	*Seat is assured and comfortable* 	*Non-AC*	*Delay ~ Low* *Speed ~ 30 km/h* 	*₹150*
C	*Long* 	*YES*	*Overcrowded* 	*Non-AC*	*Delay ~ Very low* *Speed ~ 45 km/h* 	*₹5*
D	*Medium* 	*YES*	*Crowded* 	*Non-AC*	*Delay ~ High* *Speed ~ 10 km/h* 	*₹11*
E	*Medium* 	*NO*	*Seat is assured and more or less comfortable* 	*Non-AC*	*Delay ~ Moderate* *Speed ~ 25 km/h* 	*₹30*
F	*Too long* 	*Mostly YES*	*Seat is assured but may be uncomfortable* 	*Non-AC*	*Delay ~ Moderate* *Speed ~ 25 km/h* 	*₹22*
G	*Too long* 	*YES*	*Can stand comfortably* 	*AC*	*Delay ~ High* *Speed ~ 15 km/h* 	*₹35*

Source Field survey

The following modes were observed at three metro stations during reconnaissance survey. The respondents were asked to rate all the modes irrespective of availability; so that their perception on any new mode can be captured.

(a) At Dumdum metro station, AC bus, Non-AC bus, Suburban rail, Auto-rickshaw, and App-cab are found to provide the access or egress service.
(b) At Rabindra Sadan metro station, AC bus, Non-AC bus, Meter taxi, Carpool, and App-cab play the role of access or egress mode of metro rail.
(c) At Mahanayak Uttam Kumar metro station, AC bus, Non-AC bus, Auto-rickshaw, Meter taxi and App-cab act as interface between metro rail and surface transportation.

3.2 Data Collection and Distribution

With the help of designed questionnaire, data collection was implemented through a survey of face-to-face interview by author from September 2019 to November 2019. A convenient sample of 100 regular metro users at each of three metro stations in Kolkata was interviewed directly on working days, i.e., Monday to Saturday (9 a.m.–1:30 p.m., 2 p.m.–7 p.m.). Random sampling method was adopted for sampling technique.

The observations from collected data are stated below which indicate that the respondents are true representatives for this study.

- At all of these three stations, five trips per week are made by majority of the respondents for either work or education, and their starting time of journey is within 8–10 a.m. As most of the respondents have neither a personal car nor a motorbike in their household, they are true representatives of the regular trip makers during peak hours of the day.
- Majority of the respondents are young or mid-aged people (18–35 years old) and male. Most of them completed graduation. In an average scenario, 54% trips are made by servicemen, 38% by students, 7.5% by businessman, and rest is by homemakers.
- Average access/egress trip length of respondents at Dumdum, Rabindra Sadan, and Mahanayak Uttam Kumar station is 6.4 km, 6 km, and 9 km, respectively.

The characteristics of monthly income and monthly expenditure on public transportation of respondents are stated in Table 2.

Table 2 Monthly income and monthly expenditure on public transportation of respondents

Distribution of monthly income (in terms of % of respondents)						
Station	Monthly income					
	₹0–₹10,000	₹11,000–₹20,000	₹21,000–₹30,000	₹31,000–₹40,000	₹41,000–₹50,000	>₹50,000
Dumdum	41%	15%	19%	17%	4%	4%
Rabindra Sadan	31%	18%	16%	12%	10%	13%
MH Uttam Kumar	27%	18%	19%	11%	8%	7%

Distribution of monthly expenditure on public transportation (in terms of % of respondents)									
Station	Monthly expenditure								
	≤₹500	₹600–₹1000	₹1100–₹1500	₹1600–₹2000	₹2100–₹2500	₹2600–₹3500	₹3600–₹4500	₹4600–₹6000	>₹6000
Dumdum	4%	16%	19%	20%	9%	15%	10%	4%	3%
Rabindra Sadan	2%	9%	16%	13%	15%	22%	6%	6%	11%
MH Uttam Kumar	1%	9%	18%	13%	11%	25%	8%	13%	2%

Mean of monthly income and monthly expenditure on public transportation			
Station	Dumdum	Rabindra Sadan	MH Uttam Kumar
Mean of monthly income of respondents	₹24,000	₹29,100	₹28,200
Mean of monthly expenditure on public transportation for all purpose	₹2362	₹2950	₹2958
% of monthly expenditure on public transportation for all purpose	9.84%	10.14%	10.49%
Mean of monthly expenditure on public transportation for work trip	₹1250	₹1685	₹1480
% of monthly expenditure on public transportation for work or educational trip	5.21%	5.79%	5.25%
Access + egress travel cost (Rs./10 km) in this particular trip	₹15	₹25	₹25

Source Field survey

4 Results and Discussion

4.1 Preferences of Modes from Travel Chart

Figures 3, 4, and 5 show number of commuters those have chosen a particular mode as their 1st preference, 2nd preference, and so on to access or egress the Kolkata metro.

Analysis and Inference

- At all the three stations, on an average 48% respondents have chosen Carpool as their 1st preference. Most of them (about 70%) earn ₹0–30,000/month, i.e., belong to low and middle income group [15]. As all the respondents are either work or educational trip makers, they prefer a mode by which their overall journey

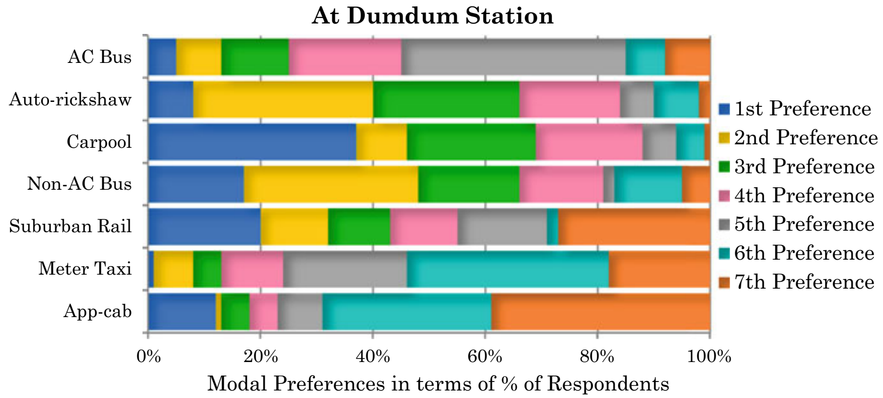


Fig. 3 Modal preferences in terms of no. of respondents at Dumdum. *Source* Field survey

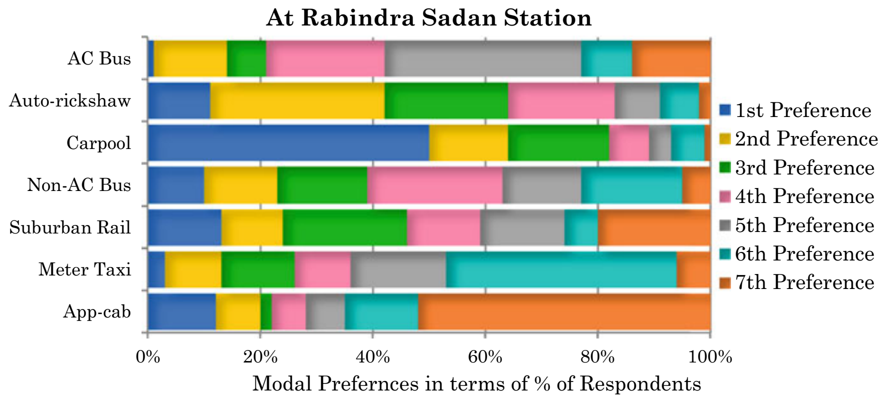


Fig. 4 Modal preferences in terms of no. of respondents at Rabindra Sadan. *Source* Field survey

can be completed in comparatively lesser time. Though Carpool has no schedule time, commuters have chosen it for its comparatively lesser travel time delay and an assured seat with an affordable travel cost.

- About 30% respondents at these three stations have chosen Auto-rickshaw as their 2nd preference. Most of them have chosen Carpool as their 1st preference. From this observation, it can be inferred that they can pay a little bit more cost to reduce waiting time and a comfortable seat. AC bus and Auto-rickshaw have same waiting time but Auto-rickshaw provides less travel time delay while AC bus provides more comfort. In this case, travel cost has no such significant role. So, it can be inferred that they prefer less travel time over air conditioned mode.
- More than 30% respondents have chosen Meter taxi as their last or 2nd last preference at these three stations, while another 30% respondents have chosen App-cab. Meter taxi and App-cab charge travel cost much higher than other five modes.

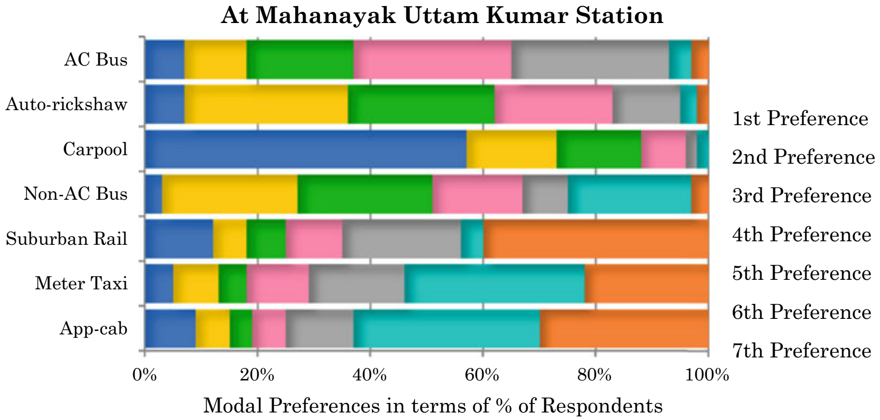


Fig. 5 Modal preferences in terms of no. of respondents at Mahanayak Uttam Kumar. *Source* Field survey

Thus, commuters do not prefer them for their daily trips. But it is notable that some commuters prefer App-cab over Meter taxi; that means to get a comfortable trip they are ready to spend some more charges when the increment in charge is not much significant.

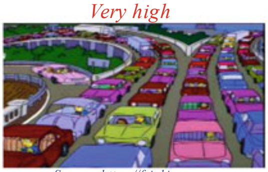
Finally, it can be stated that majority of these three stations' commuters prefer the characteristics of Carpool the most. Then they go for Auto-rickshaw; they are ready to wait in a long queue to get an assured seat and less travel time. Meter taxi and App-cab are the least preferred at all of these three stations. Though the middle ranker modes are different at these stations, the first and last two choices are same in everywhere.

4.2 Clarification of Reasons for Modal Preference

To clarify the reason of commuters' modal preference, six more stated preference questions were added in questionnaire. Every question represented a pair of situations with the help of suitable images. The pairs were decided on an assumption: Choice is a compensatory process. Commuter will compensate some lower parameter if there is sufficiently higher parameter. For example, high travel time + low cost and low travel time + high cost can be equivalent. The study find out that the majority of respondents want to compensate which parameter over which.

In Table 3, the 1st mode is charging more fare than the 2nd one with longer waiting time and lower travel time delay. The increment in fare for 1st mode is not much notable while waiting time for 1st mode is significantly longer and travel time delay is lower than the 2nd one. Here, about 60% respondents preferred the 1st situation over the 2nd one; that indicates majority of respondents are willing to pay more and

Table 3 1st pair of situation in SP survey

Waiting time (during peak hours)	Travel-time delay (during the journey)	Travel cost	% of respondents[a]
Too long  *Source: https://es.123rf.com*	*Very low*  *Source: https://www.crazyvector.com*		D → 60 R → 62 M → 58
Very short  *Source: https://www.alamy.com*	*Very high*  *Source: https://frinkiac.com*		D → 40 R → 38 M → 42

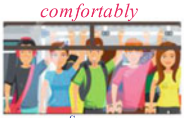
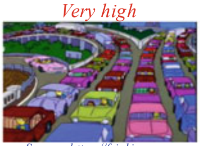
D → *Dumdum*, R → *Rabindra Sadan*, M → *Mahanayak Uttam Kumar metro station*

Source Field survey
**Respondents those preferred respective situation

ready to wait in a long queue to reduce their travel time delay. This observation can be a reason why commuters prefer Carpool and Auto-rickshaw over Non-AC bus.

In Table 4, the 1st mode is providing a significantly better service in frequency, travel time, and comfort but it is charging more fare than the 2nd one and this increment in fare is really notable. More than 65% respondents preferred 2nd situation over 1st one; that indicates those respondents do not want to invest that much increment in travel cost to get a better service of transportation in their daily trips. This

Table 4 2nd pair of situation in SP survey

Waiting time (during peak hours)	Travel-time delay (during the journey)	Seat availability (during peak hours)	Travel cost	% of respondents[a]
Very short  *Source: https://www.shutterstock.com*	*Very low*  *Source: https://www.crazyvector.com*	*Can stand comfortably*  *Source: https://www.shutterstock.com*		D → 34 R → 32 M → 29
Too long  *Source:https://www.vectorstock.com*	*Very high*  *Source: https://frinkiac.com*	*Crowded*  *Source: https://www.shutterstock.com*		D → 66 R → 68 M → 71

D → *Dumdum*, R → *Rabindra Sadan*, M → *Mahanayak Uttam Kumar metro station*

Source Field survey
**Respondents those preferred respective situation

Table 5 3rd pair of situation in SP survey

Waiting time (during peak hours)	Travel-time delay (during the journey)	Seat availability (during peak hours)	Comfort	% of respondents[a]
Too long  *Source: https://es.123rf.com*	*Very high*  *Source:* https://www.gettyimages.in	*Seat is assured and very comfortable*  *Source: https://www.canstockphoto.com*	*AC*	D → 36 R → 43 M → 56
Very short  *Source: https://www.alamy.com*	*Very low*  *Source: https://www.crazyvector.com*	*Overcrowded*  *Source: https://www.deccanherald.com*	*Non-AC*	D → 64 R → 67 M → 44

D → *Dumdum*, R → *Rabindra Sadan*, M → *Mahanayak Uttam Kumar metro station*

Source Field survey
**Respondents those preferred respective situation

observation can be a reason why commuters' preference for App-cab and Meter taxi falls down.

In Table 5, the 1st mode is providing a far better service in comfort but it is performance in frequency and travel time is very much poor than the 2nd one. At Dumdum and Rabindra Sadan metro station, more than 60% respondents preferred 2nd situation over 1st one which indicates those respondents give more importance in reduction in waiting time and travel time delay than a confirmed and comfortable seat for their daily trips.

But at Mahanayak Uttam Kumar station, majority of respondents go for the 1st mode. Those respondents cannot accept such overcrowded transport mode daily for the sack of reduction in total travel time. Though travel cost is a vital factor, this observation can also be a reason why commuters prefer Non-AC bus over AC bus and vice-versa.

Table 6 gives that in 1st situation, the mode is charging more fare than the 2nd one with an assured and comfortable seat but the increment in fare is not much notable. More than 65% respondents preferred the 1st situation over the 2nd one; that indicates majority of respondents are interested to pay more to get a comfortable journey daily. This observation can be a reason why commuters prefer Carpool and Auto-rickshaw over Non-AC bus.

In Table 7, the 2nd mode is providing far better service in travel time, seat availability, comfort, and convenience but it is charging more fare than the 1st one and this increment in fare is really notable. More than 60% respondents preferred 2nd situation over 1st one; that indicates most of the respondents do not want to invest that much increment in travel cost to get a better service in their daily transportation.

Table 6 4th pair of situation in SP survey

Comfort	Seat availability (during peak hours)	Travel cost	% of respondents[a]
AC	*Seat is assured and very comfortable*  *Source: https://www.canstockphoto.com*		D → 65 R → 65 M → 67
Non-AC	*Overcrowded*  *Source: https://www.deccanherald.com*		D → 35 R → 35 M → 33

D → *Dumdum*, R → *Rabindra Sadan*, M → *Mahanayak Uttam Kumar metro station*

Source Field survey
**Respondents those preferred respective situation

Table 7 5th pair of situation in SP survey

Travel cost	Walking time to reach next mode's stop (min)	Seat availability (during peak hours)	Comfort	Travel-time delay (during the journey)	% of respondents[a]
	*8*	*Crowded*  *Source: https://www.shutterstock.com*	*Non-AC*	*Very high*  *Source: https://www.gettyimages.in*	D → 61 R → 62 M → 69
	*2*	*Seat is assured and very comfortable*  *Source: https://www.canstockphoto.com*	*AC*	*Very low*  *Source: https://www.crazyvector.com*	D → 39 R → 38 M → 31

D → *Dumdum*, R → *Rabindra Sadan*, M → *Mahanayak Uttam Kumar metro station*

Source Field survey
**Respondents those preferred respective situation

The 2nd pair of situations is made by clubbing waiting time, travel time, seat availability, and travel cost; for this 5th pair of situations instead of waiting time, convenience is added. In 2nd pair of situations, on an average 31% respondents want to invest more travel cost to get a good service in their daily transportation and in this 5th pair average the number of interested respondents is 36%. So, it is evident that these extra 5% commuters are interested about mode's convenience.

Table 8 6th pair of situation in SP survey

Number of mode transfer	Waiting time (during peak hours)	Travel-time delay (during the journey)	Seat availability (during peak hours)	% of respondents[a]
4	*Medium*  *Source: https://www.freepik.com*	*Low*  *Source:* https://www.vectorstock.com	*Crowded*  *Source: https://www.shutterstock.com*	D → 57 R → 54 M → 42
2	*Too long*  *Source:https://www.vectorstock.com*	*Moderate*  *Source: https://images.slideplayer.com*	*Can stand comfortably*  *Source: https://www.shutterstock.com*	D → 43 R → 46 M → 58

D → *Dumdum*, R → *Rabindra Sadan*, M → *Mahanayak Uttam Kumar metro station*

Source Field survey
**Respondents those preferred respective situation

Table 8 gives that in 1st situation, the mode is providing lower total travel time than the 2nd one with more number of mode transfers and poor service in comfort. At Dumdum and Rabindra Sadan station, more than 50% respondents preferred 2nd situation over the 1st one; that indicates those respondents prefer long waiting time, moderate long travel time delay, and a little bit more comfort over more number of transfers of modes for their daily trips. At Mahanayak Uttam Kumar station, majority of respondents have chosen 2nd situation. Thus, commuters of this station mostly want a comfortable journey than lesser travel time.

The 3rd pair of situations is made by clubbing waiting time, travel time, seat availability, and comfort. For this 6th pair of situations instead of comfort, the number of mode transfer is added. In 3rd pair of situations, around 65% respondents preferred less travel time over assured and comfortable seat and in this 6th pair the number of % of respondents choosing less travel time falls down. It happens because there is not much significant difference between travel time delay of these two situations, one is low and another is moderate. In the 1st pair of situation, it is evident that commuters can stand in long queue to get their suitable mode. So, the reason behind reduced number of respondents (choosing less travel time) is more number of mode transfers. As 2nd situation provides less transfers of modes, commuters prefer this one. This observation state that not only reduction in travel time but also number of mode transfers can be considered in efficient transportation planning.

5 Conclusions

In this study, the pictorial stated preference questions explained reasons behind choosing access-egress modes from the travel chart well. The response obtained from travel chart and paired SP questions matched reasonably well. Thus, this type of survey can test the reliability and validity of the users' response.

This study concludes that the most of the commuters needs from access-egress modes are lesser travel time delay, an assured comfortable seat, and an affordable travel cost. Air conditioned mode is not their concern at all. The paired SP questions revealed that

- The commuters are willing to pay more and can accept longer waiting time to reduce in-vehicle travel time. They are also interested to pay more to get comfortable daily trips when travel time is constant.
- In both the cases, increment in travel cost was comparatively less. When this increment becomes notably high, respondents go for the mode incurring lower travel cost. Thus, it is evident when commuters choose any mode for their daily trips, travel cost remains a secondary parameter upto a certain limit of increment. This limit varies person to person depending on their perception, work characteristics, and definitely on socio-economic factors. As majority of respondents belong to low or middle income group in study area, they want to find a suitable mode along with an affordable travel cost for their daily trips.
- Travel modes' better convenience and lesser number of mode transfers get more priority than lower travel cost and reduction in travel time, respectively.

The RP survey of this study reveals that in the Indian metropolitan city Kolkata, commuters spend on an average 10% of their personal monthly income as monthly expenditure on public transportation for all purpose and around 5% only for their daily work or educational trips.

These observations can be taken account for planning of an efficient access and egress service for any new rapid transit. For example, in Kolkata a number of metro lines are in proposal and under construction. Kolkata East–West metro service would be operational in very recent times. In such cases, the outcomes of this study can be applied for access-egress service implementation. But, the land use and traffic characteristics of those new stations should be checked before the application if they are similar to the study area of this study.

References

1. Advani M, Tiwari G (2005) Evaluation of public transport systems: case study of Delhi. In: Proceeding in START conference
2. Bachok S, Ponrahono Z (2017) Modelling willingness to pay for improved public transportation system. J Malays Inst Plan 15:333–346
3. Chandra R, Sen S, Roy SK (2016) Mode choice modelling of work trips. In: Proceedings of the fourth international conference on advances in civil, structural and environmental engineering—ACSEE, pp 41–45
4. Chen X (2003) Subregional governance of bus services: an integrated study. J Public Transp 6(2):37–60
5. Directorate of Census Operations (2011) District census handbook, Kolkata, vol Series-20, Part XII-A. Directorate of Census Operations, West Bengal, Kolkata

6. Eboli L, Mazzulla G (2008) Willingness-to-pay of public transport users. Eur Transp/Trasp Eur 38:107–118
7. Goel R, Tiwari G (2015) Access–egress and other travel characteristics of metro users in Delhi. IATSS Res 39(39):164–172
8. Hess DB, Yah A, Iseki H, Taylor B (2002) Increasing transit ridership: a survey of successful transit systems in the 1990s. J Public Transp 5(3):33–66
9. Ingvardson JB, Nielsen OA (2017) Effects of new bus and rail rapid transit systems—an international review. Transp Rev 38(1):1–21
10. Karthikeyan H (2013) Kolkata tops Indian cities in public transport: study. Times of India
11. Kolkata Metro. Available at: http://www.kolkatametrorail.com/kolkata-metro-stations
12. Kolkata Metro Rail Corporation Ltd. Available at: http://www.kmrc.in/overview.php
13. Li N, Yan H, Liu XM, Zhang WP, Jin LN (2017) Analysis of traveler willingness to pay for multi-level public transport based on an ordered logit model. ASCE, pp 3775–3784
14. Metro Railway, Kolkata. Available at: https://mtp.indianrailways.gov.in/
15. Pandya P (2017) What incomes define upper, middle and rich classes in India? Available at: https://myindiamake.com/2017/07/26/what-incomes-define-upper-middle-and-rich-classes-in-india/

Estimating Regional Travel Demand from Intraregional Urban Demand Models

Raghav Tiwari, B. Anish Kini, and B. K. Bhavathrathan

Abstract Studies related to travel demand help determine future demands and travel patterns. The efforts needed for regional demand estimation are complex since the same techniques used in the urban context seldom hold. As an alternative, the regional demand could be estimated from the existing urban travel demand which could be used as building blocks. Doing so would affect the accuracy of the estimate which could be constrained to permissible levels so as to ensure the reliability of the regional travel demand estimator. This paper presents four cases and employs optimisation techniques to estimate the production/attraction growth factors utilising the traffic counts on freeways. Numerical example explains the methodology, which holds promise in constructing regional travel pattern from available urban OD matrices thus saving precious time and resources in the planning process. Regional travel demand estimation assists in highway connectivity planning and prioritisation, future demand prediction, etc.

Keywords Regional travel demand · Urban travel demand · Growth factor

1 Introduction

Urban Transport Demand modelling using the four stage modelling started in the 1960s [1] which involved Trip Generation, Trip Distribution, Modal Split and Trip Assignment. Trip Distribution is a procedure by which the trips get distributed between the Traffic Analysis Zones (TAZ) in the study area which is represented

R. Tiwari · B. K. Bhavathrathan (✉)
Indian Institute of Technology Palakkad, Palakkad, Kerala, India
e-mail: bhavathrathan@iitpkd.ac.in

R. Tiwari
e-mail: 101701025@smail.iitpkd.ac.in

B. Anish Kini
KSCSTE-National Transportation Planning and Research Centre, Thiruvananthapuram, Kerala, India
e-mail: anishkini.natpac.0@gmail.com

L. Devi et al. (eds.), *Proceedings of the Sixth International Conference of Transportation Research Group of India*, Lecture Notes in Civil Engineering 272,
https://doi.org/10.1007/978-981-19-3494-0_9

as an Origin Destination (OD) matrix with each cell (i, j) representing the number of trips originating from TAZ i destined to TAZ j in its simplest form. These aggregate OD matrices could be sub-divided based on the purpose, travel mode, time of day, etc. to obtain a far more detailed trip pattern of the study area.

However, the travel demand model was urban area centric and car-based. The same method could seldom be applied on a regional level in order to plan for the region as a whole. Using the four stage method, employed for estimating urban level travel demand, for estimating a regional model would be prohibitively expensive. Using this as motivation, this paper delves into alternatives—OD matrices of urban areas—that could be used for estimating the regional travel demand. Therefore the intraregional OD matrices were used as building blocks to estimate the regional travel demand.

The paper has been organised into the following sections: (i) Literature review covers the relevant studies that have been carried out in the area of OD estimation and the methodologies followed. It also mentions the research gap identified. (ii) Scope and Objectives specifies the major objectives of the study and the regional setting. (iii) Methodology describes the four cases and the optimisation formulations used for solving each of these cases. (iv) Numerical example hypothesises a region, its transport network connecting the cities based on which the methodology has been further discussed for better understanding. (v) Conclusion summarises the study, points out the limitations and scope for future research.

2 Literature Review

Origin Destination matrix estimation has been an active research topic since decades. Gravity model has been widely used for trip distribution which was postulated, analogous to Newtonian gravity model, such that the trips between an OD pair would be proportional to the productions from the origin and attractions to the destination whilst they would be inversely proportional to the distance between the OD pair. The initial model gave way to various modified versions, the simplest being represented in the form of a generalised cost function of distance or travel time between the OD pair [1]. Wilson [2] introduced the Entropy Maximisation approach based on which various models, Gravity model being one amongst them, could be generated.

Researchers have used different building units—Traffic counts, Automatic Number Plate Recognition (ANPR) data, Global Positioning System (GPS) tracks, Call Detail Records (CDR), Electronic Ticketing Machine/Automatic Fare Collection system records, proxy of these data or different combinations of the datasets—for the estimation of OD matrices. References [3–16] have presented various techniques to estimate OD matrices from traffic counts.

Chang and Wu [6] presented approaches to determine the dynamic OD from time varying traffic counts on a congested freeway and OD distributions; Yang et al. [7] presented heuristic algorithms for bi-level OD estimation problem satisfying the network equilibrium conditions; Parry and Hazelton [10] recommended a likelihood-based inference for estimating the OD matrices from link counts and route information of a few vehicles; Lorenzo and Matteo [12] used the approximation property of neural network model for estimating the OD matrices from link counts which could be used for dynamic traffic management given the low computing time; Cascetta et al. [13] developed a quasi-dynamic approach for estimating the OD matrices assuming that the OD shares would remain unaltered for a reference period and only the flow from each origin would change thereby reducing the number of unknowns; Ma and Qian [14] proposed a t-distributed Stochastic Neighbour Embedding (t-SNE) and k-means methods that statistically cluster the traffic data for estimating the daily dynamic OD matrices using 24 $\times$ 7 traffic counts and speed data; Krishnakumari et al. [15] proposed a data driven approach for estimating OD matrices from production/attraction time series, shortest paths between OD zones, proportionality of path flows between origins and destinations; Ashok and Ben-Akiva [17] illustrated the use of deviations in departure rates and the proportion moving to each destination for the estimation of time dependent OD matrices; Cipriani et al. [18] used a gradient approximation method, based on a modified Simultaneous Perturbation Stochastic Approximation (SPSA) optimisation, for solving the time varying OD matrices from traffic counts and speed; Caggiani et al. [19] used a Fuzzy-Generalised Least Square estimator for estimating OD matrices from traffic data and uncertain OD demand data; Foulds et al. [20] proposed a Fuzzy set-based approach for OD estimation whilst utilising incomplete or imprecise data; Tympakianaki et al. [21] proposed the cluster-based SPSA (c-SPSA) technique for estimating dynamic OD matrices with lesser bias than the conventional SPSA method; Shen et al. [22] proposed a two stage OD demand estimation by time of day over the year using traffic data; Chu et al. [23] developed a deep learning approach—Multi scale convolutional long short-term memory network—using the OD tensor for OD flow representation without geographic data loss for predicting future OD flows and Yao et al. [24] proposed a spatial interaction graph convolution network model for improving the OD flow estimation accuracy.

For large networks, Ma et al. [16] proposed a multi class dynamic OD estimation using forward–backward algorithm on computational graphs which can be used even in real world large scale networks; Perrakis et al. [25] employed a statistical Bayesian approach for estimating OD matrices from census data using Metropolis–Hastings algorithm and Frederix et al. [26] presented a hierarchical decomposition approach for offline OD estimation employing more accurate dynamic OD estimation in sub-regions which experience congestion effects.

The review of literature shows that majority of the studies concentrate on estimating the dynamic OD matrices from traffic counts, its proxies or a combination of datasets including route information, speed, OD shares, previous OD, geographic location data, etc. which were confined to small scale transport network applications. The methods used include bi-level optimisation, SPSA technique, Fuzzy set-based, neural networks, computational graphs, deep learning approaches, etc. The review points at limited studies on the development of regional demand estimation techniques that could be used on a large scale network. Therefore the focus of this study is aimed at the development of regional OD matrices from intraregional urban demand models.

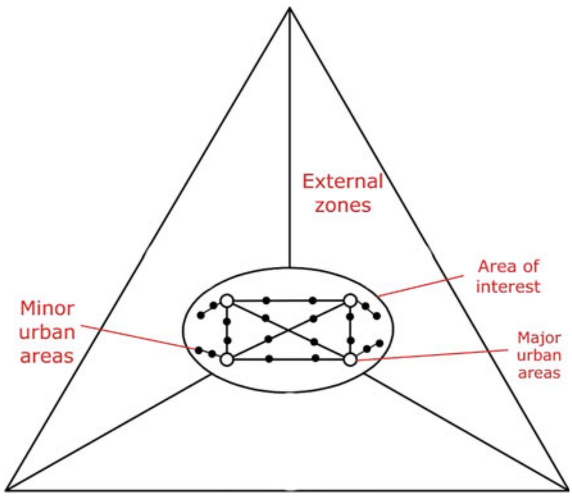
3 Scope and Objectives

The urban OD matrix can be estimated by conducting roadside interview surveys at the outer cordon points supplemented by household surveys. However collecting the same data at regional level requires tremendous and concerted efforts which makes it inconvenient and expensive. As an alternative there may be various urban level travel demand studies conducted by an agency(s) which would most probably belong to different time frames. There are possibilities that as the prime focus of such studies would be the urban area, the external origins/destinations may not be correctly identified or would be marked with the direction of the route taken, i.e. with non-specific origins/destinations.

The aim of this study is to reduce the cost and effort involved in establishing the regional travel demand by deriving it from the already available urban travel demand. Therefore the objective is the estimation of regional OD matrices using the intraregional OD matrices of urban areas, as building blocks, belonging to different time periods and having specific/non-specific origins/destinations.

In order to understand the techniques used, a regional setting is assumed as shown in Fig. 1. It contains four major urban areas within a region, 20 minor urban areas (two minor urban areas on each link joining major urban areas and 2 minor urban areas beyond each major urban area), and 3 urban areas external to the area of interest.

Fig. 1 Regional setting depiction



4 Methodology

The urban OD matrices belong to various major urban areas which are assumed to be captured at various time periods. This problem then presents four cases as follows which are detailed in subsequent sub-sections—(1) Available urban OD matrices with specific trip origins and destinations and belonging to the same time period, (2) Available urban OD matrices with specific origins and destinations but belonging to different time periods, (3) Available urban OD matrices with non-specific origins and destinations and belonging to the same time period, (4) Available urban OD matrices with non-specific origins and destinations but belonging to different time periods. In this paper, all the available urban OD matrices have been randomly generated which could be collected on site using various OD surveys.

4.1 Case 1

In this case the urban OD matrices belong to the same time period and hence it is a straight-forward compilation of all the urban OD matrices. This type of data is however hard to come by due to cost and time constraints. This can be considered as the Base case to which all the other cases have to be reduced to.

The data in the base case can be simply compiled considering that the production/attraction from different parts of an urban region contribute to the total production/attraction of that particular urban region. If there are two values for the same OD pair then an average value could be utilised in the regional OD matrix. Mathematically, it can be expressed as given in Eq. 1.

$$T_{AB} = \sum_{i=1}^{na} \sum_{j=1}^{nb} T_{ij}, \tag{1}$$

where T_{AB} = Number of trips from urban area A to urban area B, T_{ij} = Number of trips from an internal zone 'i' of urban area A to an internal zone 'j' of urban area B, na = number of internal zones in urban area A and nb = number of internal zones in urban area B.

4.2 Case 2

In Case 2 the urban OD matrices are from different time periods. All the urban OD matrices have to be brought to the same forecast year and then compiled as in Case 1. As the other urban areas fall under external zones whilst considering a particular urban area, proxies for growth factors like number of vehicle registration, population, etc. could not be used. Therefore the growth factors as variables need to be multiplied to the total production and attraction of each urban area to get forecasted productions and attractions from various zones. Forecasted total productions and attractions are then used to apply the Gravity model for the generation of forecasted travel demand data for a particular urban region. Mathematically, the gravity model can be expressed as given in Eq. 2.

$$T_{ij} = P_i\left[\frac{A_j F_{ij} K_{ij}}{\sum_j A_j F_{ij} K_j}\right], \tag{2}$$

where T_{ij} = Number of trips produced from zone i and attracted to zone j, P_i = Total number of forecasted trips produced in zone i, obtained as given in Eq. 3, A_j = Total number of forecasted trips attracted to zone j, obtained as given in Eq. 3, F_{ij} = A value which is the inverse function of travel time, K_{ij} = Calibration factor for interchange ij, which is obtained by calibration of the OD matrix for the base year to get the K_{ij} such that the total productions and attractions remain the same for any zone.

$$P_i = xp_i, \quad A_j = ya_j, \tag{3}$$

where x = Production growth factor for the corresponding urban area, y = Attraction growth factor for corresponding urban area, p_i = Total number of trips produced from zone i in the base year, a_j = Total number of trips attracted to zone j in the base year.

On applying the gravity model, the forecasted OD matrix is obtained in terms of variable growth factors. All the urban area matrices are then compiled as in Case 1 to get a regional OD matrix, which is then compared to the traffic flow data obtained from freeways to minimise the difference to obtain growth factors and eventually the final regional OD matrix using the method of optimisation as given in Eq. 4.

$$\text{Minimize } z = \sum\left(T_{ij} - f_{ij}\right)^2, \tag{4}$$

where T_{ij} = Number of trips from urban area i to urban area j as obtained by compiling forecasted OD matrices, f_{ij} = Number of trips from urban area i to urban area j obtained from the freeway.

4.3 Case 3

In Case 3, the urban OD matrices belong to the same time period but the destinations or origins are not well defined but are mentioned as the routes. To convert these matrices similar to the ones in Case 1, definite origins and destinations need to be identified for each trip.

In order to get the definite origin and destination points for each trip, methods of optimisation have been used to equate the trips going towards particular destinations with the total attractions of all the regions on the same route, using the attributes such as population, vehicle per capita and area of land for the production whereas employability of the area, commercial area and distance for the attraction. Mathematically this can be expressed as given in Eqs. 5 and 6.

$$T_{Ab} = m\sum_i E_i + n\sum_i C_i + o\sum_i \frac{1}{D_i}, \tag{5}$$

where T_{Ab} = Number of trips predicted from urban area A towards urban area B, E_i = Employability of each urban area in the link AB before urban area B and beyond urban area B, C_i = Commercial area of each urban area in the link AB before urban area B and beyond urban area B, D_i = Distance of each urban area from urban area A lying in the link AB before urban area B and beyond urban area B, m, n, o = Corresponding weight factors, to be obtained using the method of optimisation.

$$T_{aB} = q\sum_i P_i + r\sum_i V_i + s\sum_i \frac{1}{D_i} + u\sum_i A_i, \tag{6}$$

where T_{aB} = Number of trips predicted from the direction of urban area A and attracted to urban area B, P_i = Population of each urban area in the link AB before urban area A and beyond urban area A, V_i = Vehicle per capita of each urban area in the link AB before urban area A and beyond urban area A, A_i = Area of the land of each urban area in the link AB before urban area A and beyond urban area A.

These predicted trips are then compared to the values obtained from surveys to get the weight factors using methods of optimisation as expressed in Eq. 7.

$$\text{Minimise } z = \sum (T_{aB} - t_{aB})^2 \quad \text{or} \quad \sum (T_{Ab} - t_{Ab})^2, \tag{7}$$

where t_{aB} = Number of trips coming from the direction of urban area A and attracted to urban area B as obtained from the surveys, t_{Ab} = number of trips originating from urban area A and going towards urban area B as obtained from the surveys.

On obtaining the weight factors, the number of trips from A to B can be predicted using the proxies for production and attraction and the obtained urban area matrices can be compiled as in Case 1 to get the regional OD matrix.

4.4 Case 4

In Case 4, the urban OD matrices belong to different time periods and also have non-specific origins or destinations. This case could be treated as a combination of other cases explained above. Firstly the origins and destinations could be found for the base year for each urban areas which have non-specific origins/destinations as per methodology explained for Case 3 and then the data could be forecasted for the same horizon year using growth factors as per methodology explained for Case 2 to get matrices similar to Case 1 which could then be compiled.

5 Numerical Example

For better understanding of the methodology formulated, a numerical example has been hypothesised in this section. Figure 2 shows the virtual region consisting of four cities and the transport network. For the said region, the random matrices for major cities were developed and processed considering 5, 7, 9 and 6 numbers of internal zones for major cities 1, 2, 3 and 4, respectively, along with three external zones (marked e1–e3). Case 3 being complex, this has been dealt with further. The randomly generated OD matrices for the cities are shown in Tables 1, 2, 3 and 4. Furthermore, the demographic data and distances were also assumed for all major (marked c1–c4) and minor cities (marked a–t) as presented in Tables 5 and 6, respectively.

Applying Eqs. 3 and 4 and by solving the optimisation problem, the city matrices are obtained with definite origin and destination points (with trips rounded off to nearest integer value) as shown in Tables 7, 8, 9 and 10. The modified OD matrices would then be compiled as per Eq. 1 to obtain the regional OD matrix.

As detailed through the presentation of a numerical example, the methodology formulated for constructing the regional OD matrix from the intraurban OD matrices could be achieved utilising the optimisation techniques. This methodology could be adopted to estimate the regional travel demand which would help in understanding the trip patterns on the regional level and hence would be useful in highway connectivity planning and prioritisation in addition to forecasting future regional trips.

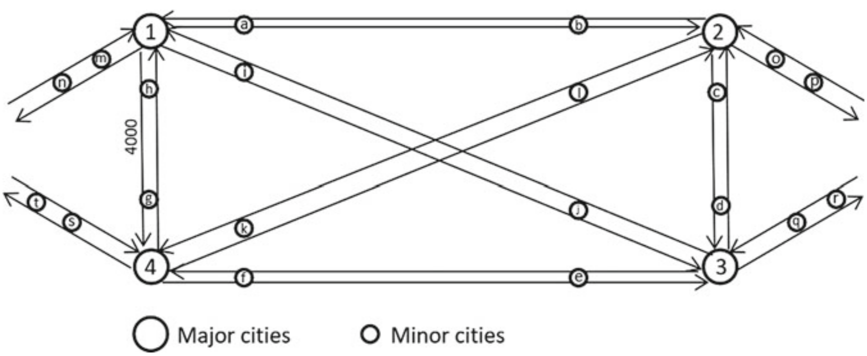


Fig. 2 Virtual region

Table 1 Generated OD matrix for city 1

	1	2	3	4	5	c2	c3	c4	e1	e2	e3
1	82	98	85	4	18	50	55	25	38	34	45
2	91	96	94	9	49	96	14	62	57	16	8
3	12	49	68	83	45	34	15	47	7	80	23
4	92	80	76	70	65	59	26	35	5	31	92
5	63	14	75	32	71	22	84	83	53	53	15
c2	9	42	39	95	76	75	25	59	78	16	83
c3	28	92	66	3	27	25	82	55	94	60	54
c4	55	80	17	44	68	51	24	92	13	26	100
e1	96	96	71	38	66	70	93	28	57	66	7
e2	97	66	3	77	16	89	35	76	47	69	44
e3	15	3	27	80	12	96	19	76	1	75	10

Table 2 Generated OD matrix for city 2

	1	2	3	4	5	6	7	c1	c3	c4	e1	e2	e3
1	80	63	13	75	49	82	17	99	55	17	19	99	16
2	2	68	72	50	16	72	39	16	95	3	43	73	18
3	93	39	10	48	98	15	83	10	42	56	48	34	42
4	73	37	66	91	71	66	81	37	99	89	12	58	9
5	49	99	49	61	50	52	6	20	30	67	59	10	60
6	58	3	78	62	47	98	40	49	70	19	22	91	47
7	23	89	72	86	6	65	53	34	67	37	38	88	70
c1	46	92	91	81	68	80	42	96	54	46	58	82	70
c3	97	80	89	58	4	45	66	92	70	99	25	26	64
c4	55	9	33	18	7	43	63	5	67	15	29	60	3
e1	52	26	70	24	52	83	29	74	17	86	62	2	6
e2	23	33	19	89	9	8	43	27	12	65	26	42	32
e3	49	68	3	2	82	13	1	42	100	38	83	31	53

Table 3 Generated OD matrix for city 3

	1	2	3	4	5	6	7	8	9	c1	c2	c4	e1	e2	e3
1	66	44	22	32	77	23	54	22	9	24	43	77	9	88	59
2	41	53	67	79	19	12	87	10	53	44	89	16	11	99	24
3	82	46	85	47	29	61	26	11	53	69	39	87	13	0	67
4	72	88	34	3	9	45	32	6	86	36	77	99	68	87	8
5	97	52	78	17	58	46	12	40	48	74	40	51	50	61	63
6	53	95	68	72	69	66	94	45	39	39	81	89	19	99	66
7	32	64	0	47	55	77	65	36	67	69	76	59	49	53	73

(continued)

Table 3 (continued)

	1	2	3	4	5	6	7	8	9	c1	c2	c4	e1	e2	e3
8	10	96	60	15	42	35	48	77	74	71	38	15	14	48	89
9	61	24	39	34	65	66	64	63	52	44	21	20	5	80	99
c1	78	68	92	61	65	42	55	77	35	1	79	41	85	23	77
c2	42	29	0	19	68	85	65	94	15	33	95	75	56	50	58
c4	9	67	46	74	64	84	54	98	59	42	33	83	93	90	93
e1	26	70	42	24	95	25	72	19	26	27	67	79	70	58	58
e2	15	6	46	92	21	61	52	14	4	19	44	32	58	85	1
e3	28	25	77	27	71	58	100	70	76	82	84	53	82	74	12

Table 4 Generated OD matrix for city 4

	1	2	3	4	5	6	c1	c2	c3	e1	e2	e3
1	87	12	12	10	99	47	90	17	75	44	30	34
2	48	20	49	14	86	65	11	36	74	1	4	29
3	85	14	86	16	79	2	99	5	56	90	51	75
4	21	19	88	62	51	85	54	52	18	19	76	1
5	55	4	27	57	17	56	71	33	60	9	63	4
6	63	64	21	5	40	86	100	17	30	31	9	67
c1	3	28	57	94	13	35	29	21	13	46	8	60
c2	62	54	64	73	3	45	41	91	21	10	78	53
c3	36	70	42	74	94	5	46	68	90	100	91	73
e1	5	50	20	6	30	17	77	47	7	33	53	71
e2	49	54	95	86	29	66	82	92	24	30	11	78
e3	19	44	8	94	33	33	10	10	5	6	83	29

Table 5 Demographic data

	Population	Vehicle per capita	Area of land (km^2)	Employability	Commercial area (km^2)
City 1	5000	1.5	60	400	2
City 2	3000	2	40	500	1
City 3	5500	2	54	300	1
City 4	3300	1	35	500	3
a	700	1	8	1000	3
b	1300	1	10	700	2
c	1200	1	9.4	200	4
d	1470	1	11	180	4
e	1690	2	12	307	5
f	2570	2	17	367	8

(continued)

Table 5 (continued)

	Population	Vehicle per capita	Area of land (km^2)	Employability	Commercial area (km^2)
g	2600	2	19	147	3
h	1460	1	11	425	5
i	1800	1	15	853	4
j	1700	1	14	743	2
k	2003	2	17	489	5
l	2469	2	20	654	7
m	2435	1	21	598	3
n	2130	2	19	456	8
o	2378	1	23	873	3
p	1320	3	10	158	5
q	1200	2	9.5	345	2
r	1400	1	11	123	6
s	1700	1	14	674	4
t	1520	1	13	356	3

Table 6 Distance matrix

	City 1	City 2	City 3	City 4
City 1	0	90	120	63
City 2	90	0	60	132
City 3	120	60	0	99
City 4	63	132	99	0
a	30	60	–	–
b	60	30	–	–
c	–	20	40	–
d	–	40	20	–
e	–	–	33	66
f	–	–	66	33
g	42	–	–	21
h	21	–	–	42
i	40	–	80	–
j	80	–	40	–
k	–	88	–	44
l	–	44	–	88
m	20	110	140	83
n	40	130	160	103

(continued)

Table 6 (continued)

	City 1	City 2	City 3	City 4
o	110	20	80	152
P	130	40	100	172
q	140	80	20	119
r	160	100	40	139
s	83	152	119	20
t	103	172	139	40

Table 7 Modified OD matrix for city 1

	1	2	3	4	5	c2	c3	c4	e1	e2	e3
1	82	98	85	4	18	6	4	9	38	34	45
2	91	96	94	9	49	6	4	9	57	16	8
3	12	49	68	83	45	6	4	9	7	80	23
4	92	80	76	70	65	6	4	9	5	31	92
5	63	14	75	32	71	6	4	9	53	53	15
c2	19	19	19	18	18	4	5	10	78	16	83
c3	24	24	24	24	24	5	4	11	94	60	54
c4	17	17	17	16	11	5	5	9	13	26	100
e1	96	96	71	38	66	70	93	28	57	66	7
e2	97	66	3	77	16	89	35	76	47	69	44
e3	15	3	27	80	12	96	19	76	1	75	10

Table 8 Modified OD matrix for city 2

	1	2	3	4	5	6	7	c1	c3	c4	e1	e2	e3
1	80	63	13	75	49	82	17	7	5	5	19	99	16
2	2	68	72	50	16	72	39	7	5	5	43	73	18
3	93	39	10	48	98	15	83	6	5	5	48	34	42
4	73	37	66	91	71	66	81	6	5	5	12	58	9
5	49	99	49	61	50	52	6	6	4	5	59	10	60
6	58	3	78	62	47	98	40	6	5	5	22	91	47
7	23	89	72	86	6	65	53	6	4	5	38	88	70
c1	16	16	16	16	16	16	16	14	15	16	58	82	70
c3	14	14	13	14	13	14	13	15	12	15	25	26	64
c4	10	9	9	10	10	10	10	11	11	9	29	60	3
e1	52	26	70	24	52	83	29	74	17	86	62	2	6
e2	23	33	19	89	9	8	43	27	12	65	26	42	32
e3	49	68	3	2	82	13	1	42	100	38	83	31	53

Table 9 Modified matrix for city 3

	1	2	3	4	5	6	7	8	9	c1	c2	c4	e1	e2	e3
1	66	44	22	32	77	23	54	22	9	6	4	9	9	88	59
2	41	53	67	79	19	12	87	10	53	6	4	9	11	99	24
3	82	46	85	47	29	61	26	11	53	6	4	8	13	0	67
4	72	88	34	3	9	45	32	6	86	6	4	9	68	87	8
5	97	52	78	17	58	46	12	40	48	6	4	9	50	61	63
6	53	95	68	72	69	66	94	45	39	6	3	9	19	99	66
7	32	64	0	47	55	77	65	36	67	6	3	8	49	53	73
8	10	96	60	15	42	35	48	77	74	5	3	8	14	48	89
9	61	24	39	34	65	66	64	63	52	6	4	9	5	80	99
c1	30	30	30	30	30	30	30	30	30	11	11	11	85	23	77
c2	21	21	21	21	21	21	21	21	21	15	15	15	56	50	58
c4	18	17	18	18	18	18	18	18	17	8	8	7	93	90	93
e1	26	70	42	24	95	25	72	19	26	27	67	79	70	58	58
e2	15	6	46	92	21	61	52	14	4	19	44	32	58	85	1
e3	28	25	77	27	71	58	100	70	76	82	84	53	82	74	12

Table 10 Modified OD matrix for city 4

	1	2	3	4	5	6	c1	c2	c3	e1	e2	e3
1	87	12	12	10	99	47	4	2	2	44	30	34
2	48	20	49	14	86	65	4	2	2	1	4	29
3	85	14	86	16	79	2	4	2	2	90	51	75
4	21	19	88	62	51	85	4	2	2	19	76	1
5	55	4	27	57	17	56	4	2	2	9	63	4
6	63	64	21	5	40	86	4	2	2	31	9	67
c1	22	21	21	21	21	21	8	9	8	46	8	60
c2	15	15	15	16	15	15	10	11	10	10	78	53
c3	20	20	20	19	20	20	10	11	10	100	91	73
e1	5	50	20	6	30	17	77	47	7	33	53	71
e2	49	54	95	86	29	66	82	92	24	30	11	78
e3	19	44	8	94	33	33	10	10	5	6	83	29

6 Conclusion

Urban travel demand has been the focus of major research initiatives primarily aimed at estimating the urban OD matrices from various building units—traffic count and its proxies the most. However regional OD matrices are required for planning purposes by the policy makers to undertake decisions considering the regional setting. Utilising the same process for building the regional OD matrix is prohibitively expensive and time consuming. This paper presents novel ways of establishing the regional travel demand based on the already available urban travel demand matrices that could belong to different time periods and have non-specific origins/destinations. This paper presents four cases/possibilities which have been addressed effectively through the formulated methodology and has been further discussed using a numerical example.

However, it needs to be seen as to how this methodology works in the real world which can be tested only with a case study of a region with the availability of several urban OD matrices. This is one of the limitations of the study which would be addressed in the future work. Nevertheless, it is evident that the methodology presented holds promise in the direction of estimating regional travel demand from intraregional travel demand models. This could be an efficient tool in the hands of policy makers dealing with regional planning and forecasting efforts.

References

1. Ortúuzar JD, Willumsen LG (2011) Modelling transport, 4th edn. Wiley, United Kingdom
2. Wilson AG (1974) Urban and regional models in geography and planning. Wiley, Chichester
3. Willumsen LG (1981) Simplified transport models based on traffic counts. Transportation 10:257–278
4. Bell MGH (1983) The estimation of an origin-destination matrix from traffic counts. Transp Sci 17(2):198–217
5. Cascetta E, Marquis G (1993) Dynamic estimators of origin-destination matrices using traffic counts. Transp Sci 27(4):363–373
6. Chang G, Wu J (1994) Recursive estimation of time varying origin-destination flows from traffic counts in freeway corridors. Transp Res Part B 28(2):141–160
7. Yang H, Iida Y, Sasaki T (1994) The equilibrium-based origin-destination matrix estimation problem. Transp Res Part B 28(1):23–33
8. Yang H (1995) Heuristic algorithms for the bilevel origin-destination matrix estimation problem. Transp Res Part B 29(4):231–242
9. Bera S, Rao KVK (2011) Estimation of origin-destination matrix from traffic counts: the state of the art. Eur Transp 49:2–23
10. Parry K, Hazelton ML (2012) Estimation of origin-destination matrices from link counts and sporadic routing data. Transp Res Part B 46(1):175–188
11. Caceres N, Romero LM, Benitez FG (2013) Inferring origin–destination trip matrices from aggregate volumes on groups of links: a case study using volumes inferred from mobile phone data. J Adv Transp 47:650–666
12. Lorenzo M, Matteo M (2013) OD matrices network estimation from link counts by neural networks. J Transp Syst Eng Inf Technol 13(4):84–92

13. Cascetta E, Papola A, Marzano V, Simonelli F, Vitiello I (2013) Quasi-dynamic estimation of o-d flows from traffic counts: formulation, statistical validation and performance analysis on real data. Transp Res Part B 55:171–187
14. Ma W, Qian S (2018) Estimating multi-year 24/7 origin-destination demand using high-granular multi-source traffic data. Transp Res Part C 96:96–121
15. Krishnakumari P, Lint H, Djukic T, Cats O (2019) A data driven method for OD matrix estimation. Transp Res Part C. https://doi.org/10.1016/j.trc.2019.05.014
16. Ma W, Pi X, Qian S (2020) Estimating multi-class dynamic origin-destination demand through a forward-backward algorithm on computational graphs. Transp Res Part C 119:102747
17. Ashok K, Ben-Akiva ME (2000) Alternative approaches for real-time estimation and prediction of time-dependent origin-destination flows. Transp Sci 34(1):21–36
18. Cipriani E, Florian M, Mahut M, Nigro M (2011) A gradient approximation approach for adjusting temporal origin-destination matrices. Transp Res Part C 19(2):270–282
19. Caggiani L, Ottomanelli M, Sassanelli D (2013) A fixed point approach to origin-destination matrices estimation using uncertain data and fuzzy programming on congested networks. Transp Res Part C 28:130–141
20. Foulds LR, Nascimento HAD, Calixto ICAC, Hall BR, Longo H (2013) A fuzzy set-based approach to origin-destination matrix estimation in urban traffic networks with imprecise data. Eur J Oper Res 231(1):190–201
21. Tympakianaki A, Koutsopoulos HN, Jenelius E (2015) c-SPSA: cluster-wise simultaneous perturbation stochastic approximation algorithm and its application to dynamic origin-destination matrix estimation. Transp Res Part C 55:231–245
22. Shen L, Shao H, Wu T, Lam WHK (2019) Spatial and temporal analyses for estimation of origin-destination demands by time of day over year. IEEE Access. https://doi.org/10.1109/ACCESS.2019.2909524
23. Chu KF, Lam AYS, Li VOK (2020) Deep multi-scale convolutional LSTM network for travel demand and origin-destination predictions. IEEE Trans Intell Transp Syst 21(8):3219–3232
24. Yao X, Gao Y, Zhu D, Manley E, Wang J, Liu Y (2020) Spatial origin-destination flow imputation using graph convolutional networks. IEEE Trans Intell Transp Syst 1–11
25. Perrakis K, Karlis D, Cools M, Janssens D, Vanhoof K, Wets G (2012) A Bayesian approach for modeling origin-destination matrices. Transp Res Part A 46(1):2002–2212
26. Frederix R, Viti F, Himpe WWE, Tampère CMJ (2014) Dynamic origin–destination matrix estimation on large-scale congested networks using a hierarchical decomposition scheme. J Intell Transp Syst 18(1):51–66

Underlying Factors and Dependencies Towards a Dynamic Vehicle Ownership Model in India: A Content Analysis Approach

Somnath Bhui and Debapratim Pandit

Abstract Most of the vehicle ownership studies have been conducted in developed countries using static models. Undermining the temporal dynamics of this phenomenon may lead to uncaptured underlying mechanisms. The rapid economic growth in India, coupled with its diversity presents a rather different context to focus upon. Considering this, a dynamic vehicle ownership model is required where various biological domains and their interactions influencing vehicle choice decisions can be investigated. In this paper, we investigate the underlying processes and various factors behind vehicle ownership of a household, employing a content analysis approach on transcribed retrospective interviews. The outcome is a pathway that depicts various trajectories of vehicle ownership one can find in India. It is a first attempt, to capture the ownership trajectories of Indian households considering both two-wheelers and cars. Apart from the known factors, the role of employer, retirement, interpersonal interaction was found impacting vehicle ownership. The findings of this study can be used in further development of dynamic models of vehicle transactions.

Keywords Vehicle ownership · Mobility biography · Content analysis

1 Introduction

Vehicle ownership has been extensively studied from economics, marketing, and transportation perspective primarily for two reasons; determination of consumer demand of various types of vehicles, and its role in the activity-travel behaviour of individuals and households [1]. The former is beneficial for the manufacturers whilst the latter is relevant for policy makers, local government, urban planners who intend to understand the underlying reasons of vehicle ownership decisions

S. Bhui (✉) · D. Pandit
Department of Architecture and Regional Planning, Indian Institute of Technology Kharagpur, Kharagpur, India
e-mail: sbhui@iitkgp.ac.in

D. Pandit
e-mail: debapratim@arp.iitkgp.ac.in

L. Devi et al. (eds.), *Proceedings of the Sixth International Conference of Transportation Research Group of India*, Lecture Notes in Civil Engineering 272,
https://doi.org/10.1007/978-981-19-3494-0_10

towards devising efficient policies for travel demand management of a particular urban area. Needless to reiterate the significance of emission modelling in today's context, vehicle ownership is a crucial input. With the adoption of electric vehicles, and tele-commuting by office goers, these studies have become even more relevant.

Most of the studies in vehicle ownership have been conducted in the context of developed countries. Due to increase in fuel cost; automobile saturation; wealth effects; ageing populations; increased urbanization; improved alternative travel options and incentives; changing consumer preferences; and increased concerns for health and environment, peak car use has resulted in stagnation or even decline in car use in developed countries [2]. On the other hand, with rise in income in developing countries including India, it is hypothesized that automobile ownership will also increase. India also being different in terms of demography and cultural diversity may have different implications on vehicle ownership. Additionally, in the wake of policies like allowing non-transport (personal use) cars to serve as pool cars; voluntary scrappage of personal vehicles etc. in India, it is of paramount importance to investigate their impacts on vehicle ownership decisions.

