Arabian Journal for Science and Engineering
https://doi.org/10.1007/s13369-025-10088-y

RESEARCH ARTICLE-CIVIL ENGINEERING

Performance Evaluation of Fused Floating Car Data (FCD) and Bluetooth (BT) Data Speed Estimation on Urban Arterials

Beyhan Ipekyuz[1,2] · **Pınar Karatas Sevinen**[2] · **Hediye Tuydes Yaman**[2]

Received: 27 April 2023 / Accepted: 24 February 2025
© The Author(s) 2025

Abstract
Access to reliable data is crucial, particularly in emerging economies, where data may be scarce. Various collection methods, such as commercial Floating Car Data (FCD) and Bluetooth (BT) traffic data, which are low-cost and widely used methods making them ideal for monitoring and analyzing networks, provide valuable insights. This study presents "Link-Based Generalized Data Fusion Method (LB-GDFM)", which integrates and fuses commercial FCD with BT data to assess and manage traffic congestion on urban arterial links. By leveraging the strengths of both data sources, this fusion method provides a more accurate speed estimation and traffic state assessment of urban links and offers a robust comparative method for determining average link speed using fused FCD and BT data, along with performance results. The method, which involves determining a BT-based fused FCD speed, was applied to selected links of an arterial road in Mersin, Turkiye. Two data collection methods were compared to assess the effectiveness of measuring speed reliability across the links over a corridor. Performance metrics such as $\mathrm{MAE}_{FCD,BT}$, $MAPE_{FCD,BT}$, and $\mathrm{RMSE}_{FCD,BT}$ were used to evaluate the accuracy of both data sources, demonstrating that BT-based fused FCD and BT data provide reliable speed estimates in both heavy and normal traffic scenarios. Results indicated that even with limited data sources, effective traffic management strategies can be developed in emerging economies. This study, therefore, provides a valuable data fusion model for traffic monitoring and emphasizes the potential to enhance the used approach by incorporating additional data sources and broader traffic patterns in developing countries.

Keywords Floating car data (FCD) · Bluetooth (BT) data · Speed estimation · Data fusion · Urban arterial link

1 Introduction

The rapid growth and increasing complexity of urban transportation networks have made accurate traffic data analysis necessary. Accurate modeling of traffic density, especially in urban arterials, is critical for effective transportation management and planning. To be able to operate the urban transportation network, collecting data from scattered urban geography composed of many road sections is very challenging, as it traditionally requires the installation of multiple traffic sensors, which is not economically feasible. Thus, a new generation of traffic monitoring data has emerged from the use of technologies and commercial products that provide traffic data. Data obtained from different sources offer an important opportunity to estimate traffic characteristics such as traffic flow, density, travel time, and speed more precisely. Both Floating Car Data (FCD) and Bluetooth (BT) data stand out as two primary data sources widely used in monitoring traffic conditions in urban areas. While they do not provide direct counts of flow or density, they still enable the estimation of traffic flow model parameters such as speed (or travel time). FCD includes dynamic data covering speed, location, and travel time obtained through GPS devices integrated into vehicles, along with detailed real-time information about traffic conditions. There are different types of FCD, such as commercial and probe-based FCD. Commercial FCD includes travel time or speed data provided by

✉ Beyhan Ipekyuz
beyhan.ipekyuz@dicle.edu.tr; beyhan.ipekyuz@gmail.com

Pınar Karatas Sevinen
pinarkaratas8@gmail.com

Hediye Tuydes Yaman
htuydes@metu.edu.tr

[1] Department of Civil Engineering, Dicle University, Diyarbakır 21280, Turkey

[2] Department of Civil Engineering, Middle East Technical University, Ankara 06800, Turkey

Published online: 10 April 2025

private companies (i.e., Be-Mobile, TomTom, INRIX, etc.). It is aggregated and processed track data from a variety of vehicles, including trucks, fleets, and taxis, while probe-based FCD has more details regarding individual vehicle characteristics in addition to those. Commercial FCD is provided for specific time intervals from predefined road segments.

FCD enables dynamic traffic detection at a lower cost compared to other data collection techniques such as loop detectors, magnetic loops, road counters, and radar [1]. However, FCD may have accuracy limitations that require further improvement. BT data use signal detection times from road-installed devices to estimate travel time and traffic density. BT-based traffic monitoring is another low-cost technique that enables BT readers (BTRs) to detect vehicles via active BT devices that have unique Media Access Control Identification Addresses (MAC-IDs). This helps detect vehicle movements and their travel speeds based on BTR locations. The increased use of BT in vehicles and mobile devices has made this technology a prominent data source for traffic studies, particularly for estimating travel time and origin–destination (OD) patterns. However, BT data requires preprocessing and filtering to eliminate signals from stationary devices or those deviating from BTR-to-BTR routes. FCD and BT data play a critical role in reducing traffic congestion, optimizing travel times, and improving overall transportation efficiency.

FCD and BT can provide real-time traffic information required for intelligent transportation system (ITS) applications to operate effectively. The widespread use of ITS applications can lead to a more efficient use of existing transportation infrastructure. ITS could revolutionize traffic management and planning by using elements such as advanced data analytics, sensor technologies, and connected vehicles that utilize traffic data. While ITS has great potential to improve traffic management in emerging economies, data collection processes can be more complex and challenging due to insufficient infrastructure, limited financial sources, and technological capabilities. Therefore, it is crucial to thoroughly analyze existing traffic data and/or integrate it with other data sources. The joint use of different data sources can be achieved through various approaches, such as combining, integrating, adapting, fusing, and comparing data. Data-fusing approaches may help mitigate the constraints of individual data sources and increase confidence, robustness, and spatial coverage of the estimates. Using information gathered from multiple sources instead of just one may have several advantages [2, 3]. Additionally, the characteristics of each data source should be carefully considered when fusing them.

Fusing transportation data obtained from different sources provides a significant opportunity to estimate traffic, speed, and travel time more precisely. Travel time is a basic metric for ITS applications [1], such as traffic management strategies, adjusting speed limits, and recommending alternative routes. Once the road section length and travel time are known, the speed can be easily determined. Thus, if estimated properly, traffic speed can help detect traffic issues, such as recurrent delays, bottlenecks, incident-related queueing, and the serviceability of the road network in both the spatial and temporal domains [4–7]. As a result, both FCD and BT data are prominent data sources widely used for the evaluation of traffic state, particularly in terms of travel time or speed.

There are many FCD and BT data studies in the literature, but none of them addressed commercial FCD with a data fusion approach. The proposed method involves fused FCD utilizing BTRs spatial data and BT data according to arterial link characteristics to provide more consistent estimates of traffic performance in urban networks. This study aims to make a contribution to improve transportation performance in urban arterials by multiple sources, which are different enablers of traffic data. The proposed method provides a fused data preparation as a systematic study using commercial FCD and BT for travel time and speed analysis. FCD and BT datasets differ in their spatial and temporal granularity; a structured data fusion methodology is applied to harmonize them. First, a preprocessing step is performed to clean, normalize, and filter both datasets, ensuring consistency in measurement intervals and spatial coverage. Next, FCD speed data are aggregated based on the distances between consecutive BT sensor locations, enabling direct comparability. The link-based speed estimation process is then carried out by computing travel times across FCD segments and aligning them with corresponding BT travel times. Additionally, a mid-to-mid travel time adjustment is introduced to account for the duration a vehicle remains within a BT detection zone. Finally, both datasets are synchronized using a common time window, allowing a refined speed estimation process that leverages the strengths of each data source.

Moreover, it should also be noted that the method could be applied in a generalized manner, making it adaptable to other networks. Another key contribution of this study is the introduction of a new approach, called the "Link-Based Generalized Data Fusion Method (LB-GDFM)", for analyzing the traffic conditions on urban links. The proposed generalized fusion method used follows an intuitive approach: commercial FCD is reliable and widely available but aggregated and pre-processed, whereas BT is low-cost but requires calibration. The proposed fusion approach not only enhances the accuracy of traffic speed estimation but also provides a more robust assessment of urban traffic conditions, particularly in regions with limited data collection infrastructure. In the data fusion process, speed calculations derived from the fused FCD and BT data contribute to more consistent estimates of traffic performance on urban links. The detailed

methodology and fusion framework are further elaborated in the subsequent sections.