There have been numerous studies on estimating vehicle ownership, and identifying influencing factors for the same using cross-sectional data [1, 3–9]. Due to lack of time dependency in such models, many underlying mechanisms remain uncaptured. Hence a dynamic vehicle ownership modelling approach is the need for the day where various biological domains and their interactions influencing vehicle choice decisions over time can be investigated [8, 10–22]. Identifying them can equip us with the knowledge of factors and conditions that lead to the changes in vehicle ownership state. Hence, we propose to develop a dynamic vehicle ownership model, with a narrative accounting for the factors and dependencies of vehicle ownership in India.

In this paper we attempt to find the underlying processes behind vehicle ownership of a household. The motivation behind this study is our endeavour to develop a framework to predict the vehicle ownership of a household, during a certain time period, based on demographic and landuse characteristics. We employed a retrospective mobility biography approach to find the processes in which we interviewed 40 households from multiple Indian cities. From the transcribed interviews, the factors have been identified using a content analysis approach.

2 Literature Review

The most common socio-economic characteristics used in transportation studies are; gender, age group, household size, household income, education level, employment status, marital status, etc. Increase in age reduces personal trips [23]. Individuals lesser than 30 years age are less likely to have cars, whereas, individuals beyond 65 years age are more likely to replace old vehicles [12]. Evidence of gender-based differences in choice making behaviour between men and women has also been found [24]. Cars have been found to be a more prized possession to men than women

[25]. In almost all the studies, household income has been consistently found to be correlated with increase in vehicle ownership [2, 12, 23, 26, 27]. Employment type, location, duration impacts vehicle allocation in a household [28, 29] and hence its role in vehicle ownership cannot be neglected. Education has not been found to have any role on vehicle ownership, but recent studies dealing with the choice of electric vehicles and autonomous vehicles, have found to have some impact on the vehicle type choice. Theory of planned behaviour highlights the importance of attitudes, subjective norms, and perceived behavioural controls behind the intention to take a decision, as the best predictor of a person engaging in a certain behaviour [30]. Psychosocial and symbolic-affective benefits also play a major role behind car ownership, and even has higher impact than instrumental benefits [6, 31–33].

With regards to the tools used to analyse data in such studies, there are instances of Multiple linear regression, Non-linear regression, Ordered probit, Ordered logistic models, Multinomial logit models, Nested logit models, and Mixed logit models, being used in identifying the static and dynamic factors behind various aspects of vehicle ownership [4–6, 23, 34–39]. Endogenous implications of vehicle ownership has been studied using techniques like Bayesian belief network, Bayesian multivariate ordered probit and tobit models, Mixed multidimensional choice modelling, Copula-based models, and Discrete-continuous models [1, 19, 40–44]. Popularization of hazard-based duration models brought another approach to account for the dynamic factors like the lead and lag effects of certain events on vehicle ownership [15, 21, 45, 46]. These studies are increasingly gaining importance with the adoption of microsimulation-based activity-travel demand models for urban areas where, both long and short term transportation decisions, life-course events, and their interrelationships are considered [11, 12, 42]. Panel data required for such longitudinal studies is non-existent in most of the urban areas in the world, which can be resolved using retrospective mobility biography approach [47]. Mobility biography research (MBR) investigates the life-course trajectories that link directly to an individual's travel behaviour such as availability of a private car, public transport season ticket, driving licence or actual travel patterns. The novelty of this approach is in its framework that allows analysis of travel behaviour of individuals with respect to choices made in other domains of life. Key events and critical incidents which are the unit of a life-course trajectory can play a role in formation of attitudes [48]. Events like car acquisition, driving licence acquisition and marriage can impact vehicle ownership and have indirect impact on travel behaviour; on the other hand, key events like childbirth, relocation, employment/education, incidents can have direct impact on travel behaviour and no or indirect impact on mobility resource [32]. Urban form, quality of transport (and built environment), preferences, activity patterns, financial and time resource are also found to be important in explaining any travel behaviour and vehicle ownership. Based on social norms and structural condition, patterns can also be identified in MBR to inductively reconstruct biography of others for generalization [33]. Prediction models of vehicle ownership using life-course events have also been undertaken to measure the impact of desire to replace existing vehicle;

anticipated change in mobility needs; characteristics of alternatives; and other socio-demographic characteristics on the acceptance of EV, electric bikes, and car sharing [49].

Although life-course-based studies mentioned above delve into dynamism, there is a lack of studies accounting for interdependencies amongst various spheres of life-course events. The need to study the interdependencies between other spheres of long term mobility decision like residential location choice and employment location choice, and impacts of social interaction has been identified in literature [50–52]. Transcending individual characteristics, very few studies have been undertaken to explain the implications of interpersonal interactions, essentially of two types; household and social [53], on vehicle ownership, which have been already validated for travel behaviour [11, 18, 50, 54, 55].

3 Methodology

With the intention to unravel the patterns in Indian households in vehicle ownership decisions, 40 in-depth interviews were conducted from multiple cities in India, using convenience sampling. A summary of the household composition, age of household head, and monthly income is shown in Table 1. Qualitative studies ought to have a sample size between 25 and 30 [56]. There are also a couple of qualitative studies in transportation domain, having sample size between 10 and 32 [32, 33, 57–61]. A meta study of 560 Ph.D. thesis found 62 and 70 being the maximum sample size considered for life-history-based and content analysis-based studies, respectively [62]. Whilst quantitative surveys rely on the concept of confidence interval; qualitative studies, by virtue of its intense labour, and the motive to understanding the process, relies more on the concept of saturation. Saturation is the condition when addition of new data does not shed any further light on the issue under investigation. Based on a method

Table 1 Descriptive statistics of the sample

Household composition	No. of families	Age of HH head	No. of families	Monthly income	No. of families
A or H-W	2	30–40	3	10,000–40,000	3
H-W, 1C	7	40–50	11	40,000–50,000	3
W, 2C	1	50–60	20	50,000–60,000	3
H-W, 2C	19	60–70	3	70,000–80,000	5
H-W, 2C, 1P	1	70+	3	80,000–100,000	12
H-W, 2C, 1D, 1G	1			100,000+	14
H-W, 3C	9				

A Alone; *H* Husband; *W* Wife; *C* Children; *D* Daughter-in-law; *P* Parent to assess saturation, this study attained saturation (0% threshold), with a base size of 6, and run length of 3, at the 27th observation [63].

The interviews were open ended and lasted for 40 min on an average, and even stretched over multiple days sometimes. We intentionally targeted respondents who are either working or pension holders, in the age group of 32–78 years, as below that age a person seldom has any vehicle ownership biography in the Indian context. Respondents were asked about their trajectory of residential mobility, household structure, employment, income, and vehicle ownership. Their aspiration to undergo any vehicle transaction in near future; and impact of interpersonal interactions on vehicle ownership was also explored. Respondent's perception about car sharing, electric vehicles, and autonomous vehicles; and impacts of COVID-19 pandemic on vehicle ownership was also recorded. Following the data collection, the interviews were transcribed and simultaneously themes were identified using content analysis approach, employing inductive and deductive thinking [64], based on their dialogues recorded during the interview.

4 Results and Discussion

4.1 Physical Factors

Evidences from the survey validating various physical aspects of vehicle ownership has been shown in Table 2. Household income arguably is the most common factor playing a role in vehicle ownership. High neighbourhood accessibility to services and infrastructure; Landuse characteristics; parking availability; policies and regulations regarding vehicle ownership also plays a significant role. Owning a motorized vehicle being a liability to the household, only gets opted if feasibility is achieved in competition to public transport (PT) and non-motorized (NMT) modes, for a given time-money budget. Provision of housing and transport service for commuting, by the employer, defers vehicle ownership of single driver households, unless addition of a new driver is anticipated in a few years. This deferral may reduce, in case the number of non-regular trips increases for an individual. Employers often provide incentive on vehicle purchase and/or fuel; or incentivize tele-commuting. Whilst the former may increase vehicle ownership, the later may lead to reduction in vehicle ownership. The lack of parking has created opportunities for spaces of occasional use to be used as paid parking space. The urban form and characteristics of an area, emerging from the coming together of organically developed settlements often results in bottlenecks for vehicle accessibility, constraining vehicle ownership. Dense and compact settlements have been found to lead to low car ownership but high two-wheeler ownership. Although less in numbers, mobility impairment due to accident or geriatric and general health condition can also lead to acquisition and disposal of certain types of vehicles. This generally leads to reduced activities with intensive driving.

Table 2 Physical factors impacting vehicle ownership

Theme	Statements	Instances
Income and employer	*"Purchasing power is the main factor behind vehicle ownership …"*	*2*
	"I used to commute in office bus"	*2*
	"We get incentives for fuel and vehicle purchase from our employer"	*3*
PT, NMT, policy and regulation	*"We never owned a motorized vehicle as everything was at walking distance, and public transport was very good … In Singapore, due to high revenues, private cars are owned only by millionaires"*	*1*
	"Lack of administrative control over laying local roads results in bottlenecks for access to big vehicles …"	*1*
	"Lack of security in the area triggered to acquire a scooty/car, even though children could ride bicycle/scooty"	*1*
Urban form	*"Schools, hospitals, market, etc., are located in and around the older part of the city … City has expanded way beyond that area, making it necessary to own a vehicle"*	*1*
	"Due to inequitable distribution of services, multiple motorized vehicles are required per household"	*1*
	"Unconducive conditions for NMT makes motorized vehicle ownership an automatic decision …"	*2*
	"Due to congested roads, narrow streets, and lack of parking, we had to buy a two-wheeler for short and frequent trips …"	*4*
Parking	*"I wish to own a car, but unable due to lack of parking space"*	*2*
	"We had to sell off our car twice due to disputes related to parking space …"	*1*
Mobility impairment	*"After my accident I had to sell my bike and buy a scooty"*	*1*
	"Sold bike after getting diagnosed with knee problems which prevented me from kick-starting my bike and lifting my leg around the seat …"	*2*
	"Father's diagnosis of backbone issues, stopped his bike riding"	*1*

4.2 Subjective Factors

Evidences from the survey validating various subjective aspects of vehicle ownership has been shown in Table 3. Instrumental benefits being important, subjective-affective benefits of vehicle cannot be overlooked. Similar to studies in literature, we found psychological factors like, attitude and self-image impact vehicle ownership [2, 6]. A pro-technology attitude tends to increase the number of vehicle transactions. Perception of various latent aspects like safety, comfort, etc. also plays a role in vehicle ownership. Despite the direct proportionality between income and vehicle

Table 3 Subjective factors impacting vehicle ownership

Theme	Statements	Instances
Pro-technology	*"I have changed many vehicles because I had the desire to experience the latest vehicle in the market and aspired to own them"*	*1*
	"I always want to possess the latest model of a certain bike, so change my bike every 3–5 years …"	*1*
Self-image	*"There are people who buy vehicles for show-off … The ethical and cultural values in a household plays a role in the consumption pattern of the household …"*	*1*
	"Manufacturers exploit the behaviour of people to maintain a self-image by charging a lot for image boosting features like alloy wheels, exquisite interiors, etc."	*1*
	"Getting bored with a colour of vehicle resulted in changing it every 3–6 years …"	*1*
	"Need to maintain a certain social status due to employment type, resulted in car acquisition …"	*2*
	"Two-wheeler vehicles are judged based on their functional aspect, but a car also has social value … a social status … attached to it …"	*3*
Comfort, safety	*"Given a financial state, safety, comfort, and lifestyle are the factors that people want to pay extra for, whilst buying car …"*	*1*
Dream	*"I always wanted to buy a car …"*	*2*
Independence	*"We travel to our hometown frequently which is not well connected with transit … We acquired car to be reduce dependence on transit …"*	*1*
	"Car could give us freedom to choose our travel time for long trips …"	*2*
Habit	*"After my bike was stolen, I bought another bike within a year because I had by then developed the habit of using it"*	*1*
	"After disposal of scooty, had to buy another one as it became an integral part of our activity scheduling"	*1*
	"After disposing car, bought a scooty for wife as she was habituated with the independence in mobility …"	*1*

ownership, ‘habit’ has been known to induce inelasticity in the relationship, and in the survey also it has been found to trigger vehicle ownership. This happens due to the exposure of household to the acquired benefits of vehicle ownership leading to escalation of the value of the vehicle for the household, and impelling to maintain a certain level of ownership, by replacing or acquiring (for e.g.: in case of theft) the vehicle as and when required. The need to maintain a certain social image (e.g. businessmen, entrepreneurs), and independence are other subjective factors found to be impacting vehicle ownership.

4.3 Life-Course

The surveys show that vehicle ownership and use trajectory of an individual starts with bicycle, as it is considered a life skill. With the growing up of the individual, a larger bicycle is acquired. Later on as the value of time of the individual increases, a motorized two-wheeler is added to the fleet. Since that point of time, the bicycle gets reserved for neighbourhood trips, and a tool for learning the skill for others, till it becomes unserviceable. Prior to first-time acquisition of a certain motorized vehicle type, in order to get acquainted with the changes in driving experience, people tend to buy a used vehicle whose retention period ranges from 2 to 10 years, depending upon the mileage, reliability, comfort, safety, speed, etc. offered, and the need of the household. Long travel times, and unavailability of reliable public transport, triggers the need to have a faster and/or bigger vehicle. With the increase in household size, travel needs become overwhelming for a person to satisfy. Hence cases of scooter being acquired for serving as a vehicle for both the spouses are plenty, allowing the female counterpart to contribute in the overall travel needs of the household.

Presence of children is a trigger in bicycle acquisition. The bicycle meant for children, if bought during the early years, is small and is meant only for training purposes. It remains in the fleet for 3–6 years and eventually, based on attitudes, and perceptions of safety and security, gets replaced by a larger bicycle for making neighbourhood trips. The older bicycle is either used by some other user, or it is disposed of.

People who are or were employed, were found to be abstaining from four-wheeler ownership, till a few years from retirement. But ownership of motorized two-wheeler was found to be a contemporaneous or lagging event to a few years of employment. Low average income in India; time taken in becoming resource surplus; aversion to being exposed to weather elements in the later years of life; perceiving car as an emergency vehicle, may cause such behaviour. Motorized two-wheelers, acquired at old age are essentially physically less demanding. Depletion of skill and appetite of driving may also result in transactions aiming to reduce physical and psychological stress.

Having no dependencies (marriage), and relocating to an undesirable place; occurring simultaneously, may restrain any big investment, till a point where owned vehicle can make access to desired location feasible. Evidences of all the above life-course event impacts are shown in Table 4.

4.4 Interpersonal Interactions

All the vehicle ownership decisions post marriage has exhibited some sort of intra-household interaction. The evidences of interpersonal interactions influencing vehicle ownership has been shown in Table 5. Mobility of the spouse; the needs of the children; differences amongst the household members, are some of the factors found. Financial independence coupled with low household size has been found to reduce the significance of intra-household interactions with respect to mobility decisions.

The functional attribute level decisions are made by the head of the household with the help of a decision support system. This system may include experienced colleagues, professional drivers, and family members as well. Although the impact of intra-household interaction has been found to dominate the decision making process, Social interaction has also been found to impact the final choice, apart from the self-bias of the individual. The exhibition of snob effect is evident from the fact that people pay extra for features that can set them apart from the mass [65]. This behaviour is capitalized by the manufacturers by packaging other features that maybe rather irrelevant for the consumer. Few cases of snob effect in consequent two-wheeler transactions have been found. Habit can also be induced by the social utility of a vehicle through appreciation from the society, reinforcing the position of the vehicle in the fleet. Although desire is something related to the personal traits of a person, people acknowledge of it being induced by social pressure as well.

4.5 Disposal

A vehicle gets disposed (scrapped/transferred/sold) when the cost of maintaining it exceeds the benefit obtained from it. Such condition may arise when a vehicle becomes irrelevant, obsolete with respect to market trends, and/or qualify as ELV (End of Life Vehicles). Irrelevance is the case when the vehicle ceases to serve the prime objective behind its acquisition. Ageing/demise of the user has been found to be a reason for a vehicle becoming irrelevant. ELVs are defined as vehicles having surpassed the prescribed age, defined by a competent authority. Often such vehicles can incur high cost of repair, increasing the opportunity cost of keeping the vehicle(s). Under such circumstances, breakdown may become a trigger for disposal of the vehicle. Obsolescence of a vehicle is realized when parts for the vehicle cease to be available in the market, or the model itself loses its relevance in the market.

First hand vehicles end up being second hand, and second hand vehicles become ELV, and end up being scrapped. Selling and transfer of ownership becomes a means of disposal when the vehicle is in operable condition. Emotional attachment to vehicle

Table 4 Life-course event impacts on vehicle ownership

Theme	Statements	Instances
Ageing	*"Father has retired, and parents are growing old, so thinking to get a car"*	*1*
	"Propensity to travel in bus is getting reduced with age"	*1*
	"Deferring car buying would reduce my chances to own one, as I will not get licence after 60; with age I may become driving averse"	*1*
	"Parents getting old, So I may buy a car for their mobility during any medical emergency or family trips"	*2*
	"… My wife sold her heavy scooty and bought a lighter one as it would be easy to use for her"	*1*
Childbirth; household size	*"Upgraded to scooter as the former was for practise only. Later upgrade to a bike as family size increased and more power was required. Again upgraded to car as kids grew up and we relocated to a place far from church …"*	*1*
	"We have a car now, but an anticipated family expansion triggers us to add a 7 seater car to the fleet"	*2*
	"After childbirth, medical trips became necessity, so bought a bike"	*1*
	"Small car was not adequate for the family with children growing"	*3*
Growth; increased travel demand	*"I rode a used scooter for 10 years, and bought a new motorized-bike after that, which I had to replace with another having a larger engine, in 3 years due to increased travel for business growth …"*	*1*
	"I bought a scooty so that my wife can escort my daughter to tuitions, and serve as a secondary vehicle for my commuting as well …"	*2*
	"With children growing, upgraded to bike as the distances of tuition increased and so did the value of time"	*1*
	"School van takes a lot of time; my child/children would starve. Added a scooty for my wife to be able to escort kid(s) to school, and for my own use as well …"	*2*

(continued)

Table 4 (continued)

Theme	Statements	Instances
	"I need to make frequent trips to my hometown, so bought a car"	*2*
	"We got our son first bicycle at the age of 6 for learning, second and third at the age of 11 and 16 for educational trips and occasional use by me also. He still uses his bicycle for short trips …"	*1*
Relocation	*"Was posted far from home … got my first vehicle after I relocated to a city near my native village/town … before getting married"*	*2*
Retirement	*"Before retirement, we replaced the old car with a new car, better in terms of mileage, comfort, and speed …"*	*1*
	"… I bought the car in the year I retired, to escort my son to and fro airport. We also thought it to be helpful in emergency situations"	*1*

may play a role in deterring disposal, and can even dictate if the vehicle is sold to a known person or an unknown person. Although unreported, it is a pretty common to find ELV vehicles being operated in rural areas, due to lack of enforcement of regulations. The various patterns in disposal are shown in Table 6.

4.6 Emerging Technology

Shared mobility and electric two-wheelers are the only two emerging transportation modes that people in the study area have been exposed to. Shared mobility was reported desirable as it can liberate households from various costs associated with ownership, and can be economical too. So far electric mobility/Electric vehicle (EV) is concerned, the willingness of people to accept it was overwhelming due to its; cost effectiveness from fuel, low maintenance, low overhead costs; safety, due to speed limit (in certain vehicles); and reduced vehicular emission. Despite these, people are sceptic due to the high initial cost; non-existent charging infrastructure; and unreliability of battery. People anticipate policy intervention by government with ardour, to ease out the adoption of electric mobility in both two-wheeler and cars. Although unknown, people, on explaining the features of autonomous vehicles (AV), pointed out heterogeneous traffic and lack of road discipline as the major challenge to it, and affirmed the deployment of AV in controlled environments like university campuses, army cantonments, etc. Also, people with driving aversion found it fascinating. The various perception towards the emerging technologies in transportation sector are shown in Table 7.

Table 5 Interpersonal interaction impacts on vehicle ownership

Factor	Statements	Instances
Intra-household differences	*"Discussion is done at household level on whether to buy a particular type of vehicle or not. The model, capacity, etc. is decided by me"*	*1*
	"I could not buy car as my family was not cooperative. By the time I could persuade them, I was about to retire"	*1*
Spouse's mobility	*"We were determined that the acquired vehicle will be driven by both of us (spouses), so scooty was the only option. The brand and all were decided by me"*	*1*
	"We got a scooty to empower my wife to contribute to household travel needs, and to get a secondary vehicle for myself"	*3*
Financial independence	*"After starting to earn, children did not discuss their mobility decisions with the rest of the household, and acquired vehicle on their own, as per their need, but we share the parking space"*	*2*
Decision support	*"Definitely relied on my colleagues for new information on any vehicle, as many of them have decent experience about vehicles. One should also talk to professional drivers as they might have handled vehicles of many manufacturers, and can point the pros and cons of each of them"*	*1*
	"For new vehicle, social network definitely helps as a decision support system by sharing their personal experience, giving test drives, etc. …"	*1*
	"Even though the primary decision on any aspect of vehicle (especially aesthetic) totally depends on family members, I would take cues from my social network for the need to buy, vehicle type, and model …"	*8*
Social pressure	*"Given the choice between two vehicles, I would choose the one that is bought more by people … This will assure me about its post-sales service reliability … People in KRVS (a social group) encouraged me to buy a motorized vehicle which reinforced my decision to buy a scooty"*	*1*
	"The more a vehicle is present in the market, the more reliable it is … in terms of servicing and availability of parts"	*2*
Social utility	*"My in-laws did not have any vehicle. So I had to take them to (long distance) places, whenever required"*	*1*
	"I am attached to an NGO in which my car finds great utility for travelling for social welfare works"	*1*

Table 6 Factors supporting/preventing vehicle disposal

Factor	Statements	Instances
End of life	*"I disposed my bike as it became increasingly unreliable day by day"*	*1*
	"Extensive use of the bike resulted in a bad condition"	*12*
Poor performance/major repair	*"Wife met with an accident whilst riding scooty … had to sell it afterwards …"*	*1*
	"Vehicle was offering a very low mileage"	*2*
Obsolescence	*"Disposed my scooter as the parts were not available. But as I was emotionally attached to it (first vehicle), I gave it to someone I know would maintain it properly"*	*1*
	"The vehicle model had become too old …"	*5*
Irrelevance	*"I sold the car as I was getting old and was nervous to drive it in busy roads these days, and my stopped visiting us"*	*1*
	"Scooter/bike became too small for the whole family travelling together"	*5*
	"After Husband's death, his vehicle was of no use … and was old anyways …"	*1*
	"Car became too small for the whole family"	*1*
	"Requirement to uphold social status resulted in disposal of two-wheeler …"	*6*
Emotional attachment	*"I sold my second hand scooter after using for 3 years. I transferred my (first hand) bike to my uncle after using for 12 years"*	*1*
	"I still possess the bike my father gave me. I maintain it to the best of my capacity and use it"	*1*
	"Even after buying new large car, we will not replace or dispose the existing one. We will use it for trips with a smaller party size"	*1*

4.7 Impact of Pandemic

The prevailing COVID-19 pandemic has a huge impact on economy worldwide. In order to investigate its impacts on the vehicle ownership behaviour, we asked people about their experiences and perceptions. Given the travel curbs in public transit modes like capacity restriction, dynamic fare, suspension of suburban trains, etc., and the highly contagious nature of the virus, people have adopted one or combination of these actions; curtailing trips, tele-commuting, job change, travelling by personal vehicles. This has led to changes in the regional economy and travel behaviour. The reported implications of the pandemic on vehicle ownership is shown in Table 8.

Table 7 Perception of emerging technology

Factor	Statements	Instances
Shared mobility	*"Shared mobility is a very sustainable and economical mode … it has potential to curtail the need to own a personal vehicle"*	*1*
Operating cost	*"EV is light on maintenance, overhead cost (licence, insurance, PUC) and hence it is pretty desirable … Will opt if I require to dispose existing one"*	*1*
	"Low fuel cost is the main reason we would opt EV …"	*10*
Sustainability	*"EVs are good for reducing air pollution, and overall climate change"*	*5*
Safety	*"… Electric 2-wheelers have speed limit, preventing from indulging into over speeding"*	*1*
	"Would get son/daughter EV for safety (low speed) … Not appropriate for my use …"	*1*
	"AVs will be much safer than CVs as it will follow rules"	*1*
Slow traffic	*"Slow EVs, slows down the traffic … People riding them often encounter undesirable incidents on road … They should have a separate bay …"*	*1*
Reliability	*"I am sceptic about the reliability of the batteries used now; their capability of keeping up the performance for long distances, and longer durations … lack of charging infrastructure is also a major constraint behind people not opting for EVs, for longer drives"*	*15*
Literacy and ethics	*"The literacy level of India is a huge challenge to adoption and success of AVs … Most people do not follow road rules"*	*5*
Controlled deployment	*"In campuses where the traffic is controlled and have ample surveillance (e.g.: army cantonments, University campuses, etc.), AVs can be adopted relatively faster"*	*1*
Accessibility to driving averse	*"I do not enjoy the act of driving, I will surely add AV to the fleet, subject to affordability. Level of education, and political temperament in the region are also factors behind faster adoption"*	*1*

4.8 Pathway Pattern

It is evident that the dimensions identified from the study influence the decision making of a household, and even plays in tandem with each other, resulting in various patterns, leading to heterogeneities in outcome. Given particular circumstances, a

Table 8 Impact of pandemic on vehicle ownership

Factor	Statements	Instances
Toned down transit	*"Flouting of COVID protocols discourage people from taking PT. The reduced frequency and dynamic pricing has further led to people opting for personal vehicles"*	*1*
	"Lack of government initiative to revamp PT, resulted in overall increase in vehicle ownership"	*4*
Regional economy	*"People whose business depended on travelling, have started selling locally. This can have an impact in the regional economy of the area"*	*1*
Two-wheeler	*"Sale of two-wheelers (motorized, NMT) has increased. On the other hand, new car sales have been low"*	*1*
	"Two-wheeler ownership increased …"	*1*
Car ownership	*"This pandemic has created the need to have a car especially for households residing far from market, and with old people in the household"*	*1*
	"Used car ownership might have increased as it has less capital investment …"	*2*
	"4-Wheeler ownership decreased …"	*1*
No. of trips	*"Vehicle ownership has not changed, only the number of trips has decreased …"*	*2*
	"Trips have reduced but vehicle ownership has increased (safety, convenience), which may be detrimental to environment in long term …"	*8*

household may end up with completely different decisions related to vehicle ownership, as compared to other households. This very fact makes it imperative to study the dynamics within the household and the environment it is subjected to, as a requisite for true unravelling of vehicle ownership behaviour. We have tried to present a generalized trajectory based on the collected interviews, in Fig. 1.

A person may have a bicycle at an adolescent age for educational and neighbourhood trips, which may be (1) replaced if required. A (2) new bicycle maybe added, or the person may (3) upgrade to a used motorized two-wheeler, keeping the bicycle as backup. Eventually, getting acquainted with new mode may have a (4) lagged effect on the disposal of bicycle, or it may get disposed due to due to lack of use. (5) Acquisition of a new two-wheeler, has a contemporaneous effect on disposal of the used one. All characteristics being the same, in future there may be (6) replacement of the same vehicle type, due to increased travel, unserviceability or theft. Further demand triggers may lead to acquisition of (7a) new motorized two-wheeler; (7b) used car; and (7c) bicycle. Although (7b) can potentially lead to disposal of motorized two-wheeler, but it seldom happens as it tends to hamper the mobility of household. (7c) can potentially be seen as a starting point of a new trajectory. Gaining competency in driving, (8) new car can be acquired, having a contemporaneous effect on disposal of the used car. In future, the (9) new car may be replaced with a similar one, owing

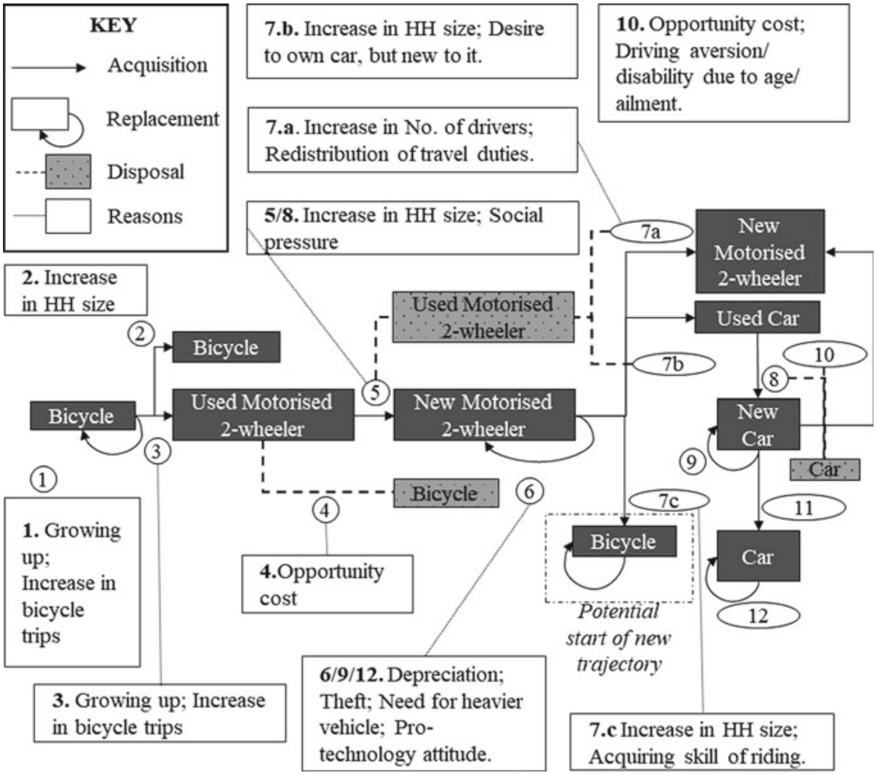


Fig. 1 Path of vehicle ownership for households

to reason similar to (6). Also, it can be disposed due to old age (of owner) and lack of utility, and (10) acquisition of motorized two-wheeler may be seen. (11) Given the demand, another car can also be added to the fleet. (12) The car may be replaced in future due to reason(s) similar to (9).

5 Conclusion

This study brings novelty in its endeavour to study the dimensions in vehicle ownership decision making by Indian households. Decisions on multiple vehicle types, connecting them with the life-course, and other mobility spheres of the households has been demonstrated in this study. The various dimensions identified, owing to their dynamic nature and interplay, results in patterns resulting in specific decision outcomes. This study presents one such generalized case; and with many other patterns yet to be elucidated, we argue the need for studies with regards to the mentioned nature of the various dimensions of vehicle ownership decision making.

Similar to the studies in literature, life-course events; increase in household size, presence of multiple workers, and relocation was found to be triggering vehicle ownership. Subjective factors that were found to be impacting are; desire, attitudes, cultural values, habit, perception of comfort, safety, and lifestyle. Physical factors like neighbourhood characteristics, urban form, parking space availability, number of employees, and household income constrain vehicle ownership. A peculiar behaviour of acquiring a two-wheeler was found when a household shifted to a denser city, in spite of having a car. Apart from that increase in propensity of car acquisition by employees approaching their retirement age, is a new finding. Also, for households with motorcycle, incident induced physical ability impairment results in replacement of motorcycle with a scooter, which demands lesser physical prowess. Intra-household interaction was found to have primary impact on all the aspects of vehicle ownership, whereas, social interaction impacted on the functional attributes of the vehicle. The employer, through provision of tele-commuting option, transport service and housing to employees, was revealed as a factor behind deferring vehicle ownership. On the other hand, incentivizing vehicle acquisition may lead to increase in vehicle ownership. Obsolescence and unserviceability were the two broad reasons behind disposal of any vehicle. Given the consistent rise on cost of fossil fuel; safety; and environmental sustainability, people are pretty optimistic about owning an electric two-wheeler in near future, although the same is not true for electric four-wheelers. Reason being, the high cost of reliable batteries, and almost non-existent charging infrastructure in India. With respect to the travel restrictions introduced due to COVID-19 pandemic, people think that the sale of two-wheelers has increased but at the same time, sale of cars has reduced. The probable reason may be the relatively higher initial cost of car, and the financial uncertainty due to pandemic.

Having recognized the underlying factors of vehicle ownership, the framework is an attempt to generalize the multiple pathways of vehicle career for a household. Eventually other nodes, and underlying reasons can be appended or inserted to get a more generalized template of the vehicle transaction behaviour of households. This understanding, coupled with appropriate survey for any study area, can be used to simulate the various life-course processes of the households in the study area, which would go on to give a more accurate estimate of vehicle acquisitions, vehicle replacements, and vehicle disposals, as compared to the current models. Taking inputs from the Residential location choice component, this model would serve as a plugin to activity scheduling component, and emission modelling component of the Landuse Transport Integration (LUTI) framework.

Although the underlying factors and phenomena have been identified from this pilot study, the magnitude of impact of each of them in vehicle ownership decisions, needs to be estimated. Also, there might be heterogeneities in the vehicle ownership behaviour of people from the same group, which have not been investigated, but can be taken up in future studies. The time lag or lead between various events, and their statistical significance, for different groups classified based on socio-economic parameters, also needs to be determined. Stated choice survey also can be coupled with this type of a study in order to understand the impact of past on future decisions. Further studies must ensure a representative sample in order to find reliable estimates.

References

1. Bhat CR, Sen S (2006) Household vehicle type holdings and usage: an application of the multiple discrete-continuous extreme value (MDCEV) model. Transp Res Part B Methodol 40(1):35–53. https://doi.org/10.1016/j.trb.2005.01.003
2. Gärling T, Friman M (2018) Economic and psychological determinants of ownership, use and changes in use of private cars. In: The Cambridge handbook of psychology and economic behaviour, 2nd edn, pp 567–594. https://doi.org/10.1017/9781316676349.020
3. Bansal P, Kockelman KM, Schievelbein W, Schauer-West S (2018) Indian vehicle ownership and travel behaviors: a case study of Bangalore, Delhi and Kolkata. Res Transp Econ 71(9). https://doi.org/10.1017/CBO9781107415324.004
4. Choo S, Mokhtarian PL (2004) What type of vehicle do people drive? The role of attitude and lifestyle in influencing vehicle type choice. Transp Res Part A Policy Pract 38(3):201–222. https://doi.org/10.1016/j.tra.2003.10.005
5. Flamm B (2009) The impacts of environmental knowledge and attitudes on vehicle ownership and use. Transp Res Part D Transp Environ 14(4):272–279. https://doi.org/10.1016/j.trd.2009.02.003
6. Steg L (2005) Car use: lust and must. Instrumental, symbolic and affective motives for car use. Transp Res Part A Policy Pract 39(2–3):147–162. https://doi.org/10.1016/j.tra.2004.07.001
7. Verma M, Manoj M, Verma A (2016) Analysis of the influences of attitudinal factors on car ownership decisions among urban young adults in a developing country like India. Transp Res Part F Traffic Psychol Behav 42:90–103. https://doi.org/10.1016/j.trf.2016.06.024
8. Wu G, Yamamoto T, Kitamura R (1999) Vehicle ownership model that incorporates the causal structure underlying attitudes toward vehicle ownership. Transp Res Rec 99:61–67
9. Zegras C (2010) The built environment and motor vehicle ownership and use: evidence from Santiago de Chile. Urban Stud 47(8):1793–1817. https://doi.org/10.1177/0042098009356125
10. Mohammadian A, Miller EJ (2003) Dynamic modeling of household automobile transactions. Transp Res Rec 1831:98–105. https://doi.org/10.3141/1831-11
11. Roorda MJ, Carrasco JA, Miller EJ (2009) An integrated model of vehicle transactions, activity scheduling and mode choice. Transp Res Part B Methodol 43(2):217–229. https://doi.org/10.1016/j.trb.2008.05.003
12. Fatmi MR, Habib MA (2018) Microsimulation of vehicle transactions within a life-oriented integrated urban modeling system. Transp Res Part A Policy Pract 116:497–512. https://doi.org/10.1016/j.tra.2018.06.029
13. Yamamoto T, Madre JL, Kitamura R (2004) An analysis of the effects of French vehicle inspection program and grant for scrappage on household vehicle transaction. Transp Res Part B Methodol 38(10):905–926. https://doi.org/10.1016/j.trb.2004.02.001
14. Mohammadian A, Rashidi TH (2007) Modeling household vehicle transaction behavior: competing risk duration approach. Transp Res Rec 2014:9–16. https://doi.org/10.3141/2014-02
15. Rashidi TH, Mohammadian A, Koppelman FS (2011) Modeling interdependencies between vehicle transaction, residential relocation and job change. Transportation (Amst) 38(6):909–932. https://doi.org/10.1007/s11116-011-9359-4
16. Hossein Rashidi T, Mohammadian K (2016) Application of a nested trivariate copula structure in a competing duration hazard-based vehicle transaction decision model. Transp A Transp Sci 12(6):550–567. https://doi.org/10.1080/23249935.2016.1162873
17. Khan NA, Habib MA (2021) When do households make vehicle transactions? What type of vehicle households choose after transaction events? Travel Behav Soc 22:32–47. https://doi.org/10.1016/j.tbs.2020.08.002

18. Oakil ATM, Ettema D, Arentze T, Timmermans H (2014) Changing household car ownership level and life cycle events: an action in anticipation or an action on occurrence. Transportation (Amst) 41(4):889–904. https://doi.org/10.1007/s11116-013-9507-0
19. Wang B, Rasouli S, Timmermans H, Shao C (2018) Relationships between consecutive long-term and mid-term mobility decisions over the life course: a Bayesian network approach. Transp Res Rec 2672(47):159–170. https://doi.org/10.1177/0361198118777604
20. Nazari F, (Kouros) Mohammadian A, Stephens T (2018) Dynamic household vehicle decision modeling considering plug-in electric vehicles. Transp Res Rec 2672(49):91–100. https://doi.org/10.1177/0361198118796925
21. Yamamoto T (2008) The impact of life-course events on vehicle ownership dynamics: the cases of France and Japan. IATSS Res 32(2):34–43. https://doi.org/10.1016/S0386-1112(14)60207-7
22. Gillingham K, Iskhakov F, Munk-Nielsen A, Rust J, Schjerning B (2015) A dynamic model of vehicle ownership, type choice, and usage. [Online]. Available: http://bschjerning.com/papers/iruc.pdf
23. Kitamura R, Mokhtarian PL (1997) A micro-analysis of land use and travel in five neighborhoods in the San Francisco Bay Area. Transportation (Amst) 24:125–158. https://doi.org/10.1023/A:1017959825565
24. McFadden D (1986) The choice theory approach to market research. Mark Sci 5(4):275–297. https://doi.org/10.1287/mksc.5.4.275
25. Dittmar H (2008) Consumer culture, identity and well-being. Psychology Press
26. Gu G, Feng T, Yang D, Timmermans H (2020) Modeling dynamics in household car ownership over life courses: a latent class competing risks model. Transportation (Amst) 0123456789. https://doi.org/10.1007/s11116-019-10078-8
27. Ha TV, Asada T, Arimura M (2019) Determination of the influence factors on household vehicle ownership patterns in Phnom Penh using statistical and machine learning methods. J Transp Geogr 78:70–86. https://doi.org/10.1016/j.jtrangeo.2019.05.015
28. Arman MA, Kalantari N, Mohammadian A (2015) Joint modeling of household vehicle and activity allocation statistical analysis and discrete choice modeling approach. Transp Res Rec 2525:121–130. https://doi.org/10.3141/2495-13
29. Zhang J, Van Acker V (2017) Life-oriented travel behavior research: an overview. Transp Res Part A Policy Pract 104:167–178. https://doi.org/10.1016/j.tra.2017.06.004
30. Ajzen I (2005) Attitudes, personality and behaviour. McGraw Hill Education (UK)
31. Ellaway A, Macintyre S, Hiscock R, Kearns A (2003) In the driving seat: psychosocial benefits from private motor vehicle transport compared to public transport. Transp Res Part F Traffic Psychol Behav 6:217–231. https://doi.org/10.1016/S1369-8478(03)00027-5
32. Lanzendorf M (2010) Key events and their effect on mobility biographies: the case of childbirth. Int J Sustain Transp 4(5):272–292. https://doi.org/10.1080/15568310903145188
33. Sattlegger L, Rau H (2016) Carlessness in a car-centric world: a reconstructive approach to qualitative mobility biographies research. J Transp Geogr 53:22–31. https://doi.org/10.1016/j.jtrangeo.2016.04.003
34. Holtzclaw J, Clear R, Dittmar H, Goldstein D, Haas P (2002) Location efficiency: neighborhood and socio-economic characteristics determine auto ownership and use—studies in Chicago, Los Angeles and San Francisco. Transp Plan Technol 25(1):1–27. https://doi.org/10.1080/03081060290032033
35. Whelan G (2007) Modelling car ownership in Great Britain. Transp Res Part A Policy Pract 41(3):205–219. https://doi.org/10.1016/j.tra.2006.09.013
36. Matas A, Raymond JLL (2008) Changes in the structure of car ownership in Spain. Transp Res Part A Policy Pract 42(1):187–202. https://doi.org/10.1016/j.tra.2007.08.005
37. Potoglou D, Kanaroglou PS (2008) Modelling car ownership in urban areas: a case study of Hamilton, Canada. J Transp Geogr 16(1):42–54. https://doi.org/10.1016/j.jtrangeo.2007.01.006
38. Rouwendal J, Borger B (2009) Multiple car ownership, fuel efficiency and substitution between cars

39. Hamed MM, Al-Eideh BM (2020) An exploratory analysis of traffic accidents and vehicle ownership decisions using a random parameters logit model with heterogeneity in means. Anal Methods Accid Res 25:100116. https://doi.org/10.1016/j.amar.2020.100116
40. Fang HA (2008) A discrete-continuous model of households' vehicle choice and usage, with an application to the effects of residential density. Transp Res Part B Methodol 42(9):736–758. https://doi.org/10.1016/j.trb.2008.01.004
41. Eluru N, Paleti R, Pendyala RM, Bhat CR (2010) Modeling injury severity of multiple occupants of vehicles: copula-based multivariate approach. Transp Res Rec 2165:1–11. https://doi.org/10.3141/2165-01
42. Pinjari AR, Pendyala RM, Bhat CR, Waddell PA (2011) Modeling the choice continuum: an integrated model of residential location, auto ownership, bicycle ownership, and commute tour mode choice decisions. Transportation (Amst) 38(6):933–958. https://doi.org/10.1007/s11116-011-9360-y
43. Oakil ATM, Ettema D, Arentze T, Timmermans H (2011) Longitudinal model of longer-term mobility decisions: framework and first empirical tests. J Urban Plann Dev 137(3):220–229. https://doi.org/10.1061/(ASCE)UP.1943-5444.0000066
44. Liu Y, Tremblay JM, Cirillo C (2014) An integrated model for discrete and continuous decisions with application to vehicle ownership, type and usage choices. Transp Res Part A Policy Pract 69:315–328. https://doi.org/10.1016/j.tra.2014.09.001
45. Yamamoto T, Kitamura R, Kimura S (1999) Competing-risks-duration model of household vehicle transactions with indicators of changes in explanatory variables. Transp Res Rec 1676:116–123. https://doi.org/10.3141/1676-15
46. Rashidi TH, Mohammadian A (2011) A competing hazard model of household vehicle transaction behavior with discrete time intervals and unobserved heterogeneity. Transp Lett 3(3):219–229. https://doi.org/10.3328/TL.2011.03.03.219-229
47. Schoenduwe R, Mueller MG, Peters A, Lanzendorf M (2015) Analysing mobility biographies with the life course calendar: a retrospective survey methodology for longitudinal data collection. J Transp Geogr 42:98–109. https://doi.org/10.1016/j.jtrangeo.2014.12.001
48. Van Der Waerden P, Timmermans H, Borgers A (2003) The influence of key events and critical incidents on transport mode choice switching behaviour: a descriptive analysis. In: Moving through nets: the physical and social dimensions of travel, pp 10–15
49. Gu G, Yang D, Feng T, Timmermans H (2019) Household vehicle holding decisions in response to life cycle events. Transp Res Procedia 37:171–178. https://doi.org/10.1016/j.trpro.2018.12.180
50. Ettema D, Arentze T, Timmermans H (2011) Social influences on household location, mobility and activity choice in integrated micro-simulation models. Transp Res Part A Policy Pract 45(4):283–295. https://doi.org/10.1016/j.tra.2011.01.010
51. Xiong Y, Zhang J (2017) Empirical evidence of behavioral interdependencies across life choices. In: Zhang J (ed) Life-oriented behavioral research for urban policy. Springer, Tokyo, pp 7–99
52. Oakil AT (2013) Temporal dependence in life trajectories and mobility decisions
53. Ho C, Mulley C (2015) Intra-household interactions in transport research: a review. Transp Rev 35(1):33–55. https://doi.org/10.1080/01441647.2014.993745
54. Zhang J, Kuwano M, Lee B, Fujiwara A (2009) Modeling household discrete choice behavior incorporating heterogeneous group decision-making mechanisms. Transp Res Part B Methodol 43(2):230–250. https://doi.org/10.1016/j.trb.2008.05.002
55. Arentze TA, Timmermans HJP (2009) A need-based model of multi-day, multi-person activity generation. Transp Res Part B Methodol 43(2):251–265. https://doi.org/10.1016/j.trb.2008.05.007
56. Dworkin SL (2012) Sample size policy for qualitative studies using in-depth interviews. Arch Sex Behav 41(6):1319–1320. https://doi.org/10.1007/s10508-012-0016-6
57. Flamm M, Kaufmann V (2006) Operationalising the concept of motility: a qualitative study. Mobilities 1(2):167–189. https://doi.org/10.1080/17450100600726563

58. Schub M, Wintersberger P, Riener A (2021) Let's share a ride into the future: a qualitative study comparing hypothetical implementation scenarios of automated vehicles. In: Conference on human factors in computing systems—proceedings. https://doi.org/10.1145/3411764.3445609
59. Clark B, Lyons G, Chatterjee K (2016) Understanding the process that gives rise to household car ownership level changes. J Transp Geogr 55:110–120. https://doi.org/10.1016/j.jtrangeo.2016.07.009
60. Jones H, Chatterjee K, Gray S (2014) A biographical approach to studying individual change and continuity in walking and cycling over the life course. J Transp Health 1(3):182–189. https://doi.org/10.1016/j.jth.2014.07.004
61. Kieslinger J, Kordel S, Weidinger T (2020) Capturing meanings of place, time and social interaction when analyzing human (im)mobilities: strengths and challenges of the application of (im)mobility biography. Forum Qual Sozialforsch 21(2). https://doi.org/10.17169/fqs-21.2.3347
62. Mason M (2010) Sample size and saturation in PhD studies using qualitative interviews. Forum Qual Sozialforsch 11(3). https://doi.org/10.17169/fqs-11.3.1428
63. Guest G, Namey E, Chen M (2020) A simple method to assess and report thematic saturation in qualitative research. PLoS ONE 15(5):1–17. https://doi.org/10.1371/journal.pone.0232076
64. Wilton RD, Páez A, Scott DM (2011) Why do you care what other people think? A qualitative investigation of social influence and telecommuting. Transp Res Part A Policy Pract 45(4):269–282. https://doi.org/10.1016/j.tra.2011.01.002
65. Leibenstein H (1950) Bandwagon, Snob, and Veblen effects in the theory of consumers' demand. Q J Econ 64(2):183–207. https://doi.org/10.2307/1882692

Mode Shift Behaviour of Commuters Due to the Introduction of a New Waterway Route

Keerthy Sabu, N. M. Sabitha, and Geeva George

Abstract Sustainability is a major goal of the policy makers, with the aim of limiting the effects of global warming and climate change. When it comes to finding ways to make the transportation sector more environmentally friendly and carbon neutral, always land transport is the main focus. But, shifting to more environmentally friendly modes of transportation, such as inland waterway transportation (IWT), offers the potential to reduce the side effects caused by road transport, including costs from emissions, noise and congestion. For passenger and freight transportation, IWT has lower operating costs and emissions than road, rail or air. This paper focusses on the mode shift behaviour of the commuters to the newly introduced water way from Kakanad to Infopark in Ernakulam district, Kerala. The identification of variables influencing mode choice behaviour, the analysis of survey respondents, and intention to use boat service were determined. The identification and ranking of the factors influencing the choice of waterway using the relative importance index were also done.

Keywords Relative importance index · Mode choice · Inland water transport

1 Introduction

Travel behaviour research is one of the most difficult to conduct because it governs multi-parameter evaluation at the same time and involves the collection of large amounts of data on human behaviour in a variety of situations. In order to analyse

K. Sabu (✉) · G. George
Department of Civil, RIT, Kottayam, India
e-mail: keerthysabu09@gmail.com

G. George
e-mail: geevageorge@rit.ac.in

N. M. Sabitha
Water Transport Division and GIS, KSCSTE NATPAC, Thiruvananthapuram, Kerala, India
e-mail: sabithanm@gmail.com

L. Devi et al. (eds.), *Proceedings of the Sixth International Conference of Transportation Research Group of India*, Lecture Notes in Civil Engineering 272,
https://doi.org/10.1007/978-981-19-3494-0_11

travel behaviour and estimate future demand for travel, significant investments have been made in transportation planning and policymaking. This forecasting needs to be incorporated into the design of transportation systems, utilising global infrastructure and knowing the travel behaviour of the studied area's population, in order to create a system that can support future travel demands.

Many countries have recently focussed on environmentally friendly transportation systems as it is an important element in the development towards a sustainable city. Transit systems such as buses and subways, are attractive means of transportation in urban areas because, when utilised effectively, they not only produce fewer emissions but also provide a high level of efficient mobility. To provide consumers with efficient and sustainable transportation networks, major cities throughout the world have actively implemented transit-oriented policies.

When considering methods to make the transportation sector greener and carbon neutral, land transport is always at the forefront. Inland water transport, on the other hand, is more energy efficient and has a greater potential for lowering greenhouse gas emissions when compared to other modes of transportation. When compared to road transportation, inland shipping consumes around 75% less energy. Moreover, IWT has hardly any noise emissions and will give a calm environment and a comfortable ride. It barely accounts for around half of the energy usage per km/tonne of carried commodities as compared to rail transport. When compared to road transport, it only uses around 17% of the energy [1].

From a study of comparison of the external cost categories (i.e., accident costs, noise costs, congestion costs, costs of habitat damage, air pollution costs, climate change costs and costs of well-to-tank emissions) of different inland transportation modes (road, rail and IWT), it was concluded that the external cost of IWT is 40% lower than the external cost of road/rail [2].

2 Literature Review on Variables Affecting the Mode Choice

From the literatures, it is clear that a wide variety of factors need to be considered to adequately capture the full spectrum of mode choice. The variables affecting the mode choice behaviour have been identified from the literatures and summarized in Table 1.

From the literature review, the variables affecting mode choice behaviour can be categorized into three sections: Socio-demographic characteristics, activity characteristics and trip characteristics. Socio-demographic characteristics includes details about gender, age, marital status, education, employment, monthly income, vehicle ownership, driving licence, whether head of the house, number of people in the household, number of working people in the household and location of residence. Activity characteristics include details regarding location of work, work duration and work start time. Trip characteristics include details about origin and destination, trip

Table 1 Different variables affecting the mode choice behaviour

Author(s)	Factors
Desai and Joshi [3]	Travel time, travel cost, convenience, comfort, safety and security and number of transfers, sending kids to school, irregular working hours, stress, personal status and privacy (while comfort, convenience, safety and security are latent variables)
Kim et al. [4]	Individual and household characteristics (age, gender, working, student, etc.), mode variables (In-vehicle time, out-vehicle time, transit fare, etc.), latent variables (water transit preference, environmental preferences, comfort oriented, etc.), Interaction effect variables (In-vehicle time * latent variables, out-vehicle time * latent variables, etc.)
Sohoni et al. [5]	Socio-economic background, travel mode, waiting and travel times, travel cost and discomfort level (old and new mode)
Chuen et al. [6]	Characteristics of the travellers such as traveller's background, household structure and income, vehicle ownership, and availability of vehicle choice; characteristics of the trips such as the purpose of the trip, time of the trip, and trip distance; characteristics of the transport facility such as travel duration, costs, quality of service and parking space availability
Rogerson et al. [7]	Analyses strategies to overcome barriers to a modal shift to inland waterway transport (IWT). Barriers identified are categorized as regulatory, financial, service quality and market characteristics

purpose, regular mode for work trip, travel duration, daily travel expense for work trip and distance in km.

3 Objectives

One of the most significant traditional models in transportation planning is mode choice of transportation. A mode choice study helps a planner or transportation engineer in calculating daily ridership and making recommendations for improvements to existing transportation planning design criteria. Understanding the most important factors influencing mode choice can help designers in attracting more riders by incorporating them during design and planning stages. The objectives of this study are to identify the variables affecting mode choice behaviour from various literatures, to identify the intension to use new boat service using cross contingency tables, and then ranking the most important mode shift attributes that affects mode shift towards waterways using relative important index.

4 Methodology

The methodology includes literature review, study area selection, questionnaire preparation, data collection and data analysis. Section 4.1 describes the study area and its features. From extensive literature review, the objectives for the study were identified and discussed in Sect. 3. Analysing the literatures, the factors influencing mode choice were identified. The variables include 12 socio-demographic, three activity and six trip related variables. The ten factors that influences mode shift towards waterways includes comfort and convenience, safety, travel time, travel cost, parking facilities at boat jetties, facilities at boat jetties, accessibility, environmental preference, quality of water and reliability are also considered in the study to determine the relative importance index (RII). A self-descriptive questionnaire was prepared and data was collected from site interview and household interview either through direct interaction or internet based survey (Google form). After conducting pilot survey, questionnaire was redefined, and the final data collection was conducted. After the data collection, the preliminary analysis was carried out. And then intension to use new boat service analysis was carried using cross contingency tables. Using relative important index, the most important factor that affects mode shift towards waterways were ranked.

Relative importance index (RII) is used to find the relative importance of the factors included, i.e. which is the most important factor that affects mode shift towards waterways. From the calculated RII values, the factors are ranked and most important to least important factors are determined. More details about RII are provided in Sect. 5.4.

Description on the study area, design of questionnaire, data collection, data analysis and finally the conclusions are presented in the sections following.

4.1 Study Area

The waterway stretch extending from Vytilla Mobility Hub to Infopark in Kochi, Kerala, is selected as the study area. Kochi city has inherent problems of traffic congestion, heavily populated central core, intense development of land uses and inadequate mass transport system. The growing demand for transportation in the city cannot be met with the existing infrastructure facilities alone. There has to be greater emphasis on developing additional transport facilities with matching infrastructure development to meet the future travel needs of the people. It is required to build waterways for navigation with the appropriate infrastructure, such as fairway, terminals, navigational aids and a fleet, so that the IWT mode becomes competitive and attracts market-driven traffic. At this context, it is important to extend the existing water transport facility available from Vyttila mobility hub to Kakkanad and further to Infopark to serve people from major institutions coming on the way.

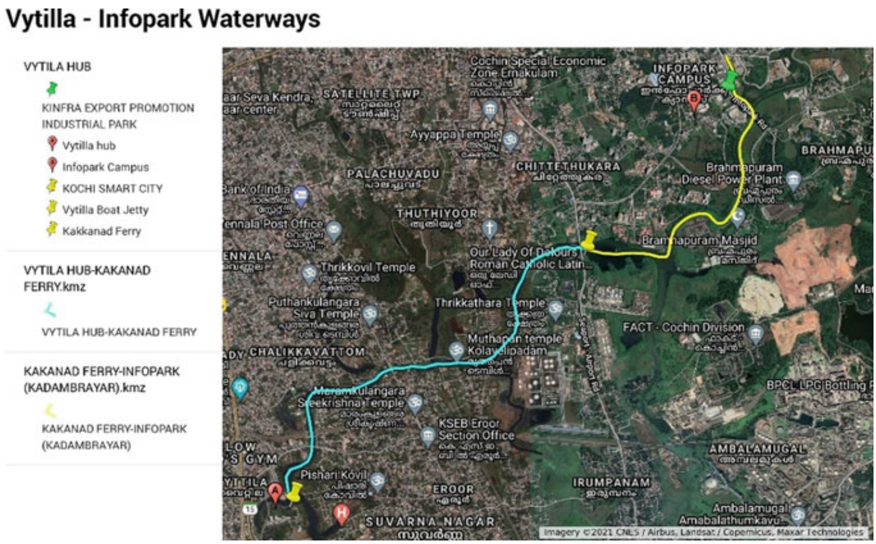


Fig. 1 Vytilla–Infopark study area

A mobility hub is a strategically located urban structure that integrates several forms of transportation. Vyttila mobility hub envisage the convergence of road transport, water transport and metro rail in the peaceful stretch of land beside the Kaniayampuzha River. By extending this waterway from Kakkanad to Infopark via the Kadambrayar River, existing passenger transportation will be boosted, and an easy and comfortable traffic facility will be provided, saving time, money, and fuel. Infopark campus includes around 48,000 IT professionals working with various companies and also includes many schools and colleges within the campus.

The selected study stretch is given in the Fig. 1.

Figure 1 shows the Google map image of the study area. The area specified with blue band represents the existing waterways from Vytilla hub to Kakkanad ferry (5 km), and the yellow band represents the waterways to be extended from Kakkanad ferry to Infopark area (3.8 km).

4.2 Design of Questionnaire

The design of questionnaire plays an important role in the quality and reliability of data collection. The variables identified from the literatures are included in the questionnaire. All together 31 variables were identified in which 12 socio-demographic variables, three activity characteristics, six trip characteristics and ten mode shift attributes are also included in the questionnaire. The details regarding the variables are given in Table 2.

Table 2 Variables selected for questionnaire survey

	Factors
Socio-demographic Variables	Gender, age, marital status, education, employment, monthly income, vehicle ownership, driving licence, whether head of the house, number of people in the household, number of working people in the household and location of residence
Activity characteristics	Location of work, work start time and work end time
Trip characteristics	Origin and destination, trip purpose, regular mode for work trip, travel duration, daily travel expense for work trip and distance in km
Mode shift attributes	Comfort and convenience, safety, travel time, travel cost, parking facilities at boat jetties, facilities at boat jetties, accessibility, environmental preference, quality of water and reliability

An importance analysis is included in the questionnaire to rank the important factor that influence mode shift towards waterways. This analysis is done to determine the relative importance index of mode shift attributes involved and to identify and prioritize those factors. The importance of each factors is done by rating in a 5 point Likert scale, five representing extremely important to one representing extremely unimportant [8].

4.3 Data Collection

A self-descriptive questionnaire was prepared which includes questions regarding behavioural factors followed by socio-demographic, activity characteristics and trip characteristics. Paper and pencil Interview (PAPI) along with telephonic or internet survey (mixed mode survey) helps to collect data during this pandemic situation. Data collection is done from site interview through direct interaction and internet-based survey (Google form). During the first phase of the work, a pilot study was conducted in order to improve the clarity and content of the questionnaire. A pilot or feasibility study is a small experiment used to test logistics and gather information prior to a larger research in order to enhance the questionnaire's quality and efficiency. A pilot study can reveal deficiencies in the design of a proposed experiment or procedure and these can be addressed before time and resources are used on large scale studies. After conducting pilot survey questionnaire was redefined, and the final data collection was conducted.

Since we have assumed that the population is normally distributed, empirical formulas (1–3) are used to determine the sample size [9],

$$N_0 = \frac{Z^2 pq}{e^2} \tag{1}$$

$$q = 1 - p \tag{2}$$

$$n = \frac{N_0}{\left[1 + \frac{N_0 - 1}{N}\right]} \tag{3}$$

where

- N_0 is the sample size from infinite population,
- e Desired error (here we have adopted as 5%),
- Z Statistical parameter corresponding to confidence level (Z is 1.96 for 95% confidence interval),
- N Total population size of study area.
- p Hypothesized true proportion of population (adopted as 0.5 to account for the worst-case scenario).
- n Sample size.

From the above equation, it can be seen that 385 number of survey samples must be obtained to get good proportion of population covered. The time taken for interviewing single respond was about 10–12 min.

5 Data Analysis and Results

After the data collection, the samples were entered into Microsoft Excel and preliminary analysis was carried out. The statistical distribution of the responses is presented in the sbelow section.

5.1 *Socio-Demographic Characteristics*

Gender distribution of survey sample is shown in Fig. 2. It shows that majority of respondents are male (57%). Figure 3 shows age distribution of respondents. Majority of respondents taking part in survey falls in the age group of 26–35 (42%). 40% of

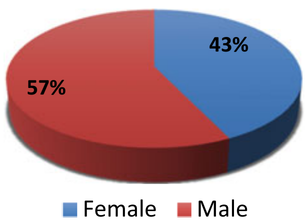
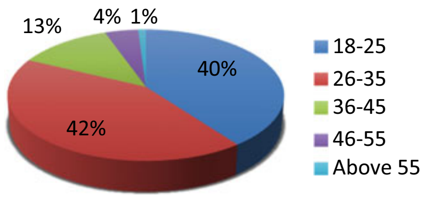


Fig. 2 Gender distribution

Fig. 3 Age distribution



the total respondents fall in the age group of 18–25, 13% belongs 36–45 category and 4% fall in 46–55 age group.

Marital statuses of the respondents are shown in Fig. 4. Unmarried respondents are the majority (58%) took part in the survey. Figure 5 shows the education status of observations. It is found that majority of respondents are graduates (59%) and 22% are postgraduates.

Employment status is depicted in Fig. 6. Majority of the respondents are employed in private sectors (89%). The income distribution amongst the responders is shown in Fig. 7. 34% of observations had monthly income in the range of Rs. 20,001–30,000. High income respondents having income greater than Rs.75,000/- is very less about

Fig. 4 Marital status distribution

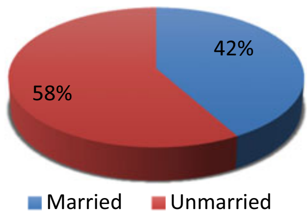


Fig. 5 Education status

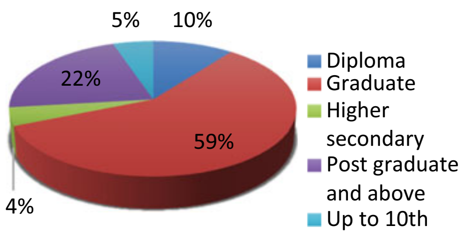
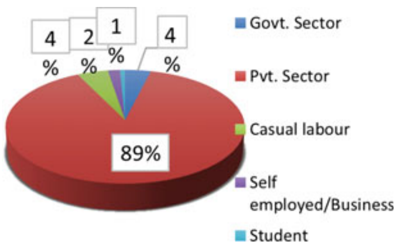


Fig. 6 Employment status



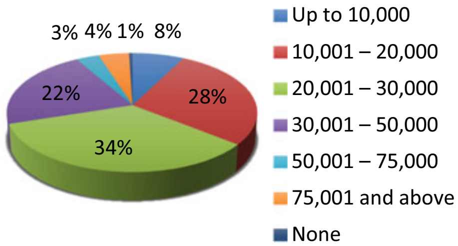


Fig. 7. Monthly income distribution

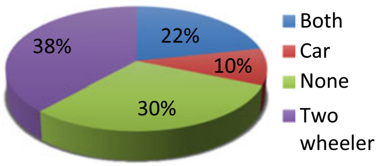


Fig. 8 Vehicle ownership distribution

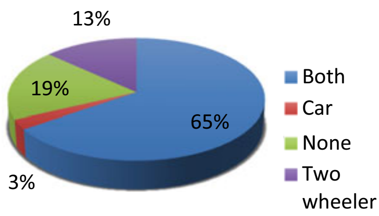


Fig. 9 Driving licence distribution

4%. Figure 8 shows the vehicle ownership distribution of the surveyed sample. About 38% of respondents have two-wheeler and 10% owns car. Respondents who have both car and two-wheeler are 22%. About 30% of respondents do not have any vehicles. Licence possession statistics is shown in Fig. 9. It shows that majority of the respondents possess driving licence. About 65% have licence for both two-wheeler wheeler and car. About 19% of respondents do not have any driving licence. About 13% of respondents have two-wheeler and 3% have car driving licence.

5.2 Mode Share of Commuter Trips

The mode share of trips of the respondents is shown in Fig. 10. Amongst the responses, 99% of trips are work trips and only 1% is for educational. Modes found to be used for work trips from the survey sample are as follows: walk, bus, car, two-wheeler, auto rickshaw, taxi and institutional vehicles. The mode selected as walk and auto was 4.4% and 4.7%, respectively, and taxi and institutional vehicles were only 1.8% and 0.5% in the survey sample. Therefore, the mode choice set considered is bus, car and two wheeler. Responses show that amongst 88.6% of work trips are

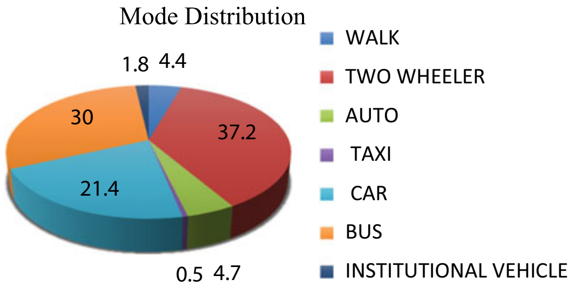


Fig. 10 Mode share of commuter trips

shared by two wheeler, bus and car. 37.2% of share is by two wheelers. The next higher portion of trips are shared by bus (30%) and car (21.4%).

5.3 *Intension to Use Newly Introduced Boat Service*

The respondents were requested to share their knowledge about the existence of Vyttila–Kakkanad waterway. Figure 11 shows the awareness of Vytilla–Kakkanad waterway. From the respondents, it is understood that 55% of people are aware of the waterway. The study shows that about 27% of people are willing to shift to IWT on daily basis. In the survey, 99% are work trips, which mean mostly daily trips only. The mode share of passenger using walking mode is found to be very low, i.e. 4.4% and remaining 95.6% of people are using motorized vehicle. Hence, 27% of daily shift will contribute a reduction in traffic of the selected study area. But, the exact prediction of motorized personal vehicle trips shifting towards the newly introduced waterway can only be ascertained by developing a mode shift model. The 30% of commuters are willing to use the boat service once in a while, and 18% are not willing to shift to boat service. Intention to use boat service is given in Fig. 12.

The relationships between the socio-demographic factors and commuter's intention to use boat service are examined using cross contingency tables. This can be done by performing crosstabs in SPSS. It is a method for displaying the relationship between two variables in tabular form by cross tabulating them. Crosstabs produce information on bivariate relationships, as compared to frequencies, which summaries information about a single variable.

The results of the cross analysis are summarized as below:

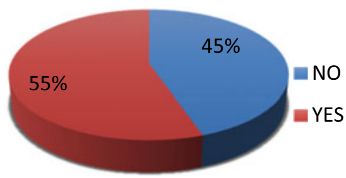


Fig. 11 Awareness of Vytilla–Kakkanad waterway

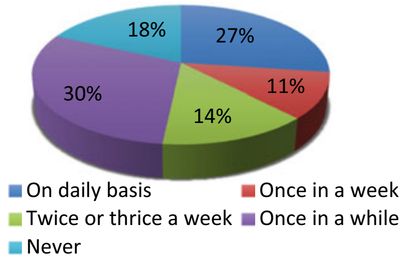


Fig. 12 Intention to use boat service

- Gender: Female (31.3%) passengers are more willing to use the newly introduced water comparing to males. Figure 13 shows gender versus intention to use boat service.
- Age: When comparing with people in age group 26–35, people in age group 18–25 are more willing to use boat service on daily basis or twice or thrice a week. Figure 14 shows age group versus intention to use boat service.
- Marital status: Married people are more likely to use boat service on daily basis and unmarried people are more likely to use boat service in once, twice or thrice a week. Figure 15 shows marital status versus intention to use boat service.
- Education: Respondents whose education level is below the graduate level (diploma, higher secondary and high school) have the high percentage intention to use boat service. Figure 16 shows education versus Intention to use boat service.
- Monthly income: Respondents whose monthly income is below 30,000 are more likely to use boat service on daily basis. Figure 17 shows monthly income versus intention to use boat service.
- Vehicle ownership: Respondents who does not have vehicle are more likely to use boat service on daily basis. Figure 18 shows vehicle ownership versus intention to use boat service.

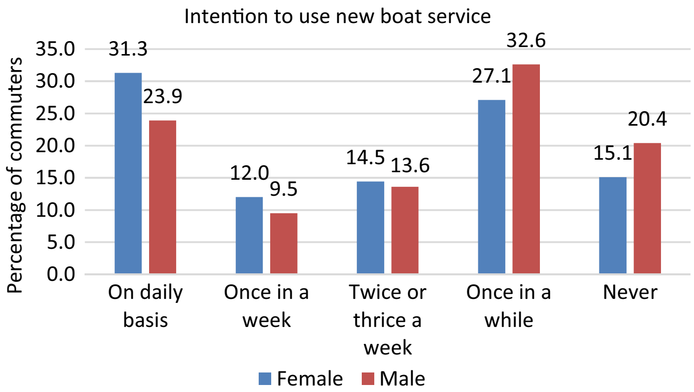


Fig. 13 Gender versus intention to use boat service

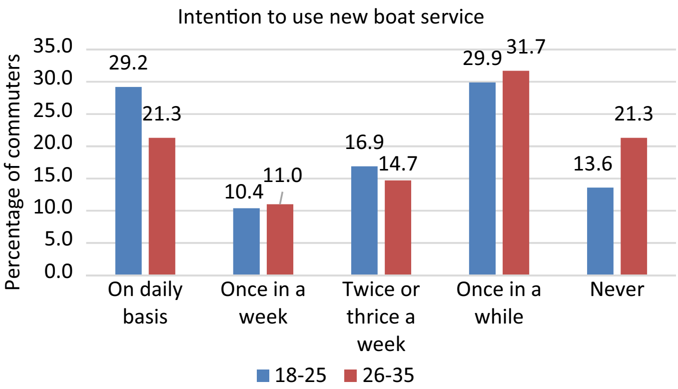


Fig. 14 Age group versus intention to use boat service

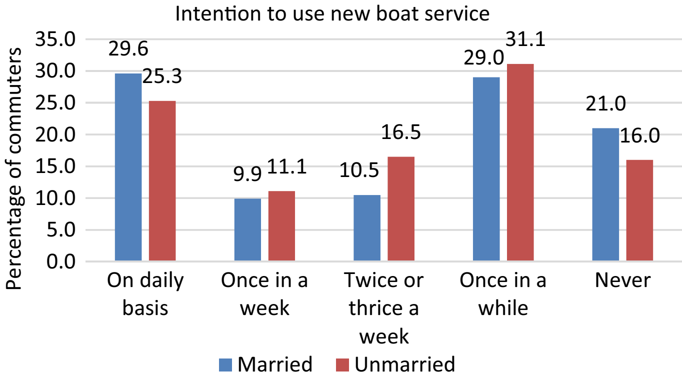


Fig. 15 Marital status versus intention to use boat service

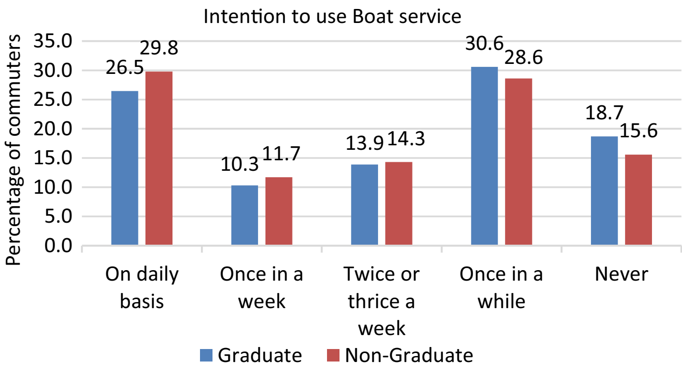


Fig. 16 Education versus intention to use boat service

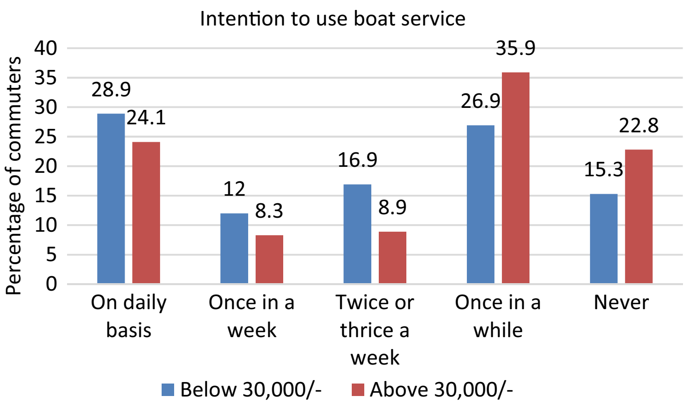


Fig. 17 Monthly income versus intention to use boat service

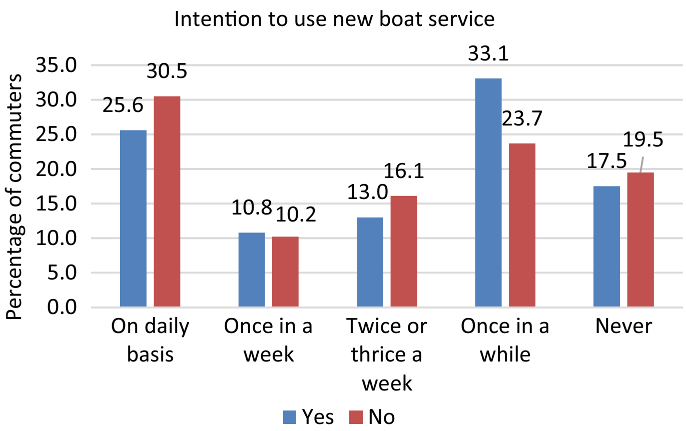


Fig. 18 Vehicle ownership versus intention to use boat service

Understanding the willingness for daily shift amongst the respondents can be utilized during infrastructure design of the terminal and boat service facilities.

5.4 Relative Importance Index (RII)

The relative importance of various factors towards the water transportation mode was examined, and the ranking of the factors in terms of their perceived criticality was done using the relative importance index (RII). The respondents were asked to rate the importance of ten identified factors using a 5 point Likert scale questionnaire where five representing extremely important to one representing extremely unimportant.

The relative importance index (RII) is calculated using the equation given below [8, 10],

$$\text{Relative Importance Index } = \frac{\sum W}{A * N} = \frac{5n_5 + 4n_4 + 3n_3 + 2n_2 + 1n_1}{A * N} \quad (4)$$

where

- W = weighting as assigned on Likert's scale by each respondent in a range from 1 to 5
- n_5 = Number of respondents for extremely important,
- n_4 = Number of respondents for important,
- n_3 = Number of respondents for neutral,
- n_2 = Number of respondents for unimportant,
- n_1 = Number of respondents for extremely unimportant,
- A = Highest weight (here it is 5)
- N = Total number in the sample.

The points of Likert scale used are equal to the value of W, weighting given to each factor by the respondent. The higher value of RII represents the most important factor and is ranked as one. The RII of all the factors were computed and ranked. The factors, their indexes and ranks are presented in Table 3.

Safety of operation is ranked 1st with the highest relative importance index value of 0.933 followed by travel time (RII value 0.929) and reliability (RII value 0.92). Quality of water (RII value 0.809) has the least importance index as per this survey and hence ranked as 10th. Hence, prior importance should be provided for safety of operation, travel time and reliability as it is the most important factors that influence the commuter's mode shift.

On board safety measures that can be provided for mechanized boats are:

Table 3 Factors influencing the selection of water transport

Factors	RII	Rank
Safety of operation	0.933	1
Travel time	0.929	2
Reliability	0.92	3
Travel cost	0.918	4
Comort and convenience	0.916	5
Accessibility	0.912	6
Facilities at boat jetties	0.857	7
Environmental preference	0.853	8
Parking facilities	0.828	9
Quality of water	0.809	10

- Storage safety precautions for fuel storage tanks should be implemented with extreme caution and safety, such as the storage container being made of metal and fitted in such a way that leaks and spills are avoided.
- For battery-powered boats, charging facilities for batteries should be provided on board so that the boat's machinery may be started and operated without difficulty.
- Proper lighting arrangements should be provided, such as providing an alternative source of lighting for emergency use when the vessel's lighting is provided by a centralized electrical system.
- Lifesaving equipment like life buoys, life jackets and life rafts should be provided with sufficient number.
- Since mechanized boats are mainly operated by electrical and mechanical components, they have a higher risk of catching fire. As a result, a well-structured fire-fighting system, such as a fire pump and fire extinguishers, must be maintained in mechanized boats.

According to the cross tabulation between gender versus intention to use boat service, 31.3% of females are willing to use the boat service on a daily basis. Women's safety should be considered at night and during off-peak hours. As a result, the jetty should have adequate lighting and sanitation. Because of the jetty's remote location and lack of security, it is recommended that CCTV surveillance cameras can be installed on the access road network to cover the jetty's influence area. To improve safe access to the jetties, installing solar-powered lights around the jetties to enable safe and active spaces even during off-peak hours and late hours of the day is suggested.

To improve connectivity and access to the boat jetties, small occupancy vehicles can be introduced.

The travel time of the boat will be constant in most of the case and the reliability can be increased by efficient time scheduling and maintenance. Also, the attractiveness amongst the riders can be enhanced if adequate feeder services are introduced.

6 Conclusions

This study identified variables influencing mode choice behaviour and analysed the survey responses to calculate the mode shift percentage towards a proposed waterway. The most important factors that affect mode shift towards the waterway are also identified. The study area is defined as the area between Vytilla Mobility Hub and Infopark in Ernakulam district, and the conclusions from the study are given below:

- The factors affecting mode choice behaviour are: All together, 31 variables were identified, and amongst them, 12 were socio-demographic variables, three activity characteristics, six trip characteristics and ten mode shift attributes were selected.
- Responses from mode shift of commuters show that 88.6% of work trips were made by two-wheeler, bus and car. Willingness survey results showed that 27%

of the respondents are willing to shift to the new boat service on daily basis which indicates that emissions can be reduced upon the introduction of this proposal.

- Cross contingency tables are used to analyse the correlations between the socio-demographic variables and commuter's intention to use boat service. The results are presented as bar diagrams in the study.
- Most important factor contributing to the mode shift towards waterway is ranked by formulating relative important index. The weightage of these factors was also computed. Amongst the 10 factors studied, it is found that the most contributing factor in making a decision of mode shift towards waterway is safety of operation followed by travel time and reliability. Incorporating these factors during the planning and design stages of boat terminal infrastructure and boat services can attract more riders.
- The weightages obtained from RII formulation can be used to calculate the performance index of various boat terminals or service facilities. This can be used as a tool to rank and compare different boat terminals or facilities and also to identify the deficiencies in the design.

Acknowledgements The authors would like to thank KSCSTE–National Transportation Planning and Research Centre (KSCSTE–NATPAC) and Rajiv Gandhi Institute of Technology, Kottayam, Civil engineering department (M-tech transportation engineering) for complete support. And also, we would like to thank CEO, Infopark Campus for their help in obtaining responses.

References

1. https://bellona.org/news/transport/2019-03-a-modal-shift-the-potential-of-inland-waterways
2. Hofbauer F, Putz LM (2020) External costs in inland waterway transport: an analysis of external cost categories and calculation methods. In Sustain (Switzerland) 12(14)
3. Desai NA, Joshi GJ (2016) Database development approach and survey design for travel mode shift behavior study with respect to mass transit in a metropolitan context. Transp Dev Econ
4. Kim JH, Chung JH, Kim T (2013) The effect of psychological traits on mode choice behaviour: an application to a new water transit system in Seoul, Korea. Transp Plan Technol 36(6):547–566
5. Sohoni AV, Thomasa M, Rao KVK (2017) Mode shift behavior of commuters due to the introduction of new rail transit mode. Transp Res Proc 25:2603–2618
6. Chuen OC, Karim MR, Yusoff S (2014) Mode choice between private and public transport in Klang valley, Malaysia. Sci World J 14, Article ID 394587
7. Rogerson S, Santen V, Svanberg MJ, Williamsson J, Woxenius J (2019) Modal shift to inland waterways: dealing with barriers in two Swedish cases. Int J Logist Res Appl
8. Tikote RH, Magdum MM, Khandare MA (2017) Delay analysis in industrial projects by using relative importance index method. Int Res J Eng Technol (IRJET) 4(6)
9. Levy PS, Lemeshow S (2008) Sampling of Populations: Methods and Applications. 4th ed, p 616
10. Gunduz M, Nielsen Y, Ozdemir M (2013) Quantification of delay factors using the relative importance index method for construction projects in Turkey. J Manage Eng 29(April):133–139

Other Transportation Modes (Including NMT) and Pedestrian

Effect of Cycle Tracks and Footpaths on Improving the PT Modal Share for the City of Nagpur

Sanket Gupta, Vishrut Landge, and Udit Jain

Abstract Nagpur is in the process of developing a significant public transport mode, the Nagpur Metro. Along with the metro, various other facilities to integrate the metro system with different modes of transport to create a seamless public transport system are being planned. Under this, bicycle and walk are the major modes of transport. All along the metro corridor and along the city's major roads, cycle tracks and footpaths have been proposed. This paper attempts to find the impact of facilities planned in the city on the ridership of the metro services. It also attempts to understand where the authority should develop these facilities first to increase the ridership at a faster pace. The results found that commercial places need footpaths and cycle tracks at the earliest, closely followed by the residential places. On the other hand, commuters from industrial areas do not seem keen on the services. The paper also stated that females were more willing to use such services than males.

Keywords Cycle tracks · Footpaths · Metro

1 Introduction

The transport sector has the second highest energy demands after industry [1]. The major part of this energy consumption is due to the private modes of transport. Therefore, the most popular method to reduce the energy needs of the transport sector is to indulge in more active modes of transport and promote public transport.

Cycling is one of the most economical, sustainable, convenient, and equitable mode of transport due to the low cost involved and flexibility [2, 3]. In tight spaces, bicycles have the highest efficiency to provide a safe and fast mode of transport. Environmentally, the use of bicycles instead of cars or any other mode of transport can significantly reduce energy demands even if bicycles are used for short trips to access public transport.

S. Gupta (✉) · V. Landge · U. Jain
Department of Civil Engineering, Visvesvaraya National Institute of Technology, Nagpur, Maharashtra, India
e-mail: sanketgupta0409@gmail.com

L. Devi et al. (eds.), *Proceedings of the Sixth International Conference of Transportation Research Group of India*, Lecture Notes in Civil Engineering 272,
https://doi.org/10.1007/978-981-19-3494-0_12

In order to increase the modal share of cycling and encourage the commuters to cycle, various policies have been promoted and designed, such as the bicycle-sharing programs [4], provision of cycle lanes [2], creating a cycle-friendly environment [5]. Also, many policies have been made to discourage the use of personal vehicles and increase the public transport modal share, such as congestion costing, peak hour pricing, etc. [6].

Even though these policies have been suggested and implemented at various locations, the policies have not received the desired response as expected from the commuters due to various factors such as lack of safety, lack of bicycle shelter/parking facilities, cultural bias, weather, length of trip, etc. [7] Another major reason that has led to commuters not preferring bicycle as a mode of transport is the lack of integration within the transit system. However, integration within the system can be a good alternative to private vehicles due to the seamless connectivity that can be provided by it [8, 9].

The integration of different modes of transport has seen a rise in recent years. For example, in 1995, the city of Los Angeles, USA, had constructed a metro line (The Green Line) along a national highway which received heavy criticism from all alike. Nevertheless, even after much criticism, the line performed better than the anticipated figures by different analysts due to better integration with the bus transit system and the line acting as a better feeder service to the blue line of the Los Angeles Metro service [10]. Though no attempts were made to integrate by the planners, a much poorly planned bus system has allowed the metro line to work satisfactorily. Another research found that metro-based commuters are the most likely to shift to cycling as a feeder once it is integrated. Commuters with travel distances higher than 15 km by private modes will benefit the most by integration of cycle [11].

Metro services are being built rapidly in the city of Nagpur. Nagpur is a tier-II city in central India. It is the second capital of Maharashtra, also known as the winter capital. The district has a population of 30.7 lakhs. 76% (24.1 lakhs) of this population resides in the urban areas, while the remaining 24% (6.6 lakhs) resides in rural areas [12]. Nagpur is spread over an area of 227 km^2. It has a radial road network, with the inner and the outer ring road as the two concentric roads circling the city. The city is located at the junction of two major national highways, namely NH 53 and NH 44.

In the year 2012, Nagpur was ranked as the cleanest and the second greenest city in the country with Bangalore being the greenest city. This has changed over the past few years due to the construction activities in the country and the increased pollution due to the increasing number of vehicles plying on the city roads. The city has slowly seen a rise in the number of private vehicles registered, and this has been steady over the past few years [12].

This paper studies the proposals that have been made by various planning authorities in the city to improve upon the cycling network and the modal share of cycling in the city. The paper also tries to understand the commuter's reaction to using cycle as a mode of transport to access the public transport facilities provided by the city's various authorities provided by the city's various authorities.

Section 2 of the paper consists of the literature that has been studied across the globe regarding the use of bicycles as a mode of transport to access public transport. The following section provides a brief about the transportation profile of the city and gives a very brief idea about the topology of the city. Finally, Sect. 2 states the attempts that have been taken by the planning authorities to integrate the different modes of transport with public transport and their current status as obtained by the planning authorities.

2 Literature Review

Researchers have said that not making any measures for integrating cycling with public transport will not encourage commuters to use it.

Dios Ortuzar [13] said that, for Santiago, Chile, if the current status changes to a denser cycle track network and a more extensive rail network, the number of cycle trips could jump three times in the future. Also, they stated that short trips are very important, and bike trips and the bike–metro integration will serve as a good option to the medium and long trips, which may help reduce the economical bias among consumers.

Wu [14] tried to determine the best feeder system, depending on the transit system and city size. They found that a shared bike feeder system can reduce generalized costs by almost 7% over the system accessed on foot when retrofitted to an existing transit system. They stated that for a small-sized city or a city with low demand, shared bikes with the combination of bus rapid transit or ordinary bus system was the best. While for cities with high demand, feeder buses were the best for operational purposes, economically, the shared bike system was considered appropriate for a metro rail transit system.

Bachand-Marleau [15] has found that in the city of Montreal, the commuters who used the train-based service for their major trips are most likely to use cycle for the first and last mile connectivity. It also suggested that this will be the most suitable, due to the economic condition of commuters living away from the city.

Zhao and Li [16] found that the transit distance was the most essential factor for cyclists in Beijing. Also, cycle-friendly environment and streets were found to be essential and relevant. They also stated that the metro passengers in suburban areas prefer to take a bus over cycling while the commuters in the city center preferred cycling. They also found that the younger commuters were more likely to use bicycles, and also, the poorer commuters would not prefer to cycle.

Souza [17], using the binary logit models, attempted to find the significant factors influencing the commuters to use bicycles as a feeder mode for two localities having low income in Rio de Janeiro. They found that in a developing country such as Brazil, where the bicycling network has not been established yet, bicycle infrastructure, parking facilities, and cycle ownership are the primary motivators for cycling. At the same time, safety is an essential barrier in such locations.

Cheng and Liu [18] tried to find the problems that may arrive due to the integration of cycling with transit systems using the Rasch model from the users' perspective in Taiwan. They found that the factors such as weather conditions, cycling infrastructure, safety are the major factors that affect the inconvenience of the commuters. Along with these, factors such as parking facilities, use of washrooms during transit, and the bike and ride methods create inconvenience among the commuters. As mentioned, all of the above factors created different inconvenience to the commuters based on their age, gender, trip purpose, and bicycle riding frequency.

Martens [9] analyzed the bike and ride system for Netherlands, Germany, and the UK. He found that the effect of travel purposes, access and egress distances, start and end location of the trips was similar in these countries despite having cultural differences. Furthermore, they found that the faster and good quality transport system has a higher share of bike and ride users than the ones that are not comfortable. The paper states that for slower modes of public transport, the commuters do not prefer to cycle for more than 2 or 3 kms while for faster modes, they can cycle 4 or 5 km. The author also stated that the share of bicycle as a feeder system to trains is almost equal to the total share of bicycle trips.

Liu [19] attempted to determine the impact of a micro-scale built environment on pedestrians' or cyclists' preferences in Tianjin, China using a stated choice experiment. They found that for their pathways, cyclists and pedestrians preferred buildings less than six floors. They also found that a green area or development of greenery along the pathway increases the commuter's willingness to walk or cycle. The commuters stated that the lack of streetlights on cycle tracks and footpaths was causing a safety issue to the commuters. They also found that shopping frontages can have a huge impact on the users' preference to cycling or walking as a mode of transport.

Zuo [20], in their study, tries to assess the advantages and potential of integrating bicycle and public transport services and the inequity among the distribution for the city of Hamilton County, Ohio. They found that using bicycle transit integration can lead to increased employment opportunities, and bicycle transit integration can also help increase transit opportunities irrespective of the race and income of groups.

Not a lot of research on the integration of different modes of transport was done in the Indian context. Research done by Singh [21] stressed the importance and the benefits of integration of different modes of transport to develop a system which shall be helpful in creating a good public transport system and increase the ridership of the public transport.

This paper attempts to determine the effects of the different built environment factors concerning the Indian context. It was seen that other countries have been studying the integration, but Indian context lags on this aspect. It was also seen that the research for tier 2 and tier 3 cities was lagging worldwide; hence, the authors tried to study this aspect.

3 Methodology

The city of Nagpur is a growing city in the region and is the only major city in the region of Vidarbha. Metro system is being constructed in the city, with multimodal integration being planned while the construction activity is going on. The integration of cycle and pedestrian modes of transport requires huge infrastructural cost and requires space to be made on roads that is not available at all locations.

The paper tries to study the effect of the cycling and pedestrian infrastructure being developed in the city on the metro ridership using the stated preference technique. First, the stated preference survey was conducted, and the data from the commuters were extracted such as their age, gender, income, total trip distance, trip characteristics, and the commuters were asked about their preference to use metro services. It was then asked whether they would access the metro services by cycling or by walking. The users were then asked about their needs for cycle tracks or footpaths.

The paper tries to determine the impact of the cycle tracks and footpaths in the locality. The analysis is done to determine the effect on metro ridership from the construction of cycle tracks and footpaths. The study tries to determine the needs of footpaths according to the age groups of the commuters, the socio-economic constraints of the commuters, the land use of the locality, gender of the region, and the metro arm of the study area.

The data were collected in May 2018 in person. Since the metro services were not operational at the time of data collection, the data was collected at the locations of the proposed metro stations. The bus stations nearest to the metro station along the haul line of the metro services were chosen as one location for conducting the survey. The survey was also done at the workplaces near the metro stations and at some homes, where the participants cooperated. A user-opinion survey was done at these locations. It was attempted to obtain 20 samples at each of the metro stations.

Total of 769 samples were collected from these locations, with an average of 20 samples per location. It was found that a lot of the surveys conducted were not proper and also had some factual errors which could not be considered for our study. Therefore, the filtered sample size reduces to 504 samples. The data collected was then further processed to find the acceptance to cycling and walking in the city.

4 Transport Characteristics of Nagpur

The City Mobility Plan (CMP) [12, 22–24] reports state that the modal share for two-wheelers dropped to 42% in 2018 as compared to 49% in 2013, share of trips by car also reduced to 5.2% from 9%, while the share of IPTs and bus transport increased from 10 and 8% to 19.8% and 15.6%, respectively, while the walk share reduced from 20% in 2013 to 10% in 2018. The reports also stated that the per capita motorized trip rate increased from 0.96 in 2013 to 1.3 in 2018. The average travel speed during the peak hours had reduced to 23.4 kmph from 27 kmph. The average trip length by cars and two-wheelers increased from 6.87 km and 5.5 km to 10.8 km and 8.6 km, respectively. The trip length by NMT modes remained more or less constant at 2 km for walk and 3 km for cycling, respectively. In 2013, 82% of the total registered vehicles in the city were two-wheelers, as shown in Table 1 [12, 22] (Fig. 1).

Nagpur has a city bus transport system which is maintained by the Nagpur Municipal Corporation (NMC). The current fleet size is 375, of which 375 are operational on 36 routes. The city also has a metro transport system under construction. Nagpur has a vast metro system being developed by the Maha-Metro Rail Corporation, for a total length of 38.215 km consisting of 38 metro stations across two metro corridors, namely the North–South corridor and East–West corridor. The metro services have started along both the corridors but in a phased manner. Out of the total length, the South and West corridors have started operations, and out of the 21 stations proposed on these corridors, 12 have already started operations.

The metro system has been divided into four reaches. Reach 1 connects Sitabuldi (Central Business District) to the airport and the MIHAN SEZ, which the government is proactively promoting for the industries to setup. The area is buzzing, and the city is expected to develop prosperously along this stretch. The new airport for MIHAN

Table 1 Number of registered vehicles in the city [12, 22]

Sr. No.	Vehicle category	2014–15	2017–18	CAGR (%)
1	Two wheelers	42,958	61,412	9
2	Cars	6524	11,157	14
3	Jeeps	1789	734	−20
4	Station wagons	0	4	–
5	Taxis	416	866	20
6	Auto rickshaws	2746	1920	−9
7	Stage carriage	0	53	–
8	Contact carriage	35	45	6
9	Minibuses	0	0	–
10	School buses	254	137	−14
11	Pvt. Buses	4	0	− 100
12	Ambulance	27	30	3
13	Multi and articulated vehicles	22	151	62
14	Trucks	144	200	9
15	Tanker	11	0	− 100
16	Del vans	1952	1579	−5
17	Tractors	30	88	31
18	Trailers	8	3	−22
19	Other tippers	0	34	–
Total		56,920	78,413	8

Fig. 1 Figure showing the metro corridor in the city of Nagpur

is also proposed near the metro line. Reach 2 connects Sitabuldi and automotive junction. The automotive junction is an upcoming industrial zone that houses the workers of the city. Reach 3 connects Hingna to Sitabuldi. Hingna is primarily an industrial zone of the region and houses significant industries from the city. Finally, Prajapati Nagar is connected to Sitabuldi by Reach 4. The major commercial zones of the city are in the vicinity of Reach 4.

As a feeder to the Metro, 38 minibus/minivan feeder routes have been proposed having an average length of 4–6 km perpendicular to the metro routes have been proposed, covering a total length of 210 km approximately. The feeder routes will be catered by different fleets consisting of 151 minivans and 42 minibuses.

Along with the modes as mentioned above of transport, the city also has a public bike-sharing scheme, developed as a feeder to the metro system. The system includes bicycle, E-bikes, E-scooters, E-rickshaw, and E-car sharing systems to be operated across the city for providing seamless public transport to the commuters.

5 Attempts to Integrate Different Modes of Transport

To integrate different modes of transport in the new metro system being developed in Nagpur, the Maha-metro has taken various steps to develop new modes and to integrate existing modes into the metro system with the primary objective of providing connectivity to the metro stations by all the existing modes of transportation available.

The significant elements proposed for the integration of metro services with the other modes of public transport available to the commuters are:

- **Development of Feeder buses:** 38 feeder bus service routes have been proposed based on the existing public transport system with an average length of 4–6 km per feeder service route. The service is to be delivered using 151 minivans and 42 minibuses that the respective agencies are procuring. The feeder service is expected to cover a corridor having a total length of 210 km. The feeder service network is planned considering the number of trips and the employment potential of a particular location.
- **Route Rationalization of the City bus Services:** City bus service has a fleet size of 375 buses operational on 36 routes across the Nagpur region and not limited to the city. The city bus service has been proposed to integrate with the new metro system to create seamless connectivity with the city bus service. Therefore, the bus service routes have been planned per the existing routes, metro routes, and feeder bus routes.
- **Development of NMT Network:** NMT masterplan has been prepared for the city, which states the need to construct NMT facilities like footpaths and cycle tracks, etc., in various parts of the city. The masterplan suggests dividing the roads as primary, having dedicated footpaths and cycle tracks at least 1.8 m each on both sides. Then, secondary roads, having dedicated footpaths of clear width 1.8 m while the cycle track can be within the shared traffic space, and tertiary with the footpath and the cycle track within the shared traffic space.
- **Development of Public Bike Sharing Scheme:** The city has developed a PBS scheme with identified parking stations. The system is a dock-less kind of system with no docking stations but with parking spaces for the bikes. The entire network is proposed to be 209 km long of which 122 km of PBS network is proposed to be developed in phase 1 and the remaining in phase 2 depending upon current and future needs of the city.

6 Integration of Cycling with Metro Services

The CMP [12, 22] and City Development Plan (CDP) [24] have proposed various measures that should be implemented throughout the city in order to give the NMT users a safe movement and also to try and increase its capacity. These proposals ranged from construction of FOB's at the CBDs and at major locations along with the development of cycle track for the cyclists. Also, in order to expand the cycling culture among the commuters in the city, it was proposed that a PBS scheme be developed in the city [12, 22–24].

The CMP's and the CDP's along with the data collected from the various junctions across the city along the metro corridor clearly show that the total modal share is around 6% for cyclists. The stated preference survey also shows that every one in three households owns a bicycle. The average trip that is made on bicycles ranges from 3–4 km across the city. Also, it can be seen that more than 80% of these trips are made for work or for educational purposes. This shows that Nagpur has commuters who prefer cycling and has huge potential to grow if nurtured scientifically.

In order to develop the cycling infrastructure in the city, various NMT corridors have been proposed and have to be developed by the respective authorities. These are divided into three categories, namely primary, secondary, and tertiary. The primary NMT corridor consists of continuously dedicated 1.8 m footpaths and cycle tracks on both sides of the roads. These are differentiated from the carriageway by height or by placing a barrier depending upon the site situation. The secondary NMT corridor are the roads that do not have enough ROW to plan these separately. These are proposed to have dedicated continuous footpaths of minimum clear width of 1.8 m, and the cycle track is in the shared carriageway space which paint markings may separate. The tertiary corridors are the internal lanes that do not have adequate widths for constructing a dedicated footpaths and cycle tracks and have very less traffic. It is proposed that these have paint markings that differentiate the footpaths but have cycle lanes shared with the multi-vehicle lanes. It is proposed that by 2027, all the primary (95 km) and secondary (134.60 km) corridors have these provisions, while the tertiary may be developed by 2041.

Also, the planning authorities have started a PBS scheme which is a pick-and-drop scheme. The bikes can be picked from the picking stations spread throughout the city, and it may be dropped at any location, which will then be picked by another commuter or picked up by the operator. A total network of 209 km is proposed out of which, 122 has been started as phase I and the phase II of 88 km is yet to be started.

In order to reduce the ridership of private vehicles and promote public transport by using bicycles as a feeder mode, the Maha-metro has proposed a green journey mobile application. The application will include all the essential features such as schedule, timings, and bookings for metro service. Along with this, the application is proposed to have features to guide the commuter about the travel time and cost for travel for different modes of travel, i.e., the public and the private modes. It also gives the travel time and cost differential for travel using public transport but by different feeder systems such as bus, cycles, walking, etc. The app also proposes displaying the carbon emission by using a particular mode of transport to encourage commuters to use green modes of travel.

As per the National Urban Transport Policy (NUTP) (2014), it is proposed that the priority is to be given to the commuters using public transport, second to the users of cycles and pedestrians, third to IPT users, and finally to the private vehicles. This has been seen in the case of the planning of metro stations. It is seen that as per the construction of the station region that is done, cyclists have priority over IPT's and two-wheelers. The cycle parking stand is located at the closest location to the entrance of the station. The IPT drop-off bay is located at some distance, while the car and two-wheeler drop-off bays and parking are located at the farthest distance. This can be seen clearly in the planning of Airport South Metro station (Fig. 2).

Fig. 2 Schematic representation of arrangements for different modes near the South Metro station area

7 Acceptance Among Commuters

The study majorly focuses on the commuter's perception to use the services if they would be using the metro services. For this reason, the survey was done at the work places near the metro stations. The commuters were asked about their preference to use bicycles or walk if they would use the metro as a primary mode for their commute. It was found that 22.06% of the commuters were willing to use the bicycle or walk as a feeder service if dedicated footpaths and cycle tracks are developed. However, the same survey found that 21.50% of the commuters were willing to use the modes if only a continuous dedicated footpath is built in the city. The city already has a footpath along 80% of the roads, but due to inappropriate heights of the footpath and its lack of continuous access, the acceptance is low. Also, at major locations, it was found that the commuters could not use the footpath due to it being encroached, and hence it was not easy to use. Hence, it was asked of the commuters about their preference.

The survey found that among the survey participants, 22.44% of the males said that they would be willing to use a bicycle or walk to the metro stations, while 30.5% of the females said that they might use it. Meanwhile, 23.72% of the women said that they may walk or cycle to the metro station if only footpaths were provided, while 20.84% of men said that they would do so if only footpaths were provided. The sample size difference in the males and females was found to be very high. This may be because not many females were interested in the survey. On the other hand, the high acceptance among the females may also be due to the good law and order situation in Nagpur during the survey period.

The research divided the participants into five age groups, less than 18 years, 18–25 years old, 25–45 years old, 45–60 years old, and more than 60 years old. It found that since the sample size for commuters less than 18 and that of more than 60 years was very less, it was removed from the study. It was found that for commuters with age group 18–25 years old, 29.29% of the commuters, in the age group of 18–25 years, would walk or cycle given the infrastructure of cycle track and footpaths is constructed, while 22.76% would use it if the only footpath was developed. Most of the commuters in this age group are students and face a lack of funds to access the metro services but require a relatively high-speed mode to access the metro station. For the age group of 25–45 years, the research found that 21.84% of the commuters might use cycle or walk if cycle tracks and footpaths are constructed, while 20.80% might use cycle or walk if the only footpath is developed. For the age group of 45–60, it was found that 30.64% of the commuters might use cycle tracks and footpaths, while only 20.96% of the commuters would prefer to use it if the only footpath was constructed. It might be due to the increased emphasis on health in the aged generations.

The researchers also asked about the profession of commuters. It divided the commuters into seven professions: business, farmers, retired professionals, government employees, private employees, students, unemployed, and laborers. The sample size for farmers, retired professionals, government employees, and unemployed commuters was very low and hence cannot be included in the study. It was found that 21.70% of the business owners preferred to use cycle or walk if footpaths and cycle tracks were constructed. While, if only footpath was constructed, only 18.49% of the commuters would use it to access the metro services. For the private employees, 28.54% would use these modes to access the metro system if cycle tracks and footpaths were developed. At the same time, 25.10% of the private employees would use it if only footpaths were constructed. Among students, 25.97% said they would walk or cycle to access the metro system if cycle tracks and footpaths were constructed, while 22.07% of the students would use it if the only footpath was constructed. It was found that for workers and laborers, if cycle tracks and footpaths were constructed, 20.37% of workers and laborers would cycle or walk to access the metro, while 17.28% of the workers and laborers would use it if only footpath was constructed.

The data was also analyzed as per the land use. The land use was divided into three major classes, namely residential, commercial, and industrial. Since the data was collected near the metro stations, the metro stations were divided as per the land use and the data collected at each was hence analyzed. It was found that if both cycle tracks and footpaths were constructed, 24.42%, 27.28%, and 13.01% of the commuters would use cycle or walk to access the metro stations in residential, commercial, and industrial zones, respectively. While, 21.26, 25.91 and 8.89% of the commuters would do so in residential, commercial, and industrial zones if only footpaths were constructed. This may be because the commuters traveling for industrial commute majorly live nearby and hence travel using own vehicles.

Along Reach 1, it was found that if the cycle tracks and footpaths were constructed, almost 25.23% would be preferring to use cycling or walking as a feeder mode of transport, while this number dropped to 21.16% if the provision of dedicated footpath was made only. This can be attributed to the stretch of Reach 1 passing through a major residential and commercial zone at Somalwada and Ajni, respectively, which may prompt the user to use its own cycle to reach the metro station.

Reach 2 is a mixed zone with government offices, automobile markets, and repair shops all along the coverage. The area is also a major market for blacksmiths and welders working and living nearby. The survey in the region found that if the cycle tracks and footpaths were constructed, almost 26.07% would prefer to use cycling or walking to reach the metro station, while this number dropped to 17.86 if the provision of dedicated footpath was made only. This is the case in the region due to the socio-economic characteristics of the region.

Reach 3, along Hingna Road, a major industrial hub in the city, it was found that 21.51% of the commuters might use cycle or walk as their mode of transport if cycle tracks and footpaths are constructed, while only 15.09% might use metro as their mode of transport if footpaths were constructed. This was reported to be lower than the city average due to the distance of industries from the metro stations. Also, few industries in the region have their own transport service for their workforce so it may be a reason for the commuters not to prefer walking or cycling to the metro station.

Reach 4 is a major commercial marketplace in the city. The location has various grain markets, cloth markets, and travels through the old city. It was found that almost 29.44% of the commuters would prefer to cycle/walk to the metro station if the provision of cycle tracks and footpaths were made, while only 19.40% of the commuters would walk/cycle if only footpaths were constructed. A major reason for the above the city average numbers is that the area is a marketplace. It is difficult to find parking spots in the vicinity. Also, since the route passes through the old city, crowded paths and narrow lanes make up for most of the roads. It might be a reason for the commuters to prefer walking or cycling to access the metro stations.

8 Conclusion

Nagpur metro has taken various steps to increase the ridership among the commuters. This ranges from developing an app that can be used for all kinds of trips within the city to developing a single payment gateway for the commuters. Different steps have been taken to integrate the existing modes of transport available with the metro services as proposed.

The metro also tries to establish and develop the existing bus system as a feeder service to the metro as can be seen in the case of Los Angeles. The development of metro services and the development of footpaths and cycle tracks has been proposed by the authority, which can be a game changer for the ridership to the metro. The study finds, that development of good infrastructure across the city, can increase the ridership of the metro services by 22.06% if dedicated, continuous and user-friendly cycle tracks and footpaths are developed. The study also finds that even though almost 80% of the roads in Nagpur have footpaths, but developing a user-friendly footpath will increase the ridership by 21.50%.

The study found that the construction of footpaths and cycle tracks will have the highest effect on the ridership in the Reach 4 of the metro stretch, while if only the footpath is constructed, Reach 1 will have the most benefit. It also stated that if the construction of footpath and cycle tracks would have the least effect on ridership in the Reach 3 of the services.

The service would be most beneficial in areas with a commercial space rather than those with an industrial background. The data also suggested, that workers and laborers, a major commuter base in the industrial region, do not prefer it the most, while the private sector employees and students, a major commuter base in the commercial and residential zones, would prefer to use it the most.

The study suggests the phasing in which the construction of the services can be done at various reaches of the metro services and their plausible effect on the metro ridership.

References

1. Government of India, 2015 (2019) India energy security scenarios 2047. http://indiaenergy.gov.in/iess/docs/Transport-doc.pdf. Accessed on 10 Jan 2019
2. Pucher J, Buehler R, Pucher J, Buehler R (2008) Making cycling irresistible: lessons from the making cycling irresistible: lessons from The Netherlands, Denmark and Germany. Transp Rev 1647(Oct). https://doi.org/10.1080/01441640701806612
3. Akar G, Clifton KJ (2009) Influence of individual perceptions and bicycle infrastructure on decision to bike. Transp Res Rec 2140:165–172. https://doi.org/10.3141/2140-18
4. Demaio P (2009) Bike-sharing: history, impacts, models of provision, and future. J Public Transp 12(4):41–56
5. Cervero R, Duncan M (2003) Walking, bicycling, and urban landscapes: evidence from the San Francisco Bay area. Am J Public Health 93(9):1478–1483
6. Buehler R, Pucher J, Gerike R, Götschi T, Buehler R, Pucher J, Götschi T (2016) Reducing car dependence in the heart of Europe: lessons from Germany, Austria, and Switzerland. Transp Rev 1647(May). https://doi.org/10.1080/01441647.2016.1177799
7. Church A, Frost M, Sullivan K (2000) Transport and social exclusion in London. Transp Policy 7
8. Rissel C (2003) Ride your bike.pdf. Health Promot J Aust 14(3)
9. Martens K (2004) The bicycle as a feedering mode: experiences from three European countries. Transp Res Part D 9:281–294. https://doi.org/10.1016/j.trd.2004.02.005
10. Mieger D, Chu C (2007) Los Angeles, California, metro green line why are people riding the line to nowhere? 50–59. https://doi.org/10.3141/2006-06
11. Meshram DS (2011) Dynamic growth of cities and tardy plan preparation and approval processes: a case study of the development plan, Nagpur. Inst Town Plan, India 21(December):1–20
12. UMTC (2018) Comprehensive mobility plan
13. De Dios Ortuzar J, Iacobelli A, Valeze C (2000) Estimating demand for a cycle-way network. Transp Res Part A 34(5):353–73
14. Wu L, Gu W, Fan W, Cassidy MJ (2020) Optimal design of transit networks fed by shared bikes. Transp Res Part B 131:63–83. https://doi.org/10.1016/j.trb.2019.11.003
15. Bachand-marleau J, Larsen J, El-geneidy AM (2011) Much-anticipated marriage of cycling and transit how will it work? J Transp Res Board 2247:109–117. https://doi.org/10.3141/2247-13
16. Zhao P, Li S (2017) Bicycle-metro integration in a growing city: the determinants of cycling as a transfer mode in metro station areas in Beijing. Transp Res Part A 99:46–60. https://doi.org/10.1016/j.tra.2017.03.003

17. De Souza F, Puello LLP, Brussel M, Orrico R, Maarseveen Van M (2017) Modelling the potential for cycling in access trips to bus, train and metro in Rio de Janeiro. Transp Res Part D 56:55–67. https://doi.org/10.1016/j.trd.2017.07.007
18. Cheng Y, Liu K (2012) Evaluating bicycle-transit users' perceptions of intermodal inconvenience. Transp Res Part A 46(10):1690–1706. https://doi.org/10.1016/j.tra.2012.10.013
19. Liu Y, Yang D, Timmermans HJP, De Vries B (2020) Analysis of the impact of street-scale built environment design near metro stations on pedestrian and cyclist road segment choice: a stated choice experiment. J Transp Geogr 82:102570. https://doi.org/10.1016/j.jtrangeo.2019.102570
20. Zuo T, Wei H, Chen N, Zhang C (2020) First-and-last mile solution via bicycling to improving transit accessibility and advancing transportation equity. Cities 99(March 2019):102614. https://doi.org/10.1016/j.cities.2020.102614
21. https://doi.org/10.1016/j.tra.2019.09.037
22. UMTC (2013) Comprehensive mobility plan for Nagpur draft final report contents
23. UMTC (2014) Detailed project report for feeder system for Nagpur metro (August)
24. CRISIL (2015) City development plan for Nagpur, 2041 Ministry of Urban Development, Government of India capacity building for urban development. 2041(March)

Investigating the Factors Affecting the Walking Dynamics of Pedestrians in Mass Gatherings

Karthika P. Sobhana and Ashish Verma

Abstract The interest in crowd modeling stems from the need to predict pedestrian flow based on current conditions and pedestrians' interaction within the crowd. Therefore, understanding the walking dynamics of pedestrians in a crowd is essential to modeling and simulating the pedestrian flow. Here, we study the effect of the geometry of location, different flow conditions, and sociodemographic characteristics on the pedestrians' walking speed in a crowd. Further, appropriate statistical tests are used to check for differences in the mean speed of pedestrians across different groups obtained by social categorization. Surprisingly, the mean speeds are not considerably different in road sections with pervious boundaries. A comparison of the mean speed of pedestrians in the field vs experimental setups indicates that there is a significant difference in the mean speed of pedestrians in the two environments. These insights can be used to model the pedestrian flow in mass gatherings more realistically.

Keywords Walking speed · Pedestrians · Mass gatherings · Crowd

1 Introduction

There has been considerable interest in measuring pedestrian speeds, especially while crossing, and investigating the different factors affecting the speed. However, pedestrian walking dynamics in mass gatherings is considerably different from the former instance. Alternatively, many studies have focused on the relation between macroscopic flow variables such as speed, flow, and density. However, it has been noted that density is not the only characteristic that affects pedestrian's walking speed. The other factors that are likely to influence walking speed include gender, age, presence of luggage, the geometry of the study section, boundaries, space headway,

K. P. Sobhana · A. Verma (✉)
Transportation Systems Engineering (TSE), Department of Civil Engineering, Indian Institute of Science (IISc), Bengaluru, Karnataka 560012, India
e-mail: ashishv@iisc.ac.in

K. P. Sobhana
e-mail: karthikaps@iisc.ac.in

L. Devi et al. (eds.), *Proceedings of the Sixth International Conference of Transportation Research Group of India*, Lecture Notes in Civil Engineering 272,
https://doi.org/10.1007/978-981-19-3494-0_13

time headway, local density, etc. Commonly used methods to manage crowds in mass gatherings are barricading some areas of the roadway or regulating the passage in narrow channels through frequent stoppages. Such strategies aim to modulate a person's natural gait by repeatedly reducing their walking speed and altering the walking direction. This study's objective is to assess if these measures influence all the pedestrians in the same way. We explore the impact of gender and height of the pedestrians on walking speed for different locations, different flow conditions, and end activity.

Generally, pedestrian data are collected from controlled experimental settings and through field investigations. These include direct interviews, recording pedestrian flows in artificial environments, recording responses through virtual reality, and recording pedestrian behavior by directly observing them in natural environments. Here, the effect of the environment on the pedestrian walking speed is explored by comparing the pedestrian walking speeds in the natural environment and experimental settings. Video graphic data of pedestrian flow at six locations in Kumbh Mela, Ujjain–2016 is used to understand the walking dynamics in the natural environment. The data include uninterrupted unidirectional flow (Panchkroshi yatra—a religious yatra/travel pilgrimage), interrupted flow conditions in barricaded roadways, unidirectional flow through narrow corridors, bottleneck sections, and approach road to the intersection. Data from pedestrians' single-file movement in controlled experimental conditions are used to compare walking dynamics across natural and experimental settings.

In this paper, we attempt to identify the different factors influencing the walking dynamics of pedestrians in mass religious gatherings. Contrary to many studies, females' mean speed is found to be higher or comparable to the mean speed of males in Kumbh Mela.

The upcoming section reviews the various studies on the walking dynamics of pedestrians. Data description is given in Sect. 3. The following section focuses on the analysis of pedestrian walking speed at different locations in the study area. The last section discusses the insights derived from the study.

2 Literature Review and Background

Understanding the walking dynamics of pedestrians is an extensively researched topic as it encompasses studies on walking speeds of pedestrians, modeling the walking dynamics of pedestrians in a crowd, and modeling the macroscopic relations between fundamental flow parameters. Table 1 gives the mean speed of pedestrians from various literature. It also highlights the differences in the speeds based on the data collection technique employed. An exhaustive review of literature on the walking dynamics of pedestrians is beyond the scope of this paper. Instead, this section consolidates the works carried out under the themes mentioned above. Focus is given to the studies conducted in India.