The method was applied to a road in the Mersin province of Turkiye, a developing country, to increase the performance of the use of the limited data available. A simple approach that has been expressed in a detailed manner models the fusion method to determine which was more effective in determining the reliability of average speed along links of an arterial road in Mersin city. In this context, after applying the fusion method to the average link speed determination model was offered for assessing traffic states on urban arterials using average link speed from fused FCD and BT. Thus, an adaptive traffic speed estimation model to adapt BT property to FCD and evaluate the difference between average link speeds coming from both was developed. The analysis was conducted over four days in September 2022. The speed analysis results generated from FCD and BT were assessed by using different performance metrics commonly used in the literature, such as mean absolute error (MAE), mean absolute percentage error (MAPE), and root-mean-squared error (RMSE). Developing traffic management on an urban arterial requires understanding the average link FCD speed from various perspectives, and the model can serve as a simple and useful technique when two traffic data types are available simultaneously. Moreover, this basic approach has the ability to be used for further testing FCD with an additional and different dataset or data source. The key contributions of the study to the field of urban traffic analysis can be summarized as follows:

- A new approach called "Link-Based Generalized Data Fusion Method (LB-GDFM)," which includes the fusion of commercial FCD and BT data, is presented. The proposed method aims to make more reliable speed estimation on the basis of arterial road sections.
- The method determines link-based average speed estimation on arterial road sections by evaluating fused commercial FCD and BT data together.
- The developed fusion method is designed to be adaptable to different road networks and alternative data sources, making it a flexible approach for future applications.
- To the best of our knowledge, no prior study in the literature has investigated the fusion of commercial FCD with BT data for traffic performance evaluation. This research aims to fill this gap.
- The proposed method has been tested on a real network. It applied to an arterial road in Mersin, Turkiye, demonstrating how limited data sources can be effectively utilized for traffic management and planning.

A review of FCD and BT studies is provided in Sect. 2, followed by the methodology presented in Sect. 3, which had a formulation for both BT-based average link speed determination and estimation of average link speeds from commercial FCD. Section 4 introduced numerical results over an urban arterial in Mersin, Turkiye. Discussion is given in Sect. 5. A summary of the findings and recommendations for future studies are given in the final section.

2 Literature Review

The collection and evaluation of transportation data have transformed over the years with the development of many different technologies. The FCD and the BT traffic data, two emerging low-cost high-coverage traffic data collection options, could provide speed (or travel time) estimation for traffic networks addressed in various published studies [8–10]. Pattern recognition and congestion detection are commonly achieved using FCD and BT. As summarized below, the FCD literature, which primarily concentrates on traffic speed evaluation, typically either uses FCD alone to derive speed profiles and determine the traffic state or combines FCD with other data types like loop detectors and radars. BT studies mostly focused on the estimation of inter-BTR travel times for major arterials while also addressing the OD estimation problem.

2.1 FCD-based Studies

FCD has become widespread as a technology that can create large data sets by determining the travel speed and instantaneous location of vehicles. The probe vehicle-based methods are effective ways to figure out how long it takes to get from one place to another on a travel segment. FCD is probe-based traffic data, enabling, managing, and monitoring traffic state over a whole network [9]. The FCD is generated by GPS-equipped vehicles and produces the fundamental information broadcasting the position and time of a moving vehicle [5, 9, 11, 12]. There are two types of FCD: i) individual probe-based FCD and ii) commercial FCD. The commercial FCD provides average travel time and speed over a given segment during a chosen aggregation epoch time only, whereas the individually probed data provide detailed trajectory information including vehicle positions with queue entry/leave, acceleration, or deceleration in a time–space diagram.