Table 1 Mean walking speeds of pedestrians from literature

Reference	Country	Description	Data collection	Mean speed (m/s)
Virkler and Elayadath [26]	USA	Walkway	Natural environment	1.23
Daamen and Hoogendoorn [7]	Netherlands	Unidirectional flows	Laboratory experiment	1.46
		Opposite flows	Laboratory experiment	1.33
		Crossing flows	Laboratory experiment	1.38
		Wide bottleneck	Laboratory experiment	1.2
		Narrow bottleneck	Laboratory experiment	0.96
Daamen [6]	Netherlands	Bottleneck	Laboratory experiment	1.58
Chandra and Bharti [3]	India	Sidewalk	Natural environment	1.25
		Wide sidewalk	Natural environment	1.36
		Precincts	Natural environment	0.97
		Carriageway	Natural environment	1.23
Gupta et al. [13]	India	Carriageway	Natural environment	1.02 (Male) 0.91 (Female)
Chattaraj et al. [5]	India	Corridor	Controlled experiment	1.27
	Germany	Corridor	Controlled experiment	1.24

2.1 Walking Speed of Pedestrians

Chandra and Bharati [3] have investigated walking speeds and crossing speeds at three different cities. It is found that the pedestrian walking speeds follow normal distribution at all locations. The crossing speeds of males (1.52 m/s) are higher than females (1.45 m/s). Yugendar and Ravishankar [28] modeled walking speed as a function of gender, age, persons with luggage, persons with children, and density. Several studies also focus on the comparison of normal walking speed and crossing speed at signalized or unsignalized intersections or midblock [3, 11, 15, 16, 21]. Common factors which have been found to influence the walking speed of pedestrians include age, gender, location of facility, season or weather conditions, group size, signal type, crosswalk markings, time of day, and roadway width [4, 18, 20]. Gupta et al. [13] explored the effect of gradient on pedestrian flow characteristics.

The effect of gender, age, and baggage on the walking speed of uphill and downhill pedestrians indicates that male pedestrians (1.02 m/s) walk faster than female pedestrians (0.91 m/s). Also, the study suggests that human intentions affect the walking speed of pedestrians besides the topography of the sidewalk.

2.2 Fundamental Flow Diagrams

The relation between the fundamental flow parameters speed, flow, and density in a pedestrian flow has been studied extensively. This could be attributed to the fact that these relations prove vital in the design and operation of pedestrian facilities such as walkways and crosswalks. Virkler and Elayadath [26] test the speed–density hypothesis of several established models, including single-regime, two-regime, and three-regime models. The Edie model was found to be the best fit model for the dataset. Over the years, several studies have been conducted to understand the fundamental diagram of pedestrian flow using data from controlled experimental setups [7, 14, 24, 30]. Relatively fewer studies have explored the relationship between fundamental flow parameters using data from field observations. Speed-flow-density relations were plotted for two locations in a mass gathering by Yugendar and Ravishankar [28]. Here, speed–density shows an inverse relationship with a maximum density of nearly 2.2 ped/m^2.

2.3 Walking Dynamics of Pedestrians in Crowd

Duives et al. [9] model the operational walking dynamics of pedestrians as influenced by the combined effect of interactions between neighboring pedestrians. The change in walking speed and direction is specified as a function of several explanatory variables current walking speed, relative positioning of nearby pedestrians, space headway, sight angle, etc. It was noted that obvious variables such as distance headway and time headway turned out to be insignificant in some models. Other studies, such as [2], explore the sensitivity of walking speed to the lateral motion of the walking platform. Moussaid et al. [19] model the walking behavior of pedestrian social groups and concludes that the speeds decrease linearly with density and group size. They model pedestrian interactions by extending the social force model using a new interaction term considering the gazing direction of a pedestrian. Antonini et al. [1] proposed a discrete choice modeling approach to model the pedestrian walking behavior. Here, the choice set for each pedestrian is a combination of speed regimes and direction.

The presence of groups can significantly influence the walking behavior of individuals, considering the interaction between group members. Many studies report that as the group size increase or as the proportion of groups in a crowd increase, the walking speed of pedestrians' decrease [8, 16, 27, 27]. A large proportion of these studies use data from controlled experiments to understand the walking behavior—walking speed and stepping behavior of pedestrians [23, 27, 29]. The participants of these experiments are asked to walk in marked sections and at times asked to adjust their speed as slow walkers, normal walkers, and fast walkers to explore the effect of different speed compositions on the collective behavior of individuals [10].

Though numerous studies have explored the walking dynamics of pedestrians, comparatively fewer studies focus on pedestrian flow in natural environments, especially in mass gatherings. Also, the effect of boundary conditions of the study stretches; as the boundaries could be rigid without allowing entry/exit of pedestrians throughout the section, it could be soft boundaries such as movable barricades and ropes or without any physical boundary such as an open roadway can considerably influence the walking dynamics. We would like to explore the above-mentioned effects on pedestrian walking speed. Additionally, the influence of gender and height on the walking speed at different flow conditions is investigated. The following section describes, in brief, the data used for the analysis.

3 Data Description

Simhasth Kumbh, as Ujjain Kumbh Mela is popularly known, is one of the largest spiritual gatherings in India. Video graphic data of pedestrian flow at five locations in Ujjain—is used to understand the walking dynamics of pedestrians in the natural environment. The data collection was carried out for 30 days starting from 22nd April 2016 to 21st May 2016 during Kumbh Mela and for a week during the Mahashivrathri festival in 2017. During the data collection, a minimum temperature of 22 °C and a maximum temperature of 44 °C was recorded on 6th May and 19th May 2016, respectively. The videos have been captured using mobile phones, Go Pro cameras, and taken from CCTV footages. We consider different walking conditions including an uninterrupted unidirectional flow (Panchkroshi yatra—a religious yatra/travel pilgrimage), interrupted flow conditions in barricaded roadways, unidirectional flow through narrow corridors, bottleneck sections, and an approach road to the intersection (refer Fig. 1). A brief description of the locations considered, and the flow conditions are presented in Table 2.

Panchkroshi yatra is a travel pilgrimage by walk undertaken by the pilgrims. The pilgrims come to the city to visit many temples to offer their prayers and walk a total distance of about 118 km by foot. Many of these pilgrims travel with head luggage. Mahakal temple, being one of the prime attractions, is visited by a large number of pilgrims. Inside the temple, the pilgrims are required to pass through narrow channels with strong iron railings on either side. The bottleneck section considered for the study is a barricaded temple queue where the width of the channel reduces from 1.6 m in the upstream section of the U-turn to 0.75 m in the downstream section. Data of pedestrians in the queues in straight sections and bottleneck sections are also extracted.

Fig. 1 Locations considered for the study

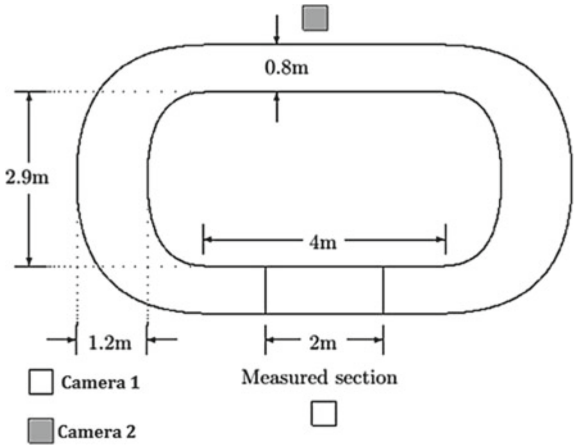
Table 2 Specifics of the study locations

Flow conditions	Locations	Sample size	Length of the study section (m)
Uninterrupted unidirectional flow	Ambodhya stretch in Panchkroshi yatra	788	5.8
	Gaughat	117	3.8
	From Harsiddhi Marg to Temple	2729	1.8
Interrupted unidirectional flow	Mahakal Temple queue—Upstream section of bottleneck	1548	2.3
Bi-directional flow	Approach road to Harsiddhi Chauraha	120	3.4

The pedestrian data from videos are processed manually by drawing the trap lengths on the screen using the Swordsoft Screenink annotation tool. To estimate the walking speed of pedestrians, the entry time (t^i) and exit time (t^j) of the pedestrian into the study section are noted. Wherever feasible, sociodemographic characteristics of the pedestrian such as age, gender, and height were noted. Since these observations were made manually, we have not considered age for the analysis due to the subjective nature of the variable. Walking speed is then estimated as

$$V_n = l/(t_n^j - t_n^i); \tag{1}$$

Fig. 2 Experimental setup adopted in Gulhare et al. 12



V_n: walking speed of pedestrian *n,* (m/s)
t_n^i: entry time of pedestrian *n* into the study section, (s)
t_n^j: exit time of pedestrian *n* from the study section, (s)
l: measured length of the study section, (m)

Data from pedestrians' single-file movement in controlled experimental conditions are used to compare walking dynamics across natural and experimental settings. The data of pedestrian walking from a controlled experimental setup is taken from the study performed by Gulhare et al. [12]. The experiment was conducted at the Indian Institute of Science, Bangalore. The participants were asked to move in a closed loop. About 40 participants in the age group of 25–50 took part in the experiment. The total length of the closed loop was 17.3 m, where the pedestrian data are recorded for the measured section shown in Fig. 2. Further details on the experimental setup can be obtained from Gulhare et al. [12].

4 Analysis of Pedestrian Walking Speed

In general, unidirectional flow refers to pedestrian flow in one direction alone. Many times, the boundary conditions of the study stretch can be different for the same flow conditions. Single-lane roadways with the restricted flow in one direction, pedestrian flow in temple queues with strong iron railings for support, separated and designated paths on roads by portable barricades/ropes fall under unidirectional flow. In addition, some of these flow conditions are regulated by event managers who stop and release the pedestrian flow at regular intervals. However, the boundary conditions, the type of activity, and regulations imposed on flow have a significant impact on the walking speed of pedestrians.

Hypothesis testing on the difference in the mean speeds of pedestrians of different socioeconomic characteristics such as gender and height was done for different study locations. The hypothesis that the mean speed of male pedestrians and that of female pedestrians is the same (Null hypothesis H0: $v_{\mathrm{m}} - v_{\mathrm{f}} = 0$ and alternate hypothesis H1:$v_{\mathrm{m}} - v_{\mathrm{f}} \neq 0$) is tested. Similarly, the hypothesis that the mean speed of pedestrians with medium height and the mean speed of short pedestrians is the same (Null hypothesis H0: $v_{\mathrm{md}} - v_{\mathrm{sh}} = 0$ and alternate hypothesis H1: $v_{\mathrm{md}} - v_{\mathrm{sh}} \neq 0$) is tested. The appropriate t-test is carried out for the various study locations.

4.1 Effect of Geometry and Boundary Conditions of the Section on Walking Speed

It is noted that the average walking speed of pedestrians in a unidirectional flow depends on the geometry of the study section (see Fig. 3). The highest mean walking speeds of 1.11–1.20 m/s are noted on open roadways without barricading, even when the densities are as high as 1.2 ped/m^2. The lowest mean walking speeds of 0.4–0.88 m/s are observed in narrow corridors with strong boundaries such as iron railings. Larger variations in pedestrian speed are also noted in temple queues compared to pedestrian flow in Panchkroshi yatra. From Table 3, it can be concluded that walking speeds are higher when the pedestrian flow happens through locations with soft boundaries. Restricting the movement by forcing the pedestrians to walk through a maze of narrow corridors enclosed by strong boundaries lowers the walking speed of pedestrians.

It is observed that when the boundaries of the study section are impervious such as iron railings, it does not allow the pedestrians to filter out or in through the boundaries except in some exceptional cases. The exceptions could be in terms of people

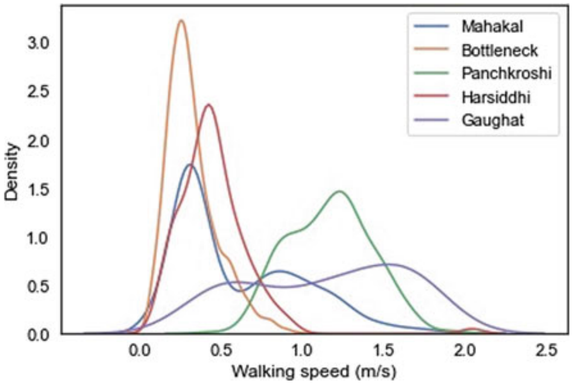


Fig. 3 Pedestrian walking speed distribution at different locations

Table 3 Average walking speed at different locations

Location	Average speed (m/s)	Walking direction
Panchkroshi yatra	1.11	Unidirectional (Continuous movement)
Bottleneck-u/s	0.5	Unidirectional (Barricaded with strong railings)
Harsiddhi Marg to Mahakal temple	0.60	Unidirectional (Continuous movement)
Gaughat	1.14	Unidirectional (Separated and demarcated roadways)
Experimental SFM	0.42	Unidirectional (Continuous movement in a loop without barricading)
Harsiddhi Chauraha	0.45	Bi-directional movement on approach roads

who attempt to jump over the rails to gain quick access. However, previous boundaries allow for faster pedestrians to overtake slower pedestrians for short sections through the sides. An example of this is pedestrian flow through a roadway with footpaths/sidewalks on either side. Faster pedestrians use these sidewalks to cut across the slow pedestrians just ahead of them and later merge into the mainstream.

Pedestrian speeds reduce drastically at bottleneck sections in narrow corridors. The walking speed of pedestrians in the upstream section of the bottleneck is 0.5 m/s. Here, the two lanes of pedestrian's merge and form a single lane, thereby lowering the pedestrian speeds considerably. Bi-directional flows also limit the walking speed due to the need to look out for pedestrians in opposite directions and resulting side friction.

In the experimental setup, pedestrians were asked to move in a loop without overtaking and without leaving the boundaries of the marked section. This setup allows the pedestrians to walk only in a queue, limiting their walking speeds as the speeds become a direct function of mostly spacing between the pedestrians. However, the average speed in an experimental setup is still lower than the average speed in the upstream section of the bottleneck in field conditions. This raises the question of relying solely on insights from pedestrian flow in controlled experimental setups to understand the walking dynamics of pedestrians.

4.2 Comparison of Walking Speeds Across Different Sociodemographic Groups of Pedestrians

To investigate the effect of gender on walking speeds, unidirectional pedestrian flow data of three locations is used: Roadway from Harsiddhi Marg to Mahakal temple, Panchkroshi yatra, and upstream section of the bottleneck (Fig. 4 shows the variations

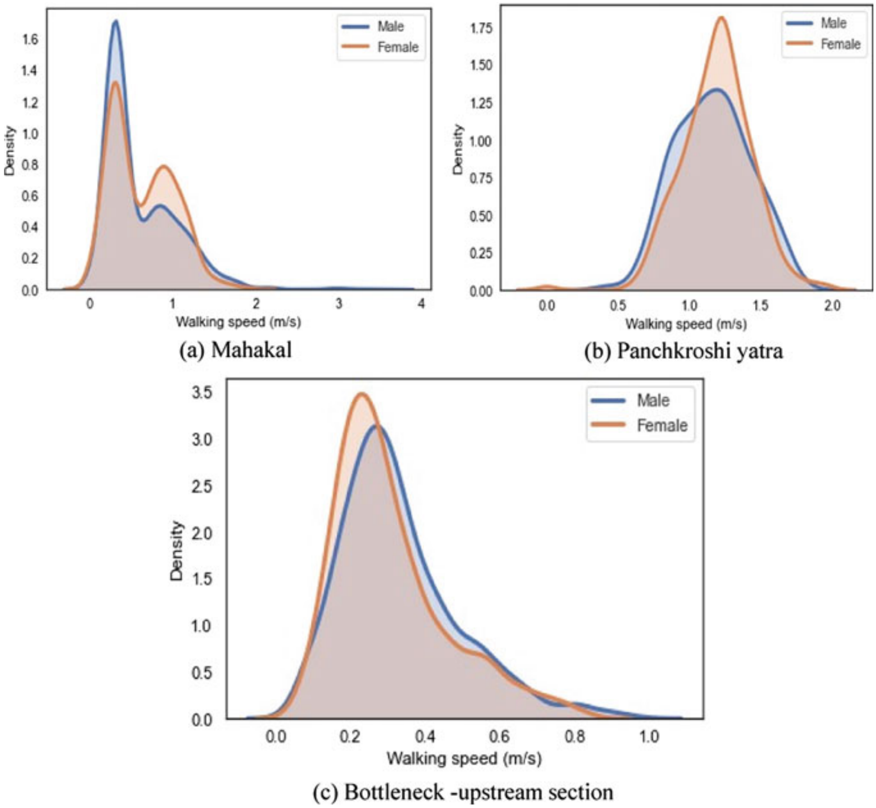


Fig. 4 Variation in walking speed distribution across gender (Mahakal, Panchkroshi, and Bottleneck-u/s section clockwise)

in the walking speeds across gender in the three locations considered). The details are presented in Table 4. The walking speeds of females are higher than the walking speeds of males in two of the study locations, which is counterintuitive as most studies report that females walk slower than males. In the upstream section of the bottleneck, the mean speed of males is higher than females.

Table 5 gives the results of the statistical analysis. The analysis reveals a significant difference in male and female pedestrians' mean speeds in temple queue at a 95% confidence level. Investigating the effect of gender on pedestrian speed in the bottleneck section reveals that the mean speeds of male and female pedestrians in the bottleneck's upstream section are significantly different at 95% confidence level. However, this difference is not evident in male and female pedestrians' mean speed in Panchkroshi yatra. It is seen that the pedestrians with shorter heights have lesser walking speed compared to pedestrians with medium height. This can be attributed to the advantage in terms of visibility, which allows them to practice better walking strategies. Shorter pedestrians are blocked by people taller than them and cannot

Table 4 Descriptive statistics of walking speed for different categories of pedestrians

Location	Description	Percent of total	Mean walking speed (m/s)	Standard deviation (m/s)
Toward Mahakal Temple	Gender: male	64.2	0.59	0.61
	Gender: female	35.8	0.63	0.38
	Height: medium	82.3	0.65	0.17
	Height: short	17.7	0.61	0.14
	Total	100	0.60	0.54
Panchkroshi yatra	Gender: male	37.2	1.16	0.26
	Gender: female	62.8	1.19	0.28
	Total	100	1.11	0.27
Bottleneck-upstream section	gender: male	64.3	0.33	0.16
	Gender: female	35.7	0.31	0.15
	Height: medium	91.8	0.33	0.03
	Height: short	8.2	0.29	0.02
	Total	100	0.32	0.16

Table 5 Analysis results

Location	Null hypothesis	t-statistic	Conclude
Toward Mahakal Temple	Mean speeds of male pedestrians are equal to the mean speeds of female pedestrians	−1.97	Reject H0
Toward Mahakal Temple	Mean speeds of pedestrians with medium height are equal to the mean speeds of short pedestrians	−2.07	Reject H0
Panchkroshi yatra	Mean speeds of male pedestrians are equal to the mean speeds of female pedestrians	−1.47	Do not reject H0
Bottleneck-u/s	Mean speeds of male pedestrians are equal to the mean speeds of female pedestrians	2.76	Reject H0
Bottleneck-u/s	Mean speeds of pedestrians with medium height are equal to the mean speeds of short pedestrians	2.26	Reject H0
Experimental setup -SFM	Mean speeds of pedestrians in experimental setup are equal to the mean speeds of pedestrians in field	−11.07	Reject H0

plan their strategy as well as them. To further validate, if the walking environment influences the walking speeds other than the sociodemographic features, we next compare the walking speed of pedestrians between experimental conditions and field conditions.

In addition, comparing pedestrians' mean speeds in controlled experimental setup and field setup can help to corroborate the applicability of using experimental data to calibrate pedestrian flow models. It is found that there is a significant difference in the mean speed of pedestrians in the experimental setup and the mean speed of pedestrians in the field setup. These differences indicate that the walking dynamics of pedestrians in the natural environment are different from the walking dynamics of pedestrians in controlled experimental conditions. It is also difficult to motivate the participants in controlled experimental setups to represent the state of pedestrians in mass gatherings.

5 Discussion

Efficient management of crowd by the event managers relies on a realistic pedestrian flow model. Pedestrian walking speed is an essential parameter in modeling pedestrian flow. It varies depending on the facility type and other infrastructural elements. A multitude of factors can influence the walking speed of pedestrians, such as age, gender, being a part of the group, presence of vulnerable segments of people in the group such as children, older adults, presence of luggage, and health issues. Here, we investigate the factors influencing the walking speed of pedestrians in mass gatherings. Statistical tests are used to check if the mean speeds of pedestrians with different sociodemographic characteristics are different. It is seen that the boundary conditions of the pedestrian flow influence the walking speeds of pedestrians. The study also shows that the speeds of females are on par or higher than the speeds of males in two of the study locations. This could be the effect of the nature of the gathering and the motivation levels as females are found to be more spiritually oriented than their male counterparts [17, 25].

To overcome the practical difficulties associated with field data collection of pedestrian movement in crowds, many studies use controlled experiments. These experiments offer much more flexibility for the researcher to ensure proper volunteer coordination, ideal camera positioning, and also allows for a more focused data collection in terms of the objectives of the study. Data of single-file movement, bottleneck situations, and evacuation scenarios have been collected using such experiments. However, these studies involve participants belonging to a homogeneous segment such as students. The situation is very different in field where the crowd is much more heterogeneous and are more enthusiastic/motivated to perform activities unlike the paid volunteers. Therefore, this study also compares the experimental data with field data to bring out the differences in walking speed of pedestrians, which further stresses on the need to calibrate the crowd models based on real empirical data.

This study based on empirical data indicates that pedestrians' walking speeds in mass gatherings are considerably slower than the generally adopted speed of 1.2 m/s. These insights can be used to model pedestrian flow in mass gatherings more realistically. Specifically, variations in walking speed and the factors affecting pedestrians' speed can help better understand the following behavior of pedestrians.

Acknowledgements The work reported in this paper is part of the project titled The Kumbh Mela Experiment: Measuring and Understanding the Dynamics of Mankind's largest crowd, funded by the Ministry of Electronics and IT, Government of India (MITO-0105), Netherlands Organization for Scientific Research, NWO (Project no. 629.002.202), and Robert Bosch Center for Cyber Physical Systems, Indian Institute of Science, Bangalore (Grant No. RBCO001). The authors also express their gratitude toward Kumbh Mela administration and government of Madhya Pradesh, India for providing constant support and official permissions to carry out research work and establish Indo-Dutch collaborative research camp at Kumbh Mela 2016.

References

1. Antonini G, Bierlaire M, Weber M (2006) Discrete choice models of pedestrian behavior. Transp Res Part B: Methodol 40(8):667–687
2. Bruno L, Venuti F (2008) The pedestrian speed–density relation: modelling and application. In: Proceedings of footbridge
3. Chandra S, Bharti AK (2013) Speed distribution curves for pedestrians during walking and crossing. Procedia Soc Behav Sci 104:660–667. https://doi.org/10.1016/j.sbspro.2013.11.160.
4. Chang C-Y, Wang S-F (2011) Analysis of pedestrian walking speeds at crosswalks in Taiwan. J Eastern Asia Soc Transp Stud 9:1186–1200
5. Chattaraj U, Seyfried A, Chakroborty P (2009) Comparison of pedestrian fundamental diagram across cultures. Adv Complex Syst 12:393–405
6. Daamen W, Hoogendoorn SP (2003) Experimental research of pedestrian walking behavior. Transp Res Record 20–30
7. Daamen W, Hoogendoorn SP (2007) Flow-density relations for pedestrian traffic. In: Schadschneider A, Pöschel T, Kühne R, Schreckenberg M, Wolf DE (eds) Traffic and Granular Flow'05. Springer, Berlin. https://doi.org/10.1007/978-3-540-47641-2_27
8. Do T, Haghani M, Sarvi M (2016) Group and single pedestrian behavior in crowd dynamics. Transp Res Record: J Transp Res Board 2540:13–19
9. Duives DC, Daamen W, Hoogendoorn SP (2017) Operational walking dynamics of crowds modeled with linear regression. Transp Res Rec 2623(1):90–97
10. Fujita A, Feliciani C, Yanagisawa D, Nishinari K (2019) Traffic flow in a crowd of pedestrians walking at different speeds. Phys Rev E 99
11. Gates TJ, Noyce DA, Bill AR, Van Ee N (2006) Recommended walking speeds for timing of pedestrian clearance intervals based on characteristics of the pedestrian population. Transp Res Rec 1982:38–47
12. Gulhare S, Verma A, Chakroborty P (2018) Comparison of pedestrian data of single file movement collected from controlled pedestrian experiment and from field in mass religious gathering. Collect Dyn 3:1–14. ISSN 2366-8539
13. Gupta A, Singh B, Pundir N (2017) Effect of gradient on pedestrian flow characteristics under mixed flow conditions. Transp Res Procedia 4720–4732
14. Jin C-J, Jiang R, Wong SC, Li D, Guo N, Wang W (2017) Large-scale pedestrian flow experiments under high-density conditions. arXiv: Physics and Society
15. Laplante JN, Kaeser TP (2004) The continuing evolution of pedestrian walking speed assumptions. ITE J (Inst Transp Eng) 74(9):32–40
16. Lu L, Chan CY, Wang J, Wang W (2017) A study of pedestrian group behaviors in crowd evacuation based on an extended floor field cellular automaton model. Transp Res C: Emerg Technol 81:317–329

17. Luqman N, Puri A, Tanwar KC (2015) Gender differences in spiritual personality. Int J Multidiscipl Current Res 3(10):719–722
18. Montufar J, Arango J, Porter M, Nakagawa S (2002) Pedestrians' normal walking speed and speed when crossing a street. Transp Res Record: J Transp Res Board 90–97
19. Moussaïd M, Perozo N, Garnier S, Helbing D, Theraulaz G (2010) The walking behaviour of pedestrian social groups and its impact on crowd dynamics. PLoS ONE 5(4):1–7
20. Rahman K, Ghani NA, Kamil AA, Mustafa A (2012) Analysis of pedestrian free flow walking speed in a least developing country: a factorial design study. Res J Appl Sci Eng Technol 4(21):4299–4304
21. Sahani R, Bhuyan PK (2019) Pedestrian service quality models at un-signalized crossings. Adv Transp Stud 47(March):61–74
22. Sangeeth K, Lokre A (2019) Factors influencing pedestrian speed in level of service (LOS) of pedestrian facilities. Transp Res Interdiscipl Perspect 3:100066
23. Seitz MJ, Köster G (2012) Natural discretization of pedestrian movement in continuous space. Phys Rev E Stat Nonlin Soft Matter Phys. PMID: 23214653
24. Seyfried A, Steffen B, Klingsch W, Boltes M (2005) The fundamental diagram of pedestrian movement revisited. J Statist Mechan: Theor Exp
25. Trzebiatowska M, Bruce S (2014) Why are women more religious than men? Oxford University Press, New York
26. Virkler MR, Elayadath S (1994) Pedestrian speed-flow-density relationships. Transp Res Record
27. Wei X, Mai XU, Lv WEI, Song W (2014) Microscopic character and movement consistency of pedestrian group: an experimental study in campus. Saf Sci 133:1103–1114
28. Yugendar P, Ravishankar KVR (2018) Crowd behavioural analysis at a mass gathering event. J Konbin 46(1):5–20
29. Zhang B, Chen W, Ma X, Qiu P, Liu F (2020) Experimental study on pedestrian behavior in a mixed crowd of individuals and groups. Physica A 556:124814
30. Zhang J, Seyfried A (2014) Comparison of intersecting pedestrian flows based on experiments. Physica A 405:316–325

Effect of Free Left-Turning Vehicles on Pedestrian Safety at Signalized Intersection

Piyush Lalwani, Mukti Advani, and Sanjay Dave

Abstract This study aims to quantify the risk associated with these conflicts using one of many surrogate safety measures, i.e., post-encroachment time (PET). Study analysis is based on the data collected at two 4-arms signalized intersection one located in two Indian cities, one located in Vadodara, and other in Delhi. At both the locations, vehicular and pedestrians' movement has been captured through videography, and the same has been used for data extraction. Data extraction included pedestrian volume, classified vehicle volume, speed of vehicles, speed of pedestrians, waiting time of pedestrian, lag/gap time and post-encroachment time, vehicle type, pedestrian gender, and age group. A linear regression model was prepared to estimate the total number of conflicts, and as expected, pedestrian volume was found to be positively correlated with number of conflicts. Further, conflict severity was calculated based on PET values which is a conflict proximity parameter and indicates how close vehicle and pedestrians were while interaction (conflict). Based on PET values, all conflicts have been categorized into three categories, i.e., severe, moderate, and for both the locations, it was observed that nearly 25–45% of the conflicts falls in severe category. This highlights the need of intersection/signal improvement at both locations to reduce conflict severity.

Keywords Intersections · Free left turning · Pedestrians · Safety · India

P. Lalwani (✉)
The Maharaja Sayajirao University of Baroda, Vadodara, India
e-mail: piyushlalwani1998@gmail.com

M. Advani
Transportation Planning and Environment Division, CSIR-Central Road Research Institute, New Delhi, India
e-mail: mukti.crri@nic.in

S. Dave
Civil Engineering Department, The Maharaja Sayajirao University of Baroda, Vadodara, India
e-mail: smdave-ced@msubaroda.ac.in

L. Devi et al. (eds.), *Proceedings of the Sixth International Conference of Transportation Research Group of India*, Lecture Notes in Civil Engineering 272,
https://doi.org/10.1007/978-981-19-3494-0_14

1 Introduction

Traffic safety is a major concern due to increased number of crashes in India. India has a road crash death rate of 22.6 per 1 lac of population WHO, 2019 [1]. Figure 1 presents the crash (accident) data for India for the period of five years, i.e., year 2014–2018.

From, Fig. 1, it can be observed that there was decrease in number of accidents from 2014 to 2017, but however, the number of fatalities was still increasing. Apart from the count of accidents, Road Accidents in India report MoRTH, 2018 also provides the classifications of these accidents and deaths based on road user category, location type [2].

- As extracted from report, pedestrian accounts for 15% of its which is the third highest after motorcycle and four-wheeler
- There has been increase of 10.75% in pedestrian death from 2017 to 2018. Moreover, 33% of the road accidents takes place at junctions (Intersections) of which 12.1% is at T-junction and 7.4% is at four-armed junction. The road accidents occurring at signalized junctions are 2.9% (13,726) of total crashes at junctions. These 13,726 accidents result in deaths of 3325 people which is 21% of the total people involved, which shows that the accidents occurring at the signalized intersection have significant severity.

The signalized intersections are provided to reduce the number of conflicts by reducing the number of movements taking place at an intersection at a given time and to enhance the capacity of the intersection. This is possible only if the signal phases are designed efficiently, and the rules are followed by the road user. Any error in phase designing or indiscipline behavior by road users will cause delays and increase in conflicts. Garder [3] studied crash data of 120 intersections in Sweden and found that majority of cases of accidents were due to a turning vehicle hitting a green walking pedestrian and a red walking pedestrian hit by a vehicle. Another study by Xu et al. [4] found that major factors contributing to the increased crashes at signalized intersection are as follows: pedestrian above 65 years of age, heedless crossing, also, he found that places with low pedestrian volume have higher number

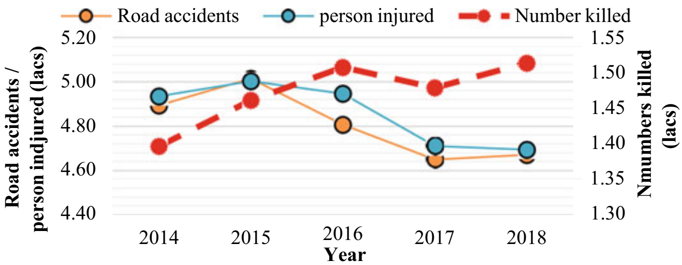


Fig. 1 Trend of road accidents, person injured, and numbers killed from year 2014–2018

of pedestrian crashes. Abdel [5] concluded that pedestrian and driver demographic factors, road geometry, traffic conditions, environment, are the factors affecting the road severity and frequency of road crashes. With increase in the traffic volume, there was increase in the number of accidents. In research conducted by Rankavat et al. [6] on pedestrian accident analysis using GIS, they found that the factors affecting the number of pedestrian road fatalities are population density of the area, the category of the road, and the type of the vehicle. The crashes with cars and buses were found to be more fatal than other categories of vehicles. Anjana et al. [7] carried out study under mixed traffic conditions and identified the factors that affect the safety at signalized intersections. Study identified median width, green time, proportion of two wheelers, and volume of the traffic stream as the significant parameters. Same study also found that the provision of the channelized left-turn lanes can help to reduce the number of crashes. More than of intersections with channelized left turning, most commonly observed are signalized intersections with free left turning. However, limited research on the effect of the free left turns on the pedestrian safety under heterogeneous conditions is available in literature. Though the subjective discussion is included in few studies regarding increase in number of free lefts-turning vehicles and its effect on pedestrians' safety, quantification of the same is missing gap in literature. This study aims to study the pedestrian behavior while crossing against the free left-turning vehicles under the heterogeneous traffic conditions and their risk of getting into a conflict using time gap and post-encroachment time (PET) as key performance indicators. To strengthen the method of analysis, existing studies with respect to pedestrians' crossing behavior and pedestrian-vehicle interaction have been studied and discussed in the next section of literature review.

2 Literature Review

Most commonly, pedestrian-vehicle interaction is termed as conflict. As per Swedish conflict technique Hyden et al. [8], "Conflict is a situation in which two road users approach each other in time and space in such way that a crash is highly probable if their paths remain unchanged." The surrogate safety measures such as gap time (GT), encroachment time (ET), time to collision (TTC), post-encroachment time (PET), time to accident (TA), and deceleration rate (DR) are most widely used in the studies for determining the risk of getting involved in crashes. Encroachment time and post-encroachment time indicate the actual conditions on the field, and their measurement is possible only after the conflict has occurred, but both differ from one another because ET considers the boundary of right of way of the road user for defining the conflicts, whereas post-encroachment time considers a collision point, which is the actual point of intersection of two conflicting road users. Time to accident and time to collision are based on predicted trajectories and hence can be calculated before the conflict actually happens. Time to accident is different from time to collision since it considers the beginning of the measurement after the evasive action is taken, and time to collision can be measured from the beginning the two road users can be seen to be on a collision path till the conflict has taken place (not evasive action); multiple TTC are possible, and the minimum TTC is the one that is to be taken into consideration.

This section deals with the studies that have used traffic conflict technique for determining the pedestrian safety at intersections and crosswalks and determine the factors responsible for decreased safety of pedestrians. Chandrappa et al. [9] conducted the stated preference survey and revealed behavior survey; the revealed behavior and the stated behavior had huge variation. Also, the 85th TTC observed was 1.92 s for intersections and 2.25 s for midblock, showing that there is a higher risk of collision at intersections than midblock. Perumal et al. [10] found that approaching vehicle direction, position, type along with pedestrian age and speed are the parameters that have a significant impact on the PET category of the pedestrian. The authors found 15th percentile to be 2 s and 50th percentile PET to be 5.5 s. Also, turning movements were less dangerous than the through movement, which was because the turning drivers were reducing the speed of the vehicle. Kumar et al. [11] took into consideration the pedestrian dominance and vehicle dominance at 12 sites in India and found that the near-side pedestrians show higher coercive dominance as compared to that of both far-side pedestrian, and that the yielding preference of the driver depends on the type of maneuvering they take. The left-turning driver yield more often than through and right-turning drivers. Chaudhari et al. [12] from the PET analysis found 15th percentile PET to be 1.5 s and considered interactions with PET < 1.5 s as dangerous conflicts, 50th percentile PET as 5 s, and any conflict between 1.5 and 5 s was considered as a normal conflict. Nearly, 8–32% was dangerous conflicts, and 2–35% was normal conflicts. Lord [13] observed that X-intersection is 2–4 times safe than the T-intersection in terms of pedestrian safety. He also added that in case of T-intersection, there is higher risk of conflict if pedestrian starts crossing in earlier stage of green phase, while for X-intersection, it was opposite; later, the pedestrian starts crossing the road higher is the chance of conflict. Zhao et al. [14] prepared models for four different turning radii, viz., 2, 10, 22, and 30 m. The models had R square value above 0.7 which shows a good correlation between traffic conflicts and pedestrian and vehicular flow. In addition, the numbers of severe conflicts were more at large turning radius, and there were no severe conflicts found for 2 m turning radius. Ling et al. [15] from their study on pedestrian and right-turning vehicle conflicts found that individual pedestrians, pedestrian crossing from near side, high right-turning vehicle volumes, higher distance of pedestrian from conflict point, and smaller time difference between the vehicle and pedestrian arriving lead to increased probability of the pedestrian yielding. Sha et al. [16] in their conflict analysis at a signalized T-intersection found that there were 31, 41, and 28% conflicts between 0–1, 1–2, 2–3 s, respectively. The authors observed that nearly 65% of the conflicts was due to right-turn movements, and rest 35% was because of other movements. The author concluded that the right turn on red maneuvers poses a threat to the pedestrian, and moreover, pedestrians' spatial and temporal violation lead to increased conflicts. Cheng et al. [17] observed that 57% of PET < 3 s occurred inside the cross walk, and 43% outside the cross walk. It was also observed that drivers tend to adopt smaller turning radius, and their trajectory was within half the crosswalk length from the curb side. Jiang et al. [18] from a comparative test on three different sites, viz., non-channelized right-turn lane, non-channelized right turn through lane, and channelized right-turn lane, authors found that the channelized intersection posed had higher risk index, higher mean risk index for crossing decisions, reduced PET value, higher mean crossing speed, high-rate failures in interactions (evasive behaviors).

3 Objective and Study Methodology

The objective of this study are as follows:

- To study the safety of pedestrians against free left-turning vehicles using surrogate safety parameters.
- To find out relationship of pedestrian characteristics (gender, age, walking speed, waiting time, gap acceptance) and vehicular characteristics (vehicle type, turning speed) with PET.
- To find out the factors affecting the number of conflicts and develop a suitable mathematical model.

Figure 2 presents the methodology adopted for this study. Based on comprehensive review of the available literature, the research gap was identified, which was followed by setting of objectives as mentioned above. Once the objectives were set, the study site selection was carried out by going through the available video data and finding the site that fits the objective, and two signalized intersections were selected, one located in Vadodara and other in Delhi. For these two sites, video data were collected; this video data were later used for extracting various parameters such as vehicle characteristics and pedestrian characteristics. This extracted data were then analyzed using the SPSS statistical software for different categories. Final analysis focuses on preparation of linear regression model for prediction of total conflicts in an hour and conflict severity measurement.

4 Study Area Characteristics

This study considers two four-legged signalized intersections; one is located in Khodiyar Nagar of Vadodara city (location-1), and another location is in Shakti Nagar, Delhi (location-2). Both the locations have channelization for left turns at each leg. This study is based on data collection at one of the four arms at both the locations. Figure 3 shows the intersection layout of the selected signalized intersection.

Table 1 provides geometric details of site in Vadodara and Delhi. The measurements “A” and “I” are of importance to us since those are the approaches whose left turn has been taken into consideration, and ‘I’ denotes the crosswalk length. The approach width “A” of Khodiyar Nagar intersection excludes the additional width provided for channelized left turn; the additional lane for left turn is 4.5 m and is

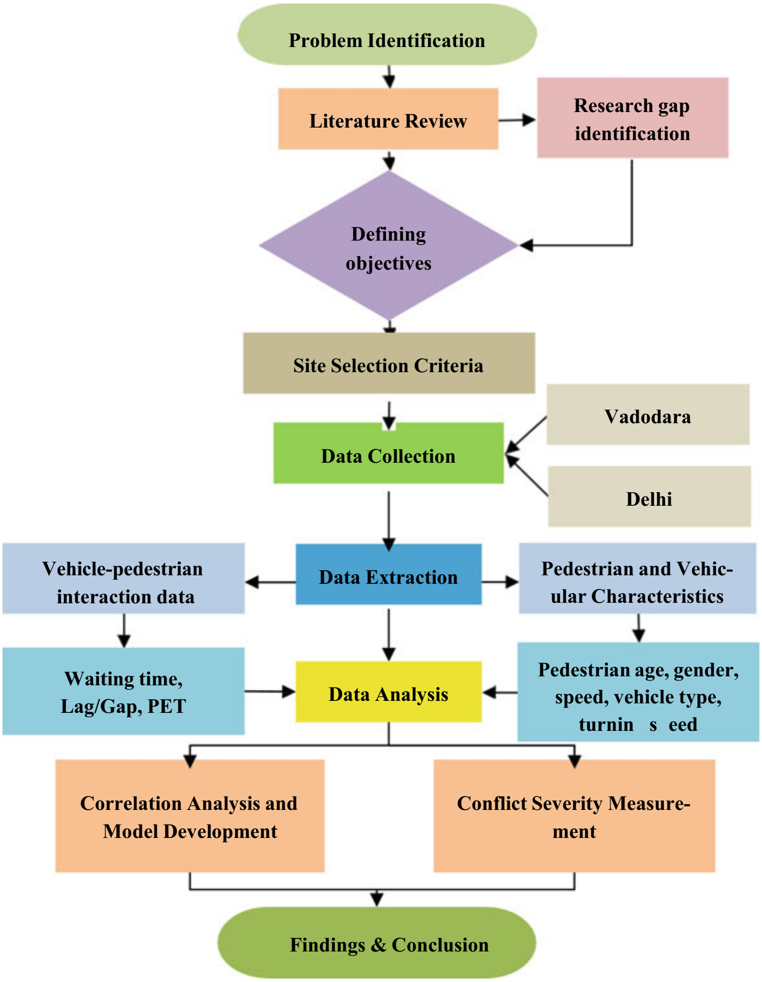


Fig. 2 Study methodology

made available for a length of 50 m (paved), and the remaining part is provided as a shoulder, whereas approach width A for Shakti Nagar intersection of Delhi includes the width of channelized left turn in it. It can be observed that the approach taken into consideration for Vadodara is two-lane road and that of Delhi is a three-lane road. The crosswalk length of both the left turns is of same length, i.e., 5 m.

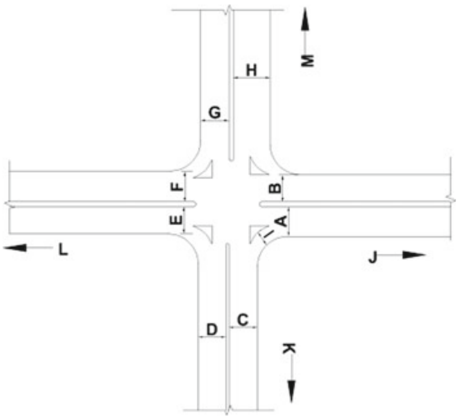


Fig. 3 Intersection layout

Table 1 Showing the geometric details of the site

Notation	Vadodara	Delhi	Notation	Vadodara	Delhi
A	8 + (4.5 m)	12.2	H	10 + (6 m)	10.2
B	8 + (6 m)	9.4	I	5 m	5
C	11 + (6 m)	9	J	Khodiyar Nagar	Jhangirpuri
D	10 + (6 m)	10.5	K	Sardar Estate	Bhagat Singh Chowk
E	8	9.5	L	Santaram Park	TIS Hazari Court
F	8 + (6.4 m)	9.2	M	Manek Park	Nangia Park
G	10 + (5 m)	9.2	('X' m) indicates shoulder width		

5 Data Collection and Data Extraction Methodology

5.1 Data Collection

Data collection for location-1 and location-2 was done in September 2020 and November 2018, respectively. Data collection includes 16 and 3.5 h of video data for locations 1 and 2, respectively. Camera view captures the left-turning vehicles and crossing pedestrians. Data extraction included following:

- Classified vehicle volume (left turning vehicles)
- Pedestrian crossing volume
- Turning and exit speed of vehicles turning left
- Pedestrian crossing speed
- Pedestrian waiting time
- Lag/gap time
- Post-encroachment time.

Along with these parameters, various characteristics of pedestrians such as age, gender, group size, direction of crossing, crossing area used, and vehicle type have also been extracted. Data extraction from collected videos is carried out with the help of open-source software, i.e., Avidemux 2.76. Process considered for data extraction is semiautomatic. Avidemux is a video playing software which allows to play the videos at varying speed as well as play it frame by frame. Prior to data extraction, a grid was laid over the captured video. The grid was prepared using the AutoCAD 2016 software and was rendered in video using the Movavi Video Editor Plus 2021. Four control points for grid calibrations were obtained by comparing the video data with the satellite image from Google Earth Pro software, and once the points from the video data and Google Earth Pro were matched, grid was laid using these points as control points. Figure 4 shows the grid overlaid on the video, and Fig. 5 shows the details of crosswalk area.

Table 2 shows the scale variables and the categories associated with it. These categories have been used in analysis in this study, to study factors affecting the scale variable. Column-1 of the table shows category name, column-2 describes what that category represents. For example, the row-1 in Table 2 signifies that the

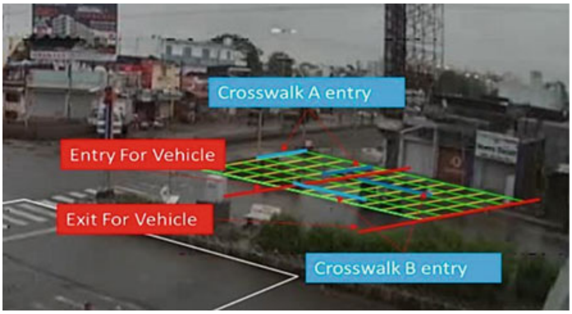


Fig. 4 Snapshot from video captured for location-1 with overlaid grid

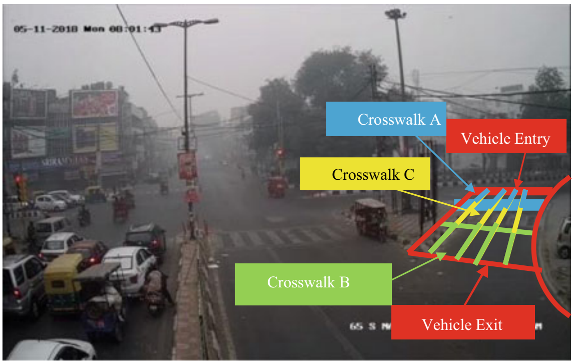


Fig. 5 Details of crosswalk area and stretch considered for vehicle speed at location-2

Table 2 Various categories used in analysis of post-encroachment time

Category	Definition
Motorized two-wheeler (2 W)	Vehicle class compromising of motorcycles and scooters
Motorized three-wheeler (3 W)	Vehicle class compromising of auto-rickshaw, cars or three-wheeler, and four-wheeler goods vehicle
Cycle/cycle rickshaw	Vehicle class compromising of cycles and cycle rickshaw
Bus/truck	Vehicle class compromising of the buses, trucks, light commercial vehicle (LCV)
Male	Pedestrian identified as male from video data
Female	Pedestrian identified as female from video data
< 30 years	Pedestrian with age less than 30 years
> 30 years	Pedestrian with age greater than equal to 30 years
Single	Pedestrian crossing alone
Group	Pedestrian crossing in a group
C2M	Pedestrian crossing from curb to median
M2C	Pedestrians crossing from median to curb
Crossing Area-A	Pedestrian crossing against the left turning vehicles entering the left turn
Crossing Area-B	Pedestrian crossing against the left turning vehicles exiting the left turn
Crossing Area-C	Pedestrians crossing from the area in the middle of the channelized left turn, generally where crosswalk is provided
< 15 Kmph	Vehicle speed group consisting of vehicles whose speed is less than 15 kmph
> 15 kmph	Vehicle speed group consisting of the vehicles whose speed is greater than equal to 15 kmph
Didn’t wait	Pedestrians who didn’t wait before crossing the road/crossed as soon as they reached the crosswalk
Waited	Pedestrians who waited before crossing the road

“Vehicle type” has a category by the name of “Motorized Two-Wheeler (2 W)” which provides the details about the vehicle type included in this category, and this category is used in both the study locations.

5.2 Post-Encroachment Time (PET)

Post-encroachment time is a measure of the proximity of a road user of getting into a crash. It indicates the time gap, by which pedestrian was saved from getting into a crash. PET value are obtained by drawing the actual trajectory of the conflicting road users in this case the pedestrian and the conflicting vehicle and determine their conflict point that is where their trajectories intersect. Once the conflict point (CP) is known, the time when the first road user left the CP ($tc1$) is noted down and the time when the second road user arrives at the CP ($tc2$) is noted down, and thus, the PET can be obtained as,

$$\mathrm{PET}(s) = tc2 - tc1;\ \ \text{seconds} \tag{1}$$

There are two scenarios in PET; one is called pedestrian pass first (PPF) in which the pedestrian first crosses against the vehicle, leaves the conflict point, and then, the conflicting vehicle enters the conflict point, and the other is vehicle pass first (VPF) scenario, in which the pedestrian hasn't yet crossed the road but has started to cross, and the driver increases the speed of the vehicle and passes first in front of the pedestrian; thus, it is a scenario in which the vehicle first leaves the conflict point, and then, the pedestrian reaches the conflict point.

6 Analysis and Results

As discussed in the previous section, various parameters were extracted from the video data. The descriptive statistics were performed on these parameters (Vehicle speed, pedestrian speed, waiting time, lag/gap time, and PET), and mean, median, standard deviation, maximum, minimum values were obtained. After the descriptive statistics, t-test was carried out to determine whether there is some significance difference between categories of the same group, for example, whether there is significant difference between the PET value of male and female of gender group. To determine the correlation between the various continuous variables such as vehicle speed, pedestrian speed, waiting time, lag/gap time and PET, Pearson's correlation analysis was carried out, and to determine the correlation between the categorical variables and PET, point-biserial analysis was carried out.

6.1 Pedestrian/Vehicle Volume Versus Conflicts

As presented in Fig. 6, left-turning vehicular flow was highest during 10–11 am. Observed traffic composition has 80% motorized two-wheelers, and remaining 20% vehicles constituted of cars, auto-rickshaw, and 3 W/4 W goods vehicles.

Also, pedestrian volume was highest during the same period, and it was 144 pedestrian/hour. A total of 1022 pedestrians crossed during 10 h duration. From the total pedestrian, there were 538 pedestrians who walked from the crossing area (non-violators), and there were 484 pedestrians who crossed away from the crossing area (violators). A total of 588 conflicts were observed of which 406 conflicts were between vehicle non-violators, and 182 were between vehicle and violators. Scope of this study is limited to pedestrians that are non-violators.

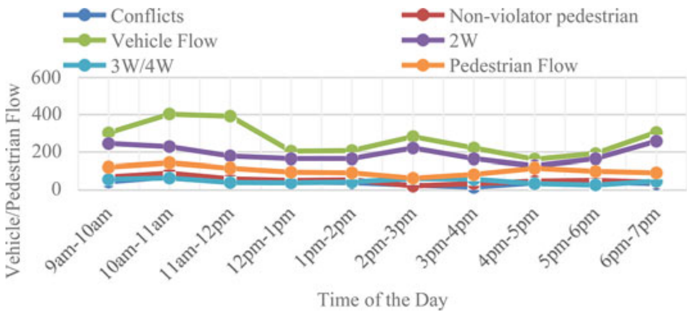


Fig. 6 Hourly variation in pedestrian volume, vehicular volume, and pedestrian-vehicle conflicts for location-1

Figure 7 shows vehicular and pedestrian flow variation for every 15 min at Shakti Nagar intersection of Delhi.

Left-turning vehicular flow at Shakti Nagar was lower than that of the pedestrian flow as seen from Fig. 7. Cars and auto-rickshaw constituted 51% share of the total traffic; two wheelers share was found to be 39%; trucks/buses share was 3%, and the slow-moving vehicles that are cycle and cycle rickshaw had a share of 7% in the total traffic. A total of 1050 pedestrians crossed in 3.5 h against the left turning vehicles. A total of 418 conflicts were observed in 3.5 h. of video data.

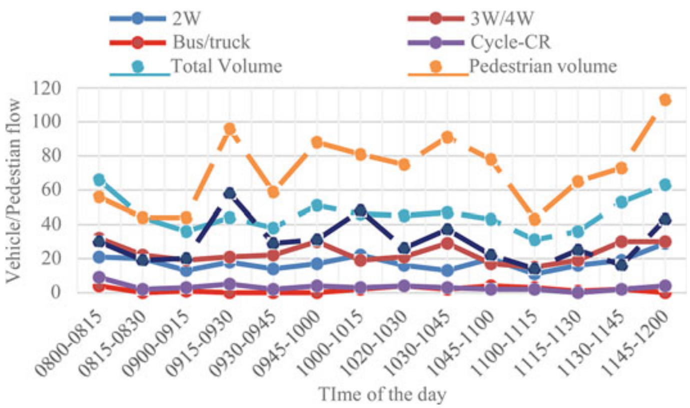


Fig. 7 15-min variation in pedestrian volume, vehicular volume, and pedestrian-vehicle conflicts for Shakti Nagar, location-2

6.2 Correlation Analysis and Linear Regression Model

Correlation analysis is used to understand the strength of interrelationship among variables. From correlation analysis, total number of conflicts (conflicts/hour) was found to be strongly correlated with the pedestrian flow, with "r" value of 0.89; vehicular flow had a correlation coefficient of 0.60, whereas number of 2 W and 3 W/4 W (per hour) had weak correlation. For location-2, it was observed that the total number of conflicts is strongly related to pedestrian volume. Other variables such as 2 W, 3 W/4 W, bus/truck are not strongly correlated. Using total conflicts/hour as dependent variable and pedestrian volume and vehicle volume as independent variable, a linear regression model (without constant, since number of conflicts can be zero, when pedestrian flow and vehicle flow is zero) was developed using SPSS stepwise method for both the locations. The regression analysis for both locations gave multiple models each having different variables, coefficient, and r square values. For a selecting the final model, following points were taken into consideration:

- The variables included in the model should be significant
- The variables should not exhibit multi-collinearity since multi-collinearity may improve the R square, but it will give over-estimated predictions
- The standard error of the variables and standard error of estimate of the model should be minimum.

Using the above criteria, the below models shown were finalized for the respective locations:

Equation 2 represents the linear regression model for location-1

$$Cn = 0.7639 * p_f (\text{Adj. } R^2 = 0.9806), (\text{RMSE} = 5.68) \quad (2)$$

Equation 3 represents the linear regression model for location-2

$$Cn = 0.4167 \, * \, P_f; (\text{Adj.} R^2 = 0.9032), (\text{RMSE} = 11.82) \quad (3)$$

where

CnP_f	Number of conflicts
VTFW	Number of vehicles of type two/three-wheelers
VBT	Number of vehicles of type bus/truck.

From Eqs. 2 and 3, it can be interpreted that at location-1 the total conflicts in an hour are strongly related to the pedestrian volume, and with increase in pedestrian volume, the total conflicts increase. For location-2, also, it was observed that number of conflicts are dependent on the pedestrian flow.

6.3 Vehicle Speed

The turning speed of vehicles was extracted from the video data, and descriptive statistics regarding it were obtained for both the locations and for vehicle groups. Observed average speed of 2 W at location-1 is 18.26 kmph and that of 3 W/4 W group 13.24 kmph, and the difference is found to be statistically significant ($p = 0.00$). For location-2, the 2 W group had a speed of 18.22 kmph; 3 W/4 W group had a speed of 15.85 kmph; bus/truck group had a speed of 17.03 kmph, and cycle/cycle rickshaw had a speed of 7.95 kmph, and the difference between the speed of different vehicle categories was statistically significant ($p = 0.00$). The two-wheelers were traveling at higher speed as compared to 3 W/4 W because 2 W can maneuver easily without slowing down to a great extent because of smaller physical dimensions.

6.4 Pedestrian Crossing Speed

Average speed of pedestrians at location-1 was 1.03 m/s, whereas avg. speed of pedestrians at location-2 was 1.08 m/s. At location-1, the males had a speed of 1.05 m/s, and females have a speed of 1.0 m/s, and at location-2, the males were found to be at a significantly ($p = 0.00$) faster speed than females, with average speed of 1.11 m/s as compared to 1.02 m/s of females. The average crossing speed for pedestrian below age of 30 and above the age of 30 years was found to be significantly different ($p = 0.00$) for both the locations. The people below age of 30 years had a speed of 1.11 m/s at location-1 and 1.17 m/s at location-2, and the people above the age of 30 years had a speed of 0.96 m/s and 1.01 m/s at locations-1 and 2, respectively. It was also observed that at location-2 the pedestrian walking alone was walking at a significantly ($p = 0.00$) higher speed of 1.09 m/s as compared to those pedestrians who were walking in group at an average speed of 1.02 m/s; for location-1, this difference was insignificant.

6.5 Pedestrian Waiting Time

From the data extracted, it was found that majority of the pedestrians preferred crossing against the left turning vehicles without waiting. At location-1 out of 538 pedestrians, just 81.58% didn't wait at all before crossings; 13.16% waited for ≤ 5 s, and 5.26% waited for more than 5 s, and at location-2, 90.31% of the pedestrians was those who didn't wait at all; 7.59% waited for ≤ 5 s, and 2.09% waited for more than 5 s. The average waiting time was less than 1 s for both the locations-1 and 2. There wasn't any significant difference between the waiting times of male and female, pedestrian below 30 years and greater than 30 years, or pedestrian crossing alone or in a group, and the average waiting time was ≤ 1 s.

6.6 Pedestrian Lag/gap Rejection and Acceptance

From the descriptive statistics of the rejected lag/gap, it was observed that the total number of the average rejected gap by the pedestrians at location-1 was 1.15 s, and for location-22, it was 1.15 s. For male and female, the average rejected gap was 1.14 and 1.14 s, respectively. For location-2 too, there wasn't any significant difference between the gap rejection of male and female. At location-1, the pedestrian who was <30 years old had an average gap rejection of 1.10 s, and those above the age of 30 years rejected an average gap of 1.18 s; the difference was statistically insignificant. The average value of accepted gap for pedestrians at location-1 was found to be 4.72 s, and for location-2, it was 4.61 s. The gap acceptance behavior of the pedestrian at locations-1 and 2 was similar since for locations-1 and 2; male accepted an average gap of 4.77 and 4.61 s; females accepted an average gap of 4.72 and 4.72 s, respectively; pedestrian below 30 years accepted an average gap of 4.52 and 4.53, respectively, and people above 30 years accepted an average gap of 4.98 and 4.75 s, respectively. There wasn't any significant difference in the gap acceptance of two locations as well as in gap acceptance of pedestrians within the gender group and age group.

6.7 Post-Encroachment Time and Conflict Severity

Along with the number of conflicts, severity of the conflict is of prime importance. PET is a measure that can be extracted easily and is accepted widely in the research community as a measure of conflict severity. The severity of conflict can be affected by many factors such as vehicle speed, vehicle type, pedestrian speed, pedestrian gap acceptance, pedestrian age, gender, and group size. In order to determine the correlation between PET and various factors as mentioned above, Pearson's correlation test was carried out using SPSS software, and to determine whether there is difference in the severity of conflict among categories of various groups, nonparametric t-test was carried out. The PET values less than or equal to 8 s were just compared for determining the descriptive statistics since all the conflicts greater than 8 s were defined as outliers by SPSS, and using these values, descriptive statistics for different categories were obtained as shown in Table 3 (Table 4).

From descriptive statistics, the 15th percentile, 50th percentile, and the 85th percentile PET for location-1 were obtained as 1.066, 2.486, and 5.005 s, respectively. For location-2 statistics, the 15th percentile, 50th percentile, and the 85th percentile PET were obtained as 0.834, 1.668, and 3.603 s, respectively. Further, from the Table 3 of descriptive statistics, it can be observed that the average conflict at location-1 is significantly lower than that of location-2. After reviewing the literature, it was observed that the PET less than 1.5 s denotes dangerous conflicts for pedestrians; a PET between 1.5 and 5 s denotes moderate conflict, and PET greater

Table 3 Descriptive statistics of post-encroachment time for different categories

Category	Sample	Mean	Median	Std. Dev
Location-1	353	2.73	2.47	1.65
Location-2	365	2.09	1.67	1.38
2 W-Location-1	288	2.69	2.39	1.65
3 W/4 W-Location-1	65	2.92	2.50	1.67
2 W-Location-2	161	2.05	1.54	1.36
3 W/4 W-Location-2	172	2.20	1.68	1.44
Bus/truck-Location-2	21	1.66	1.44	0.94
Cycle/CR-Location-2	11	1.80	1.30	1.26
< 15 kmph-Location-1	104	2.72	2.22	1.76
≥ 15 kmph-Location-1	249	2.74	2.50	1.61
< 15 kmph-Location-2	133	2.04	1.67	1.40
≥ 15 kmph-Location-2	232	2.12	1.68	1.37
Male-Location-1	221	2.65	2.17	1.71
Female-Location-1	74	2.86	2.65	1.49
Male-Location-2	275	2.11	1.67	1.42
Female-Location-2	54	1.93	1.67	1.14
< 30 yrs-Location-1	146	2.48	1.87	1.59
≥ 30 yrs-Location-1	149	2.93	2.50	1.70
< 30 yrs-Location-2	176	2.13	1.67	1.49
≥ 30 yrs-Location-2	153	2.02	1.67	1.24
Single-Location-1	295	2.71	2.34	1.66
Group-Location-1	58	2.88	2.52	1.61
Single-Location-2	330	2.08	1.67	1.39
Group-Location-2	35	2.17	1.86	1.36
Didn't wait-Location-1	282	2.66	2.31	1.62
Waited-Location-1	71	3.02	2.50	1.77
Didn't wait-Location-2	323	2.14	1.67	1.41
Waited-Location-2	42	1.70	1.40	1.04
C2M-Location-1	156	2.75	2.50	1.65
M2C-Location-1	197	2.72	2.24	1.66
C2M-Location-2	165	2.03	1.67	1.38
M2C-Location-2	200	2.15	1.67	1.38
A-Location-1	253	2.76	2.50	1.66
B-Location-1	100	2.66	2.31	1.64
A-Location-2	88	1.86	1.67	1.15
B-Location-2	231	2.22	1.67	1.48

(continued)

Table 3 (continued)

Category	Sample	Mean	Median	Std. Dev
C-Location-2	46	1.91	1.57	1.18
PPF-Location-1	279	2.94	2.50	1.72
VPF-Location-1	74	1.97	1.72	1.07
PPF-Location-2	250	2.18	1.67	1.48
VPF-Location-2	115	1.84	1.67	1.02

than 5 s indicates a safe crossing. Therefore, using these values, three categories of conflict severity were created which are as follows:

- PET < 1.5 s—dangerous conflicts
- PET ≥ 1.5 and ≤ 5 s—moderate conflicts
- PET > 5—safe crossing.

Using these severity categories', the percentage conflicts for each severity for different categories were obtained using the crosstabs function of SPSS. Crosstabs were used to prepare the stacked column charts. Figures 8, 9, and 10 show the stacked column chart that provide the split of the conflict in different severity categories for various groups. In those figures, Vad indicates location-1, and Del indicates location-2.

Conflict Severity Based on Vehicle Type and Vehicle Speed

At location-1, since the share of 2 W was higher, 82% of the conflicts is between 2 W pedestrians, and remaining are between pedestrian and 3 W/4 W; from t-test, it was found that there wasn't significant difference in PET distribution of these categories. For location-2, the conflicts due to 2 W and 3 W/4 W were nearly equal, and although the mean PET value for 2 W is less, but it wasn't significantly different as per t-test. From Fig. 8, it can be concluded that the pedestrians crossing at location-2 had higher chances of getting into a dangerous conflict since higher number of pedestrians were under the category of severe conflicts as compared to that of location-1. PET reduces with increase in speed of the vehicle generally, but here, due to lower number of samples, this thing can't be verified. The PET value for vehicles traveling at less than 15 kmph was little lower than that for vehicles traveling at a speed greater than 15 kmph, but the t-test gave a significance value of >0.05, and hence, it can be concluded that there is no significant difference between two categories of speed group.

Conflict Severity Based on Pedestrian Gender

Although the males at location-1 were subjected to smaller PETs, but there wasn't any significant difference between male and female. For locaton-2, females had lower mean PET value than males, but again, t-test provides a significance value >0.05. At location-2, higher percentage of male and females was under scenario of severe conflicts as compared to the males and females of location-1.

Table 4 T-test statistics for comparison of PET across different categories

Category-1	Category-2	*p*-value
Site-1	Site-2	**0**
2 W-Site-1	3 W/4 W-Site-1	0.566
2 W-Site-2	3 W/4 W-Site-2	0.345
Bus/truck-Site-2	Cycle/CR-Site-2	1
2 W-Site-2	Bus/truck-Site-2	1
2 W-Site-2	Cycle/CR-Site-2	1
3 W/4 W-Site-2	Bus/truck-Site-2	1
3 W/4 W-Site-2	Cycle/CR-Site-2	1
2 W-Site-1	2 W-Site-2	**0.003**
3 W/4 W-Site-1	3 W/4 W-Site-2	**0.031**
<15 kmph-Site-1	≥ 15 kmph-Site-1	0.434
<15 kmph-Site-2	≥ 15 kmph-Site-2	0.114
<15 kmph-Site-1	<15 kmph-Site-2	**0.026**
≥ 15 kmph-Site-1	≥ 15 kmph-Site-2	**0.001**
Male-Site-1	Female-Site-1	**0.041**
Male-Site-2	Female-Site-2	0.539
<30 yrs-Site-1	≥ 30 yrs-Site-1	0.145
<30 yrs-Site-2	≥ 30 yrs-Site-2	0.926
Single-Site-1	Group-Site-1	0.303
Single-Site-2	Group-Site-2	0.431
Didn't wait-Site-1	Waited-Site-1	0.176
Didn't wait-Site-2	Waited-Site-2	0.064
Male-Site-1	Male-Site-2	**0.015**
Female-Site-1	Female-Site-2	**0.007**
<30 yrs-Site-1	<30 yrs-Site-2	0.226
≥ 30 yrs-Site-1	≥ 30 yrs-Site-2	**0**
Single-Site-1	Single-Site-2	**0.001**
Group-Site-1	Group-Site-2	0.154
Didn't wait-Site-1	Didn't wait-Site-2	**0.01**
Waited-Site-1	Waited-Site-2	**0**
C2M-Site-1	M2C-Site-1	0.312
C2M-Site-2	M2C-Site-2	0.898
A-Site-1	B-Site-1	0.748
A-Site-2	B-Site-2	0.805
A-Site-2	C-Site-2	0.932
B-Site-2	C-Site-2	0.808
PPF-Site-1	VPF-Site-1	**0**

(continued)

Table 4 (continued)

Category-1	Category-2	*p*-value
PPF-Site-2	VPF-Site-2	0.412
C2M-Site-1	C2M-Site-2	**0.001**
M2C-Site-1	M2C-Site-2	**0.04**
A-Site-1	A-Site-2	**0**
B-Site-1	B-Site-2	**0.032**
PPF-Site-1	PPF-Site-2	**0**
VPF-Site-1	VPF-Site-2	0.33

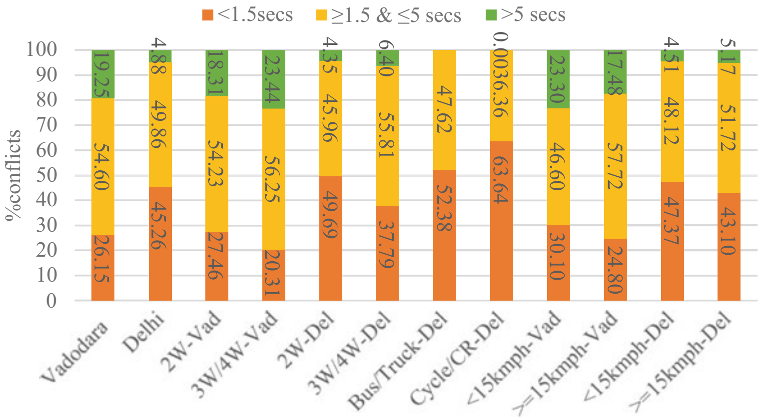


Fig. 8 Stacked column chart showing conflict severity distribution in various categories related to vehicle

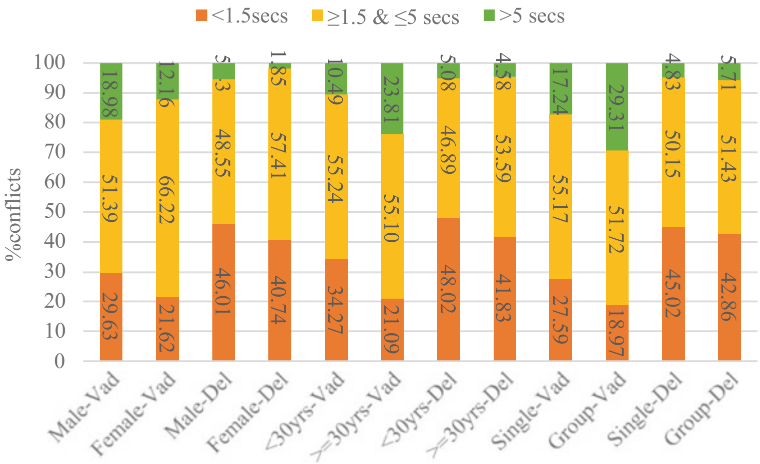


Fig. 9 Stacked column chart showing conflict severity distribution in various categories related to pedestrians

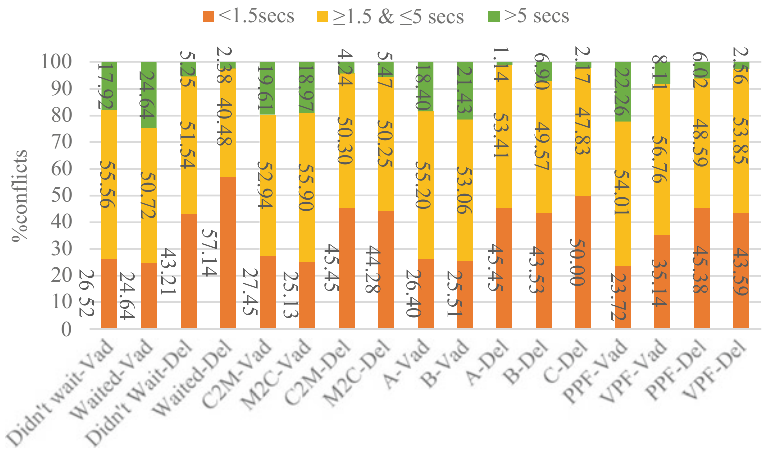


Fig. 10 Stacked column chart showing share of conflict severity categories across categories related to direction of crossing, area of crossing

Conflict Severity Based on Pedestrian Age

At locations-1 and 2, there wasn't found to be any significant difference between the pedestrian of the age below 30 years and above 30 years. From clustered bar chart, it can be observed that higher percentage of pedestrians below the age of >30 years is under severe risk as compared to pedestrians of age ≥30 years in case of location-1, whereas for location-2, there isn't much difference between the two categories.

7 Conclusion and Recommendations

This study focused on the effect of free left turning of vehicles on safety of pedestrians at signalized intersections. This study is based on two four-armed signalized intersections located in Vadodara (Location-1) and Delhi (Location-2). Number of conflicts between vehicles and pedestrians have been observed from the captured video and counted to be as 0.58 and 0.42 conflicts per pedestrians at location-1 and location-2, respectively. With respect to wait time of pedestrians, further, it is observed that 80–90% of the pedestrians didn't wait before crossing the road at these locations. This shows that the free left turn allows for less waiting time before crossing, but also, it subjects the pedestrian to high risk. For all the observed conflicts, PET values have been calculated. Considering PET values less than 1.5 s to be threshold for safety, observed conflicts have been categorized as severe. Percentage pedestrians that were categorized to be severe were 26.15 and 45.26% at locations-1 and 2, respectively. This indicates that the probability of a pedestrian getting into a severe conflict is high at location-2. Based on PET value, it is concluded that 20–50% of the pedestrians has the probability of getting to a severe conflict. Since this risk level is very high at studied locations, free left-turning movement needs to be controlled through traffic signal. Even though the free left turns are known to improve the capacity of the intersections, they were seen to be affecting the safety of pedestrians by subjecting them to severe risk interactions. Though this study is based on two locations only, some key learning is free left turning for vehicles at signalized intersections with high number of pedestrians needs to be checked carefully for both the options (i.e., free/controlled free left turning) before considering "free" as a default strategy. Depending on the pedestrian and vehicular flow, level of control on left turning movement can be managed by alternate traffic management/calming strategies, i.e., (i) proper signage (warning sign for drivers such as "turning traffic must yield to pedestrian," (ii) provision of speed hump/elevated crosswalks so as to slow down the vehicles and allow the pedestrians a larger time gap for crossing, and (iii) provision of separate pedestrian signal, allowing the pedestrians to safely cross against the left-turning vehicles.

References

1. WHO (2018) Global Status Report on Road Safety. https://www.who.int/publications/i/item/9789241565684
2. MoRTH (2018) Road Accident in India. https://morth.nic.in/road-accident-in-india
3. Gårder P (1989) Pedestrian safety at traffic signals: a study carried out with the help of a traffic conflicts technique. Accident Anal Prevent 21(5):435–444. ISSN 0001-4575. https://doi.org/10.1016/0001-4575(89)90004-3
4. Xu X, Xie S, Wong SC, Xu P, Huang H, Pei X (2016) Severity of pedestrian injuries due to traffic crashes at signalized intersections in Hong Kong: a Bayesian spatial logit model. J Adv Transp 50:2015–2028. https://doi.org/10.1002/atr.1442
5. Abdel-Aty M, Keller J (2005) Exploring the overall and specific crash severity levels at signalized intersections. Accident Anal Prevent 37(3):417–425. ISSN 0001-4575. https://doi.org/10.1016/j.aap.2004.11.002
6. Rankavat S, Tiwari G (2013) Pedestrian accident analysis in Delhi using GIS. J Eastern Asia Soc Transp Stud 10:1446–1457
7. Anjana S, Anjaneyulu MVLR (2015) Safety analysis of urban signalized intersections under mixed traffic. J Safety Res 52: 9–14. ISSN 0022–4375. https://doi.org/10.1016/j.jsr.2014.11.001
8. Hyden C (1977) A traffic conflicts technique for examining urban intersection problems. In: Proceedings of the first workshop on traffic conflicts, pp 87–89
9. Chandrappa AK, Bhattacharyya K, Maitra B (2016) Estimation of post-encroachment time and threshold wait time for pedestrians on a busy urban corridor in a heterogeneous traffic environment: an experience in Kolkata. Asian Transp Stud 4(2):421–429.https://doi.org/10.11175/eastsats.4.421
10. Kathuria A, Vedagiri P (2020) Evaluating pedestrian vehicle interaction dynamics at un-signalized intersections: a proactive approach for safety analysis. Accident Anal Prevent 134:105316. ISSN 0001-4575. https://doi.org/10.1016/j.aap.2019.105316
11. Kumar A, Ghosh I (2020) Analysis of spontaneous order of pedestrian–vehicle conflicts at signalized intersections. Transp Res Record https://doi.org/10.1177/0361198120945994
12. Chaudhari A, Gore N, Arkatkar S et al (2020) Exploring pedestrian surrogate safety measures by road geometry at midblock crosswalks: a perspectiv.... IATSS Research. https://doi.org/10.1016/j.iatssr.2020.06.001
13. Lord D (1996) Analysis of pedestrian conflicts with left-turning traffic. Transp Res Record 1538(1):61–67. https://doi.org/10.1177/0361198196153800108

14. Zhao Y, Bai Y, Yang X (2011) Modeling conflicts between Rightturn vehicles and pedestrians at signalized intersection, pp 529–536. https://doi.org/10.1061/41177(415)67
15. Ling Z, Ni Y, Li K (2012) Modeling interactions between pedestrians and right-turn vehicles at signalized intersections. In: CICTP 2012: multimodal transportation systems—convenient, safe, cost-effective, efficient, pp 2072–2084
16. Sha D, Guo Y, Ding Y (2016) Pedestrian safety diagnosis at signalizedcrosswalks using traffic conflict techniques. In: CICTP 2016, pp 1811–1823. https://doi.org/10.1061/9780784479896.166
17. Chen P, Zeng W, Yu G, Wang Y (2017) Surrogate safety analysis of pedestrian-vehicle conflict at intersections using unmanned aerial vehicle videos. J Adv Transp 2017:1–12. Article ID 5202150. https://doi.org/10.1155/2017/5202150
18. Jiang C, Qiu R, Fu T, Fu L, Xiong B, Lu Z (2020) Impact of right-turn channelization on pedestrian safety at signalized intersections. Accident Anal Prevent 136:105399. ISSN 0001-4575. https://doi.org/10.1016/j.aap.2019.105399

Pedestrian Level of Service of Un-signalised Intersections Using an Integrated Approach of Pedestrian-Risk-Perception Survey and Structural Equation Modelling

Ankit Bansal, Tripta Goyal, and Umesh Sharma

Abstract Pedestrian's wide choice of freedom in crossing manoeuvres and traffic complexity at un-signalised intersection crosswalks deems it very difficult to develop the standards for the evaluation of Pedestrian Level of Service (PLOS). To achieve the desired objective, 15 un-signalised crosswalks of Chandigarh city, India, has been studied using a questionnaire survey (pedestrian-risk-perception survey based on 5-point Likert Scale). Initially, EFA extracts four constructs from the statements namely ACCESSIBILITY, COMFORT, SAFETY and LOS which together contributes to 77.24% of total variance as confirmed by CFA model. Moreover, gender ($p = 0.049$) and age ($p = 0.018$) significantly associate with the mean perceived LOS score. Further, SEM assesses and predicts the impact of constructs (Accessibility, Comfort, and Safety) on the perceived PLOS with 93.3% accuracy. The authors thereby recommend taking up proactive measures such as making crosswalk accessible for all users and promote awareness about safety regarding traffic rules and regulations at the un-signalised crosswalks.

Keywords Accessibility · Comfort · Safety · Pedestrian level of service · Pedestrian-risk-perception · Un-signalised intersection crosswalks

1 Introduction

In the start of twentieth century, the advent of new techniques and facilities shifted the focus of planning of transportation facilities from user's view to planner's view with the greater dependency on the travel time and capacity. This increase in automobile usage resulted in urban sprawl with people moving farther away from the cities and deterioration of the local environment due to vehicular pollution, etc. Therefore, the start of twenty-first century earmarked the massive strides forward made by developed nations to make their cities pedestrian and bicyclist friendly and encourage their citizens to travel non-motorised short trips. However, most people in developing

A. Bansal (✉) · T. Goyal · U. Sharma
Department of Civil Engineering, Punjab Engineering College, Chandigarh 160012, India
e-mail: ankitbansal.phdcivil16@pec.edu.in

L. Devi et al. (eds.), *Proceedings of the Sixth International Conference of Transportation Research Group of India*, Lecture Notes in Civil Engineering 272,
https://doi.org/10.1007/978-981-19-3494-0_15

countries like India, Iran, Qatar, etc., still prefer driving and riding than walking because of weather conditions (heat, dust and air pollution), deteriorated conditions of pedestrian facilities (may be uninterrupted or interrupted like sidewalks, crosswalks, etc.), distance of a trip and many other factors unfavourably affecting the decision of road user to walk [1–3].

Thus, it can be inferred that the availability of well-connected pedestrian facilities is vital for ensuring the adequate pedestrian safety.

In developing countries, the policies (e.g. roadway design and traffic management) and fund investment aspects favour the motorised vehicle over non-motorised modes of transport (pedestrians and bicycles). Due to heterogeneity of traffic conditions and operational characteristics, it is a very complex task for traffic engineers and planners to design diverse pedestrian user facilities. Moreover, the available guidelines in pedestrian safety context do not cater to the present needs of road users [4]. Among all road users, pedestrians are most susceptible to road accidents at un-signalised crosswalks. Therefore, there is a need to study the pedestrian's perceived sense of safety and security at the crosswalks. In order to achieve this, evaluation of LOS of the existing pedestrian facilities with reliable methods and reasonably measurable parameters such as Measure of Effectiveness (MOE) is necessary.

2 Literature Review

Evaluation of the prevailing service quality of pedestrian facilities is an effective way of improving the standards of the facility as well as laid the basis for designing new facilities. Literature has depicted the service quality of existing facilities, i.e. the service offered to the user in terms of the Level of Service (LOS). The measurement of PLOS is a tool which ensures that existing land-use, motorised and non-motorised facilities are catering to the needs of pedestrian and warrants their safety. The LOS for pedestrian facilities is a function of numerous factors and difference in pedestrian perceptions [5].

At un-signalised crosswalks, pedestrian flow largely gets affected by motor vehicles; due to which, these are considered to be more critical locations [6]. Studies found pedestrian LOS of un-signalised crosswalks on the basis of pedestrian waiting time, holding area and pedestrian flow characteristics adopting gap acceptance as the key index measure by most of the researchers. Due to continuous movement of vehicular flow at these crosswalks, pedestrians either wait for a long time to get the required gap or cross the road unsafely with the gaps available in traffic [7–10].

The factors influencing LOS vary with respect to the type of facility available and their location. Highway Capacity Manual (HCM 2010) and Indo-HCM (2017) analysed LOS of un-signalised intersections and defined six service categories (A–F) based upon pedestrian gap size [11, 12]. Several qualitative and quantitative measures were used to identify the other possible environmental factors affecting the PLOS of un-signalised intersection crosswalks [13, 14]. Studies have also been carried out for pedestrian LOS at un-signalised intersections in view of the pedestrian's perception of comfort and safety and psychological limitation [15–17].

Also, in case of roundabouts, factors significantly affecting PLOS were vehicle volume, vehicle speed, carriageway width, pedestrian refuge, crosswalk marking condition, crosswalk surface condition and lighting at crossing area [18]. In a recent study, Sahani and Bhuyan proposed Pedestrian Level of Service (PLOS) model using ridge regression technique on pedestrian real-time perceived satisfaction ratings obtained using pedestrian delay as well as pedestrian gap acceptance. Results indicated that delay value of less than 15 s provided PLOS score $\leq$ 1.5 that represents PLOS 'A' category and delay more than 60 s with PLOS score value of > 5.5 represents PLOS 'F' category [17].

Although un-signalised intersections are very common in developing countries, there are not plentiful of studies for setting out the standards of pedestrian LOS of the unsignalized intersection crosswalks with heterogeneous traffic flow conditions. In this context, due to complicated pedestrian movements, it is necessary to evaluate pedestrian sense of satisfaction and develop a proper methodology to estimate the service quality of un-signalised intersection crosswalks.

3 Methodology

3.1 Study Area

The present study aims at determining the pedestrian LOS of signalized crosswalk facilities (interrupted pedestrian facilities) in Chandigarh city. Pedestrians constitute the largest single group of people dying in road accidents in Chandigarh. During the past five years (2014–2018), out of total 615 lives claimed during road accidents, 196 were pedestrians accounting for 31.87% of the total deaths [19]. This indicates that most of the fatal accidents took place at crosswalks which makes it imperative to analyse the pedestrian LOS of un-signalised facilities so as to ascertain the safety of pedestrians. Total 15 un-signalised crosswalks situated along the prominent arterial road of Chandigarh, i.e. **Madhya Marg** (also known as V-2 roads), were studied, and data of 300 pedestrians (bi-directional flow) with different behavioural characteristics were collected and analysed using statistical software SPSS and AMOS (built-in software of SPSS). The unsignalized crosswalks were designated from U1 to U15.

The present study is Exploratory and Descriptive in nature which includes the exploration of factors/latent variables affecting the LOS of crosswalks, determining the Perceived LOS of the crosswalks on the basis of questionnaire survey.

3.2 Data Collection Through Questionnaire Survey (Pedestrian Opinion Survey)

For formulating the Perceived PLOS model of the crosswalks, Pedestrian Opinion survey was carried out with the help of questionnaire using roadside interview method. Total 300 responses were collected in the months of December 2020 and January 2021 on weekdays. In the absence of the availability of a well-structured and established scale, the questionnaire was self-framed keeping in mind the need and objectives of the study.

The questionnaire includes two parts (refer Appendix 1): Part I comprises profile of respondents based on the socio-demographic variables like gender, age, education, city, use of crosswalks and purpose of trip and part II includes total 16 statements/items/variables that give insight about the pedestrian's perception about the location and condition of the crosswalks, comfort and convenience to the pedestrians while using the crosswalk, safety aspect of the pedestrians and crosswalks service quality.

The questionnaire was developed based on previous studies like Lee et al., Shi et al., Bian et al. and Papadimitriou et al. [16, 20–22]. The statements/questions were based on the Likert scale (Point Value System) (Strongly Disagree-1, Disagree-2, Neutral-3, Agree-4 and Strongly Agree-5). Questionnaire data was further analysed using descriptive and inferential statistics to determine the pedestrian perceived LOS of crosswalks. Based on the collected data, PLOS threshold range was defined and the perceived LOS model of the crosswalk was formulated.

Threshold Values and Limits of Perceived LOS. The 5-point Likert scale questionnaire (where '5' represents the 'strongly agree' and '1' represents the 'strongly disagree' category) was converted into six categories of perceived LOS from A to F using equal interval [(5–1)/6 = 0.67]. 'A' corresponds to strongly agree and represents the excellent LOS, whereas 'F' represents the very poor LOS, i.e. strongly disagree category. The ranges so obtained are presented in Table 1. These six LOS ranges were further validated through single LOS grade representation (Table 2) [23].

The frequently single-value measures used to represent a distribution are either the mean or the mode. The mode is appealing as it embodies the common LOS response, but at the same time could give irregular results in case of different frequent

Table 1 Ranges for perceived LOS

Range	LOS	Characteristic
$1.00 \leq X \leq 1.67$	F	Very poor
$1.67 < X \leq 2.33$	E	Poor
$2.33 < X \leq 3.00$	D	Fair
$3.00 < X \leq 3.67$	C	Good
$3.67 < X \leq 4.33$	B	Very good
$4.33 < X \leq 5.00$	A	Excellent

Table 2 Representation of results by single LOS grade: distributions and mean of LOS

Results by distribution						
LOS	1.00	1.67	2.33	3.00	3.67	4.33
F	59%	13%	NA	NA	NA	NA
E	31%	57%	15%	NA	NA	NA
D	10%	30%	54%	21%	NA	NA
C	NA	NA	31%	53%	32%	10%
B	NA	NA	NA	26%	51%	31%
A	NA	NA	NA	NA	17%	59%
1	F	E	D	C	C	B
2	F	F	D	C	B	B
3	F	E	D	C	B	A

Note NA = Not available, Mean for distributions (1–4.33), respectively: 1.34, 1.78, 2.44, 3.03, 3.57, 3.99

responses; due to which. single LOS result cannot be reported. There is no such issue with the mean distribution. Although in few cases, the extreme end results only can be obtained in case of total consensus in responses. It can be observed in Table 2 that for Distribution 1, 59% of people choose LOS A (8th Row, 7th Column) and the mean is 3.99 (0.59 * 4.33 + 0.31 * 3.67 + 0.1 * 3.00), but as per ranges defined in Table 1, LOS comes out to be B in spite of LOS A; therefore, ranges are re-defined accordingly based on the mean of distribution and converting the mean values into letter grades.

- The row representing LOS 1 has been converted into letter grades using the similar mean of distribution as of values of LOS A through LOS F (Tables 2 and 3). However, it has been observed that for distributions 4.33 and 3.67, the desired results could not be obtained, i.e. LOS A and LOS B, respectively.
- To overcome the above-stated discrepancy, shifted set of thresholds (higher ranges for LOS A and F, shown in Table 3) have been used to convert the mean values to a letter grade as represented into LOS 2. Therefore, LOS A ranges from a

Table 3 Representation of results by single LOS grade: LOS mean value threshold schemes

Numeric value	LOS	LOS 1, straight thresholds	LOS 2, thresholds shifted to mid-points	LOS 3, compressed ranges
1.00	F	$1.00 \le X \le 1.67$	$1.00 \le X \le 2.00$	$1.00 \le X \le 1.75$
1.67	E	$1.67 < X \le 2.33$	$2.00 < X \le 2.50$	$1.75 < X \le 2.25$
2.33	D	$2.33 < X \le 3.00$	$2.50 < X \le 3.00$	$2.25 < X \le 2.75$
3.00	C	$3.00 < X \le 3.67$	$3.00 < X \le 3.50$	$2.75 < X \le 3.25$
3.67	B	$3.67 < X \le 4.33$	$3.50 < X \le 4.00$	$3.25 < X \le 3.75$
4.33	A	$4.33 < X \le 5.00$	$4.00 < X \le 5.00$	$3.75 < X \le 5.00$

Table 4 Final LOS mean value threshold range

Range	LOS
$1.00 \leq X \leq 1.75$	F
$1.75 < X \leq 2.25$	E
$2.25 < X \leq 2.75$	D
$2.75 < X \leq 3.25$	C
$3.25 < X \leq 3.75$	B
$3.75 < X \leq 5.00$	A

mean of 4.0–5.0 and LOS F ranges from a mean of 1.0–2.0. Regrettably, this tactic also results in inconsistencies in getting the desired LOS for different mean distributions (LOS F for distribution 1.67, in spite of the anticipated LOS E and LOS B instead of LOS A for distribution 4.33).

- The row labelling LOS 3 establishes the results with a broader threshold range for distribution 4.33 (for the conversion into letter grades) to accommodate the large number of responses at the extreme. This results in squeezing the thresholds for LOS B to LOS F (1.00–3.75) and having the much wider range for LOS A (3.75–5.0).

The LOS 3 threshold schemes depicted in Table 3 were further confirmed using the pedestrian field data. Results obtained for the mean LOS values for field data represent the desired distributions and produced a reasonable range of LOS A through LOS F. Therefore, the final LOS mean value threshold range along with the Perceived LOS (U) categories (letter grades) adopted for modelling the LOS of the crosswalks is depicted in Table 4.

Sampling of Questionnaire Data at Crosswalks. For un-signalised crosswalks, total 300 (15 * 20) pedestrian responses were collected for total population (P) of 1166 pedestrians. The value of statistical sample size (tested at 95% confidence interval with a margin error (e) of 5% and a response rate of 50%) using the formula given by Nargundkar [24] (Eqs. 1 and 2) came out to be 289 pedestrians which is less than real-time respondents' size (300 pedestrians). Hence, the collected sample for developing PLOS models at un-signalised crosswalks is significant.

$$n = p(1 - p)\frac{z^2}{e^2} \tag{1}$$

where n = Sample Size, p = Percentage picking a choice, expressed as decimal, Z = Z value from the standard normal distribution for 95% confidence level, e = Error Term.

For finite population, correction is to be applied to calculated sample size (n). The corrected sample size (N) is given as per 'Eq. 2'.

$$N = \frac{n}{1 + \frac{n-1}{P}} \quad (2)$$

where N = Corrected Sample Size, P = Population Size.

4 Results and Discussion

For formulating the Perceived Pedestrian LOS model of the un-signalised crosswalks, the questionnaire survey was distributed among 300 (15 * 20) pedestrians. Validation of questionnaire was done with the help of CFA, and SEM was applied to predict the perceived pedestrian level of service of the crosswalk.

4.1 Profile of the Respondents of Questionnaire Survey

This section provides a general background (city, gender, age, education, purpose of trip and crosswalk usage-wise distribution) of the 300 respondents surveyed on the basis of the socio-demographic variables. The analysis of the data was carried out using Descriptive analysis (frequencies and percentages) as shown in Fig. 1.

It was found that 63.3% of the respondents belonged to Chandigarh city while 36.7% belonged to cities other than Chandigarh. Among all respondents, 54.7% were males, while 45.3% were females. Age-wise distribution indicates that 40.3% of the respondents were in the age group of 18–30 years and 9.7% above 60 years. The education-wise analysis shows that the 38.0 and 34.3% of the respondents possessed the education up to Senior Secondary and Graduate level, respectively, followed by 14.0% Post Graduate, 10.7% Matric and only 3.0% belonged to others category (which may include illiteracy, etc.). The purpose of trip-wise distribution shows that 39.7% of the respondents were using the crosswalk for the work purpose, followed by 27.3% for health, 19.7% for education, 11.0% for recreation and remaining 2.3% for shopping. Crosswalk-Usage distribution indicates that 49.7% respondents were observed to frequently use the crosswalks whereas 25.3% pedestrians generally use the crosswalks. 21.0% pedestrians had rarely used the crosswalks, and rest 4.0% had never used the crosswalks before.

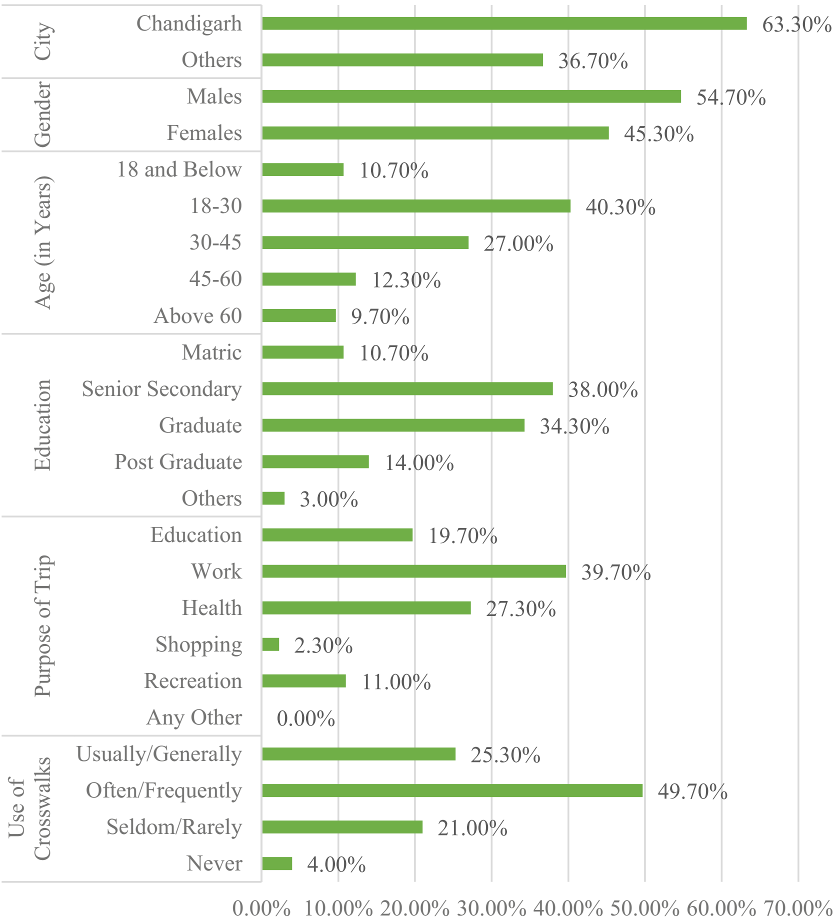


Fig. 1 Profile of respondents based on socio-demographic variables at un-signalised crosswalks

4.2 *Descriptive Analysis of Statements/Items of Questionnaire*

The responses of 300 pedestrians using the un-signalised crosswalks were compiled and Descriptive analysis (mean, median, mode, std. deviation, variance, skewness and kurtosis) was carried out (refer Appendix 2). The mean of the responses varied between 2 and 4, with the maximum mean of 3.48 for the statement q6 (Adequate Use of Pedestrian Refuge Islands/Central Islands/Medians) and minimum mean of 2.34 for the statement q10 {Suitable for disabled pedestrians (persons using wheelchair)}.

The Cronbach's Alpha coefficient value for the scale with 16 statements was found to be 0.832. This indicates the high level of internal consistency in the statements and reliability of scale for further analysis. The skewness values indicate that all the responses were close to be symmetrically distributed. Kurtosis for all statements/variables was below 1 and Skewness for all the variables was less than 0.5, far smaller than the lower bound of four or five. Thus, both Kurtosis and Skewness provide indication that the data is normally distributed.

4.3 Exploratory Factor Analysis for Un-signalised Crosswalks Questionnaire

After checking the normality, Exploratory Factor Analysis (EFA) was employed to reduce data to a smaller set of summary variables and to explore the underlying theoretical structure of the phenomena (Table 5).

The Bartlett 's Test of Sphericity value was 1470.587 with significance level <0.000, which indicate that correlation matrix has significant correlation among other variables. Also, the Kaiser–Meyer–Olkin (KMO) was found to be 0.807 (nearer to 1), which justifies the appropriateness of factor analysis. Principal Component Analysis was applied that extracted four factors whose Eigen values are greater than 1 and together contribute to 77.24% of total variance. Further, rotation by Varimax with Kaiser Normalization was carried out to distribute the variance evenly among the components.

For better data reduction, those variables that have the factor loadings more than 0.50 were considered under each factor. Two variables q10 {**Suitable for disabled pedestrians (persons using wheelchair**)} and q14 {**No Illicit Crossing (disregarding traffic rules/regulations) at the time of hurry}** were dropped out because of less than 0.4 factor loading. These variables are statistically independent and cannot be combined with other variables or factors. Therefore, the four constructs/factors extracted are named as ACCESSIBILITY, COMFORT, SAFETY and LOS. These are termed as the measures of effectiveness with high reliabilities (Cronbach's $\alpha > 0.70$) which was further validated by CFA.

4.4 Confirmatory Factor Analysis for Un-signalised Crosswalks Questionnaire

The results obtained by EFA were then imported in AMOS (built-in software of SPSS) to further check the reliability and validity of the factor structure of the perceived LOS. The relationship between observed variables (items/statements) and latent constructs (also known as CFA model) is represented diagrammatically as shown in Fig. 2. It represents the correlation between constructs and factors (also known as CFA loadings) as well as covariance's for all the constructs. The chi-square value was observed to be $\{\chi^2(72) = 83.657\}$ with probability level, $P = 0.164$. It was observed that critical ratio obtained for all the items in a construct was greater than

Table 5 Exploratory factor analysis for variables of un-signalised crosswalks

Factors	Statements		Component			
			1	2	3	4
Accessibility	Acc_1	Appropriate location of crosswalk (i.e. near to bus stop or any transit facility) q1	0.840			
	Acc_2	Appropriate waiting area/space with proper access to the crosswalk q2	0.836			
	Acc_3	Proper sidewalks at entry and exit of crosswalks for easy accessibility to and from crosswalk q3	0.814			
Comfort	Comfort_1	Good surface condition of the crosswalk q4		0.815		
	Comfort_2	No inconvenience while crossing the road due to presence of obstruction or any obstacle q5		0.822		
	Comfort_3	Adequate use of pedestrian refuge islands/central islands/medians q6		0.827		
	Comfort_4	Proper provision of signs and markings for the comfort of pedestrians while crossing q7		0.759		
	Comfort_5	Adequate visibility (street lighting) for the convenience of pedestrians to cross during night q8		0.745		
	Comfort_6	Proper provision of curb ramps and guard rails at entry/exit/medians q9		0.728		
Safety	Safety_1	Smooth and orderly movement of traffic q11			0.783	
	Safety_2	Availability of suitable and appropriate gaps in vehicular stream q12			0.894	
	Safety_3	Least interaction/interference with other non-motorised traffic (cyclists, etc.) q13			0.895	
LOS	LOS_1	Satisfactory usability of the crosswalk q15				0.776

(continued)

Table 5 (continued)

Factors	Statements		Component			
			1	2	3	4
	LOS_2	Overall ease in crossing the crosswalk q16				0.782
Extraction Method: Principal Component Analysis Rotation Method: Varimax with Kaiser Normalization		% Variance explained	33.619	22.789	13.516	7.317
		Eigen value	5.018	3.316	1.907	1.104
		Cronbach's alpha	0.772	0.874	0.827	0.854
		Kaiser–Meyer–Olkin measure of sampling adequacy (KMO) = 0.807, Bartlett's Test of Sphericity Chi-square = 1470.587, $p < 0.000$				

Note Statements q10 {**Suitable for disabled pedestrians (persons using wheelchair)**} and q14 {**No Illicit Crossing (disregarding traffic rules/regulations) at the time of hurry}** are removed because of less than 0.4 factor loading

2 and all item results were significant as p-value < 0.05. The validity and reliability of CFA results are shown in Table 6.

Composite Reliability (CR), i.e. internal consistency between the constructs, was higher than 0.7 for all the constructs [25]. All constructs were observed to have AVE > 0.5. CR values were also greater than AVE for all constructs. Thus, the convergent validity rules are fulfilled [25]. It was also observed that AVE was greater than ASV and MSV, therefore fulfilling the discriminant validity standards. Hence, all four latent constructs had no reliability and validity issues.

Moreover, the correlation between the constructs was also found to be significant. Further, the model fit indices were indicated in Table 7. The summary of all model fit measures revealed the goodness of fit in the CFA model [26, 27]. Therefore, after CFA, it was evident that Accessibility, Comfort, Safety and LOS are the measures of effectiveness of perceived LOS. Before constructing the conceptual model for perceived LOS, mean of the extracted factors and its association with the socio-demographics were checked out.

4.5 Mean Score of the Extracted Factors of Perceived LOS of Un-signalised Crosswalks

The mean score was derived by taking the grand mean value of variables clubbed in each factor. It was found that for Factor 2, the mean score was highest, i.e. 3.41. This indicates that the factor named 'Comfort' is the influencing factor with maximum weightage. This factor was followed by the mean score 3.30 of Factor 3, i.e. 'Safety', which shows that while crossing, pedestrians prefer safety. The third factor influencing LOS was Factor 1: 'Accessibility' with a mean score of 3.11. Factor 4, i.e. LOS with a mean score of 3.27 was further modelled using the three factors as the statements directly give indication towards perceived LOS [LOS(*Per*)]. In the end, overall mean of accessibility, comfort and safety was calculated which comes out to be 3.27, similar to that of Factor 4 mean.

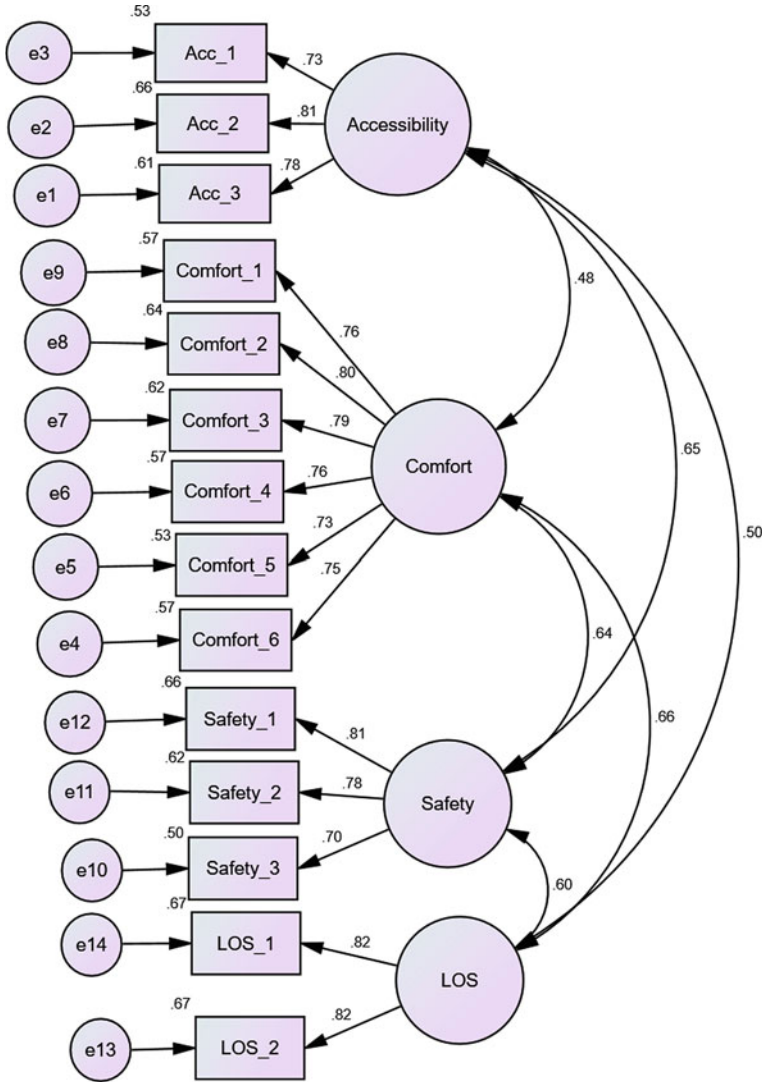


Fig. 2 Confirmatory factor analysis for crosswalks questionnaire

Table 6 Validity and reliability results of CFA

	CR	AVE	MSV	ASV	Accessibility	Comfort	Safety	LOS
Accessibility	0.818	0.601	0.421	0.302	**0.775**			
Comfort	0.893	0.583	0.438	0.360	0.484***	**0.764**		
Safety	0.811	0.590	0.421	0.397	0.649***	0.639***	**0.768**	
LOS	0.804	0.672	0.438	0.350	0.500***	0.662***	0.602***	**0.820**

Note CR Composite reliability, *AVE* Average variance explained, *MSV* Maximum shared variance, *ASV* Average shared variance. significance of co-relations † $p < 0.100$, * $p < 0.050$, ** $p < 0.010$, *** $p < 0.001$[26]

Table 7 Model fit indices of CFA model for un-signalised crosswalks

Measure	Current model	Recommended values	Threshold limit	Interpretation
Chi-square/df (CMIN/DF)	1.162	>1	1–3	Excellent
P-value for the model	0.164	>0.05	>0.05	Excellent
CFI	0.994	>0.90	>0.80	Excellent
GFI	0.963	>0.95	>0.95	Excellent
AGFI	0.945	>0.80	>0.80	Excellent
SRMR	0.020	<0.08	<0.08	Excellent
RMSEA	0.023	<0.05	<0.05	Excellent
PCLOSE	0.993	>0.05	>0.05	Excellent

4.6 Association Between Socio-demographics and Mean Perceived LOS Score of Un-signalised Crosswalks

The association between mean score of perceived LOS and various socio-demographics of 300 respondents such as City, Gender, Age, Education, Purpose of Trip and Crosswalk Usage. Results of T-test reveal that city has no significant association with the mean perceived LOS score ($p = 0.576$, $p > 0.05$), whereas gender is significantly associated with the mean perceived LOS score ($p = 0.049$, $p < 0.05$). The ANOVA statistics indicate that age (F-value $= 3.017$, $p = 0.018$, < 0.05) is significantly associated with the mean perceived LOS score. However, no significant difference exists in the mean perceived LOS value for education ($p = 0.709$, > 0.05), purpose of trip ($p = 0.178$, > 0.05) and use of crosswalks ($p = 0.255$, > 0.05).

4.7 Structural Equation Modelling for Perceived LOS at Un-signalised Crosswalks

After CFA, a conceptual model assessing the impact of constructs of Accessibility, Comfort and Safety on the perceived LOS and predicting the perceived LOS scores of respondents was formulated using SEM in AMOS. The estimates of the model are further depicted in Table 8, and the model is represented diagrammatically in Fig. 3, which represents both CFA and path analysis results. The impact of constructs on the perceived LOS was given by standardised regression weights that were represented on the arrows emerging from Accessibility, Comfort and Safety towards perceived LOS. The regression equation for predicting the pedestrian perceived LOS is given in 'Eq. 3'.

$$\text{LOS(Per)} = 0.208 * \text{Accessibility} + 0.443 * \text{Comfort} + 0.354 * \text{Safety} \tag{3}$$

Further, the assumptions were checked using the residual statistics (Table 9) and regression plots for accurately predicting the perceived PLOS. The partial regression plots and a plot of studentized residuals against the predicted values (Fig. 4) assessed the linearity between the pedestrian perceived LOS and independent variables namely accessibility, comfort and safety. This indicates that all constructs have linear relationship with the perceived LOS. Durbin-Watson statistics value of 2.104 indicates that there is independence of residuals. The visual inspection of a plot of studentized residuals against the unstandardized predicted values indicates the homoscedasticity in the data. VIF values for all the independent variables were found to be less than 2; hence, no multicollinearity was observed. Moreover, LOS predicted using SEM Model has a mean value of 3.27 which is equal to the overall mean of 3.27, with no studentized deleted residuals exceeding $\pm$ 3 standard deviations, no leverage values greater than 0.2, and less than 1 value for Cook's distance [28].

A Normal P-P plot of standardised residual and Normal Q-Q plot of studentized residuals are shown in Figs. 5 and 6, which indicates that the assumption of normality is met. The multiple regression model significantly predicts the pedestrian's perceived LOS, F (3296) = 392.588, $p = 0.000$, adj. $R^2 = 0.917$, an extremely large effect. All three variables add statistically significantly to the prediction, $p < 0.05$.

All three factors are linearly related with the lateral construct Perceived LOS which reflects satisfactory usability of the crosswalk with ease in crossing manoeuvres. The factor named 'Comfort' is the most influencing factor for determining the LOS of crosswalks, as its coefficient is highest. This shows that while using the crosswalk, pedestrians feel comfortable and, therefore, give maximum weightage to the factors that ensure their comfortability. The crosswalks have good surface condition, adequate refuge islands and proper signs and markings to facilitate pedestrians while crossing. For factor safety, the results indicate that pedestrians feel safe while crossing.

Although the crosswalks are unsignalized, pedestrians still able to find suitable gap in between vehicular stream with least interference and interaction from motorised

Table 8 Estimates of SEM model for un-signalised crosswalks

Items ← constructs	Unstandardized estimates	Standardized regression weights*	S.E**	CR***	*p*****
LOS ← Accessibility	**0.208**	0.136	0.091	2.286	0.035
LOS ← Comfort	**0.443**	0.453	0.058	7.638	0.000
LOS ← Safety	**0.354**	0.224	0.074	4.784	0.000
Acc_1 ← Accessibility	1.000	0.731			
Acc_2 ← Accessibility	1.185	0.810	0.097	12.171	0.000
Acc_3 ← Accessibility	1.247	0.782	0.104	11.940	0.000
Comfort_1 ← Comfort	1.000	0.784			
Comfort_2 ← Comfort	1.118	0.801	0.080	14.041	0.000
Comfort_3 ← Comfort	1.077	0.787	0.078	13.778	0.000
Comfort_4 ← Comfort	0.982	0.757	0.075	13.186	0.000
Comfort_5 ← Comfort	0.963	0.727	0.076	12.607	0.000
Comfort_6 ← Comfort	0.918	0.752	0.070	13.102	0.000
Safety_1 ← Safety	1.209	0.812	0.101	11.997	0.000
Safety_2 ← Safety	1.289	0.784	0.110	11.739	0.000
Safety_3 ← Safety	1.000	0.704			
LOS_1 ← LOS	1.000	0.821			
LOS_2 ← LOS	1.000	0.818			

Note * Standardised loading estimates calculated from CFA using AMOS software. ** Standard Error *** Critical Ratio should be greater than 2. *** *p-value should be <0.05

and non-motorised users. The third factor influencing LOS was ‘Accessibility’. The results indicate that the pedestrian face difficulty in accessing the crosswalks. This is due to improper access from sidewalks to crosswalks with inadequate space area at crosswalk ends. Also, the bus stops or transit stops are at greater distance from the designated crosswalks, thus limiting the use of crosswalk space by pedestrians.

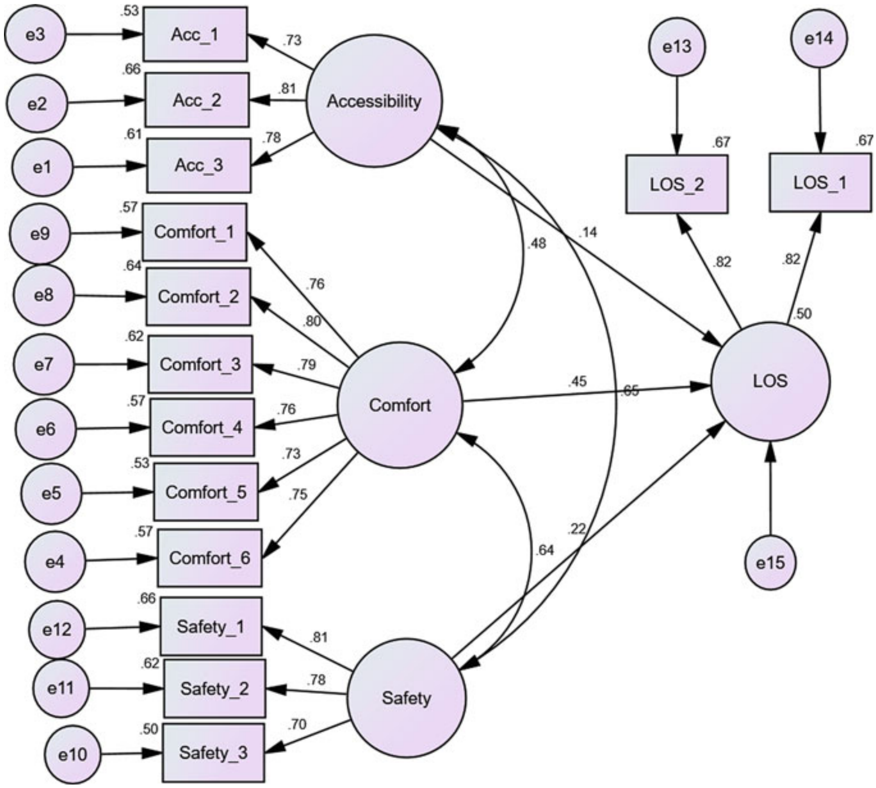


Fig. 3 SEM model for perceived LOS at un-signalised crosswalks

Table 9 Residual statistics of SEM model for un-signalised crosswalks

Statistics	Minimum	Maximum	Mean	Std. Deviation	*N*
Predicted value	2.086	4.537	3.271	0.579	300
Std. predicted value	−2.218	2.015	0.000	1.000	300
Std. error of predicted value	0.017	0.064	0.032	0.010	300
Residual	−0.676	0.805	0.000	0.290	300
Std. residual	−2.318	2.760	0.000	0.995	300
Studentized residual	−2.336	2.775	0.000	1.001	300
Cook’s distance	0.000	0.056	0.003	0.006	300
Centred leverage value	0.000	0.044	0.010	0.009	300

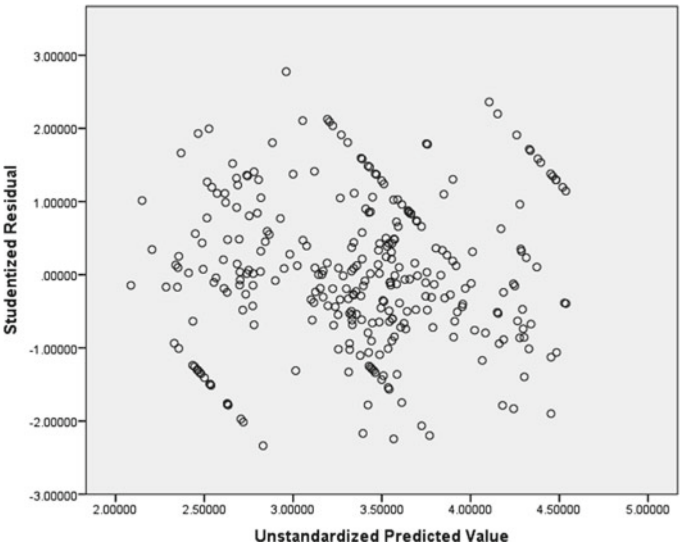
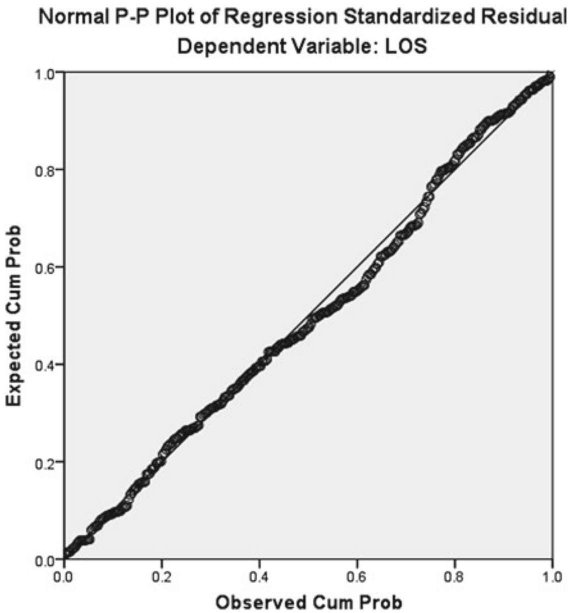


Fig. 4 Plot between studentized residuals and unstandardized predicted values

Fig. 5 Normal P-P plot of regression standardised residual



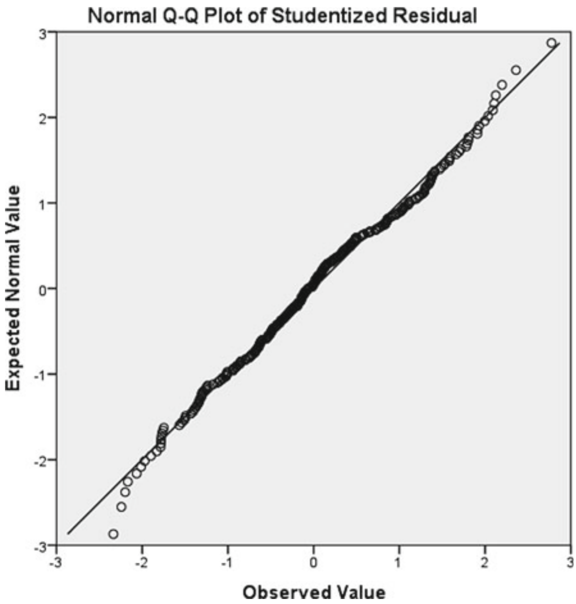


Fig. 6 Normal Q-Q plot of studentized residuals

4.8 Results of Un-signalised Questionnaire Analysis

After analysing the questionnaire data, three different sets of perceived LOS scores (respondent-wise) were obtained. First set represents the score obtained from the mean of the statements categorised under factor LOS. Second denotes the overall mean of statements of accessibility, comfort and safety. Third signifies the predicted scores obtained using SEM. Further, these scores were averaged crosswalk-wise (20 respondents per crosswalk) to obtain the crosswalk-wise LOS scores (refer Table 10). The crosswalk-wise LOS scores were further categorised into six different categories from 'LOS A to LOS F' on the basis of defined prescribed limits and threshold ranges obtained (Table 4). SEM results stipulated that accessibility, comfort and safety of pedestrians were the effective measures of Pedestrian Perceived LOS and predicted perceived LOS with 93.3% accuracy (14 cases were correctly predicted out of 15, i.e. 14/15 * 100). Further, the mean perceived LOS scores for each crosswalk were further converted into letter grades representing perceived LOS (U).