FCD could be provided commercially, such as by INRIX, Be-Mobile, and TomTom [13]. While comparing three sources of traffic data, FCD speeds from INRIX with 5-min aggregation during 20 weekdays showed non-systematic inconsistencies caused by the time lag in recovery from slow speeds, which affected the speed–flow relationship. INRIX speeds were analyzed over a two-month period, taking into account both regular and atypical occurrences. The loop detector readings were frequently at least 6 min behind

INRIX speeds, and the effective sampling durations were between 3 and 5 min [14]. These results highlight the importance of filtering and quality control of the FCD speeds prior to any intervention to decrease non-systematic error. [15] provided a method for network-wide traffic state estimation, leveraging data from both loop detectors and FCD. [16] conducted a statistical analysis of the gaps between data from loop detectors and FCD in Lille, France. To foresee motorway congestion locations, [17] built a model that combined FCD and loop detector data. [18] offered a comparison between loop detectors and FCD to determine which is able to identify a traffic event first. [19] proposed combining FCD measurements taken at crossings with loop detector data to estimate queue times and traffic flows.

Some research has compared commercial FCD speed to ground truth (GT) speed to gauge its potential usefulness. [20] compared commercial FCD speed with video-based GT speed recorded on a specific urban arterial in Ankara, Turkiye, demonstrating a nonlinear connection with a robust correlation value of 0.82. Additionally, [21] suggested 42 km/h as a speed threshold to detect queues under uninterrupted traffic conditions, based on the commercial FCD speed in Ankara.

FCD, which is frequently used in the derivation of fundamental diagrams, is generally fused with other data types. To fuse data from loop detectors and probe vehicles, [22] looked at seven multi-sensor data fusion-based estimation algorithms and compared their accuracy. [23] suggested a way to make fundamental diagrams by fusing FCD with data from loop detectors. They found that by using two data sources at the same time, the fundamental diagram estimation error could be cut down to 0.04 for the 25% of drivers that produce FCD. [24] discovered that 30% of the connections were required to build a fundamental diagram consistent with the one generated from the whole network by merging data from loop detectors and probe vehicles. [25] proposed an FCD-based method that considers geographical and temporal differences in traffic states to enhance the fundamental diagram. They defined the uncertainty of the flow estimation and examined the effect of aggregating the FCD and traffic flow data. Also, they used the distribution features of the uncertainty in the FCD-based hourly traffic flow prediction to estimate vehicle emissions.

The most prevalent FCD comes from taxis and fleet vehicles. [26] conducted an investigation into a data fusion system that combines FCD from taxis and buses with GPS data gathered by taxi routing systems. The increase in GPS-equipped probe vehicles in traffic allows for the use of fusion. [27] implemented a data fusion procedure for a distinct traffic model, which is utilized to extract traffic status information for intersections where public transportation and individual traffic coexist, based on the registration and deregistration of public transportation vehicles at designated intersections. The probes, consisting of 1300 taxis and private fleet data from PTV and BMW, gather information to provide journey times and incident alerts. Using taxi data, the algorithm calculates both the current and previous travel times for each route connection.

2.2 BT-based Studies

Bluetooth® is a low-power wireless connectivity technology that allows devices to stream audio, transfer data, and broadcast information [28]. Ericsson invented this technology in 1994, and it uses the 2.4-GHz short-range radio-frequency spectrum. The working principle of the BT is to capture the specific MAC-ID of any device, which enables monitoring its movement by localizing multiple BTRs. Class-1 through Class-3 BTR types, with ranges varying from 1 to 100 m, determine the capture zone. Additional antennas on receivers can help to increase the capture zone. A MAC-ID is a 48-bit electronic identifier that is unique to each device and takes the form of "AB:CD:**:**:**:**". Each BT system produced data sets that included both travel time and segment speed [5].