5 Conclusion and Recommendations

At 15 un-signalised crosswalks, total 300 pedestrian's data (bi-directional flow) were observed and analysed for determining the perceived pedestrian LOS of crosswalks. The descriptive analysis of the responses of the statements of questionnaire indicates that crosswalks are not suitable for disabled pedestrians because of inadequate space

Table 10 Comparison among different perceived LOS scores of un-signalised crosswalks

Intersection	Name	Designation	LOS 1 (from statements mean)	LOS 2 (Overall Mean)	Predicted LOS 3 (from SEM results)	Perceived LOS
9-10-16-17	MU12	U1	3.63	3.61	3.64	B
	MU13	U2	3.50	3.45	3.48	B
	MU14	U3	3.36	3.37	3.35	B
	MU15	U4	3.48	3.47	3.49	B
10-11-15-16	MU8	U5	3.43	3.40	3.43	B
	MU9	U6	3.38	3.38	3.40	B
	MU10	U7	3.30	3.31	3.29	B
	MU11	U8	3.28	3.29	3.32	B
11-12-14-15	MU4	U9	2.76	2.76	2.69	C
	MU5	U10	3.03	3.01	3.04	C
	MU6	U11	3.46	3.44	3.47	B
	MU7	U12	3.48	3.47	3.49	B
12-14-Khuda Jassu	MU1	U13	2.94	2.96	2.92	C
	MU2	U14	3.13	3.03	3.08	C
	MU3	U15	2.95	3.05	2.94	C

for a wheelchair to cross and absence of curb ramps for accessing the crosswalk. Also, pedestrians engage in illicit crossing behaviour in order to reach their respective destinations as soon as possible. Also, no significant association was observed between respondent's socio-demographics (city, education, purpose of trip, use of crosswalks) and mean perceived LOS score. However, gender ($p = 0.049$, <0 .05) and age ($p = 0.018$, < 0.05) were significantly associated with the mean perceived LOS [2, 29, 30]. Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA) of the questionnaire extracted four factors from the statements namely ACCESSIBILITY, COMFORT, SAFETY and LOS which together contributed 77.24% of total variance. Modelling of Perceived Pedestrian LOS using Structural Equation Modeling (SEM) predicts three constructs such as Accessibility, Comfort and Safety as the measures of effectiveness with 93.3% accuracy.

The findings of the study inculcated the need to make the crosswalks suitable for all kind of pedestrian users (especially disabled pedestrians using wheelchair). For preventing the pedestrians to engage in illicit crossing behaviour, a traffic calming campaign should be promoted to advise drivers to slow down or yield to pedestrians (the vulnerable road users) and 4 E's (Education, enforcement and training, engineering, and enthusiasm) should be put into practice. Also, educate the pedestrians about the ill effects of erratic crossing and rolling behaviour while crossing. A pedestrian signal along with the countdown timers should be installed at un-signalised crosswalks to ensure the safety of pedestrians by minimising the pedestrian–vehicle interaction.

Acknowledgements The authors would like to thank the Chandigarh Administration for their help and support in the collection of data.

Appendix 1: Un-signalised Crosswalk Questionnaire

Location:	Questionnaire No.

Section A: Respondent's Profile

1. Name (optional): ____

2. City: Chandigarh ☐ Other ☐

3. Age: 18 Years and Below ☐ 18-30 Years ☐
30-45 Years ☐ 45-60 Years ☐
Above 60 Years ☐

4. Gender: Male ☐ Female ☐

5. Education: Matric ☐ Senior Secondary ☐
Graduate ☐ Post-Graduate ☐
Others (Specify) ___________________

6. Purpose of Trip: Education ☐ Work ☐
Health ☐ Shopping ☐
Recreation ☐ Any Other ☐

7. Use of Crosswalks/Zebra Crossing: Usually/Generally ☐ Often/Frequently ☐
Seldom/Rarely ☐ Never ☐

Section B: Analysis of Crosswalk (Zebra Crossing)

Listed below are the parameters used to assess the quality of service offered to the pedestrians by crosswalk/zebra crossing of the street. Please rate your agreement/disagreement on a scale of 1-5.

S. No	Particulars	1 **Strongly Disagree**	2 **Disagree**	3 **Neutral**	4 **Agree**	5 **Strongly Agree**
1.	Appropriate Location of Crosswalk (i.e., Near to Bus Stop or Any Transit Facility)					
2.	Appropriate Waiting Area/Space with Proper Access to the Crosswalk					
3.	Proper sidewalks at Entry and Exit of Crosswalks for Easy Accessibility to and from Crosswalk					
4.	Good Surface Condition of the Crosswalk					
5.	No Inconvenience while Crossing the Road due to Presence of Obstruction or Any Obstacle					
6.	Adequate Use of Pedestrian Refuge Islands/Central Islands/Medians					
7.	Proper Provision of Signs and Markings for the Comfort of Pedestrians while Crossing					
8.	Adequate Visibility (Street Lighting) for the Convenience of Pedestrians to Cross during Night					
9.	Proper Provision of Curb Ramps and Guard Rails at Entry/Exit/Medians					
10.	Suitable for disabled pedestrians (persons using wheelchair etc.)					
11.	Smooth and Orderly Movement of Traffic					
12.	Availability of Suitable and Appropriate Gaps in Vehicular Stream					
13.	Least Interaction/Interference with Other Non-Motorized Traffic (Cyclists etc.)					
14.	No Illicit Crossing (disregarding traffic rules/regulations) at the time of hurry					
15.	Satisfactory Usability of the Crosswalk					
16.	Overall Ease in Crossing using the Crosswalk					

Any Suggestions:

…………………………………..…ThankYou……………………………….

Appendix 2: Descriptive Analysis of Statements of Perceived LOS at Un-signalised Crosswalks

Statements	N*	Mean	Std. Dev	Variance	Skewness	Kurtosis
Appropriate location of crosswalk (i.e. near to bus stop or any transit facility) q1	300	3.16	0.789	0.623	0.427	−0.339
Appropriate waiting area/space with proper access to the crosswalk q2	300	3.12	0.836	0.699	0.193	−0.596
Proper sidewalks at entry and exit of crosswalks for easy accessibility to and from crosswalk q3	300	3.05	1.008	1.015	0.243	−0.549
Good surface condition of the crosswalk q4	300	3.41	0.952	0.906	−0.030	−0.915
No inconvenience while crossing the road due to presence of obstruction or any obstacle q5	300	3.44	0.951	0.905	−0.209	−0.585
Adequate use of pedestrian refuge islands/central islands/medians q6	300	3.48	1.003	1.006	−0.015	−1.067
Proper provision of signs and markings for the comfort of pedestrians while crossing q7	300	3.39	0.956	0.913	0.025	−0.928
Adequate visibility (street lighting) for the convenience of pedestrians to cross during night q8	300	3.38	0.938	0.879	0.072	−0.243
Proper provision of curb ramps and guard rails at entry/exit/medians q9	300	3.36	0.915	0.837	0.047	−0.243
Suitable for disabled pedestrians (persons using wheelchair) q10	300	2.34	0.955	0.912	0.324	−0.094
Smooth and orderly movement of traffic q11	300	3.30	0.893	0.797	−0.157	−0.397
Availability of suitable and appropriate gaps in vehicular stream q12	300	3.31	1.026	1.053	−0.204	−0.675
Least interaction/interference with other non-motorised traffic (cyclists, etc.) q13	300	3.29	1.090	1.187	−0.073	−0.961
No illicit crossing behaviour (disregarding traffic regulations) at the time of hurry q14	300	2.36	0.77	0.6	0.12	0.05
Satisfactory usability of the crosswalk q15	300	3.27	0.448	0.195	−0.275	−0.201
Overall ease in crossing using the crosswalk q16	300	3.26	0.442	0.201	−0.292	−0.220

Note Cronbach Alpha = 0.832
* *N* represents total respondents

References

1. Kadali BR, Vedagiri P (2013) Modelling pedestrian road crossing behaviour under mixed traffic condition. Eur Transp 55(3)
2. Behbahani H, Gilani VNM, Samet MJ, Salehfard R (2017) Analysis of crossing speed of the pedestrians in marked and unmarked crosswalks in the signalized and un-signalized intersections (case study: Rasht city). In: IOP conference series: materials science and engineering, paper no. 245
3. Muley D, Kharbechea M, Alhajyaseena W, Al-Salem M (2017) Pedestrians' crossing behavior at marked crosswalks on channelized right-turn lanes at intersections. Proc Comp Sci 109:233–240
4. IRC (Indian Road Congress): 103 (2012) Guidelines for Pedestrian Facilities. New Delhi, India
5. Gowda A, Kumari RE (2013) Analysis of pedestrian level of service for crosswalk at intersections for urban conditions. Int J Students Res Tech Manag 1(6):604–609
6. Muraleetharan T, Adachi T, Hagiwara T, Kagaya S (2005) Method to determine overall level-of-service of pedestrian walkways based on total utility value. J Infra Plan Rev 22(3):685–693
7. Khatoon M, Tiwari G, Chaterjee N (2013) Pedestrian crossing behaviour at different types of crosswalks. Urban Mobility India, Research Symposium, New Delhi, India
8. Boroujerdian AM, Nemati M (2016) Pedestrian gap acceptance logit model in unsignalized crosswalks conflict zone. Int J Transp Eng 4(2):87–96
9. Pawar DS, Patil GR (2016) Critical gap estimation for pedestrians at uncontrolled mid-block crossings on high-speed arterials. Saf Sci 86:295–303
10. Sahani R, Saipriya M, Bhuyan PK (2018) Application of gap acceptance concept to investigate crossing behaviour of pedestrians at unsignalised intersections. Transp Dev Eco 4(15):1–8
11. TRB (Transportation Research Board) (2010) Highway capacity manual. National Research Council, Washington, DC
12. Indo-HCM (Indian Highway Capacity Manual) (2017) Sponsored by Council of Scientific and Industrial Research (CSIR), New Delhi
13. Miller JS, Biglow JA, Garber NJ (2000) Calibrating pedestrian level-of-service metrics with 3-D visualization. Transp Res Rec: J Transp Res Board 1705:9–15
14. Chandrappa AK, Bhattacharyya K, Maitra B (2015) Measures for improving pedestrian crossing facilities based on perceptions of urban commuters: an experience in Kolkata. In: 94th annual meeting of transportation research board
15. Chen KL, Luo X, Ji H, Zhao Y (2010) Towards the pedestrian delay estimation at intersections under vehicular platoon caused conflicts. Sci Res Essays 5(9):941–947
16. Bian Y, Lu J, Zhao L (2013) Method to determine pedestrians level of service for unsignalized intersections. Int J App Mech Mat 253(1):1936–1943
17. Sahani R, Bhuyan P (2019) Pedestrian service quality models at un-signalised crossings. Adv Transp Stud XLVII(1):61–69
18. Vijayawargiya V, Rokade S (2017) Identification of factors affecting pedestrian level of service of crosswalks at roundabouts. Int Res J Eng Tech 4(6):342–346
19. Chandigarh Traffic Police (2019) A study of fatal road accidents in UT Chandigarh. Chandigarh Black Book, Chandigarh
20. Lee JYS, Goh PK, Lam WHK (2005) New level-of-service standard for signalized crosswalks with bi-directional pedestrian flows. J Transp Eng 131(2):957–960
21. Shi J, Chen Y, Ren F, Rong J (2007) Research on pedestrian behavior and traffic characteristics at unsignalized midblock crosswalk: case study in Beijing. Transp Res Rec: J Transp Res Board 2038:23–33
22. Papadimitriou E, Lassarre S, Yannis G, Tselentis DI (2016) Road, traffic, and human factors of pedestrian crossing behavior: integrated choice and latent variables models. Transp Res Rec: J Transp Res Board 2586:28–38
23. Ye X, Chen J, Jiang G, Yan X (2015) Modeling pedestrian level of service at signalized intersection crosswalks under mixed traffic conditions. Transp Res Rec: J Transp Res Board 2512:46–55

24. Nargundkar R (2007) Marketing research-text & cases, 3rd edn. Tata McGraw-Hill Education, New Delhi
25. Hair JF, Ringle CM, Sarstedt M (2013) Partial least squares: structural equation modeling: rigorous applications, better results and higher acceptance. Long Range Plan 46(1–2):1–12
26. Gaskin J, Lim J (2016) Master validity tool. AMOS Plugin
27. Hu LT, Bentler PM (1999) Cut-off criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct Eq Model 6(1):1–55
28. Hair JF, Hult GTM, Ringle CM, Sarstedt M (2014) A primer on partial squares structural equation modeling (PLS-SEM). Sage Publications
29. Zhuang X, Wu C (2012) The safety margin and perceived safety of pedestrians at unmarked roadway. Transp Res Part F: Traf Psych Beh 15(2):119–131
30. Kadali BR, Vedagiri P (2016) Proactive pedestrian safety evaluation at unprotected mid-block crosswalk locations under mixed traffic conditions. Saf Sci 89:94–105

Socio-Demographic Variations in Mode Choice Preferences of Peri-Urban and Urban Areas—A Case Study of Bangalore

T. M. Rahul and Ashish Verma

Abstract Peri-urban regions are located in the fringes of an existing city and are important in the urban transportation planning process of the city. The present research analyzes the differential impact of socio-demographic factors on the travel behavior of urban and a peri-urban areas. Specifically, the current study, at first, performs a descriptive multivariate mode share comparison between motorized two-wheelers and Non-Motorized Transport (NMT). Further, it estimates a mode choice model and a trip distance model to determine the marginal effects of socio-demographic and transportation system characteristics. Interaction terms are introduced in the utility function of the mode choice model using a peri-urban indicator to delineate the differential impact of socio-demographic factors in urban and peri-urban areas. The estimated mode choice model gave a comparatively good fit with the data (47 and 27.5%). The significance of the interaction terms indicated a difference in the influence of travel characteristics between urban and peri-urban areas. Gender has a significant influence on the mode choice with females in both urban and peri-urban region having a positive disposition toward NMT. In both regions, an increase in the travel distance reduced the use of non-motorized modes and increased the use of private vehicles. From a social equity perspective, there was a huge scope for promotion of public transport and non-motorized transport in the peri-urban areas. Further, paratransit could be contemplated as a solution to overcome the poor connectivity in the radial routes of peri-urban areas.

Keywords Socio-demographics · Mode choice · Non-motorized transport · Motorized two-wheelers · Peri-urban areas

T. M. Rahul (✉)
Department of Civil Engineering, Indian Institute of Technology Ropar, Nangal Road, Rupnagar, Punjab 140001, India
e-mail: tm.rahul@iitrpr.ac.in

A. Verma
Department of Civil Engineering, Indian Institute of Science (IISc), Bangalore, Karnataka 560012, India
e-mail: ashishv@iisc.ac.in

L. Devi et al. (eds.), *Proceedings of the Sixth International Conference of Transportation Research Group of India*, Lecture Notes in Civil Engineering 272,
https://doi.org/10.1007/978-981-19-3494-0_16

1 Introduction

Indian cities are facing an unprecedented increase in the private vehicle ownership and usage in recent times. This increase in private vehicle usage combined with urban sprawl has resulted in a deterioration of the urban environment resulting in issues like pollution and congestion. Transportation planners and policy makers in developing countries are often faced with the stiff challenge of instantly addressing these issues. This many times has led them to adopt solutions from developed countries. However, solutions proposed for a developed country may not be ideal for a developing country because of their different transportation requirements [1]. For example, many Indian cities have a higher usage of powered motorized two-wheelers and non-motorized modes compared with western counterparts. Hence, there is a need for understanding the complex travel behavior of people before arriving at any solution.

Peri-urban areas are located in the periphery of cities [2–4]. Despite being an important trip generator, the existing urban transportation planning focuses only on the core urban area. The transportation requirement of peri-urban areas may vary from that of a core city because of their geographical and demographical differences. A comparative insight into the travel behavior of urban and peri-urban areas would give an indication of these differences.

Many factors in the existing urban regions including a wide transportation network, poor enforcement of planning policies, desire of people to live in low dense areas, absence of affordable housing near city centers, etc., encourage peri-urbanization [5]. Further, various policy considerations with respect to large-scale industrial investments, urban growth distribution, cheap labor availability, residential development, and slum resettlement also steer peri-urbanization [6, 7].

Non-motorized modes and motorized two-wheelers are two important elements in the transportation system of Indian cities [8]. Both of them cater the middle-income and lower-income sections of the society and are thus significant from an equity perspective. Motorized two-wheelers are cheaper and have better mileage compared with private cars. They also have a good running speed and are easy to maneuver. The sale of two-wheelers in India showed an increasing trend from 13.4 million in 2011–12 to 17.5 million in 2016–17 [9].

Non-motorized Transport (NMT) including walking and cycling do not have the pollution and congestion issues related with motorized modes [10–12]. Further, NMT also provide various health and environmental benefits [13–16]. However, the non-motorized usage in India has been showing a decreasing trend over the years in spite of these advantages [17]. There is an immediate need to revamp the scenario by making NMT more attractive and accessible among the people [18]. The above context eliciting a difference in urban and peri-urban transportation preference, and the importance of two-wheelers and NMT from a sustainability perspective motivates us to set the objectives of the current study as follows.

- To estimate two separate models for motorized two-wheelers and NMT that determines the difference in mode choice preference across urban and peri-urban areas using interaction variables.

- To model the trip distances of individuals in peri-urban areas from a policy perspective especially for promotion of NMT.

The current objective is achieved using a household data collected for the city of Bangalore. The mode preference is modeled using a binary logit model, and the differential variable impact is incorporated using an interaction term in the utility function of the mode choice model. The current study aims to contribute to the literature elucidating the travel patterns in peri-urban areas by eliciting their differential variable impact compared with urban areas.

2 Literature Review

The regional mode choice variation of NMT and motorized two-wheelers depends on the transportation system characteristics and socio-demographics. The current section focuses on studies that elicit the impact of various variables, depicting socio-demographics and the transportation system, on the preference of these modes in an urban and peri-urban context.

2.1 Socio Demographics

Income and car ownership are two socio-demographic factors that has been used in many studies to elicit the differences in mode choice across regions [19–21]. High income individuals were observed having an affinity toward private modes and rail mode compared with NMT [22, 23]. Further, vehicle ownership was found to increase the private vehicle usage even though for travelers who have newly arrived in a city it was the environment that affected the mode choice [24, 25]. However, Kunert and Lipps [26] questioned the role of these variables in eliciting the differences across developed countries because of their already high car ownership rate. They emphasized on a greater relevancy for demographic variables including age and gender in determination of mode choice preferences. However, income and car ownership were important factors in the region-specific mode choice models developed for bicycle mode choice. Ortuzar et al. [27] and Parkin et al. [28] observed a negative effect for vehicle ownership on bicycle usage in urban areas.

The results found with respect to the impact of age on bicycle mode choice differed depending on the spatial context of the study [27, 29]. A rise in age reduced the propensity of choosing bicycles in Britain [30, 31] while increased the propensity of choosing bicycles in United States [15].

In an Indian context, most of the studies done on an urban scale determined a significant impact for the socio-demographic factors. A higher usage of NMT among females and a decrease in the usage of NMT with increase in age were determined for the city of Bangalore [32, 33]. Similarly, Arora [34] found males in the age group of 20–40 to be the main users of bicycles in the peri-urban Delhi.

One of the major factors that differentiated the motorized two-wheeler choice between a developed country and a developing country was the purpose. In developing countries, powered two-wheelers are mainly used for commuting while in developed countries they are used for recreation [35]. Globally, there were very few studies that explored the mode choice preference of two-wheelers. One such study was done in the city of Tiruchirapalli, India, and the study determined a significant influence for the age and gender factor on the two-wheeler usage. Two-wheeler usage increased with age and was higher among males.

2.2 Transportation System Characteristics

Webster [36] emphasizes the need for comprehensively understanding the changes in the peripheries of cities. An unplanned development in peripheries may result in an increased private vehicle usage, which subsequently may negatively influence the urban traffic. Many studies done in an Indian context have attributed the rise in the motorized two-wheeler usage to an increase in the travel distance [37] and a low public transport accessibility [38]. Expanding urban public transportation network to the peri-urban areas improved the quality of life among low-income residents [39].

An increase in travel distance reduced the bicycle usage in urban areas and peri-urban areas. The peri-urban areas in the cities of Accra [40] and Dar es Salaam [41] had a smaller NMT preference for their daily commute because of the large trip distance to the central city. A similar observation was made in an Indian context for the city of Bangalore [32] and for the city of Agartala [42].

3 Data Collection

The travel survey data used in the current study was obtained for the year 2009 from Bangalore Metropolitan Regional Development Authority (BMRDA). Bangalore has a population of 8 million, and an area of 2191 km^2 under the urban land category and 5809 km^2 under the rural land category. The household travel data was collected for the entire Bangalore metropolitan region that comprised Bangalore urban, Bangalore Rural and Ramanagaram areas. The latter two regions, which constituted the peri-urban category in the current study, were further divided into Kanakapura, Ramanagaram, Channapattna, Magadi, Nelamangala, Doddaballapura, Devanahalli, Hoskote, and Anekal (Fig. 1). These towns had a huge development potential, and many key investments are proposed along their corridors [43]. Currently, there are many trips that traverse between the urban and peri-urban area with buses provided by Bangalore Metropolitan Transport Corporation being the major mode of connectivity.

Fig. 1 Bangalore Metropolitan Region. (*Source* Comprehensive Traffic and Transportation Study, 2009)

The current study adopted a random sampling procedure with household as a sampling unit. The detailed survey procedure is reported in the Field Survey Report of BMRDA [43]. A face-to-face interview was done with respondents, and details regarding their trip, mode, and trip context were gathered. Variables including age, gender, occupation, education level, marital status, income, number of earners, type of residence, in vehicle travel time, mode used, purpose, and travel distance were collected. There were 172 zones from the urban area of Bangalore for which household data was collected.

The study adopted several screening procedures to ensure the completeness of the survey including the removal of the unanswered questions and the check on the reported NMT trip distance. The data was finally obtained for 4490 households in the peri-urban area and 8147 households in the urban area. Compared with the peri-urban Bangalore, urban Bangalore had a higher number of households where more than 1 person worked. This might be a result of the higher job accessibility in the urban area compared with the peri-urban area. Currently, most of the income generating opportunities including the industries, offices, commercial and retail establishments, etc., are located in the urban Bangalore. The variation in the percentage of workers across age categories showed a similar trend for both urban and peri-urban areas.

It showed an increasing trend till the age group '26–35' after which it showed a decreasing trend.

4 Methodology

The study adopted a three-tier methodology in which the first step consisted of determination of a significant difference between the urban and peri-urban characteristics of Bangalore. This was done using a 2-sample Z-test that checked the existence of a significant difference between population proportions. The 2-sample test checked a null hypothesis that the proportions in the populations were the same. A Z-value of 1.64 at 5% significance level (two-tailed test) was set as the cut-off value for rejecting the null hypothesis.

The second step adopts a logistic regression approach for determination of impact of different variables on mode choice. The logistic regression estimates the impact of independent variable on a dependent variable that is discrete in nature. As a generalized regression variant, this approach uses a canonical parameter as defined below.

$$\ln\left(\frac{p_n}{1-p_n}\right) = \beta_0 + \sum_{k=1}^{k} \beta_k X_{nk} \tag{1}$$

K: number of unknown parameters.
X: variables influencing the probability.
P_n: probability of success for individual 'n'.
β: Parameters which can be estimated using maximum log likelihood.

The estimated maximum log likelihood estimator would give the impact of each variable used in the specification of the canonical parameter. A positive sign for the parameter would indicate a positive impact for the variable on the probability, and a negative sign for the parameter would indicate a negative impact for the variable on the probability. The canonical parameter also corresponds to the utility concept in accordance with the utility theory developed for discrete mode choice scenarios [44].

The model uses the peri-urban interaction variables to elicit the differential impact of independent variables in the urban and peri-urban areas. The interaction variables consisted of an additional variable for each individual attributes that were specified in terms of a peri-urban indicator.

The third step consisted of estimation of a linear regression model for the peri-urban area. The linear regression model is a class of generalized linear models. In linear regression, the dependent variable is a continuous variable that will have a normal distribution with mean '*Y*.' The canonical parameter is linear regression model is defined as below [45].

$$Y_n = \beta_0 + \sum_{k=1}^{k} \beta_k X_{nk} \tag{2}$$

Yn: dependent variable for observation '*n*'
K: total number of unknown parameters.
X: variables influencing the dependent variable.
β: Parameters which can be estimated using the maximum log likelihood.

5 Urban and Peri-Urban Data Comparison

Table 1 shows a comparison of different socio-demographic characteristics between urban and peri-urban areas. Most of the variables indicated a significant difference in the percentages between urban Bangalore and outer Bangalore. Urban Bangalore had a higher percentage of households owning two vehicles or more. Aptly, urban Bangalore also had a higher percentage of households having more than 1 worker. This higher private vehicle ownership might be linked to their larger household income (Table 1) and subsequent affordability. The percentage of owned houses was great in the peri-urban area. This might be attributed to the availability of cheap land at peri-urban regions.

Most of the respondents in the peri-urban Bangalore came under the employee category. In comparison with individuals involved in agriculture, who came under the self-employed or daily wage category, the large proportion of employees suggested an ongoing urbanization in these areas with less dependence on agriculture [46].

Economically Weak Section and Low-Income group were defined by The Ministry of Housing and Urban Poverty Alleviation, India (2012), using an income ceiling of 0–10,000 Indian Rupees. Urban area of Bangalore had a great number of households earning more than 10,000 Indian Rupees per month, and peri-urban areas had a higher percentage of people earning less than 10,000 Indian Rupees per month. Latter might be an indicator of residential preference among low-income groups. In a developing country context, outer regions are often preferred by low-income group because of availability of land at cheap prices and houses at low rent [40].

Peri-urban Bangalore had a lower number of short distance trips less than 1.4 km—the trip distance preferred by pedestrians and a higher number of long distance trips greater than 5 km. The latter aspect may be attributed to the comparative lower densities existing at peri-urban areas of Bangalore [8] and to the dependence of peri-urban residents on the urban area for jobs.

Table 1 Data description

	Bangalore peri-urban	Bangalore urban	*P*-value
Number of workers			
% of households with 1 worker	95	0.77	0.00
% of households with 2 workers	4	0.17	0.00
% of households with 3 or more workers	1	0.06	0.00
Age			
% of individual working in 18–25 group	17	18	0.29
% of individual working in 26–35 group	53	42	0.00
% of individual working in 36–45 group	19	25	0.00
% of individual working in 46–60 group	8	14	0.00
% of individual working in >60 group	3	1	0.00
Vehicle ownership			
% of households with 0 vehicle	54	43	0.00
% of households with 1 vehicle	44	43	0.31
% of households with 2 or more vehicle	2	14	0.00
Type of house ownership			
% of owned households	34	28	0.00
% of rented households	66	72	0.00
Mode share			
Car	0.5	4	0.00
Motorized Two-wheeler	29	26	0.00
Walking and Bicycling	58	50	0.00
Gender			
Percentage of workers among males	0.84	85	0.25
Percentage of workers among females	0.16	15	0.28
Occupational status			
Employed under others	46	54	0.00
Self-employed	18	11	0.00
Daily wage	24	8	0.00
Student	11	27	0.00
Monthly household income (Indian Rupees)			
<10,000	92	82	0.00
>10,000	8	18	0.00
Trip distance in km			
<1.4	32	38	0.00
1.4–5	37	32	0.00
>5	31	29	0.08

5.1 *Motorized Two-Wheeler and NMT Variation Across Categories*

Table 2 shows the percentage-wise distribution of their total numbers across different categories of income and travel attributes. Motorized two-wheelers are used for long trip distances, and NMT are preferred for short trip distances. In the higher income group, the variation in the share of motorized two-wheelers and NMT followed a similar trend in both urban and peri-urban areas. The NMT usage decreased with an increase in the distance, and the two-wheeler usage increased and then declined with an escalation in the distance. The higher percentage of NMT users in the lowest trip distance band implied a NMT preference for short trip distances. Aptly, a smaller trip distance induced by a good street connectivity was highlighted as an important component for promotion of walking in the study of Galpern et al. [47].

The percentage of motorized two-wheelers in the distance bands '5–10' kilometers and '>10 km was greater for peri-urban areas compared with urban areas. The percentage of two-wheelers in the time bands '20–40 min' and '>40 min' was also greater in the peri-urban areas. These two observations suggested a greater two-wheeler usage for longer trips in peri-urban areas. This higher usage might be a

Table 2 Motorized two-wheeler and NMT distribution

Urban				Peri-urban	
		NMT	Motorized two-wheeler	NMT	Motorized two-wheeler
Income <10,000 (Indian Rupees)	*Trip distance*				
	2 km	0.87	0.17	0.83	0.02
	2–5 km	0.10	0.29	0.15	0.31
	5–10	0.02	0.33	0.01	0.51
	>10	0.01	0.21	0.01	0.17
	Travel time				
	0–20 min	0.77	0.50	0.76	0.27
	20–40 min	0.19	0.30	0.23	0.50
	>40 min	0.04	0.19	0.01	0.23
Income >10,000 (Indian Rupees)	*Trip distance*				
	2 km	0.83	0.12	0.84	0.05
	2–5 km	0.13	0.27	0.14	0.15
	5–10	0.02	0.36	0.01	0.46
	>10	0.02	0.25	0.01	0.34
	Travel time				
	0–20 min	0.76	0.34	0.84	0.84
	20–40 min	0.18	0.36	0.16	0.15
	>40 min	0.06	0.30	0.01	0.01

manifestation of the poor public transport connectivity and the subsequent captive private vehicle ridership in the peri-urban areas. Currently, in the peri-urban Bangalore, public transport is provided only along the arterial roads, thus resulting in an inadequate connectivity of the area [43]. Even among the higher income group, a similar observation is made with respect to the trip distance.

6 Results and Discussion

6.1 *Mode Choice Model*

Table 3 shows the variables used in the estimation of mode choice model. The variable was selected for the study based on a review of literature, intuitive judgment, and deliberation with experts. The limited data availability because of the secondary data usage also played a role in the variable selection. Variables were categorized

Table 3 Variables used in the mode choice models

	Variable name	Description	Representation in model
1	Peri-urban indicator (dummy variable)	Whether the respondent belonged to the peri-urban area	1-0—Do not belong to a peri-urban area 1—Belongs to the peri-urban area
2	Type of residence	Whether the respondent is residing in an apartment or individual home	0—Residing in an apartment 1—Residing in an individual household
3	Employed people in household	Number of employed people in the household	–
4	Private vehicles in household	Number of private vehicles in the household	–
5	Age	Age of respondent	–
6	Gender	Gender of the respondent	0—Female 1—Male
7	Marital status	Marital status of the respondent	0—Unmarried 1—Married
8	Educational level	Whether the respondent is literate or illiterate	0—Literate 1—Illiterate
9	Occupational status	Whether respondent has a full-time or part-time job	0—Part time job 1—Full time job
10	Household income	Income of individuals	–
11	Trip distance	Trip distance of individuals	–
12	Children below 18 years P	Whether there are children below 18 years of age	0—Absent 1—Present

in to two sets: one denoting the socio-demographics and the other representing the transportation system characteristics. The latter set consisted of the variables trip distance and residence type. A city planning authority could control these factors for influencing the mode choice decision among people.

The socio-demographics was represented using the following factors: the number of employed people, the number of private vehicles owned, age, gender, marital status, education level, occupational status, household income, and children below 18 years of age. The variables, type of residence, gender, marital status, education level, occupational status, and existence of children below 18 years of age, were taken as dummy variables in the model estimation. The other variables including travel distance, age, and income were taken as continuous variables. The dependent variable for the model estimation was kept a binary variable—whether the respondent chose a two-wheeler/NMT or not. The canonical parameter corresponding to the scenario where a two-wheeler/NMT is not chosen is considered as zero. Hence, the impact of variables can be interpreted corresponding to a scenario where the mode would be chosen.

Two binary regression models are estimated to separately understand the mode choice scenarios of motorized two-wheelers and NMT. This distinction reduced the number of estimated parameters, especially corresponding to the interaction variables, thus reducing the extent of correlation. In the estimated model, the indicators capturing the model accuracy gave satisfactory results. The likelihood ratio statistics that represented the statistical significance of entire variable array elicited reasonable P-values well below 1%. The Rho-square values, which gave an indication of the goodness of fit of the estimated model, were 0.471 and 0.275, respectively, for two-wheelers and NMT. The validation procedure that was done on a set of sample not used in the model estimation also educed satisfactory results. The model results are interpreted in terms of the sign of the variable parameters, their magnitude, and their significance. The sign of the parameters indicated the nature of influence of the corresponding variables on the choice of a mode, the magnitude of the parameters represented the extent of influence, and the significance of the parameters defines the contribution of specific variables to the predictive capability of the model.

Motorized two-wheeler mode choice

Table 4 presents the results of mode choice model. The negative parameter value for the peri-urban indicator displayed a reduction in the probability for two-wheeler mode choice in the peri-urban region. If all other variables were kept the same, the propensity that a person in the peri-urban area would select a two-wheeler was smaller compared with the urban area.

Living in a single home in the urban area increased the probability that a respondent would choose a motorized two-wheeler. If the type of residence is taken as an indicator of density—with a single home representing areas of low-density and the apartments representing areas of high-density—it may be inferred that a reduction in density would increase the two-wheeler usage. A greater influence for this variable in the peri-urban area could be attributed to the poor connectivity of the existing public transportation system.

Table 4 Estimated parameters of mode choice model

Variable	Mode of transport	
	Motorized two-wheeler	Non-motorized transport
Constant	−3.43 (0.00)	4.060 (0.00)
Peri-urban indicator	−2.38 (0.03)	1.660 (0.11)
Type of residence	0.06 (0.59)	0.237 (0.06)
Type of residence P	0.790 (0.25)	0.116 (0.85)
Employed people in household	−0.420 (0.00)	0.224 (0.00)
Employed people in household P	−0.144 (0.63)	0.048 (0.84)
Private vehicles in household	0.864 (0.00)	−0.507 (0.00)
Private vehicles in household P	1.90 (0.00)	−0.362 (0.02)
Age	0.0007 (0.84)	−0.009 (0.02)
Age P	−0.035 (0.00)	0.014 (0.20)
Gender	0.743 (0.00)	−0.44 (0.00)
Gender P	0.597 (0.31)	−0.348 (0.56)
Marital status	0.282 (0.00)	−0.122 (0.26)
Marital status P	0.403 (0.18)	−0.340 (0.31)
Educational level	0.618 (0.00)	−0.539 (0.00)
Educational level P	0.589 (0.00)	−0.151 (0.47)
Occupational status	0.716 (0.00)	−1.08 (0.00)
Occupational status P	−0.443 (0.02)	1.02 (0.00)
Household income	0.0002 (0.96)	−0.04 (0.00)
Household income P	0.097 (0.00)	−0.142 (0.00)
Trip distance	0.114 (0.00)	−0.597 (0.00)
Trip distance P	−0.008 (0.53)	−0.193 (0.00)
Children below 18 years	−0.06 (0.08)	−0.02 (0.65)
Children below 18 years P	−0.329 (0.01)	0.231 (0.10)
Null log likelihood	−5857.787	−5857.787
Final log likelihood	−3074.747	−4220.642
Adj. Rho-square	0.471	0.275
Likelihood ratio estimator	5566 ($P \ll 0.01$)	3274 ($P \ll 0.01$)

An increase in the number of people working from a household reduced the propensity of choosing motorized two-wheelers. Males had a greater preference for two-wheelers compared with females. The odds corresponding to the gender variable was 1.82 times greater in the peri-urban areas compared with the urban areas. In urban areas, the odds of a male choosing a two-wheeler was 2.1 times ($e^{0.743}$) greater than females.

Getting married, getting educated, and having a full-time job increased the probability that an individual would choose a motorized two-wheeler. This effect on probability was higher for the first two variables and smaller for the last variable in the peri-urban areas.

Respondents in both urban and peri-urban area preferred motorized two-wheelers for long distance trips. However, the sensitivity of the trip distance parameter was greater for the urban residents with the odds of choosing a two-wheeler being 1.8 times higher. Both household income and age had an insignificant effect on the two-wheeler mode choice probability. The existence of two-wheelers across the entire income levels and the subsequent lack of variation during model estimation may be a reason for this income effect.

NMT mode choice

The positive value of the peri-urban indicator in the estimated model indicated a positive disposition among the peri-urban residents toward the NMT. All other factors being same, the odds that a peri-urban resident may choose a NMT was 57.9 time greater ($e^{4.060}$) than an urban resident.

Being from a single home increased the affinity of respondents toward a non-motorized mode. This effect was greater in urban areas compared with peri-urban parts. A similar observation was made for the peri-urban Chennai in the study of Srinivasan et al. [1] and suburban Helsinki in the study of Salonen et al. 37. Srinivasan et al. [1] attributed the increased NMT usage among peri-urban residents to intra-zone commute trips.

In contrast to the results obtained for motorized two-wheelers, an increase in the number of employed persons increased the chances of using NMT in peri-urban areas. Similar results were obtained for urban areas also. This may be a result of the reduction in the private vehicle accessibility with increase in the number of employees.

An increase in the level of education and the number of private vehicles owned reduced the usage of non-motorized modes. This reduction was more prominent in the peri-urban areas. Further, being a female increased the odds of using a NMT in both urban and peri-urban areas. This might be an indication of the natural inclination of females toward sustainable modes like walking and bicycling [48].

Respondents having a full-time job had a lower preference for NMT compared with respondents having a part-time job. This effect might be a consequence of the attitude among respondents having a full-time job that associates NMT usage with lower social status [49]. However, this reduction in preference among individuals with a full-time job was smaller in peri-urban areas compared with urban areas.

An increase in the trip distance reduced the probability of NMT mode choice. This result was consistent with common perception that higher trip distances often reduced the NMT usage [50]. However, this decrease was 1.21 times more in the peri-urban areas. This indicated the higher sensitivity of the peri-urban population to the changes in trip distances.

6.2 Trip Distance Model

The trip distance model was estimated to determine the partial effect of the independent variables including occupational status, age, gender, income, etc., on the dependent variable trip distance (Table 5).

The extent of linear relation between the dependent variable and the independent variables was indicated by the R^2 value. Even though the R^2 value was 0.11, the statistical significance of the independent variables prompted authors to proceed on with the estimated model for the analysis of the results. The determined F-value of 54.00 justified the fit of the estimated regression model.

Individuals employed under others had the highest positive impact on the trip distance followed by self-employed and daily wage laborers. The magnitude of influence of occupational status on the trip distance was ten times higher for the individuals employed under others compared with individuals working for daily wages. This lower trip distance combined with the higher usage of non-motorized modes among daily wage laborers may indicate the intra-zone trips undertaken [1]. The 'employed under others' category may be depending on urban areas where in more number of office and service-oriented jobs are available.

Table 5 Trip distance model for peri-urban area

Variable	Parameter	t-value	P-value
Intercept	1.19	2.02	0.044
Occupational status			
Student	0.00	–	–
Daily wage	0.41	0.70	0.48
Self-employed	2.10	3.46	0.00
Employed under others	4.38	7.77	0.00
Income	0.245	8.17	0.00
Age	−0.02	−0.16	0.87
Gender	0.07	0.13	0.89
Type of house ownership (Apartment-1, Independent-0)	−0.26	−0.46	0.64
Type of residence (owned-1, Rented-0)	−0.77	−3.53	0.00
Adjusted R^2	0.11		
F-value (entire model)	54.00(9° of freedom)		

An increase in income significantly increased the commute trip distance of peri-urban travelers. One of the reasons for this effect may be the increased emphasis on better housing compared with travel expenditure with increase in income. Further, being in an apartment reduced the trip distance compared with an independent home. An increase in age and being a female decreased the trip distance.

7 Conclusions

Peri-urban areas may be considered as future urban regions that are located in the fringes of an existing city. These areas generate a considerable amount of trips to the urban areas and are significant from a transportation planning perspective. The present study performs an initial descriptive analysis and further estimates a comparative mode choice model for the urban and peri-urban areas. It uses an interaction term to elicit the differences in the extent of impact for urban and peri-urban areas. The choice model is estimated for motorized two-wheelers and NMT, and it employs an array of independent variables depicting the socio-demographics and transportation system characteristics.

Motorized two-wheelers are an important mode of transportation for long distance trips in both urban and peri-urban areas. However, as observed in Table 2, there was a greater percentage of motorized two-wheelers undertaking long distance trips in peri-urban areas compared with urban regions. This might be an outcome of the poor public transport connectivity in the peri-urban areas. In future, as the city expands, one may expect a considerable increase in the peri-urban population, and a major contributor to this population would be the low- and middle-income residential seekers who would travel to urban centers for their daily livelihood. An absence of a good public transportation network may make this population captive to motorized two-wheelers because of its low cost compared with private cars. In such a scenario, a public mass transportation system connecting these peri-urban areas to urban areas would be an absolute necessity to reduce the motorized two-wheeler usage. The need for a coordinated transport system was highlighted in the study of Maneepong and Webster [51]. Aptly, it has been proposed to expand the existing Bangalore Metro rail connectivity to these areas [43]. One of the issues that have to be overcome while providing metro connectivity may be the poor connectivity to the radial routes [40]. This issue may be tackled to an extent by promotion of paratransit. Kumar et al. [52] had highlighted the role played by paratransit modes including auto-rickshaw and mini-buses in providing mobility to the population when public transport infrastructure provide by the government is inadequate.

The positive value of the peri-urban indicator variable in the mode choice model revealed an existing higher affinity for peri-urban residents toward NMT. Transportation planning of peri-urban areas should be done to perpetuate this positivity by focusing on these modes or else there would be a reduction in their users in future. A three-pronged approach consisting of infrastructural provision, institutional integration, and activity influence may be adopted for this. However, currently, many cities in India, including Bangalore, face a shortage of segregated bicycle paths and footpaths [17, 53]. Further, the unscientific construction practices have made the already existing bicycle paths and footpaths non-usable [54].

A reduction in the overall trip distances in urban and peri-urban areas can have a positive impact on the promotion of sustainable modes including NMT and public transport. The positive effect of the trip distance parameter on the motorized two-wheeler usage points in this direction. Various policies including mixed land-use and high-density settlements that improves the accessibility of individuals to their employment and other non-work opportunities could be adopted for this trip distance reduction. An up gradation of the existing urban planning process may be done at the state level for an integrated land-use and transport planning [55]. However, in an Indian scenario, these strategies should be adopted with care because of the pertinent economic stratification of the society. Embracing these strategies without considering the requirements of the lower economic strata may create a bias from a social equity perspective. An exploration of this equity impact can form the scope for future research. Further, additional variables specifically pertaining to the infrastructure and built environment could be included to enrich the model developed in the present study. The current study included only two modes—two-wheelers and NMT—in the mode choice estimation because of the significance of these modes in a developing country context. Further, the data limitation in the current study prevented authors from using an accessibility variable for public transport in the mode choice models. Future works may further extend the present study to include more modes and the accessibility component.

Acknowledgements Authors would like to thank UMTC (Urban Mass Transit Company) for sharing the household travel survey data of Bangalore. The support of officials from BMRDA and KUIDFC are highly appreciated.

References

1. Srinivasan KK, Lakshmi Bhargavi PV, Ramadurai G, Muthuram V, Srinivasan S (2007) Determinants of changes in mobility and travel patterns in developing countries. Transp Res Record 2038:42–52
2. Audirac I (1999) Unsettled views about the fringe: rural-urban or urban-rural frontiers. In: Furuseth OJ, Lapping MB (eds) Contested countryside: the rural urban fringe in North America, Ashgate, Aldershot
3. Friedberger M (2000) The rural-urban fringe in the late twentieth century. Agric Hist 74(2):502–514
4. The Energy and Resources Institute (2006) Sustainable settlements in peri-urban areas: with special reference to impacts of transport and energy on natural resources management. http://www.ivv.tuwien.ac.at/fileadmin/mediapool-verkehrsplanung/Bilder/Forschung/Projekte/Periurban/wp3_institutional_mechanism_background_01.pdf. Last accessed 12 May 2021
5. Bhatta B (2010) Causes and consequences of urban growth and sprawl. In: Analysis of urban growth and sprawl from remote sensing data. Advances in geographic information science. Springer, Berlin, Heidelberg

6. Webster D, Muller L (2009) Peri-urbanization: zones of rural-urban transition. Human Settlement Development, UNESCO, vol I
7. Webster D, Cai J, Muller L (2014) The new face of peri-urbanization in East Asia. J Urban Aff 36:315–333
8. Iyer NV, Badami MG (2007) Two-wheeled motor vehicle technology in India: evolution, prospect and issues. Energy Policy. https://doi.org/10.1016/j.en-pol.2007.02.001
9. Society of Indian Automobile Manufacturers (2017) http://www.siamindia.com/statistics.aspx?mpgid=8&pgidtrail=14. Last accessed 12 May 2021
10. Banister D (2011) Cities, mobility and climate change. J Transp Geogr 19:1538–1546
11. Hanssen JU (1995) Transportation impacts of office relocation—a case study from Oslo. J Transp Geogr 3:247–256
12. Verhetsel A, Vanelslander T (2010) What location policy can bring to sustainable commuting: an empirical studyin Brussels and Flanders, Belgium. J Transp Geogr 18:691–701
13. Litman T (2003) Integrating public health objectives in transportation decision-making. Am J Health Promot Inc 18:103–108
14. Liu X, Wei H, Guan H, Ma J (2003) Repositioning bicycling in transportation policies and strategies from the perspective of mode choice changes in large Chinese Cities. In: Proceedings of the 82nd annual meeting of TRB, Washington, DC (2003)
15. Replogle M (1991) Non-motorized vehicles in Asia: lessons for sustainable transport planning and policy. Non-Motorized Vehicles in Asian Cities, World Bank Technical Report 162
16. Rietveld P (2000) Non-motorised modes in transport systems: a multimodal chain perspective for The Netherlands. Transp Res Part D 5:31–36
17. Ministryof Urban Development, Government of India (2008) Study on traffic and transportation policies and strategies in urban areas in India
18. Badami MG (2013) The new challenges of rapid urban growth. Imperatives and strategies for transport in India. In: Ensuring a sustainable future: making progress on environment and equity. Oxford University Press, New York, pp 167–188
19. Dargay J, Gately D (1999) Income's effect on vehicle ownership, worldwide: 1960–2015. Transp Res Part A 33:101–138
20. Ingram KG, Liu Z (1999) Determinants of motorization and road provision. The World Bank, Washington, DC
21. Schafer A, Victor DG (2000) The future mobility of the world population. Transp Res Part A 34:171–205
22. Bhat CR, Srinivasan S (2005) A multidimensional mixed ordered-response model for analyzing weekend activity participation. Transp Res Part B Methodol 39:255–278
23. Liu Y, Chen J, Wu W, Ye J (2019) Typical combined travel mode choice utility model in multimodal transportation network. Sustainability 11:549
24. Koppelman FS, Bhat C (2006) A self instructing course in mode choice modeling: multi-nomial and nested logit models. In: Prepared for U.S. Department of Transportation Federal Transit Administration
25. Weinberger R, Goetzke F (2019) Automobile ownership and mode choice: learned or instrumentally rational? Travel Behav Soc 16:153–160
26. Kunert U, Lipps O (2005) Measuring and explaining the increase of travel distance: a multi-level analysis using repeated cross sectional travel surveys. DIW, Berlin
27. Ortuzar JD, Willumsen LG (2005) Modelling transport. Wiley, Chichester, West Sussex
28. Parkin J, Wardman M, Page M (2008) Estimation of the determinants of bicycle mode share for the journey to work using census data. Transportation 35:93–109
29. Agrawal AW, Schimek P (2007) Extent and correlates of walking in the USA. Transp Res Part D 12:548–563
30. Wardman M, Hatfield R, Page M (1997) The UK national cycling strategy: can improved facilities meet the targets. Transp Policy 4:123–133
31. Wardman M, Tight M, Page M (2007) Factors influencing the propensity to cycle to work. Transp Res Part A 41:339–350

32. Rahul TM, Verma A (2013) Study of impact of various influencing factors on NMT mode choice. Procedia—Soc Behav Sci 104:1112–1119
33. Rahul TM, Megha M, Meenakshi S, Madhavendra K, Verma A (2020) Gender differences in work trip generation: evidences from Bangalore. Transp Res Procedia 48:2800–2810
34. Arora A (2013) Non-motorized transport in peri-urban areas of Delhi, India. Case study prepared for global report on human settlements
35. World Health Organization (2017) Powered two and three-wheeler safety—a road safety manual for decision-makers and practitioners
36. Webster D (2011) An overdue agenda: systematizing East Asian peri-urban research. Pac Aff 84(4):631–642
37. Pai M, Gadgil R, Mahendra A, Vernekar S, Heywood R, Chanchani R (2014) Motorized two-wheelers in Indian Cities—a case study of the city of Pune. EBARQ, India. http://www.wrirosscities.org/sites/default/files/Motorized-Two-Wheelers-Indian-Cities-Pune-EMBARQ-India.pdf. Last accessed 12 May 2021
38. Two and Three Wheelers in India (2009) Innovative Transport Solutions (iTrans) Pvt. Ltd., TBIU, IIT Delhi, New Delhi. Available at: https://www.itdp.org/wp-content/uploads/2014/07/2_and_3_wheelers_in_India.pdf. Last accessed 12 May 2021
39. Peters SC (2011) Bringing in the peri-urban poor. Sanford J Publ Policy 2:37–56
40. Agyemang E (2017) Mode choice for long distance trips: evidence from the Greater Accra Metropolitan Area of Ghana. J Transp Geogr 64:150–157
41. Andreasen MH, Møller-Jensen L (2017) Access to the city: mobility patterns, transport and accessibility in peripheral settlements of Dar es Salaam. J Transp Geogr 62:20–29
42. Sarkar PP, Mallikarjuna C (2013) Effect of land use on travel behaviour: a case study of Agartala City. Procedia Soc Behav Sci 104:533–542
43. Comprehensive Traffic and Transportation Study for Bangalore (2010) Bangalore Metropolitan Regional Development Authority
44. Ben-Akiva M, Lerman SR (1985) Discrete choice analysis. Theory and application to travel demand. The MIT Press, Rogers Street Cambridge, pp 02142–1209
45. Hosmer D, Lemeshow S (2000) Applied logistic regression. Wiley, Hoboken
46. Shaw R, Das A (2018) Identifying peri-urban growth in small and medium towns using GIS and remote sensing technique: a case study of English Bazar Urban Agglomeration, West Bengal, India. Egypt J Rem Sens Space Sci 21:159–172
47. Galpern P, Ladle A, Uribe FA, Sandalack B, Doyle-Baker B (2018) Assessing urban connectivity using volunteered mobile phone GPS locations. Appl Geogr 93:37–46
48. Mahadeviaa D, Advani D (2016) Gender differentials in travel pattern—the case of a midsized city, Rajkot, India. Transp Res Part D 44:292–302
49. Verma M, Rahul TM, Vinayak P, Verma A (2018) Influence of attitudinal perceptions in childhood and at present on bicycle usage in the Bangalore City. J Transp Geogr 72:94–105
50. Verma M, Rahul TM, Reddy PV, Verma A (2016) The factors influencing bicycling in the Bangalore city. Transp Res Part A (89):29–40 (Elsevier).
51. Maneepong C, Webster D (2008) Governance responses to emerging peri-urbanisation issues at the global–local nexus: the case of Ayutthaya, Thailand. Int Dev Plan Rev 30(2):133–154
52. Kumar M, Singh S, Ghate AT, Pal S, Wilson SA (2016) Informal public transport modes in India: a case study of five city regions. IATSS Research
53. Rahul TM, Manoj M, Tahlyan D, Verma A (2021) The influence of activities and socio-economic/demographic factors on the acceptable distance in an Indian scenario. Transp Dev Econ 7(3)
54. Rahul TM, Manoj M (2020) Categorization of pedestrian level of service perceptions and accounting its response heterogeneity and latent correlation on travel decisions. Transp Res Part A 142:40–55
55. Charaniya S (2018) A state-level framework for integrated land use and transport. In: Sethi M, Puppim de Oliveira J (eds) Mainstreaming climate co-benefits in Indian cities. Exploring urban change in South Asia. Springer, Singapore

User Perception Study of Pedestrian Comfort Including Thermal Effects in an Educational Campus

S. Amalan Sigmund Kaushik, P. Gopalakrishnan, and G. Subbaiyan

Abstract Walking is invariably one of the most sustainable modes of transportation. Pedestrian comfort has been assessed quantitatively and qualitatively in a number of ways. The shortcoming of the existing assessing methods is the lack of focus on users' thermal comfort. Based on the literature review, the parameters that affect the perception of pedestrian comfort have been identified and a questionnaire survey was conducted. Parameters were tested for significance using Spearman correlation, and the parameters thermal comfort, shading, safety, and facility were found to be significant. The thermal comfort perception of the user kept changing through the day even when other physical characteristics and pedestrian facilities remained constant. The perception of pedestrian comfort model was developed using ordinal logistic regression method with significant independent variables thermal comfort, shading, and facilities. Policy guidelines for the campus were proposed based on respondents' preference for walking.

Keywords Pedestrian comfort · Thermal comfort · Pedestrian · Ordinal regression

1 Introduction

Walking is one of the most sustainable modes of transportation and one of the most important ways to minimize greenhouse gas emissions in the environment. In addition to improved living environment, the users also benefit from a healthier lifestyle. After the onset of industrial revolution, most of the cities and campuses were designed with vehicular mode of transport as the primary means of travel. In the last decade,

S. Amalan Sigmund Kaushik (✉) · P. Gopalakrishnan · G. Subbaiyan
Department of Architecture, National Institute of Technology Tiruchirappalli, Trichy, India
e-mail: kaushik@nitt.edu

P. Gopalakrishnan
e-mail: gopal@nitt.edu

G. Subbaiyan
e-mail: subbaiah@nitt.edu

L. Devi et al. (eds.), *Proceedings of the Sixth International Conference of Transportation Research Group of India*, Lecture Notes in Civil Engineering 272,
https://doi.org/10.1007/978-981-19-3494-0_17

a conscious effort is taken to pedestrianize campuses and cities especially due to the health issues faced by people due to their sedentary lifestyle. It is observed that in small Cities walkability needs like mobility, accessibility, and user satisfaction can be compared to large university campuses that engage variety of users like academicians, students, and technical staff.

When walking on their allotted right-of-way, pedestrians, unlike cars, are not always in motion. They come to a halt, sit down, converse with others, window shop, and browse. None of these tasks can be quantitatively set to create a perfect walkway. Culture, current climate conditions, land use influences, trip type, and all other factors influence pedestrian behavior.

The interaction between climate and the built environment is a two-way relationship. Climate impacts urban development which in turn creates a microclimate or a small area that has different climatic conditions than its surroundings. Urban form heavily influences microclimates based on the following factors: an area's basic climate conditions, street orientation and width, distribution of mass or building density within a city, Shading, Solar heat, Anthropogenic heat and Vegetation.

Thermal comfort as defined by ASHRAE 55 is the condition of mind that expresses satisfaction with the thermal environment and is assessed by subjective evaluation The thermal comfort of pedestrians in a given space varies throughout the day and is heavily impacted by the weather. Few major parameters that aid in thermal comfort include shading provided by trees and shrubs, and shading from adjacent buildings.

The National Institute of Technology, Tiruchirappalli (NIT Trichy) campus, which spans 800 acres, is one of India's largest academic campuses. The main entrance, which faces National Highway 67, is at the southern end of the campus. The academic facilities of the institute are located in the southern half of the site, while the staff residential complex is in the southern-eastern half. NIT Trichy has also implemented a non-motorized vehicle policy for the undergraduate and postgraduate students of the institute as a sustainable initiative.

Accordingly, the study was undertaken as per usage of areas and density. Keeping the above context in mind, various areas were analyzed and the pedestrian's opinions and ideas were collected. An understanding of various factors creating disturbances and hindrances to the walkability of the campus was also collected and compiled, along with suitable measures to improve the existing state. Facility considered here is walkways. "Walkway" means a hard or soft surfaced area intended and suitable for use by pedestrians, including sidewalks, plazas, and surfaced portions of Access ways.

2 Need for Study

Pedestrian comfort has been studied quantitatively and qualitatively in a number of ways. Initially, pedestrian comfort was measured using a quantitative walkway measurement. Flow density, road environment quality, traffic flow operations, road physical characteristics, adjacent land use type, and pedestrian facilities (including availability of benches, guard rail, trees and dustbins) have all been identified as impacting criteria other than walkway width in later studies. The qualitative study to assess pedestrian comfort includes user satisfaction and user perception studies. The shortcoming of the existing assessing methods is the lack of focus on users' thermal comfort.

Several studies of the literature have gone into great depth about walkability, but climatic conditions have received little or no attention. Saelens, Sallis, and Frank [1] synthesized planning, transportation, and public health literature covering the topic of active transportation, which includes walking and cycling. With the exception of a brief discussion on the seasonal variation in rates of cycling to work or school, climate is rarely discussed in the review. The authors also acknowledge that a more thorough research of non-built environmental impacts is certainly needed. Kashef [2] compiled research on community design and walkability from the domains of planning, design, transportation, and public health. His comprehensive review, covering 92 sources, made no mention of climatic conditions. Ewing et al. [3] examined classic urban design literature to develop a list of factors thought to influence the walkability of a community. The resulting list of perceptual qualities did not address issues of climate or weather. This study focuses on user perception of pedestrian comfort including thermal effects in a pedestrianized campus.

3 Literature Review

The focus of the literature review is on the elements that influence user perceptions of walkability in terms of pedestrian comfort, as well as the existing approaches for assessing the same. Here are some of the conclusions drawn from some of the most popular existing studies.

South worth and Ben-Joseph [4] and Mehta [5] claim that the safer the neighborhood, the more people will walk. Li et al. [6] discovered that designing built form to provide protection from traffic and criminality contributed to increased walking rates when examining the effects of built environment on walking. One of the most important requirements is safety. While safety encompasses both traffic safety [7] and crime safety [8], this article focuses on the latter.

Buys and Miller [9] conducted 24 qualitative interviews with residents of high-density housing in Brisbane, Australia's sub-tropical heartland. The interviews revealed a widespread belief that Brisbane's climatic conditions made walking more challenging, particularly when long walks were required in hot weather. Combining walking with public transportation was also seen unsettling. On a humid and hot sub-tropical summer day, a "sweaty walk" to the public transportation stop was painful and "too energetic." Thermal environment can be assessed by quantitative and qualitative methods. ASHRAE 55 describes the process to subjectively evaluate thermal comfort.

Evaluation criteria used in Outdoor thermal comfort including measurement, interview, attendance counting, frequency distribution, and questionnaires [10–13] have assessed thermal comfort by using simulations. Outdoor Thermal Comfort was used to assess human thermal comfort on the micro-scale, to find mitigating strategies that can create a more thermally comfortable pedestrian environment [14]. Outdoor thermal comfort largely depends on parameters such as the open space, ground cover, street geometry, and built form. Taleghani et al. [15] studied five different urban forms in the Netherlands and found that the thermal comfort was dependent on urban form.

Increasing the effect of shading in outdoor places is an essential method for improving pedestrian comfort. Buildings, trees, and other landscape elements can diminish solar radiation incidence, lower ground surface temperature, and minimize long and short-wave radiation impacts by shading [11, 16–18]. This helps to alleviate heat-related discomfort for pedestrians. The height-to-width ratio and street orientation [16, 19] are commonly used to assess the shading impact. Furthermore, the street orientation has an impact on solar radiation and wind flow, according to the study. In Colombo, for example, the street running north–south is shielded from solar radiation in the morning and afternoon, unlike the one running east–west.

Pedestrian Facilities are also found to have a direct effect on users' perception on walking [20, 21]. Seating, trash cans, street lights, and drainage are all examples of pedestrian facilities. Physical Characteristics of pedestrian facility are footpath surface, footpath width, obstructions, potential for vehicular conflict, and continuity. The Quality of Service (QOS) indicator highlights the environmental attributes of pedestrian space and is used to define pedestrian facility requirements. Pedestrian areas should be designed with human comfort in mind and tailored to the needs of pedestrians [22].

Regression analysis can be used to determine the relationship between the above factors, and it has been performed in several research. Polus, Schofer, and Ushpiz [23] explored the correlation between pedestrian speed, flow, and density on sidewalks. The statistical analysis of pedestrian walking speed is dependent on pedestrian gender and direction. A review of existing walkability indices reveals that current models may be unable to effectively predict pedestrian comfort in terms of thermal comfort. Simultaneously, existing research lack pedestrian comfort models for walkability.

4 Methodology

For conducting the user perception survey, initial set of parameters were identified from the review of existing literature and standards. This was followed by open-ended questionnaire survey which was conducted to understand the users' perception and their preference of parameters for comfortable walking and issues faced by them. This was validated by author's observation of the sites. The final set of parameters were identified based on the review of literature, open-ended questionnaire survey, and observation. Table 1 lists the identified parameters and characteristics. Shading from landscape and shading from building have been identified through observation, and corresponding characteristics have been mentioned in Table 2.

Table 1 Parameters and characteristics identified from literature

Parameters	Characteristics
Thermal comfort	Thermal sensation
	Thermal comfort rating
Shading	Shade from building
	Shade from landscape
Sense of safety	Safety in daylight
	Safety during night
	Safety with presence of pedestrians
	Safety with vehicular movement
	Safety preference
Facilities	Available facilities
	Facility preference
Walkability	Pedestrian comfort
Orientation	Orientation of the walkway

Table 2 Site characteristics

Site	Orientation	Surrounding land use characteristics	Surrounding landscape characteristics
Site 1	East–West	North-Mechanical department, 75 m away. South-Sports center, 30 m away	North- and South-Dense Evergreen trees—closely placed (Minimum 8 m apart); Average diameter 10 to 15 m
Site 2	North–South	West-Lecture Hall complex, 15 m away. East-Computer center-22 m away Southeast-Cafeteria	East- and West-Dense Evergreen trees—closely placed (Minimum 8 m apart); Average diameter 10 to 15 m
Site 3	East–West	North-Hostels, 45 m away. North-Food kiosk, 2 m away South-LHC—20 m away	North-Dense Evergreen trees—closely placed (Minimum 6 m apart); Average diameter 10–15 m. South-Dense Evergreen trees-loosely placed (Minimum 20 m apart); Average diameter 7–12 m

Based on the identified parameters, questionnaire was designed. The first section contained basic demographic details like age, sex, year of study, etc. Sections 2–5 consist of question regarding the four parameters as identified in this research. Section 6 consists of questions on walkability. The questionnaire survey was conducted across 3 sites within the campus. The user responses were first analyzed in a descriptive way followed by correlation analysis through SPSS IBM software and later an ordinal logistic regression was developed. Based on the analysis and

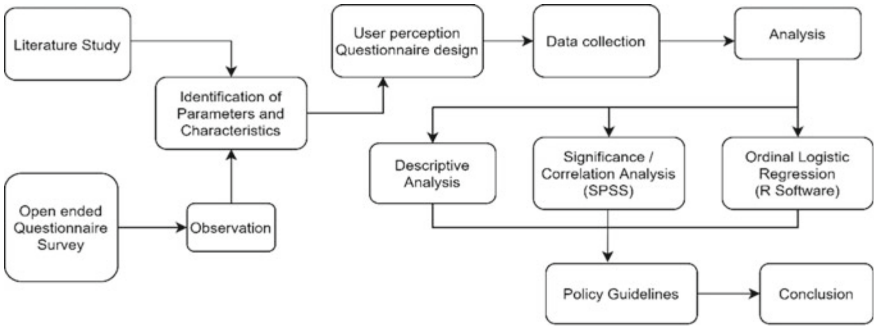


Fig. 1 Methodology

user preferences, policy guidelines are proposed for improving the walkability in the campus (Fig. 1).

5 Data Collection

5.1 Study Locations

NIT Trichy offers ten undergraduate programs and 26 postgraduate programs across 17 departments. Apart from individual departments, the campus consists of sports center, lecture hall complexes, hostels, auditoriums, cafeterias, and shopping complexes.

Three sites were identified within the campus based on landuse, orientation of the walkway, pedestrian density, pedestrian facilities, and constant characteristics (width of walkway, surface characteristics) of the walkway throughout the length of the stretch. Site 1 was surrounded by Mechanical department and sports center, site 2 was surrounded by Lecture Hall complex, cafeteria and computer center, and site 3 was surrounded by boys' hostels and food kiosk. The length of the stretch is taken as 50 m. Detailed characteristics of site is listed in Table 2.

5.2 Climate

With an annual mean temperature of 28.8 °C and a relative humidity (RH) of 67.5%, Tiruchirappalli is classified as a warm and humid climatic zone [24]. The maximum average monthly temperature is 31.6 °C and RH of 63.5% in the month of April followed by 31.3 °C and RH of 64% in the month of May (Table 3). In Tropics such as Tiruchirappalli, heating due to sun primarily influences the thermal comfort

Table 3 Monthly mean temperature and relative humidity for Tiruchirappalli

Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
DBT (°C)	25.1	27.1	29.5	31.6	31.3	31.0	30.4	31.0	29.4	27.7	26.6	25.4
RH (%)	72.4	67.6	60.8	63.6	64.0	58.0	59.3	58.9	65.7	80.1	79.7	80.9

DBT Monthly mean dry bulb temperature, *RH* Relative humidity

of user. The survey was conducted in the month of April to understand the users' perception during summer.

5.3 Data Collection

The respondents were randomly selected and agreed to participate in the survey. A number of 117 students have completed the survey. Site 1 had 52 responses, site 2 had 31 responses, and site 3 had 34 responses. The response was collected in 5-point Likert scale for each of the parameter related questions.

6 Analysis and Discussions

6.1 Respondents' Characteristics

The 117 respondents comprised 55 (47%) males and 62 (53%) females. The respondents belonged to undergraduate (93.2%) and postgraduate (6.8%). The age group ranged from 25.6% for 15–20 age, 70.08% for 21–25 age, and 4.3% for 26–30 age.

6.2 Thermal Comfort

In site 1, at T1 25% of users felt uncomfortable and slightly uncomfortable followed by 84.6% in T2, 36.5% in T3, and 3.8% in T4. (Table 4). The uncomfortableness increased from T1 to T2 as the ambient air temperature increases. Also, at T2 0% of users felt comfortable. At T1 and T3, more than one third of the users felt thermally neutral. The percentage of change in uncomfortableness between T1 and T2 is 53.8% (Table 5). Thus, change in users' thermal comfort is significant as time changes (Figs. 2, 3 and Table 4).

In site 2 at T1, 9.7% of users felt uncomfortable and slightly uncomfortable followed by 51.6% in T2, 16.1% in T3, and 6.5% in T4. Also, at T2 only 3.2% felt comfortable. At T1, 58.1% and, at T3, 61.3% felt comfortable and slightly comfortable. This behavior is attributed to the north–south orientation of this site.

Table 4 Percentage of thermal comfort

	Time	Uncomfortable	Slightly uncomfortable	Neutral	Slightly comfortable	Comfortable
Site 1	T1	5.8	19.2	34.6	26.9	13.5
	T2	59.6	25.0	15.4	0.0	0.0
	T3	9.6	26.9	36.5	25.0	1.9
	T4	3.8	0.0	11.5	32.7	51.9
Site 2	T1	0.0	9.7	32.3	25.8	32.3
	T2	16.1	35.5	29.0	16.1	3.2
	T3	0.0	16.1	22.6	41.9	19.4
	T4	0.0	6.5	6.5	32.3	54.8
Site 3	T1	5.9	32.4	23.5	38.2	0.0
	T2	70.6	20.6	2.9	5.9	0.0
	T3	14.7	32.4	35.3	14.7	2.9
	T4	5.9	2.9	23.5	41.2	26.5

T1: 8–10 am, T2: 12–2 pm, T3: 4–6 pm, T4: 7–9 pm

Table 5 Percentage change in comfort across sites

	% of change in uncomfortableness between T1 and T2	% of change in comfortableness between T1 and T2
Site 1	53.8	−13.5
Site 2	16.1	−29.1
Site 3	64.7	0

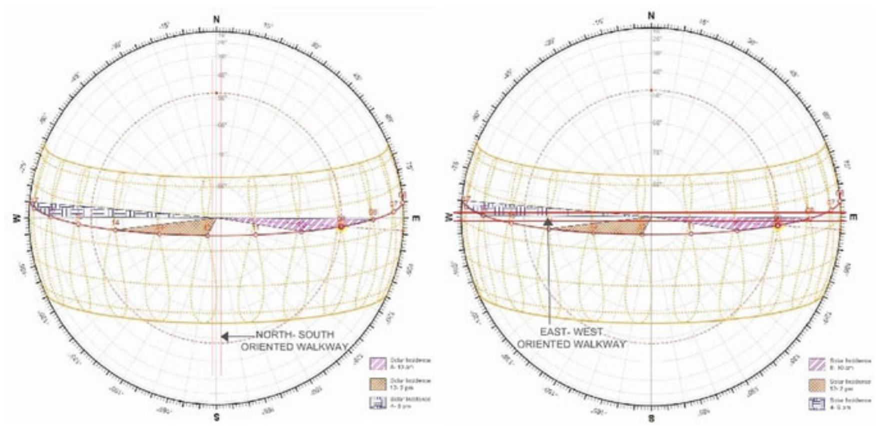


Fig. 2 Sun path diagram for Tiruchirappalli on April 3rd for all the 3 sites with its orientation

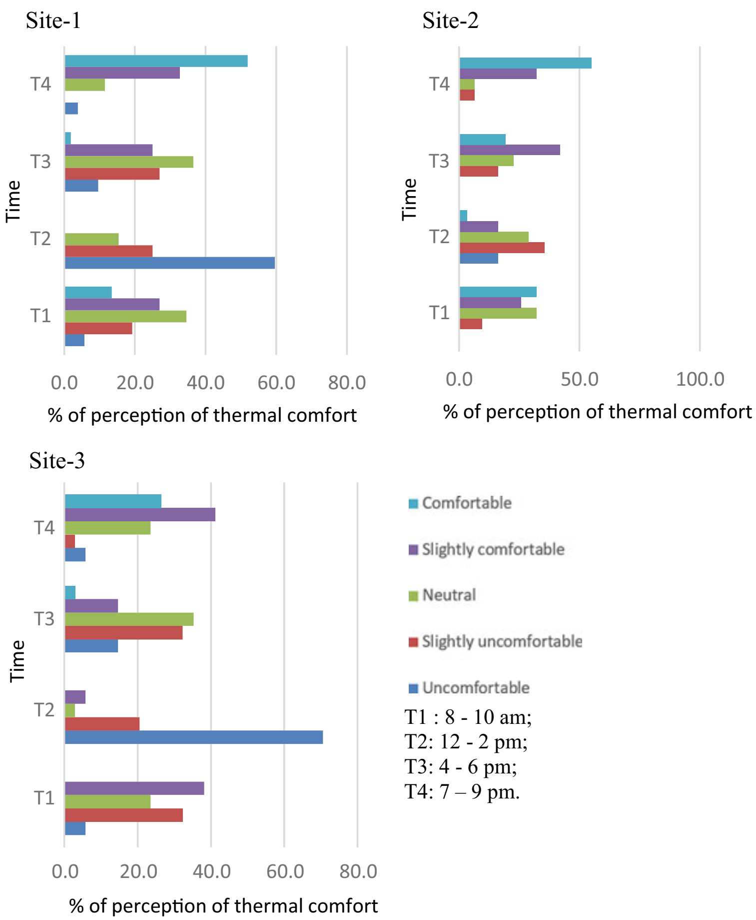


Fig. 3 Percentage of perception of thermal comfort at T1, T2, T3, and T4 for 3 sites

The percentage of change in uncomfortableness between T1 and T2 is 16.1%. At T1 and T3, 0% respondents felt uncomfortable (Tables 4 and 5).

In site 3, at T1, 38.3% of users felt uncomfortable and slightly uncomfortable followed by 91.2% in T2, 47.1% in T3, and 8.8% in T4. Also, at T1 and T2 0% of users felt comfortable. At T3, more than one third of the users felt thermally neutral. The percentage of change in uncomfortableness between T1 and T2 is 64.7%. This behavior of very high change of uncomfortableness is attributed to the east–west orientation of the stretch and also less shading of walkway by tress along the south side of the stretch.

In all the 3 sites, at T4 more than 60% felt comfortable and slightly comfortable. Thus, when there is no sun, orientation of the site does not have much effect on the thermal comfort. Among the respondents who have rated T2 as uncomfortable, 81.6% of the respondents have expressed thermal sensation as “Hot” and 18.4% as “Warm.” This also validates the pedestrian discomfort experienced.

From Table 5, it is found that % of change in uncomfortableness between T1 and T2 is more in site 1 and 3 than site 2, and this is attributed to the north–south orientation of site 2. Both sites 1 and 3 are east–west oriented but there is a noticeable change in % of uncomfortableness, and this is attributed to the less amount of vegetation to the south of site 3.

From the above analyses, it is inferred that the thermal comfort of the respondents for all sites changes with time, especially uncomfortableness increases during afternoon. Also, it is inferred that east–west oriented sites are prone to more uncomfortableness than north–south. Also, within the same orientation shading (Fig. 2) also affects the thermal comfort (Sites 1 and 3).

An important aspect to note in this regard is that even though the parameters such as pedestrian facilities, physical characteristics, and quality of the walkway remained the same, and the thermal comfort perception of the user kept changing through the day.

6.3 Shading

Shading from surrounding building is rated as 26.9% for site 1, 61.2% for site 2, and 8.8% for site 3. From Table 2 site characteristics, it is found that for site 2 the adjacent building on both sides is 15 and 22 m away from walkway, whereas in sites 1 and 3 the buildings are far apart; thus, the characteristics of surrounding buildings have an effect on the shading rating. Shading from surrounding landscape is rated as 82.69% for site 1, 96.7% for site 2, and 35.2% for site 3. Even though site 1 and site 2 is surrounded by dense vegetation (Table 2), the landscape shading rating is higher for site 2 because of the relationship between north–south orientation and sun path of Tiruchirappalli (Fig. 2) on the day of survey. 79.4% have rated site 3 as not shaded.

6.4 Sense of Safety

Mann–Whitney test was applied to check whether there is gender-based variation between the users toward the different characteristics of sense of safety. It was found that male students had positive rating of comfort for safety during night and female students have positive rating toward safety during day, safety with pedestrians, and safety with vehicular traffic. This could be attributed to the gender perception on safety (Table 6).

Table 6 Sum of ranks value of Mann–Whitney test for different safety characteristics

	Safety during night	Safety during day	Safety with pedestrians	Safety with vehicular traffic
Male	3760	3211	3199	3114
Female	3143	3692	3704	3789

6.5 *Statistical Analysis*

Kruskal–Wallis H test was done to check whether there was statistical significance between age and pedestrian comfort. It was found that the distribution of pedestrian comfort is the same across categories of age implying that age has no significance on the pedestrian comfort. The descriptive analysis of the 4 parameters along with pedestrian comfort is listed in Table 7. The median, minimum, and maximum of the ordinal data are tabulated.

To identify the significance of the 4 parameters (Thermal comfort, Shading, Sense of safety, and facilities) on the pedestrian comfort, Pearson correlation analysis was done using SPSS. The results are tabulated in Table 8. All the 4 parameters are found to be significantly correlated to pedestrian comfort.

Table 7 Descriptive statistics of the 4 parameters with pedestrian comfort

Pedestrian comfort		Thermal comfort		Shading		Sense of safety		Facilities	
Min	1	Min	1	Min	1	Min	1	Min	1
1st Qu	3	1st Qu	2	1st Qu	2	1st Qu	3	1st Qu	3
Median	3	Median	3	Median	3	Median	4	Median	4
Mean	3.32	Mean	3.02	Mean	3.04	Mean	3.61	Mean	3.47
3rd Qu	4	3rd Qu	4	3rd Qu	4	3rd Qu	4	3rd Qu	4
Max	5	Max	5	Max	5	Max	5	Max	5

Table 8 Correlation of 4 parameters with pedestrian comfort

	Pearson correlation		Kendall correlation		Spearman correlation	
Independent Variables	Coefficient	Significance (p)	Coefficient	Significance (p)	Coefficient	Significance (p)
Thermal comfort	0.568[a]	0.000	0.456[a]	0.000	0.508[a]	0.000
Shading	0.598[a]	0.000	0.496[a]	0.000	0.546[a]	0.000
Sense of safety	0.373[a]	0.000	0.346[a]	0.000	0.378[a]	0.000
Facilities	0.679[a]	0.000	0.582[a]	0.000	0.620[a]	0.000

[a] Correlation is significant at the 0.01 level

Table 9 The coefficients and intercepts from ordinal logistic regression

	Value	Std. error	*t* value	*p* value
Thermal comfort (TC)	0.73	0.30	2.425	0.015
Shading (SH)	0.56	0.28	1.992	0.046
Safety	0.21	0.28	0.765	0.444
Facilities (FA)	1.71	0.35	4.905	0.000
Intercepts 1|2	4.18	1.22	3.440	0.001
Intercepts 2|3	7.43	1.32	5.611	0.000
Intercepts 3|4	11.03	1.57	7.014	0.000
Intercepts 4|5	16.44	2.02	8.147	0.000

6.6 Regression Modeling

All 4 parameters and the pedestrian comfort are recorded by questions in Likert scale of 1–5. Ordinal logistic regression modeling is used since the dependent variable is ordinal in nature. R software is used to run the ordinal logistic regression modeling. The coefficients and intercepts are tabulated in Table 9.

From the ordinal logistic regression, it is found that except safety, coefficients of all other parameters (thermal comfort, shading, and facilities) and intercepts are statistically significant at 0.05 level. Accordingly, the following models were developed:

$$\begin{aligned}
\text{Logit}(P(Y \le 1)) &= 4.18 + 0.73(\text{TC}) + 0.56(\text{SH}) + 1.71(\text{FA}) \\
\text{Logit}(P(Y \le 2)) &= 7.43 + 0.73(\text{TC}) + 0.56(\text{SH}) + 1.71(\text{FA}) \\
\text{Logit}\ (P(Y \le 3)) &= 11.03 + 0.73(\text{TC}) + 0.56(\text{SH}) + 1.71(\text{FA}) \\
\text{Logit}(P(Y \le 4)) &= 16.44 + 0.73(\text{TC}) + 0.56(\text{SH}) + 1.71(\text{FA})
\end{aligned}$$

where TC = Thermal Comfort, SH = Shading, FA = Facilities.

This equation models the perception of pedestrian comfort with significant independent variables thermal comfort, shading, and facilities.

The above table (Table 9) can be interpreted as follows for Thermal comfort, it would be said that for a one unit increase in Thermal comfort (i.e., going from 0 to 1), we expect a 0.73 increase in the ordered log odds of being in a higher level of Perception of pedestrian comfort, given all of the other variables in the model are held constant.

For Shading, it would be said that for a one unit increase in shading, we expect a 0.56 increase in the ordered log odds of being in a higher level of Perception of pedestrian comfort, given all of the other variables in the model are held constant.

For Facilities, it would be said that for a one unit increase in facilities, we expect a 1.71 increase in the ordered log odds of being in a higher level of Perception of pedestrian comfort, given all of the other variables in the model are held constant.

7 Policy Guidelines

The respondents were asked to select the factors that would aid their preference for walking. The results are as follows: 70% Protection from sun, 60.6% safe environment, 58.1% Proper pedestrian facilities, 47.8% Active area along the pathway, and 29.9% pathway without obstacles.

Protection from sun is the top factor that is perceived to aid their preference of walking. This could be addressed by efficient design of landscape along the walkways, mutual shading of built environment or by providing covered walkways. The design process should take into consideration the orientation of the walkway as orientation and sun path relationship will affect the shading design.

Public lighting at the pedestrian scale and active frontages are key to attributes of safe environment. Active frontage is meant for active visual engagement between those in the street and those on the ground and upper floors of buildings. This quality is assisted where the front facade of buildings, including the main entrance, faces, and opens toward the street. Active frontages can provide informal surveillance opportunities and often improve the vitality and safety of an area. In this campus, the buildings are in general far away from the walkways, and thus alienating the walkways from active pedestrian movement, this could be mitigated by placing food kiosks, interactive arts, and student-oriented facilities.

Proper pedestrian facilities include seating, lighting, dustbins, drainage, signages, barrier friendly walkways, refreshments, etc. Integrated design of pedestrian facilities is more suitable in campuses. Seating designs which are integrated as a part of the pathway and do not exist as another element but more so as a coherent design. The design becomes more collective if facilities like waste bins and cycle parking are also included. Efficient signages help in better way finding within the campus. The ease to locate and understand signages during night decreases the sense of fear among users and increases their safety. Incorporating LED lights over the existing signages enables better identification.

Disabled-friendly infrastructure needs to be designed better for universal accessibility. Elements need to be strategically placed such as Curb ramps, Tactile surfaces, and Low-angle running slope. Walkway surface material needs to be consistent, stable, slip-resistant, and designed to facilitate drainage.

8 Conclusion

Based on the literature review and open-ended questionnaire survey, the parameters and its characteristics that will affect the perception of pedestrian comfort have been identified. The parameters thermal comfort, shading, safety, and facility are found to be significantly correlated to the pedestrian comfort using Spearman correlation test using SPSS. Also, the effect of orientation on the perception of thermal comfort of the users was analyzed using descriptive analyses. The thermal comfort perception of the user kept changing through the day even when other physical characteristics and pedestrian facilities remained constant. Also, gender-based variation between the users toward the different characteristics of sense of safety was tested using Mann–Whitney test and found that male students had positive rating of comfort for safety during night. Further ,an ordinal logistic regression model was done and equations that model the perception of pedestrian comfort with significant independent variables thermal comfort, shading, and facilities were derived. This model is used to evaluate pedestrian comfort of existing walkways and will help designers in redesigning the walkways. These findings can assist planners and architects in making evidence-based decisions when creating pedestrian walkways and would also help in retrofitting existing walkways, allowing for higher pedestrian flow and improved pedestrian comfort. Policy guidelines for the campus were proposed based on respondents' preference for walking.

References

1. Saelens B, Sallis J, Frank L (2003) Environmental correlates of walking and cycling: findings from the transportation, urban design, and planning literatures. Ann Behav Med 25:80–91
2. Kashef M (2010) Neighborhood design and walkability: a synthesis from planning, design, transportation and environmental health fields. Alhosn Univ J Eng Appl Sci 3:87–105
3. Ewing R, Handy S, Brownson R, Clemente O, Winston E (2006) Indentifying and measuring urban design qualities related to walkability. J Phys Act Health 3:S223–S239
4. Southworth M, Ben-Joseph E (2003) Streets and shaping of towns and cities. Island Press, Washington DC
5. Mehta V (2008) Walkable streets: pedestrian behavior, perceptions and attitudes. J Urban Int Res Placemak Urban Sustain 1:217–245
6. Li F, Fisher K, Brownson R, Bosworth M (2005) Multilevel modelling of built environment characteristics related to neighbourhood walking activity in older adults. J Epidemiol Commun Health 59(7):558–564
7. Kearns R, Collins D (2003) Crossing roads, crossing boundaries: empowerment and participation in a child pedestrian safety initiative. Space Polity 7(2):193–212
8. Kelling G, Coles C (1996) Fixing broken windows: restoring order and reducing crime in our communities. Touchstone, Simon and Schuster, New York
9. Thérèse SA, Buys L, Bell L, Miller E (2010) The role of land use and psycho-social factors in high density residents' work travel mode choices: implications for sustainable transport policy. World Rev Intermodal Trans Res 3(1–2):46–72
10. Nikolopoulou M, Baker N, Steemers K (2001) Thermal comfort in outdoor urban spaces: understanding the human parameter. Solar Energy 70(3):227–235
11. Lin TP, Matzarakis A, Hwang RL (2010) Shading effect on long-term outdoor thermal comfort. Build Environ 45(1):213–221
12. Thorsson S, Lindqvist M, Lindqvist S (2004) Thermal bioclimatic conditions and patterns of behaviour in an urban park in Göteborg, Sweden. Int J Biometeorol 48(3):149–156
13. Zacharias J, Stathopoulos T, Wu H (2001) Microclimate and downtown open space activity. Environ Behav 33(2):296–315
14. Rosheidat A, Hoffman D, Bryan H (2008) Visualizing pedestrian comfort using ENVI-met. Proc SimBuild 3(1):198–205
15. Taleghani M, Kleerekoper L, Tenpierik M, Van Den Dobbelsteen A (2015) Outdoor thermal comfort within five different urban forms in the Netherlands. Build Environ 83:65–78

16. Ali-Toudert F, Mayer H (2006) Numerical study on the effects of aspect ratio and orientation of an urban street canyon on outdoor thermal comfort in hot and dry climate. Build Environ 41(2):94–108
17. Hwang RL, Lin TP, Matzarakis A (2011) Seasonal effects of urban street shading on long-term outdoor thermal comfort. Build Environ 46(4):863–870
18. Martinelli L, Lin TP, Matzarakis A (2015) Assessment of the influence of daily shadings pattern on human thermal comfort and attendance in Rome during summer period. Build Environ 92:30–38
19. Emmanuel R, Rosenlund H, Johansson E (2007) Urban shading—a design option for the tropics? A study in Colombo, Sri Lanka, International Journal of Climatology, pp 1995–2004
20. Park S (2008) Defining, measuring, and evaluating path walkability, and testing its impacts on transit users' mode choice and walking distance to the station. PhD Dissertation, University of California Transportation Center, Berkeley
21. Park S, Deakin E, Lee JS (2014) Perception-based walkability index to test impact of microlevel walkability on sustainable mode choice decisions. Transp Res Rec 2464(1):126–213
22. Chandra S, Gangopadhyay S, Velmurugan S, Ravinder K (2017) Indian highway capacity manual (Indo-HCM)
23. Polus A, Schofer JL, Ushpiz A (1983) Pedestrian flow and level of service. J Transp Eng 109(1):46–56
24. I.S.P. (2005) National building code of India. Bureau of Indian Standards, New Delhi

Pedestrian Movement at an Urban Uncontrolled Intersection: A Case Study of Bhopal

Chetan Mahajan, Prikana Das, and Dungar Singh

Abstract Walking is the basic mode of transportation in India. It is considered as one of the feasible and eco-friendly mode of transportation. Pedestrian crossing at an urban uncontrolled intersection is unsafe for pedestrians because of dynamic movement of pedestrians. The objective of the study is to analyse the pedestrian behaviour at an urban uncontrolled intersection considering quantitative and qualitative parameters by estimating critical gap and also considering the accepted and rejected gaps by the pedestrians. The impact of traffic characteristics as well as demographic characteristics of pedestrians for instance gap size, type of vehicle, age, group size, gender, etc., is studied to analyse the crossing behaviour of pedestrians at an uncontrolled intersection. The estimation of critical gap has been done using Raff's method, which is a deterministic approach. Furthermore, the machine learning approach (ANN) is used to model the pedestrian decision making behaviour at an uncontrolled intersection.

Keywords Pedestrian characteristics · Uncontrolled intersection · Critical gap · Artificial neural network · Crossing behaviour

1 Introduction

Walking is treated as one of the feasible and eco-friendly mode of transportation which helps in reducing the traffic congestion and also improving the health of a person on the urban roads. Due to improper management and less knowledge of traffic rules and regulation the number of accidents increases rapidly day by day. According to MORTH-2019 report, the number of pedestrians killed in road accidents in 2018 is 22,656 while in 2019, this reaches to 25,858 which is an increase of about 14.13%. In 2019, 17.1% of pedestrians were killed in road accidents. Statistics show about 37% of road accident victims in the city are pedestrians. The causes of pedestrian

C. Mahajan (✉) · P. Das · D. Singh
Department of Civil Engineering, Maulana Azad National Institute of Technology Bhopal, Bhopal, India
e-mail: mahajanchetan945@gmail.com

L. Devi et al. (eds.), *Proceedings of the Sixth International Conference of Transportation Research Group of India*, Lecture Notes in Civil Engineering 272,
https://doi.org/10.1007/978-981-19-3494-0_18

injuries are vehicular speed, the behaviour of driver and pedestrian, the inadequacy of infrastructure and low visibility. The absence of infrastructure or poor maintenance leads to more serious problems to the pedestrians. So, it necessary in country likes India to provide safe and sustainable walking environment to the pedestrians.

Since last ten years, the modernisation and development of infrastructure increase rapidly which result in the increase of vehicular traffic as well, every road in the urban area deals with the tremendous traffic problems which lead to accident risk for pedestrians as well as vehicular traffic. India is a developing country which has exception amount of heterogeneous traffic in nature. The pedestrian vehicle interaction is more while crossing an uncontrolled intersection; this leads to safety issues and higher risk of accidents. At unsignalised intersections, pedestrians are one of the most vulnerable road users because the road environment is primarily designed for motor vehicles with little regard for people travelling on foot. Therefore, pedestrian road crossing is a serious threat to pedestrians at unsignalised crossing locations under mixed traffic flow conditions in country like India.

At the intersection, vehicles moving in different direction want to occupy same space at the same time make the traffic intersection most troublesome locations on any type of road. And the pedestrians also seek same space for crossing the roads. Considering speed, intersection geometry, direction and route of the different vehicles, etc., driver needs to make blink of an eye decision at uncontrolled intersections. A small error in judgement leads to severe accidents. This results in delay and depend on type, geometry and type of control. The study of intersection is very much necessary from both the accident and the capacity aspect because the traffic flow on any type of road network depends on the performance of the intersection and also affects the capacity of the road.

This study focuses on the analysis of pedestrian crossing behaviour at an uncontrolled intersection considering quantitative and qualitative parameters by estimating critical gap and also considering the accepted and rejected gaps by the pedestrians. Also, analyse the impact of traffic characteristics as well as demographic characteristics of pedestrians for instance gap size, type of vehicle, waiting time, age, group size, gender, etc., on the crossing behaviour of pedestrians at an uncontrolled intersection. ANN technique is used to model the pedestrian decision-making behaviour, i.e. acceptance or rejection of vehicular gaps at an uncontrolled intersection.

2 Literature Review

Pedestrian movement analysis in any type of road network is very much important in development of pedestrian facilities as well as the safe and secure operation of traffic function on the road network. Many researchers have studied about the pedestrian movement in the past. Laxman et al. [1] concluded in their study that there is significant variations in the pedestrians mean speed for age, gender of the pedestrian and for the flow conditions like group movement or moving with baggage. Rastogi et al. [2] found that the crossing speeds of pedestrians were affected by the traffic volume, size of the study area, number of traffic lanes, land uses of the surrounding area and width of the road and demographic characteristics such as gender and age and their movement in a group. Marisamynathan and Perumal [3] determined that male pedestrians crossing speed is faster than female. Impact of departure signal phase and pedestrian's age on the variation of crossing speed is more significant. Pedestrian compliance behaviour is significantly affected by factors like group size and gender of pedestrians. The parameters which influence the vehicle–pedestrian interactions are type of approaching vehicle and vehicle pedestrian gap. According to Banerjee et al. [4], there is variation in speeds of the pedestrians based on the age group. The walking speed of children is greater as compared to adults and elderly, it is concluded that female pedestrians walk slowly as compared to male pedestrians, and the walking speed of pedestrians carrying luggage is less as compared to pedestrians without luggage. Das and Katiyar [5], in this study, back-propagation algorithm with the activation of hidden neurons is used for fast learning procedure. Performance of model shows its accuracy, and statistical comparison of both approaches shows the validation of models. Best performance considering statistical measures are given by ANN as compared to conventional approach. The performance of ANN models is entirely based upon the data set for development of good ANN model. Zheng et al. [6] concluded that the main performance measures are walking speed and space available to the pedestrians for pedestrian movement operation evaluation along urban segments. Dinakar and Kumar [7] found that the vehicular speed is affected by the pedestrian crossing and found that the yielding to pedestrians in terms of speed is more for car and auto rickshaw drivers as compared to the two wheeler drivers. This results decrease in vehicular speed, increase in travel time and also fuel consumption of vehicles increases. Reddy and Teja Tallam [8] used Greenshield's macroscopic stream model and statistical model and soft computing ANN and found out that most efficient and flexible method for analysis of data is Greenshield's model. For statistical regression analysis, Chi square value generated is more compared to that of ANN method. This shows that ANN method is more suitable for the considered locations and collection of pedestrian data. The highway capacity manual proposed that the speed of a pedestrian is 1.2 m/s if the elderly population have less than or equal to 20% ratio, or otherwise 1.0 m/s transportation research board (TRB) [9]. And it affected by different parameters which are discussed in many previous studies.

Crossing behaviour of the pedestrians is mainly depending on the gap acceptance. Pedestrians and drivers perform the task of gap acceptance so frequently that it occurs nearly at subconscious level. But it is essential that the task for driving safely and crossing should be successfully completed. Same gap acceptance behaviour is not shown by all pedestrians and drivers, and even for the distinct condition and different locations, same pedestrian and driver could react non-identical. Researchers have always sought to better understand this behaviour. Amin et al. [10] found that the age of the pedestrian is most effective on decision process and least effect on decision process is by type of conflicting vehicle as compared to other variables. Kadali et al. [11], in their study, used ANN technique for modelling the gap acceptance behaviour of pedestrian under mixed traffic condition, various feed-forward back-propagation models of ANN were used to predict the gap acceptance behaviour, using vehicle characteristics, traffic characteristic and pedestrian behavioural characteristics as input variables. The correct prediction of pedestrian gap acceptance behaviour is shown by the performance of developed ANN model. Li et al. [12], this study shows that the pedestrian arrival rates were nonuniform for entire cycles; during green phase, pedestrians might get delay in arriving; delay were greatly affected by the pedestrians signal non-compliance, and non-complying pedestrian might still receive delays; the average delay and arrival subphase relationships were different for different direction walking pedestrians but the overall average delays were almost the same. Chao et al. [13], in this study, for simulating vehicle–pedestrian mixed flow, a new method has been introduced, which consist of feedback-intrigued action process and decision-making process for vehicle–pedestrian interaction. Based on the condition of incoming vehicles, a gap acceptance is judged by the pedestrian and will not start walking until the sufficient safe gap available. Maurya et al. [14] found that there are many parameters which influence gap acceptance behaviour of drivers like driver's age, gender, occupancy, speed of approaching vehicle, type of oncoming vehicle, waiting time and number of rejections at stop line. Varying gap acceptance behaviour is shown by various driver crowds, especially different gender and age groups. Approaching vehicle speed and passenger attendance also looked to affect driver's gap acceptance decisions. Kadali and Perumal [15] used MLE technique, Logit method, HCM method and Raff's method for estimating critical gap and concluded that the logit model shows best performance from all other methods and also shows that there is influence of age and gender of the pedestrians on the critical gap. Thakur and Biswas [16] concluded that the effect of influence varies based on the country and the priority regulations in place. When making a crossing choice, the incoming vehicle's distance from the crossing place takes precedence over the approaching vehicle's speed. In addition, as traffic volumes increase, pedestrians waiting to cross the road experience longer delays. Mukesh and Yashwan Katpatal [17], this study concludes that the traffic violation by pedestrians is more in numbers, and this shows the unawareness of pedestrians towards the road safety. Along the most important transport corridor within Nagpur urban area, pedestrian safety is very low. Due to both the infrastructure characteristics and pedestrian behaviour, the PSI is low. Vinayaraj et al. [18] from this review, it is concluded that in gap acceptance is the most important parameter during the analysis of pedestrian safety. This study discussed various methods used for the analysis of pedestrian critical gap which are: RMS, MLM, Raff method, Logit method and PEM. There is slight difference in the critical gap values estimated using Raff's method and RMS, and this slight difference is evident only at a particular location. The concept of critical gap has developed over time. According to highway capacity manual [9], the critical gap is "the minimum time interval in the major-street traffic stream that allows intersection entry for one minor-street vehicle". There are more than 20 models used for estimating critical gaps worldwide, but Raff model and Troutbeck models are the most common model among them.

From above review, it is concluded that many research has been done on sidewalks, wide-sidewalks and precinct and signalised intersection but very few research on the unsignalised intersection. Likewise many studies have been reported on gap acceptance behaviour of driver, but there are very limited studies on the pedestrian behaviour. From the available literature, many literatures are of tier 1 cities and very few literatures are of tier 2 cities. The present study attempts to analyse the pedestrian behaviour considering quantitative and qualitative parameters by estimating critical gap and also considering the accepted and rejected gaps by the pedestrians. And machine learning technique artificial neural network (ANN) is used to understand the decision-making process of pedestrians for tier 2 city Bhopal.

3 Methodology

3.1 Data Collection

Data for the study work is collected on three uncontrolled intersections of Bhopal city. Namely Bagsewniya, Piplani and Ratnagiri are the three uncontrolled intersections that are selected for the videographic survey. All of the intersections under the study area are 3-legged intersections. All the intersections were found to be suitable as the intersection comprises market areas as well as a residential area. The intersections have varying demographic, traffic and pedestrian characteristics. Videographic survey is done at all three locations for the duration of 8 h from 10 am to 6 pm with the help of video cameras. From videographic survey data, morning peak of 1 h (10–11 am) and evening peak of 1 h (4–5 pm) have been identified to analyse crossing behaviour of pedestrians. Figure 1a, b and c show the photographs of all the three locations Bagsewniya, Piplani and Ratnagiri, respectively. The video data was collected in weekdays from all three locations in the month of January 2021. Table 1 gives the details of the three uncontrolled intersection that are considered in the study. The videograhic survey is done to find out various parameters like age, gender and group size of pedestrians, speed of pedestrians, movement of pedestrians, i.e. towards curb/median, speed change condition of pedestrians while road crossing, change in crossing path of pedestrians, accepted or rejected gaps, gap size, rolling gap, type of conflicting vehicle and type of gap, i.e. near gap or far gap. Table 2 gives all the details of the parameters that are considered in the research work.

3.2 Data Extraction

The video data is used for extracting the parameters mentioned in the Table 2. For determining the above parameters, the video is included with perspective grid lines with help of Kinovea software as shown in Fig. 1. The grids are placed in such a way that one row of the grid is aligned with zebra crossing, i.e. road width and zebra crossing width. This provides help in calculation of speed (distance and time). And also helps in determining the parameters effortlessly and rapidly. Numerical

a Bagsewniya b Piplani

c Ratnagiri

Fig. 1 Images of location with perspective grid lines

Table 1 Details of study locations

S. No.	Location of study area	Road classification	Road width (m)	Crosswalk width (m)
1	Bagsewniya	Six-lane divided	9.1	3
2	Piplani	Six-lane divided	9.6	4
3	Ratnagiri	Six-lane divided	9.6	3

notations are used as given in Table 2 for rapidly extraction of data from the video. And excel sheet is prepared for the extracted data. Statistical Package for the Social Sciences (SPSS) software is used for preliminary analysis of the data. The data extraction is resulted in 312 pedestrian samples having 636 sample of total gaps (Both accepted and rejected). Figure 2 shows the demographic characteristics of pedestrians with number and percentage share at the intersections. This shows that

Table 2 Details of parameters considered

S. No.	Parameter	Detail
1	Gap size	The time difference between first and second vehicle measured in seconds
2	Age	Age of the pedestrians (Older (>50) = 0; Middle (31–50) = 1; Young (15–30) = 2; Children (<15) = 3)
3	Gender	Gender of the pedestrian (Male = 1, Female = 2)
4	Group size	Number of pedestrian in a group (Single = 1, Two = 2 and Three or more = 3)
5	Movement of pedestrians	Whether pedestrian move towards curb or median (Towards Curb = 1 and Towards Median = 2)
6	Rolling gap	Whether pedestrian rolls over the small gaps (Yes = 1 and No = 2)
7	Speed change condition	Whether pedestrian changes speed while crossing the road (Yes = 1 and No = 2)
8	Crossing path change condition	Whether a pedestrian changes crossing path while crossing the road (Yes = 1 and No = 2)
9	Type of vehicle	Type of vehicle (Two wheeler = 1, Three wheeler = 2, Car = 3 and Heavy vehicle = 4)
10	Type of gap	Type of gap (Near = 1, Far = 2)
11	Speed of pedestrians	Distance travelled by pedestrian per unit time measured in m/s
12	Accepted or rejected gaps	Gap accepted or rejected by pedestrian (Accepted = 1 and Rejected = 2)

male is having higher percentage (68.59%) than female (31.41%), young pedestrian is having maximum share (56.41%) among the all other age group and pedestrian mainly used to travel as single (77.88%) at all the study locations.

3.3 Sample Size

A sample size refers to a portion of the population selected for a survey or investigation. In statistics, a sample is a percentage of the overall population. A sample's data can be used to draw inferences about the population as a whole.

Cochran's Sample Size Formula Cochran [19] formula helps in calculating the ideal sample size for the given or desired level of precision, confidence level and the estimated proportion of the feature present in the population. In case of large populations, Cochran's formula is considered especially appropriate for calculation of ideal sample size. A sample of any given size provides more information about

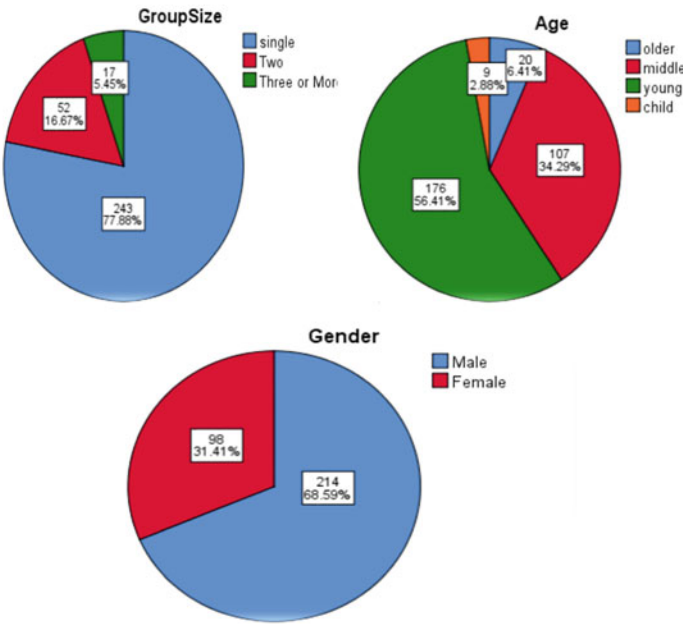


Fig. 2 Pie charts showing number and percentage of demographic characteristics

a smaller population than a larger one, so there is a "correction" through which the number given by Cochran's formula can be reduced if the whole population is relatively small. The Cochran formula is:

$$n_0 = \frac{z^2 pq}{e^2} \tag{1}$$

where:

- n_0 is the ideal sample size,
- e is the desired level of precision (i.e. the margin of error),
- p is the (estimated) proportion of the population which has the attribute in question,
- q is $1-p$,
- The z-value is found in a Z table.

Let us consider:

- $e = \pm 5\%$ is the desired level of precision
- $p = 0.5$, i.e. half of the population has the attribute
- $q = 1-0.5 = 0.5$
- Confidence level $= 95\%$, for 95% confidence level, the z-value is 1.96

$$n_0 = \frac{1.96 * 0.5 * 0.5}{0.05^2} = 385$$

So a random sample of 385 pedestrians in our target population should be enough to give us the confidence levels we need.

4 Analysis and Results

4.1 *Pedestrian Movement Analysis*

Pedestrian's vulnerability leads the pedestrian movement analysis in every study. The extracted data from the videographic survey is used for analysis. In this study, pedestrian movement analysis is done on basis of demographic characteristics, i.e. age, gender and group size and location is also considered for the analysis. The statistical parameters like mean, standard deviation, variance and range are given in the Table 3. Based on different demographic characteristics, variations can be seen in the mean speed of pedestrian which show the different approach of pedestrians towards the crossing of the road. The analysis shows that the data is normally distributed. Normal curve signifies that the maximum frequencies of values lie near the mean value and frequency of occurrence of far value is less. The normal curve is of bell shape. From our analysis, we can say that where the sample size is greater ($n > 30$)

Table 3 Details of statistical parameters of pedestrian movement analysis

S. No.	Category		Number of samples (Pedestrians)	Mean speed (m/s)	Standard deviation	Variance	Range
1	Total		312	0.9375	0.2246	0.05	1.7867
2	Age	Older	20	0.8415	0.1816	0.033	0.6114
		Middle	107	0.8965	0.1860	0.035	0.7623
		Young	176	0.9724	0.2466	0.061	1.7867
		Children	9	0.9560	0.1452	0.021	0.3425
3	Gender	Male	214	0.9532	0.2119	0.045	1.7333
		Female	98	0.9031	0.2476	0.061	1.5572
4	Group Size	Single	243	0.9592	0.2385	0.057	1.7867
		Two	52	0.8548	0.1443	0.021	0.6847
		Three or more	17	0.8800	0.1436	0.021	0.4808
5	Locations	Bagsewniya	92	0.9369	0.2224	0.049	1.6877
		Piplani	114	0.8803	0.1671	0.028	0.9835
		Ratnagiri	106	0.9995	0.2628	0.069	1.6451

Table 4 Results of Pearson correlation and ANOVA test on the factors influencing the pedestrian movement

S. No.	Factors	Pearson correlations		ANOVA			Remark
		Coefficient	*P* value	*F* value	*F* crit	Sig	
1	Age	0.182	0.001	3.950	2.63	0.009	Significant diff
2	Gender	−0.104	0.068	3.364	3.87	0.068	Not Significant difference
3	Group size	−0.161	0.004	5.367	3.03	0.005	Significant diff

data is following the normal curve/normal distribution. But in case of less sample size ($n < 30$), it does not follow the normal curve, i.e. for older, children, and three or more pedestrians category the distribution does not follow normal curve or normal distribution.

Referring to Table 3, it is determined that the mean crossing speed of pedestrians varies based on the demographic characteristics. The mean crossing speed of male is greater as compared to the female pedestrians. Young pedestrian moves faster as compared to children, middle and old age pedestrians. And the mean speed of single travelling pedestrian is greater as compared to groups of two, three or more pedestrians.

To determine the main factor influencing the pedestrian's movement at an urban uncontrolled intersection statistical hypothesis testing is done. Pearson correlation coefficient and ANOVA test are performed between age, gender and group size to find the factors which affects the crossing speed of the pedestrians. From ANOVA hypothesis testing, it is stated that the F value is greater than F critical for group size and age and less in case of gender referring from Table 4. It signifies that there is significant difference in the age and group size data; whereas, there is no significant different in the gender-based speed data. From Pearson correlation coefficient, test is determined that the P value for age and group size is less than 0.05 and greater for the gender. It indicates that there is significant difference in the age and group size data; whereas, there is no significant difference in the gender data at 5% significance level or 95% confidence interval.

4.2 Critical Gap Analysis

Critical gap is the time in seconds below which the pedestrians will not try to cross any intersection or road. In gap acceptance process, critical gap is the most important parameter. The average minimum time gap between the approaching vehicles that permits pedestrians to safely cross the road is known as critical gap. The calculation of critical gap is not possible in fields. It is measured in the form of rejected or accepted gap. There are several methods for the estimation of critical gaps like Raff, lag, maximum likelihood method (MLM), Harder, Wu, Greenshield, Logit model,

Table 5 Statistical details of gaps size (both accepted and rejected gaps)

S. No	Category		Number of samples (gaps)	Mean gap size (sec) (both accepted and rejected)	Standard deviation	Variance
1	Total		636	3.8011	3.1312	9.804
2	Age	Older	41	4.1829	4.0999	16.809
		Middle	213	3.6056	2.6192	6.860
		Young	357	3.9412	3.3156	10.994
		Children	25	2.8400	2.4226	5.869
3	Gender	Male	440	3.6318	2.9847	8.909
		Female	196	4.1811	3.4147	11.661
4	Group size	Single	467	3.8951	3.1489	9.916
		Two	139	3.3058	3.0123	9.074
		Three or more	30	4.6333	3.1811	10.120
5	Location	Bagsewniya	149	3.9866	3.1338	9.821
		Piplani	214	4.6659	3.6417	13.262
		Ratnagiri	273	3.0220	2.4377	5.943

acceptance curve, Ashworth method, clearing behaviour approach, root mean square method (RMS) and probability equilibrium method (PEM). For this study, Raff's method is used for estimation of critical gaps due to its simplicity. Descriptions of statistical details of gaps for both accepted and rejected for different categories for an urban uncontrolled intersection are given in Table 5. And from this, it is analysed that in this study, minimum gap is of 0.3 s and maximum gap of 16.8 s. There are very much variation in the gaps because of traffic characteristics, pedestrians characteristics, land-use, surroundings and locations.

Estimation of Critical gap by Raff's Method According to Raff, critical gap is the size of gap for which number of rejected gaps greater than it is equal to the number of accepted gap smaller than it. For determining the critical gap from Raff's method, an empirical formula is used, which is the relationship between accepted gaps and rejected gaps. And the relationship is shown by Eq. 2. It can be simply estimated as the intersecting point of $F_a(t)$ and $1{-}F_r(t)$ where $F_a(t)$ and $F_r(t)$ are the cumulative percentage of accepted gap and rejected gap, respectively [20].

$$F_a(t) = 1 - F_r(t) \tag{2}$$

The intersecting point between $F_a(t)$ and $1{-}F_r(t)$ is shown in Fig. 3 where blue line shows the cumulative percentage of accepted gaps while the red one is cumulative percentage of rejected gaps. And this intersecting point is the estimated critical gap for that category. Figure 3 shows the critical gap for total samples collected from all the three uncontrolled intersection selected for the study work and the estimated

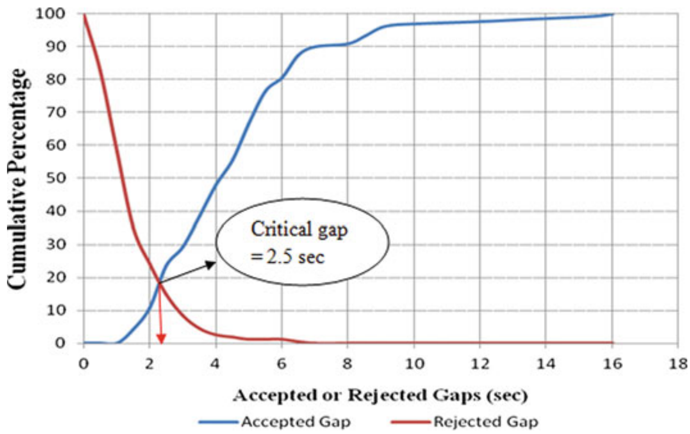


Fig. 3 Critical gap for total samples

critical value for this is 2.5 s. Table 6 gives the estimated critical gap for different categories of pedestrians considered in the study. And it is observed that the critical gap varies with the different demographic characteristics of the pedestrians.

From the Table 6, it is observed that the estimated critical gap for older age pedestrians is higher as compared to children, middle and young age pedestrians. Children has minimum value of critical gap 2.2 s among the age group category which shows that children gap acceptance behaviour is highly risky as compared to older, middle and young age pedestrians. Male has the critical gap of 2.4 s which is less than female pedestrians 3 s; this shows the aggressive behaviour of male as

Table 6 Estimated critical gap for different categories of pedestrians by Raff's method

S. No.	Category		Critical gap (s)
1	Average		2.5
2	Age	Older	2.7
		Middle	2.5
		Young	2.4
		Children	2.2
3	Gender	Male	2.4
		Female	3
4	Group size	Single	2.4
		Two	2.3
		Three or more	3.3
5	Location	Bagsewniya	2
		Piplani	3
		Ratnagiri	2.5

compared to female while crossing the uncontrolled intersection. When pedestrian travel in groups, their behaviour towards the traffic is different from the pedestrian travelling alone. It is observed that the critical gap for single travelling pedestrian is 2.4 s while the critical gap for pedestrians travelling in group of two and group of three or more is 2.3 s and 3.3 s, respectively, referring Table 6. This implies that the pedestrian travelling in the group of two has the most aggressive nature towards the crossing behaviour as compared to single and group of three or more pedestrians.

Based on locations considered in our study, there is variation in the critical gap for the location of Bagwesniya, Piplani and Ratnagiri, the critical gap is 2 s, 3 s and 2.5 s, respectively. This shows that the critical gap varies from place to place and person to person based on traffic characteristics, pedestrian's characteristics, surroundings and land use.

4.3 Pedestrians Decision-Making Behaviour Using Artificial Neural Network (ANN)

The crossing behaviour of pedestrians varies based on the situations, i.e. traffic characteristics, pedestrians characteristics and the surroundings of the area. For accessing the decision-making behaviour of the pedestrians, in this study, we are using ten parameters, the input variables are gap size, age, gender, group size, movement of pedestrians, rolling gap, speed change condition, crossing path change condition, type of vehicle and type of gap. And accepted or rejected gaps are the output variable. For accessing the pedestrians, decision-making behaviour artificial neural network (ANN) is used.

The data used for accessing the pedestrian decision-making behaviour is collected from three different locations in Bhopal city. So, to determine the significance of the data Pearson correlation coefficient and ANOVA, test is performed between the accepted and rejected gap data from all the three locations, and the results of the test is tabulated in the Table 7. The results show that the gap acceptance data from all three locations have no significant difference. So the combined data of all the locations can be used to model the pedestrian decision-making behaviour with the help of binary logistic regression model and artificial neural network (ANN).

Artificial Neural Network (ANN) Artificial neural network is a machine learning techniques widely used in many engineering fields due to its anticipating capability and ability to learn system behaviour. It mainly consists of three layers input, hidden and output layers which are interlinked with each other as shown in Fig. 4. For this study, 636 gaps are used, which consists of 348 accepted and 288 rejected gaps. For network training and testing, 70% and 30% data is used. The input layer consists of 10 units which represent the 10 input variables considered in the study. The hidden layer consists of 7 units and uses hyperbolic tangent as an activation function and output layer consist of 2 units and uses identity as an activation function. Table 8 gives the input variable with their importance and normalised importance. The normalised

Table 7 Pearson correlation and ANOVA test results for the accepted and rejected gaps

		Gap accepted or rejected at		
		Bagsewniya	Piplani	Ratnagiri
Gap accepted or rejected at Bagsewniya	Pearson correlation	1	0.108	−0.008
	Significance	–	0.188	0.926
	F value	–	1.751	0.009
Gap accepted or rejected at Piplani	Pearson correlation	0.108	1	−0.33
	Significance	0.188	–	0.636
	F value	1.751	–	0.225
Gap accepted or rejected at Ratnagiri	Pearson correlation	−0.008	-0.33	1
	Significance	0.926	0.636	–
	F value	0.009	0.225	–

importance is calculated by dividing the considered importance with the maximum value of importance for the data.

This study uses connection weights approach to determine the relative importance of input variables. In Fig. 4, blue lines show the synaptic weight less than 0 while the white lines show the synaptic weight greater than 0, and the thickness of line shows the strength of connections between the two neurons, i.e. input-hidden and hidden-output layers. From the analysis of 636 gaps with the help of SPSS tool, it is observed that 87.9% of training data is correct and 91.4% testing data is correct. Figure 5 shows the graph for normalised importance of the input parameter considered in the study.

Model 1, Model 2 and Model 3 are for location Bagsrewniya, Piplani and Ratnagiri and Model 4 is for the overall data. From the analysis of all the models, the percentage of correct prediction and area under the ROC curves are determined and this is tabulated in the Table 9.

The ROC curves plot the proportion of false positive (FP) and true positives (TP). TP represents sensitivity, and FP (1—specificity) represents true negatives. By examining the area under the ROC curve, which ranges from zero to unity, we collect further evidence as to which model was optimal and should be chosen over other models. It is interpreted that if the area under the ROC curve lies between 0.9−1, 0.8−0.9, 0.7−0.8, 0.6−0.7 and 0.5−0.6 it is excellent, good, fair, poor and fail, respectively. Figure 6 represents the ROC curve for the Model 4, and the value of area under the ROC curve for both accepted and rejected gap for all the models are in the range of 0.9−1 which is excellent according to the study of Department of Math of the University of Utah, and this shows that the model considered in the study is optimal and best suited over the other model based on the area under the ROC curve.

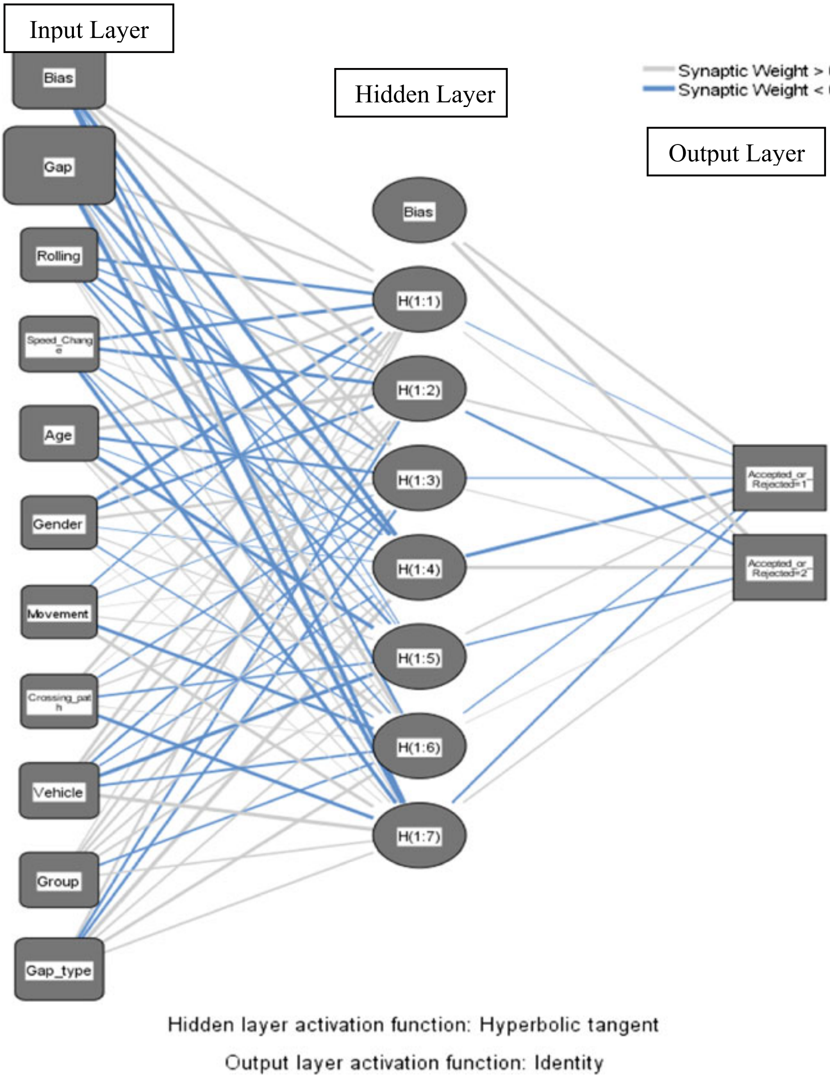


Fig. 4 Layers of ANN architecture with synaptic weight

5 Conclusion

This study concluded that the percentage share of male pedestrians is higher than that of female pedestrians, and young pedestrians acquire maximum percentage share among the age group category and pedestrian mainly used to travel as single at all the study locations. The speed of pedestrian varies based on the demographic characteristic, i.e. male used to travel at a speed of 0.95 m/s which is greater than female

Table 8 Details of input variable with normalised importance

Independent variables	Importance				Normalised importance			
	Model 1	Model 2	Model 3	Model 4	Model 1 (%)	Model 2 (%)	Model 3 (%)	Model 4 (%)
Gap size in seconds	0.282	0.387	0.483	0.474	100	100	100	100
Type of gap	0.173	0.166	0.116	0.183	61.3	42.9	24.0	38.6
Type of vehicles	0.090	0.095	0.077	0.065	31.8	24.6	16.0	13.7
Age group	0.122	0.072	0.073	0.054	43.4	18.7	15.0	11.4
Speed change condition of pedestrian	0.053	0.068	0.068	0.066	18.6	17.5	14.0	13.9
Group size	0.062	0.080	0.059	0.057	22.0	20.7	12.3	12.1
Gender	0.050	0.031	0.047	0.019	17.8	7.9	9.6	4.1
Movement of Pedestrians	0.068	0.029	0.035	0.019	24.2	7.5	7.2	4.0
Crossing path change condition	0.056	0.049	0.025	0.031	19.8	12.6	5.2	6.6
Rolling gap	0.045	0.022	0.019	0.031	16.1	5.8	3.9	6.5

Table 9 Percentage of correct prediction and area under ROC curve

		Model 1	Model 2	Model 3	Model 4
Overall Percentage of correct prediction	Training (%)	98.1	91.8	92.9	87.9
	Testing (%)	83.7	82.1	91.1	91.4
Area under the ROC curve	Accepted	0.984	0.960	0.961	0.955
	Rejected	0.984	0.960	0.961	0.955

speed 0.90 m/s. Young pedestrians travelling speed is 0.97 m/s, and this is higher than that of children speed 0.95 m/s, middle-aged pedestrians speed 0.89 m/s and old age pedestrians speed 0.84 m/s. Single travelling pedestrians moves at the speed of 0.96 m/s which is faster as compared to groups of two having speed of 0.85 m/s and three or more pedestrians having 0.88 m/s. Pearson correlation and ANOVA test show that the main factor influencing the pedestrian movement among the three demographic characteristic is age and group size having the value of significance level of 0.009 and 0.005, respectively, where as there is no significance difference between the speed data for the gender category having value of significance level of 0.068.

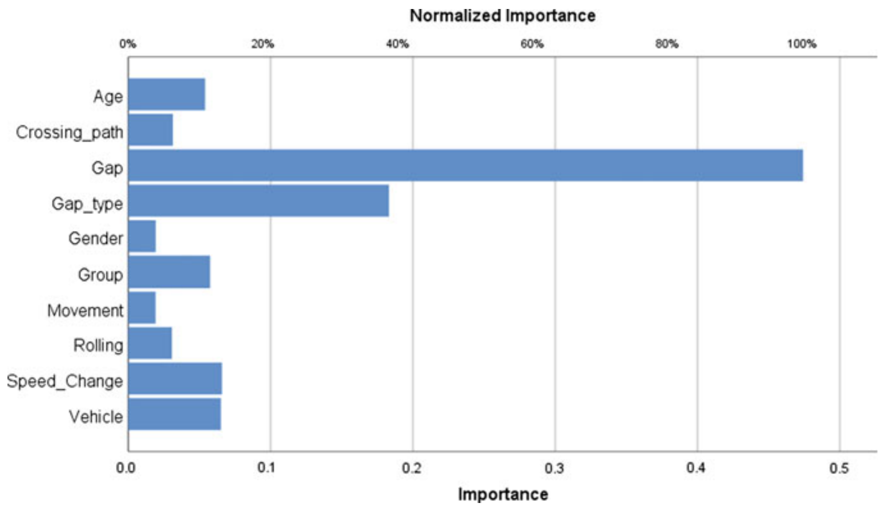
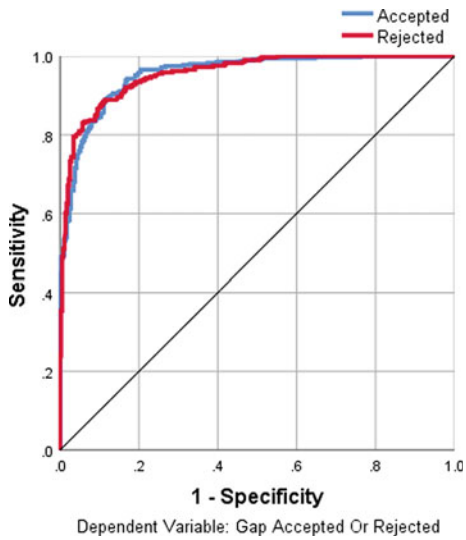


Fig. 5 Graph showing normalised importance of the input parameters

Fig. 6 ROC curve of the neural network



The critical gap for children is 2.2 s which shows that the children are the most risk taking pedestrian among the age group category which leads to severe accidents or fatality. And the critical gap for male is 2.4 s; this is less than female pedestrian's critical gap of 3 s which shows the aggressive behaviour of male towards the traffic while crossing the road at an urban uncontrolled intersection. Pedestrian travelling alone has the critical gap of 2.4 s while the critical gap for pedestrians travelling in group of two has critical gap of 2.3 s and group of three or more pedestrians has the critical gap of 3.3 s. This shows that when pedestrian travels in the group of two they show more aggressive behaviour towards traffic as compared to single and group of three or more pedestrians. According to the current analysis in Bhopal, the critical gap is assessed to be 2.5 s when all three junctions are considered, which is lower than the 3.6 s calculated using Raff's method in Pawar and Patil's [21] study. The ANN-based pedestrian decision-making behaviour model shows that the gap size is the most important parameter having importance value of 0.474 (normalised importance 100%) and movement of pedestrians, i.e. towards curb or median is least important parameter having the importance value of 0.019 (normalised importance 4.0%) among the considered 10 parameter. And it shows that age and group size has given more importance 11.4% and 12.1%, respectively than gender having importance of 4.1% only. It shows that age and group size of pedestrians matters while making the decision of road crossing regardless of gender of the pedestrians.

References

1. Laxman KK, Rastogi R, Chandra S (2010) Pedestrian flow characteristics in mixed traffic conditions. J Urban Plann Dev 136(1):23–33. https://doi.org/10.1061/(asce)0733-9488(2010)136:1(23)
2. Rastogi R et al (2011) Parametric study of pedestrian speeds at midblock crossings operational performance measure for intercity roads view project study on the effects of car restrictive policies on travel demand view project parametric study of pedestrian speeds at midblock C. https://doi.org/10.1061/(ASCE)UP.1943-5444.0000083
3. Marisamynathan Perumal, V (2014) Study on pedestrian crossing behavior at signalized intersections. J Traffic Transp Eng (Engl Ed) 1(2):103–110. https://doi.org/10.1016/S2095-7564(15)30094-5
4. Banerjee A, Roy SK, Maurya AK (2015) 'Study of pedestrian movement on functionally different types of Sidewalks in Kolkata, (December)
5. Das P, Katiyar MPVK (2015) Analysis of interrelationship between pedestrian flow parameters using artificial neural network. J Modern Transp 23(4):298–309. https://doi.org/10.1007/s40534-015-0088-9
6. Zheng Y et al (2016) Pedestrian traffic operations in urban networks. Transp Res Proc 15:137–149. https://doi.org/10.1016/j.trpro.2016.06.012
7. Dinakar, RJSJ, Kumar S (2017) Impact of vehicle pedestrian interaction on traffic flow: midblock & intersections. Int J Innovative Res Sci Eng Technol (An ISO), 6(2). https://doi.org/10.15680/IJIRSET.2017.0602168
8. Reddy SP, Teja Tallam DCNK (2019) Study of pedestrian flow parameters for urban signalized midblock 6(5):278–286
9. Transportation Research Board (TRB) (2000) Highway capacity manual. Special Rep. No. 209, National Research Council. Washington, DC
10. Amin HJ, Desai RN, Patel PS (no date) Modelling the crossing behavior of pedestrian at uncontrolled intersection in case of mixed traffic using adaptive neuro fuzzy inference system. https://doi.org/10.12720/jtle.2.4.263-270
11. Kadali RB, Rathi N, Perumal V (2014) Evaluation of pedestrian mid-block road crossing behaviour using artificial neural network. J Traffic Transp Eng (Engl Ed) 1(2):111–119. https://doi.org/10.1016/S2095-7564(15)30095-7
12. Li Q et al (2005) Pedestrian delay estimation at signalized intersections in developing cities. Transp Res A Policy Pract 39(1):61–73. https://doi.org/10.1016/j.tra.2004.11.002

13. Chao Q, Deng Z, Jin X (2015) Vehicle-pedestrian interaction for mixed traffic simulation. Comput Anim Virtual Worlds. https://doi.org/10.1002/cav.1654
14. Maurya AK, Amin HJ, Kumar A (2016) Estimation of critical gap for through movement at four leg uncontrolled intersection. Transp Res Proc. 203–212. https://doi.org/10.1016/j.trpro.2016.11.076
15. Kadali RB, Perumal V (2016) Critical gap analysis based on pedestrian behavior during peak-hour traffic at unprotected midblock crosswalks. Asian Transp Stud ATS 4(1):261–277. https://doi.org/10.11175/eastsats.4.261
16. Thakur S, Biswas S (2019) Assessment of pedestrian-vehicle interaction on urban roads: A critical review. Arch Transp 51(3):49–63. https://doi.org/10.5604/01.3001.0013.6162
17. Mukesh MS, Yashwan Katpatal B (2020) Evaluation of pedestrian safety in fast developing Nagpur city, India. J Urban Environ Eng 14(1): 2020. https://doi.org/10.4090/juee.2020.v14n1.052060
18. Vinayaraj S et al. (2020) Examining pedestrian critical gap analysis at un-signalized midblock crosswalk sections in India. Transportation Research Procedia. Elsevier B.V., pp. 2230–2250. doi: https://doi.org/10.1016/j.trpro.2020.08.280.
19. Cochran WG (1963) Sampling techniques, 2nd edn. John Wiley and Sons Inc., New York
20. Raff MS, Hart JW (1950) A volume warrant for urban stop signs, Eno Foundation for Highway Traffic Control. Saugatuck, Connecticut
21. Pawar DS, Patil GR (2016) Critical gap estimation for pedestrians at uncontrolled mid-block crossings on high-speed arterials'. Saf Sci 86:295–303. https://doi.org/10.1016/j.ssci.2016.03.011

Development of Pedestrian Level of Service (PLOS) for Foot Over Bridges and Skywalks

Arunabha Banerjee and Akhilesh Kumar Maurya

Abstract The performance evaluation and expansion/redesigning of a pedestrian facility depend on the Pedestrian Level of Service (PLOS). Indo-HCM (2017) has recently developed LOS for foot over bridges (FOBs) considering five locations across different Indian cities. However, no study has been conducted over elevated walkways such as skywalks. Hence, the current study tries to develop PLOS standards for overpass facilities (FOBs and skywalks). To achieve the above objective, videography data was collected over seventeen FOB and seven skywalk locations across six different Indian cities. Subsequently, the collected data was processed in the lab to extract the most significant parameters (flow rate, speed, density, and area module) along with microscopic factors (age, gender, luggage condition, use of hand-held devices, and group size). The *t*-test confirmed that the data for both the facilities should be kept segregated. Microscopic analysis revealed that majority of the users were male and young adults (21–40 years). The walking speed over skywalks was 6.5 m/min greater than that over FOBs. To develop the LOS ranges, equal data binning technique was applied. Thereafter, actual ground conditions were considered to define the LOS ranges. Finally, two separate LOS standards were developed which could replicate the actual condition of the existing overpass facilities. The outcome of the study would be beneficial to planners and government authorities in developing standards for grade-separated facilities such as FOBs and skywalks. This would help in developing improved and comfortable facilities for pedestrians.

Keywords Pedestrian level of service · Foot over bridge · Skywalk · Videographic survey

A. Banerjee (✉) · A. K. Maurya
Department of Civil Engineering, IIT Guwahati, Guwahati, Assam, India
e-mail: arunabhabanerjee77@gmail.com

A. K. Maurya
e-mail: akmaurya@gmail.com

L. Devi et al. (eds.), *Proceedings of the Sixth International Conference of Transportation Research Group of India*, Lecture Notes in Civil Engineering 272,
https://doi.org/10.1007/978-981-19-3494-0_19

1 Introduction

Pedestrians being the most vulnerable road users' have to face plenty of challenges while using different at-grade facilities such as sidewalks, walkways, and crosswalks. Majority of the facilities, are either curbed or unavailable, have poor connectivity or occupied by vendors/parked vehicles. This leads to the pedestrians using either the carriageway (instead of the sidewalks) or crossing at-grade illegally at mid-block opening instead of using proper crosswalk facilities. Thus, in order to protect the pedestrians, it is extremely important to construct proper grade-separated facilities in the form of foot over bridges (FOBs) and skywalks. The FOBs allow pedestrians to cross from one side of the road to the other without any interaction with the vehicular traffic. On the other hand, the skywalk facilities which are elevated walkways which allow pedestrians to travel over long distances (from a few hundred meters to a few kilometers) both across and along the length of the road. The skywalks can be the best substitute to the at-grade sidewalks and crosswalks with no interaction between the pedestrians and the motorized traffic.

Level of service (LOS) is a term closely related to the capacity of a facility; where LOS defines the qualitative measures and capacity defines the quantitative measures. The qualitative measures are regarded as the measures of effectiveness (MOE). The MOEs (speed, density, flow rate, space, delay, etc.) are the key measurable parameters which indicate the level of service of a particular facility. The LOS in general is defined by six ranges: A to F. LOS A defines the best operating condition while LOS F defines the worst operating condition. The best LOS operating condition indicates the free flow condition (at lower density) and worst LOS indicates congested condition (at jam density). Moreover, HCM defines LOS as "quantitative stratification of a performance measure or measures that represent the quality of service, measured on an A–F scale, with LOS A representing the best operating conditions from the traveler's perspective and LOS F the worst". The performance of a facility and the need to redesign it is defined by the LOS of a particular facility. Figure 1 shows the different LOS levels for sidewalk and crosswalk as per IRC-103 [1].

The earliest study related to quantitative LOS was conducted by Fruin [2] in the USA over stairway facilities where he tried to measure the space and flow rate. Studies over sidewalk facilities were conducted by Polus et al. [3] in Israel and Mori in Tsukaguchi [4] in Japan using videography technique. Thereafter, further studies [5–11] developed LOS standards for sidewalk facilities using density, space, flow rate, and pedestrian speed as the primary factors defined by four to six LOS ranges. Gerilla et al. [12] and Sahani and Bhuyani [13, 14] studied pedestrian flow rate and space over walkway facilities and defined six LOS ranges. The LOS levels were also defined by IRC-103 [1], US-HCM [15], and Indo-HCM [16] for sidewalk and stairway facilities. However, Indo-HCM [16] defined quantitative LOS for 5 FOB facilities across India based on flow rate and speed.

The lack of studies conducted for overpass facilities (FOBs and skywalks) encourages thorough data collection, analysis, and development of LOS norms, especially under Indian scenario. Thus, in the present study, an attempt is made to develop

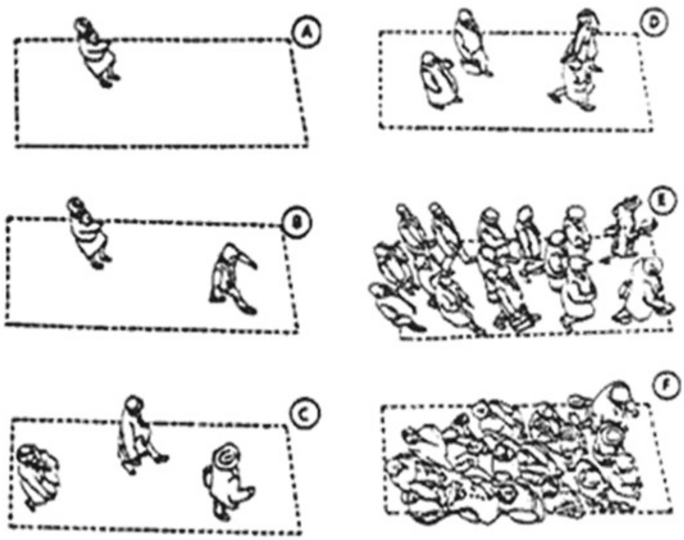


Fig. 1 Pedestrian LOS (Table 1 *Source* IRC: 103–2012)

the LOS standards for seventeen FOB facilities and seven skywalk facilities across six different Indian cities using quantitative (videographic) data collection technique. The outcome of the study would be beneficial in updating the existing LOS standards defined by Indo-HCM [16].

2 Data Collection and Analysis

In the present study, twenty-eight FOB sites and twelve skywalk sites were initially visited across six different cities (Delhi, Bangalore, Mumbai, Kolkata, Guwahati, and Gangtok). After thorough investigation, only those locations were finalized which were presently being used and had different flow levels throughout the day.

2.1 Data Collection Locations

Post thorough investigation, videographic data collection technique was used to collect data over the FOB and skywalk locations from the selected cities. All the skywalk locations were selected from Mumbai city; while the FOB locations were selected across all the six cities. The data were collected across different land uses types: commercial, public transport terminal, residential, educational, and shopping locations. During videographic data collection, the camera was installed at

Table 1 Geometric details of FOB and skywalk sections

Type of facility	Effective width (m)			Length (m)			No. of locations
	Min	Max	Mean	Min	Max	Mean	
FOB	1.2	4.7	1.95	18.5	88.1	41.7	17
Skywalk	2	3.4	2.71	315	1287	600	7

Fig. 2 **a** Skywalk facility and **b** FOB facility

a high vantage point of approximately 10–12 feet above ground level to capture the maximum distance covered by the pedestrians. The data were collected either during morning (8–11 am) or evening (3.30–6.30 pm) peak hours to capture the variability among the pedestrians. The FOB locations were selected such that no at-grade crossing was available within 50 m of the facility. Table 1 gives the aggregated details of the FOB and skywalk locations.

Figure 2 shows the two locations of skywalk and FOB facilities.

2.2 *Data Extraction and Analysis of Extracted Data*

Manual data extraction technique was used to extract the collected field data. Manual method, even though time-consuming, helps in extracting the macroscopic factors (speed, flow rate, density, and space), and the microscopic factors (such as age, gender, luggage carrying condition, use of mobile, and size of group). The total extracted videographic data was approximately 51 h for FOBs and 21 h for skywalks. The aggregation level for data extraction was 1 min.

In order to check whether to combine the data for both the facilities and segregate them, initially a t-test was conducted on the free flow speed data and it was observed that *t*-critical value was more than the *t*-statistical value. Hence, for LOS model development, the data for both the facilities were kept segregated. Table 2 gives the demographic characteristics of the pedestrians using the two different facilities.

From Table 2, it is observed that majority of the overpass users' were male pedestrians in the age group of 21–40 years across both the facilities. The percentage of pedestrian with luggage was higher over skywalks (66.4%) in comparison with FOB facilities (51.8%). Moreover, there were more pedestrians using hand-held devices (mobile phones) while using skywalk facilities. Also, the group size shows that majority of the pedestrians (74–79%) were walking alone or in a group of two pedestrians (15–17%).

The variation in flow rate and speed across both the facilities is given in Table 3. The table also gives the variation of minimum, maximum, mean, and standards deviation values for skywalk and FOB facilities.

From Table 3, it is observed that flow rate varies between 2.96–118.51 ped/min/m and 0.48–80.83 ped/min/m over skywalk and FOB facilities, respectively. The average per minute flow rate is observed to be significantly higher in case of skywalks than in case of FOBs. The speed ranges between 50–98 m/min across both the facilities, with pedestrians using skywalks having higher walking speed (by 6.5 m/min) than the ones using FOBs.

2.3 Development of Pedestrian Level of Service (PLOS) for FOB and Skywalk

In the present study, speed and flow rate were used as measures of effectiveness (MOE) to define the LOS ranges for the overpass facilities. The different techniques which can be used for data categorization are clustering techniques (such as Hierarchical clustering and *K*-means clustering) and equal data binning technique. However, if data is unavailable in the moderate to higher density ranges, the classification into different clusters becomes difficult. Moreover, previous literature shows that majority of the past studies preferred using equal data binning technique to cluster the data into different ranges and thereafter adjusted the ranges of the bins as per the field conditions [17]. In the present study, unsupervised equal data binning/equal width discretization technique was applied to define the LOS ranges. Apart from equal data binning technique, actual ground conditions (i.e., visualization as an expert) were also used to come up with the best ranges. In certain locations, it was observed that the ground condition and the classified levels did not match. In such cases, the width of the bins were adjusted to replicate the actual ground conditions. For example, after applying equal width discretization method, skywalk 1 was observed to have a service level of *C*; however, visual observation/ ground condition indicated a worse service level (say *D*) for the same location. In such cases, the bins

Table 2 Microscopic features of pedestrians using overpass facilities

Facility	Age (%)					Gender (%)		Luggage condition (%)		Hand-held device (%)		Group size (%)			
	< 10	– 20	– 40	– 60	> 60	Male	Female	With	Without	With	Without	1	2	3	≥ 4
FOB	1.9	9.8	62.4	21.6	4.3	69.7	30.3	51.8	48.2	7.9	92.1	79.5	15.5	3.7	1.3
Skywalk	1.1	11.0	64.9	18.7	4.3	77.1	22.9	66.4	33.6	14.7	85.3	74.2	17.2	5.5	2.6

Table 3 Variation in flow rate and speed across skywalk and FOB facilities

Measure	Type of facility	Minimum	Maximum	Mean	Standard deviation
Flow rate (ped/min/m)	Skywalk	2.96	118.51	25.31	22.28
	FOB	0.48	80.83	12.22	9.83
Speed (m/min)	Skywalk	54.15	94.12	78.65	5.85
	FOB	50.14	97.52	72.04	6.84

were adjusted such that the visual observation was taken into consideration as well while fixing the ranges of the service levels. Tables 4 and 5 give the range of the quantitative LOS for different FOBs and skywalks combined together.

Tables 4 and 5 give the variation of LOS for FOB and skywalk facilities over six different LOS ranges. For skywalk, the facility is expected to perform at a LOS *A* at a flow rate of ≤ 23 ped/min/m and speed ≥ 73.4 m/min. On the other hand, when the flow rate is above 118 ped/min/m and the speed is lower than 47.3 m/min, the facility is expected to perform at the worst level of service (LOS *F*). In case of FOBs, the facility is expected to perform best at LOS *A* (when flow rate is below 16 ped/min/m and average walking speed is greater than 64.1 m/min) and perform worst at LOS *F* (when flow rate is above 78 ped/min/m and average walking speed is lower than 41.6 m/min).

Table 4 Skywalk LOS

LOS	Speed (m/min)	Flow rate (ped/min/m)
A	≥ 73.4	≤ 23
B	$> 65.6– < 73.4$	$> 23– \leq 43$
C	$> 59.5– \leq 65.6$	$> 43– \leq 68$
D	$> 53.5– \leq 59.5$	$> 68– \leq 92$
E	$> 47.4– \leq 53.5$	$> 92– \leq 118$
F	≤ 47.4	Variable

Table 5 FOB LOS

LOS	Speed (m/min)	Flow rate (ped/min/m)
A	≥ 64.1	≤ 16
B	$> 58.5– < 64.1$	$> 16– \leq 29$
C	$> 52.8– \leq 58.5$	$> 29– \leq 47$
D	$> 49.9– \leq 52.8$	$> 47– \leq 63$
E	$> 41.6– \leq 49.9$	> 63
F	≤ 41.6	Variable

3 Discussion and Application

In India, the pedestrian facilities are highly overlooked and mismanaged. Thus, development of proper overpass facilities (FOBs and skywalks) can encourage pedestrians to avoid risky at-grade facilities (sidewalks and crosswalks). The performance evaluation and its need for expansion of a facility depends on the level of service (LOS) of the facility. In the present study, an attempt was made to develop LOS standards for two different overpass facilities (FOB and skywalk). Videography data was collected over seventeen FOB and seven skywalk locations across six different Indian cities. In order to capture the variability among the pedestrians using the facilities as well different geometric dimensions, a wide variety of overpass locations were surveyed and only those which had significant pedestrian flow were selected for the study.

Subsequently, the data were extracted using manual data extraction technique for macroscopic factors (speed, flow rate, density, and space) and microscopic factors (age, gender, luggage carrying condition, mobile use, and group size). Thereafter, a *t*-test was conducted to check whether to combine the data for both the facilities or keep them segregated. The result of the *t*-test showed that significant different exists between the pedestrian behavior while using skywalk and stairway facilities. Thus for analysis, the two facilities were studied separately. It was observed that pedestrians using skywalks had higher walking speed (by 6.5 m/min) and the average speed ranged between 50 and 98 m/min across both the facilities. The microscopic analysis of the data revealed were male between 21 and 40 years was the dominant group. The proportion of pedestrians carrying luggage was lower in case of FOBs (51.8%), in comparison with the skywalk facilities (66.4%). The use of hand-held devices was more visible over skywalk facilities (14.7%). The analysis also revealed that higher proportion of pedestrians using both the facilities walked alone (74–79%), while 15–17% pedestrians walked in group size of two. Further analysis presented that the average per minute flow (or flow rate) was higher in case of skywalks (3–119 ped/min/m) in comparison with FOB facilities (0.5–81 ped/min/m).

In order to develop the LOS standards for both the overpass facilities, unsupervised equal data binning/equal width discretization technique was applied. Apart from equal data binning technique, actual ground conditions (i.e., visualization as an expert) was used to come up with the best ranges. In certain locations, it was observed that the ground condition and the classified levels did not match. In such cases, the width of the bins were adjusted to replicate the actual ground conditions. For example, equal data binning indicated a LOS of B for location 1 while visual observation indicated a LOS *C* for the same location. Hence, the ranges of the service levels were adjusted such that both the data binning technique as well as visual observation were taken into consideration. After the development of the quantitative LOS tables for skywalks and FOBs, each location was compared with the developed ranges. It was observed that majority of the skywalk were performing at LOS *C*–*D*; while in the case of FOBs, the variation of LOS ranged between *A* and *D* with majority of the locations performing at LOS *B*–*C*. In general, the facilities are designed based on LOS *C*. Hence, it is essential to redesign/expand the facilities in near future which are plying at LOS *C* or *D*. This would encourage more pedestrians to use such overpass locations.

Further, Table 6 gives the comparison between the existing LOS standard developed in the present study with Indo-HCM [16] for FOB facilities. From the table, it can be observed that variability exists in speed ranges defined for different LOS levels in both studies. As the present study considered more number of locations (in comparison of Indo-HCM), the developed ranges are expected to give a better understanding of the current scenario of the overpass facilities. The outcome of the study would be beneficial to planners and government authorities in developing standards for grade-separated facilities such as FOBs and skywalks. This would indirectly help in developing better facilities with more comfort and convenience to pedestrians.

Moreover, in order to calculate the quantitative LOS for FOB or skywalk facilities, the following steps can be used (refer to Fig. 3).

The procedure used and the LOS ranges defined in the present study can be useful in estimating the LOS for the elevated facilities (skywalks and FOBs). This can be used as guidelines for study over the respective facilities.

Table 6 Comparison between PLOS for FOB facilities

Study	Measure of effectiveness	Pedestrian level of service (PLOS) ranges					
		A	*B*	*C*	*D*	*E*	*F*
Present study	Speed (m/min)	≥ 64.1	> 58.5– < 64.1	> 52.8– ≤ 58.5	> 49.9– ≤ 52.8	> 41.6- ≤ 49.9	≤ 41.6
	Flow rate (ped/min/m)	≤ 16	> 16–29	> 29–47	> 47–63	> 63	Variable
Indo-HCM [16]	Speed (m/min)	≥ 56.8	> 55.1–56.8	> 51.7–55.7	> 45.6–51.7	> 30.9–45.6	< 30.9
	Flow rate (ped/min/m)	≤ 12	> 12–17	> 17–27	> 27–38	> 38–52	Variable

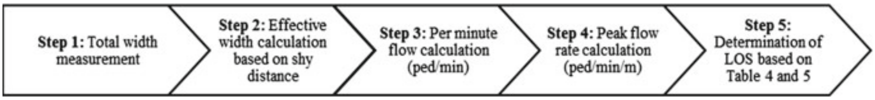


Fig. 3 Steps in determining LOS for overpass facilities

References

1. Indian Rroads Congress (IRC): 103 guidelines for pedestrian facility. Indian roads congress, India (2012).
2. Fruin JJ (1971) Pedestrian planning and design. Metropolitan Association of Urban Designers and Environmental Planners, New York, pp 2–6
3. Polus A, Schofer JL, Ushpiz A (1983) Pedestrian flow and level of service. J Transp Eng 109(1):46–56
4. Mōri M, Tsukaguchi H (1987) A new method for evaluation of level of service in pedestrian facilities. Transp Res Part A: Gen 21(3):223–234
5. Sisiopiku VP, Byrd JE (2006) Comparison of level-of-service methodologies for pedestrian sidewalks (No. 06-0898). Transportation Research Board
6. Jensen SU (2007) Pedestrian and bicyclist level of service on roadway segments. Transp Res Rec 2031(1):43–51
7. Al-Azzawi M, Raeside R (2007) Modeling pedestrian walking speeds on sidewalks. J Urban Plann Dev 133(3):211–219
8. Televska JB, Todorova MM (2012) Contribution to sidewalk pedestrian level of service analysis. Transp Prob 7(3):5–13
9. Rastogi R, Chandra S, Mohan M (2014) Development of level of service criteria for pedestrians. J Indian Roads Congr 75(1):61–70
10. Marisamynathan S, Lakshmi S (2016) Performance analysis of signalized intersection at metropolitan area. J Adv Res Appl Sci Eng Tech 2(1):19–29
11. Cepolina EM, Menichini F, Rojas PG (2018) Level of service of pedestrian facilities: modelling human comfort perception in the evaluation of pedestrian behaviour patterns. Transp Res F Traffic Psychol Behav 58:365–381
12. Gerilla GP, Hokao K, Takeyama Y (1995) Proposed level of service standards for walkways in Metro Manila. J East Asia Soc Transp Stud 1(3):1041–1060
13. Sahani R, Bhuyan PK (2017) Pedestrian level of service criteria for urban off-street facilities in mid-sized cities. Transport 32(2):221–232
14. Sahani R, Praveena D, Bhuyan PK (2016) Use of multinomial logit model in evaluation of service levels of pedestrian facilities. In: International conference on transportation planning and implementation methodologies for developing countries (12th TPMDC), IIT Bombay
15. Highway capacity manual (US-HCM): Transportation Research Board, National Research Council, Washington, DC (2010).
16. Indian highway capacity manual (Indo-HCM): Council of Scientific and Industrial Research, New Delhi, India (2017)
17. Nag D, Goswami AK, Gupta A, Sen J (2020) Assessing urban sidewalk networks based on three constructs: a synthesis of pedestrian level of service literature. Transp Rev 40(2): 204–240

Modelling the Effect of Variables Influencing the Usage of Zebra Marking by Pedestrians at Urban Midblock

Sathya Prakash and Krishnamurthy Karuppanagounder

Abstract A pedestrian using a facility like a zebra crossing to cross reduces the potential risk of accidents because using zebra marking to cross is the legal way for pedestrians over drivers. The study aims to compare the behaviour of pedestrians for locations with and without a facility. Further, the paper identifies the variables influencing the usage of zebra marking using a binary logit model. The analysis shows a significant influence of gender, age, platoon and crossing behaviour of pedestrians in choosing zebra marking area for road crossing. The model explained 17.8% variance in choosing zebra and no zebra marking area and was able to identify 72.5% of cases accurately. The results show that for every unit increase in platoon size, the odds of being using zebra marking area are 1.128 times, respectively. A unit decrease in stage number of crossing increases the chance of using zebra marking by 1.32 times. Compared to pedestrian making path changing behaviour, the pedestrian making straight crossing behaviour is 5.65 times likely to use zebra marking area for road crossing. The result identifies variables like the female, straight crossing behaviour and increase in platoon size positively influence using zebra marking. However, the other variables like male, adult and stage of crossing negatively influence using zebra marking. Also, by comparing driver non-yielding behaviour, the driver non-existence is higher at the zebra marking area. This study helps in understanding the behaviour of pedestrians that will help in improving the crossing facility effectively.

Keywords Pedestrian safety · Midblock crosswalk · Safety · Pedestrian facility

1 Introduction

Pedestrians are one of the elements in the traffic system; predicting the behaviour of pedestrians is challenging compared to other traffic elements. During their travel activity, the exposure to the road is high when they try to cross one side of the road at the other end. Further, the risk is very high at unsignalised locations. There are several

S. Prakash (✉) · K. Karuppanagounder
Civil Department, National Institute of Technology Calicut, Kerala, India
e-mail: sathyaprakash.civil@gmail.com

L. Devi et al. (eds.), *Proceedings of the Sixth International Conference of Transportation Research Group of India*, Lecture Notes in Civil Engineering 272,
https://doi.org/10.1007/978-981-19-3494-0_20

factors that influence on pedestrians to decide on crosswalks. Also, the behaviour varies based on demographic and psychological characteristics. A vulnerable wrong decision by pedestrians will expose road users at risk. An earlier study shows on average a pedestrian is killed every 88 min in a traffic crash in the USA in 2017 [1]. In India, 15% of pedestrians were killed in a road accident in 2018 [2]. Also, the study shows that pedestrians and cyclists are more vulnerable road users [2]. Among the road user type, pedestrian has the highest share of road death in India, 30–60% [3]. The most pedestrian collision occurs at a road crossing [4]. As a result of the high risk faced by the pedestrian, the study on effective usage of facility is required.

2 Literature Review

This section reports an earlier study on pedestrian behaviour towards the crossing facility. In a developing country like India with mixed traffic, pedestrians are the key elements in using the local transport system. Walking constitutes 35% of average commute trips in India [5]. From the previous literature study, the at-grade crossing is preferred by pedestrians over the grade-separated crossing [6–10]. The designed crossing facility for a pedestrian will be effective only when there are at the right place; inappropriate location of facility may not help in increasing the road user safety [11]. The pedestrian accident risk at a particular section depends on previous accident decisions [12]. Study on pedestrians at crosswalks helps understand the behaviour and further helps in developing the facility at right place. A study in Malaysia by [13] shows that an increase in vehicle speed increases the chance of pedestrians using zebra marking to cross. Also, the increase in age has a positive influence on using zebra marking. The study by [13] uses variables like speed, gender, age and baggage have influencing factors in deciding to use the zebra crossing. Another study by [14] shows there is a reduction in pedestrian crossing speed with an increase in a vehicular gap at unprotected midblock.

Further on studying pedestrians using zebra marking, the driver's willingness to give way for a pedestrian at zebra marking location is low [15]. A study by [16] at the Ghanaian metropolitan area shows pedestrians observed using zebra marking were alone and they are using mobile phones while crossing the road. This shows group behaviour is less at zebra marking locations. The pedestrian crossing prediction model was developed by [17] using context model tree for early crossing prediction. However, stage crossing, platoon size (pedestrian group size), driver yielding behaviour and crossing pattern of pedestrian (straight or path changing) can have some influence on using a crossing facility. Also, a study by [18] using simulation shows male proportion using single stage is higher compared to female. For the present study, the other variables mentioned above can be included to understand the usage of crossing facilities better. The next section describes the data collection techniques, extraction, basic statistical analysis and model development.

Table 1 Roadway and traffic characteristics of selected sites

Characteristics		Name of selected sites	
	Mavoor road	Kallai road	I.G. road
Land-use type	Bus terminal	Hospital	Hospital
Sidewalk	Yes	Yes	Yes
	Zebra marking—1	Zebra marking—1	Zebra marking—1
Type of facility	No zebra marking—1	No zebra marking—1	No zebra marking—1
Road width in '*m*'	14	11	14

3 Data Collection and Extraction

For the present study, the data collection is done at Kozhikode, Kerala, India, from the selected three mid-block sites, videographic survey was conducted simultaneously at zebra and no zebra marking area. Among the selected sites, two are near to hospitals and one at a bus terminal, total sites including zebra and no zebra marking area are six, detail characteristics of sites given in Table 1. The image of selected site locations is shown in Fig.1. The survey was conduction 6 h (peak hour) in a day with 3 h in the morning and 3 h in the evening, for three days on all six sites. The data extraction was done from the recorded video using custom made video player and extracted in excel sheets. The variables included in the study given in Table 2, where the pedestrian using zebra marking is dependent variable and independent variables are crossing behaviour, gender, age, platoon size, stage of crossing and driver yielding behaviour.

4 Descriptive Statistical Analysis

From the sample size of 12,990, basic statistical analysis was conducted and the proportion of pedestrians using zebra and no zebra marking is 72.3% and 27.7%, respectively. Among the gender, male using zebra marking and no zebra marking is 67.7 and 32.3%. Similarly, the proportion of females using zebra marking is 82.2%, and no zebra marking is 17.8%. Further, the combination of males and females moving in a group is 79% for zebra marking and 21% for no zebra marking. The age group proportion of pedestrians using zebra and no zebra were marking given in Table 3, and other traffic elements and the proportion of platoon size using zebra and no zebra marking is shown in Fig. 2.

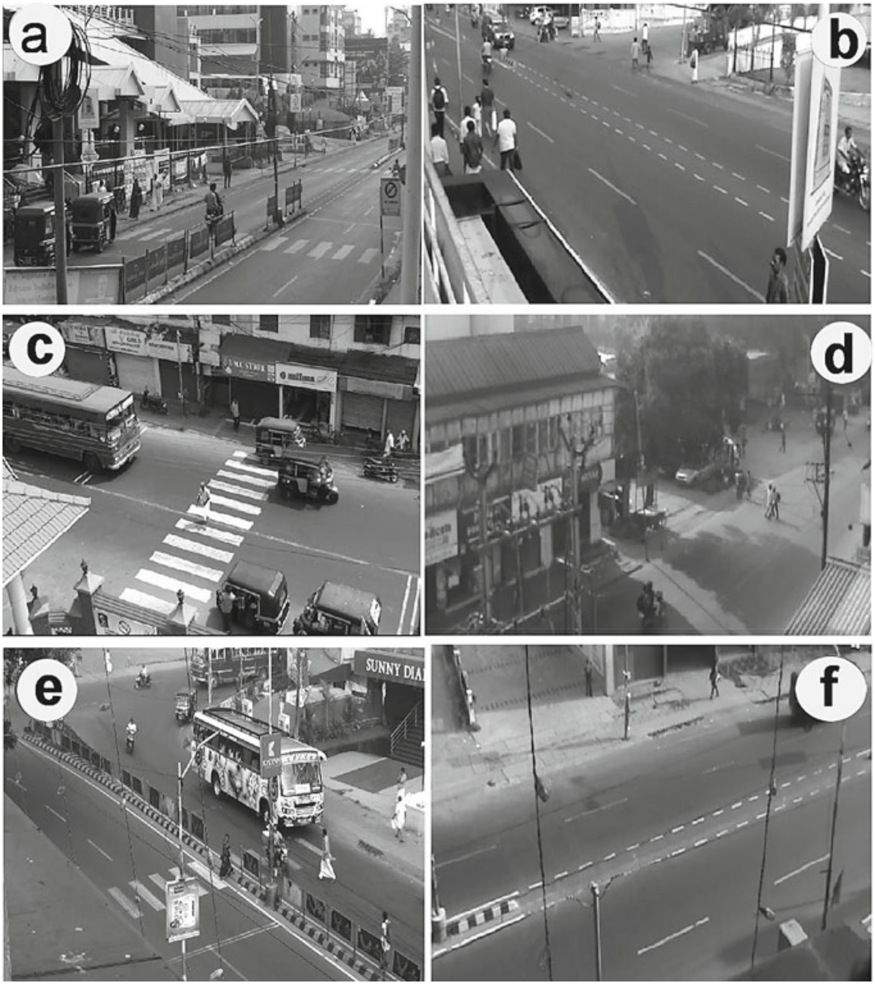


Fig. 1 **a** Mavoor road with zebra marking, **b** Mavoor road with no zebra marking, **c** Kallai road with zebra marking, **d** Kallai road with no zebra marking, **e** I.G. road with zebra marking and **f** I.G. road with no zebra marking

5 Pedestrian Choice of Selecting Zebra Crossing Using Binary Logit Model

From the data set of 12,990 pedestrians, a binary logit model analysis has been done using SPSS to predict the pedestrian zebra marking usage. Table 4 gives the results of logistic regression, based on the analysis, and the data does not have any outliers.

The logistic regression analysis shows there is a significant influence of gender, age, platoon behaviour and crossing behaviour of pedestrians in choosing zebra

Table 2 Variables considered in the pedestrian study

Variables	Type of variable	Code or units	Description
Facility type (dependent variable)	Discrete	No zebra marking—0 zebra marking—1	A pedestrian using zebra and no zebra marking area
Crossing behaviour	Discrete	Path changing—0 straight—1	Pedestrian making an irregular pattern to cross the road is path changing Pedestrian uses a straight path to cross the road is straight crossing behaviour
Gender	Discrete	Male—1 Female—2 Male + female—3	Male + female represents the group behaviour when both male and female moves in a group
Age group	Discrete	Old—0 Adult—1 Young—3 Child—4 Mixed—5	Mixed—represents more than one age group is involved in road crossing, and they move in a group
Platoon size	Continuous	No unit	More than one pedestrian crossing the road (group behaviour)
Stage crossing	Continuous	No unit	Pedestrian stops at different levels during road crossing, each stop is considered to be a stage
Driver yielding	Discrete	Driver yielding—yes Driver yielding—no observed driver behaviour—null	D.Y yes represents driver yield to pedestrian during road crossing D.Y no represents driver non-yielding behaviour to pedestrian during road crossing Null—no observed driver behaviour

Table 3 Proportion of pedestrian different age group using zebra and no zebra marking

Facility	Age Group %				
	Old	Adult	Young	Child	Mixed
Zebra marking	15.7	34.9	31.2	0.9	17.3
No zebra marking	18.4	37.1	32.9	0.8	10.8

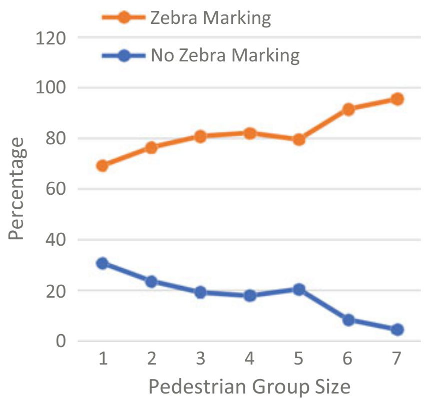


Fig. 2 Pedestrian group size in proportion with zebra and no zebra marking location

Table 4 Results of logistic regression analysis

	B [95% C.I.B]	S.E.(*B*)	Wald	Odd ratio
Crossing behaviour (straight)	1.732[5.06, 6.31] ☨	0.056	943.051***	5.653
Gender			153.749***	
Gender (male)	−0.197[0.69, 0.97]	0.088	4.965*	0.821
Gender (female)	0.515[1.39,2]	0.093	30.303***	1.673
Age group			29.200***	
Age group (old)	0.016[0.84–1.23]	0.097	0.026	1.016
Age group (adult)	−0.256[0.65–0.92]	0.086	8.784**	0.774
Age group (young)	−0.041[0.8–1.1]	0.091	0.201	0.960
Age group (child)	−0.129[0.54–1.4]	0.243	0.280	0.879
Platoon size	0.121[1–1.2]	0.037	10.612**	1.128
Stage	−0.280[0.67–0.84]	0.056	25.278***	0.756
Driver yielding			161.594***	
Driver yielding (null)	−1.441[0.16–0.34]	0.193	55.645***	0.237
Driver yielding (Yes)	−0.336[0.46–1]	0.216	2.428	0.715
Constant	1.868	0.241	60.280***	6.478

Omnibus $\chi 2$ (11) = 1708.22, $p < 0.001$, $R^2 = 0.123$ (Cox and Snell), 0.178 (Negelkerke)
*$p < 0.05$, ** $p < 0.01$, ***$p < 0.001$, ☨—95% C.I. for EXP(*B*)
Reference category
Crossing behaviour—path changing behaviour gender—(male + female) moving in a group age group—(mixed) moving in a group
Driver yielding—no

marking area for road crossing (χ^2 (11) = 1708.22, p<001). The model explained 17.8% variance in choosing zebra and no zebra marking area (Negelkerke R^2) and was able to identify 72.5% of cases accurately. The sensitivity of the model was 99.4%, and the specificity of the model was 2.2%. The results show that for every unit increase in platoon size, the odds of being using zebra marking area are 1.128 times, respectively. A unit decrease in stage number of crossing increases the chance of using zebra marking by 1.32 times. As compared to pedestrian making path changing behaviour, the pedestrian making straight crossing behaviour is 5.65 times likely to use zebra marking area for road crossing.

Further, compared to gender (male + female) moving in a group, the odds of choosing a zebra crossing to cross by the female pedestrian are 1.673 times higher. However, as compared to gender (male + female) moving in a group, for a unit decrease in male pedestrians, the chance of using zebra marking is increased by 1.22 times. Also comparing the mixed age group (moving in a group), a unit decrease in adult pedestrians will increase the chance of using zebra marking by 1.29 times. As compared to the non-yielding behaviour of the driver, a unit decrease in the non-existence of driver on the road increases the chance of using zebra marking by 4.22 times.

6 Discussion

A pedestrian using zebra marking to cross the road is a legal and safer way against the accident risk. Using a location other than zebra marking to cross is unsafe and increase the potential risk of a crash with vehicles. For a better understanding of behaviour, the zebra and no zebra marking compared with basic statistical analysis. The proportion of pedestrians using zebra marking is 72.3% over no zebra marking, earlier stated preference survey in Tanzania [9] reports more than half of the pedestrians prefer to the cross-level road over the grade-separated crossing. The prediction accuracy of the binary logit model is 72.5%, and also the study demonstrates better relation among the variables.

6.1 Effect of Pedestrian Platoon Size on Using Zebra Marking

The pedestrian platoon size shows the group behaviour of pedestrians. The result of the model shows a positive influence on platoon size in using zebra marking, which demonstrates pedestrian using no zebra marking is mostly of lesser group size or a single pedestrian. Also, the proportion of pedestrians using no zebra marking area is lesser compared to the zebra marking location.

6.2 Effect of Stage of Crossing on Using Zebra Marking

The stage of crossing is about how many times the pedestrian stops on the road in the course of making road crossing. From the present study, the stage of crossing has a negative influence on using zebra marking locations, which indicates that the increase in stages number is higher when they use no zebra marking. Because they have to wait for a suitable gap in every section of the road, but when they use zebra marking, the stages are lesser because it is the legal place for a pedestrian to cross and the drivers yield better for them.

6.3 Effect of Driver Yielding on Using Zebra Marking

When pedestrians using zebra marking area, the chance for the presence of driver presence (vehicle conflict) is very less. Based on the study, the driver behaviour impact on conflicting with a pedestrian is very less. Hence when using no zebra marking, the chance of driver non-yielding is higher for pedestrians. The model confirms the driver non-yielding behaviour is higher at no zebra marking.

6.4 Effect of Age and Gender on Using Zebra Marking

The female pedestrian has a positive influence on using zebra marking indicates a higher proportion of them are using zebra marking when compared to male. A similar study reported at Malaysia shows that female is more likely to use zebra marking over male pedestrians [13]. However, when considering males, they have a negative influence on using zebra marking. From the earlier study, males are likely to have more risk-taking behaviour while crossing the road [19]. Also, the adult age group has a negative influence to use zebra marking for road crossing.

6.5 Effect of Crossing Behaviour on Using Zebra Marking

Pedestrians making straight crossings have a positive influence on using zebra marking for road crossing. The study identifies pedestrians using zebra marking always make a straight pattern of the road crossing. However, pedestrians using no zebra marking location mostly make path changing patterns of the road crossing.

7 Conclusion

The analysis shows that the proportion of females using zebra marking is higher than males. Further, when the group size is increasing, usage of no zebra marking is increasing positively. The binary logit model describes the pedestrian behaviour predicting the utilisation of zebra marking. Also identifies the variable influencing the effective usage of the pedestrian facility, and the prediction accuracy of the model is 72.5%. Negelkerke R^2 in the model explains a 17.8% variance in using and not using zebra marking. Among the different variables, the straight crossing behaviour, platoon size and female variables have a positive influence on using zebra marking. However, the stage of crossing, male and adult variables have a negative influence on using the facility. Also, the chance of driver existence is lesser at no zebra marking, which makes pedestrians comfortable to use the facility effectively.

References

1. Pedestrian Safety, Administration N.H.T.S. (2019). https://www.nhtsa.gov/road-safety/pedestrian-safety. Accessed 16 Nov 2020
2. Road Accident India 2018, Ministry of Road Transport & Highways. Government of India (2018). https://morth.nic.in/sites/default/files/Road_Accidednts.pdf. Accessed 16 Nov 2020
3. Mohan D, Geetam T, Kavi B (2015) Road safety in India: status report transportation research and injury prevention programme. Delhi
4. Pedestrian safety: a road safety manual for decision-makers and practitioners. World Health Organization, 2004. https://doi.org/10.4271/pt-112
5. B-28 "Other Workers" by distance from residence to place of work and mode of travel to place of work-2011, census of India (2011). https://censusindia.gov.in/2011census/B-series/B_28.html. Accessed 16 Nov 2020
6. Moore RL, Older SJ (1965) Pedestrians and motor vehicles are compatible in today's world, traffic engineering
7. Tanaboriboon Y, Jing Q (1995) Chinese pedestrians and their walking characteristics: case study in Beijing. Transp Res Rec. 1441(1995):16–26
8. Räsänen M, Lajunen T, Alticafarbay F, Aydin C (2007) Pedestrian self-reports of factors influencing the use of pedestrian bridges. Accid Anal Prev 39:969–973. https://doi.org/10.1016/j.aap.2007.01.004
9. Mfinanga DA (2014) Implication of pedestrians stated preference of certain attributes of crosswalks. Transp Policy 32:156–164. https://doi.org/10.1016/j.tranpol.2014.01.011
10. Cantillo V, Arellana J, Rolong M (2015) Modelling pedestrian crossing behaviour in urban roads: A latent variable approach. Transport Res F Traffic Psychol Behav 32:56–67. https://doi.org/10.1016/j.trf.2015.04.008
11. Olszewski P, Szagała P, Wolański M, Zielińska A (2015) Pedestrian fatality risk in accidents at unsignalized zebra crosswalks in Poland. Accid Anal Prev 84:83–91. https://doi.org/10.1016/j.aap.2015.08.008
12. Yannis G, Golias J, Papadimitriou E (2007) Modeling crossing behavior and accident risk of pedestrians. J Transp Eng 133:634–644. https://doi.org/10.1061/(ASCE)0733-947X(2007)133:11(634)
13. Al Bargi WA, Daniel BD (2020) Modelling pedestrians' utilization of crossing facilities along urban streets. Case Studies Transp Policy. https://doi.org/10.1016/j.cstp.2019.12.004

14. Raghuram Kadali B, Vedagiri P (2020) Evaluation of pedestrian crossing speed change Patterns at unprotected mid-block crosswalks in India. J Traffic Transp (English Edition). https://doi.org/10.1016/j.jtte.2018.10.010
15. Varhelyi A (1998) Drivers' speed behaviour at a zebra crossing: a case study, accident analysis and prevention
16. Ojo T, Adetona CO, Agyemang W, Afukaar FK (2019) Pedestrian risky behavior and safety at zebra crossings in a Ghanaian metropolitan area. Traffic Injury Prevent. https://doi.org/10.1080/15389588.2018.1555372
17. Bonnin S, Weisswange TH, Kummert F, Schmuedderich J (2014) Pedestrian crossing prediction using multiple context-based models. In: IEEE 17th international conference on intelligent transportation systems
18. Zhang C, Zhou B, Qiu TZ, Liu S (2017) J Saf Res. https://doi.org/10.1016/j.jsr.2017.12.018
19. Marisamynathan S, Vedagiri P (2013) Modeling pedestrian delay at signalized intersection crosswalks under mixed traffic condition. Proc Soc Behav Sci 104:708–717. https://doi.org/10.1016/j.sbspro.2013.11.165