In Indianapolis, USA, researchers first attempted to estimate freeway and arterial travel times using time-stamped MAC-ID from BTRs [29]. The analysis of directional travel times for both freeways and arterial corridors revealed that arterial data had a larger variance due to the impact of signals and noise introduced when motorists briefly diverted from the network. For quantifiable travel mobility metrics for a rural interstate highway in Indianapolis, travel time profiles were collected from three BTRs over a 12-week period [30]. The study focused on traffic volumes, travel times, and hours of delay, which exceed 10 min for study periods and also investigated the effects of crashes and opening/closing lane situations. Researchers found that travel time delay serves as a significant data source for validating the volume and capacity thresholds that trigger queuing conditions. [31] conducted a comparison between BT readings and loop detector data collected over a few kilometers of ground area in Toronto. A multi-layer simulation model called Traffic and Communication Simulation (TCS) was created to examine the theoretical features of Bluetooth Media Access Control Scanner (BMS) data, revealing that BT data may provide realistic travel time estimations even in real-time traffic monitoring [32].

BT data were used to find the average speed, travel time, and OD estimate for a Portland, Oregon, 2.5-mile suburban signalized arterial over 27 days [33]. To evaluate general reliability, the traditional GPS floating car method was employed, showing that larger data sets of BT recordings could effectively capture the performance characteristics. The calculation of travel time is essentially based on the time difference between two detections for the same MAC-ID by consecutive BTRs, which is also referred to as segment travel

Table 1 FCD attributes in the static table

Segment ID	Optimal tt (sec)	Segment length (m)	Road class	Optimal speed (km/h)	Max speed (km/h)	Center coord	Starting coord	Ending coord
1327614	1.33	30.27	1	82.06	80	36.758747, 34.544224	34.544353, 36.758835	34.544094, 36.758659
1327609	1.86	42.34	1	82.03	80	36.758957, 34.544535	34.544717, 36.759079	34.544353, 36.758835
1327608	1.86	42.34	1	82.03	80	36.759201, 34.544899	34.545082, 36.759323	34.544717, 36.759079
1327597	1.78	40.64	1	82.00	80	36.760492, 34.546814	34.546988 36.76061	34.54664 36.760374

Table 2 FCD attributes in the dynamic table

Segment ID	Travel time (sec)	Speed (km/h)	Timestamp
1327614	2.90	38	9/6/2022 0:16
1327609	3.81	40	9/6/2022 0:16
1327608	3.65	42	9/6/2022 0:16
1327597	3.05	48	9/6/2022 0:16

time [34]. Four different ways of measuring travel time were compared: first-to-first (F2F), last-to-last (L2L), average-to-average (A2A), and received signal strength indicator (RSSI). A t-test showed that the RSSI-based method was significantly different from the others. The accuracy of travel time estimation using BT technology was examined using four distinct estimators: minimum BT (Min-BT), maximum BT (Max-BT), median BT (Med-BT), and average BT (Avg-BT), which were employed as outlier detection logic [35]. The results indicated that the Min-BT and Med-BT logic provides more precise travel time estimations than the other two. The reliability of BT technology for travel time estimation is the primary focus of a follow-up study, which is defined as the percentage of devices captured per trip during the experiment [36]. This research demonstrated that BTRs have short-range antennas that detect devices that are closer to the receiver, thereby enabling them to provide more precise travel time estimations. However, the penetration rates decreased as the detection zone narrowed. The study of [37] focused on the extent of errors in detection time and travel time measurement and concentrated on measurement errors in arterial travel times using BTRs. The travel time measurements were modeled using standard deviation analysis and multiple regression, yielding reliable results. BT data was collected to analyze the prediction of near-future travel times on signalized intersections using the k-nearest neighbor (k-NN) as a statistical technique by identifying historical data on a 3.1-km arterial roadway in Ontario, Canada [37]. For real-time arterial travel time prediction, the most important technical problems are defined as the presence of outliers and inconsistencies in the order of available travel time data, which should be detected. A methodology was developed to distinguish vehicular movements and travel characteristics (specifically OD and corridor travel time estimation) from BT data in a case study of Ankara, Turkiye [38]. However, the study found no effectiveness in estimating OD. To estimate reliable OD matrices, the data must be compared with any other supportive data source. Further studies should enhance the data for validation purposes. The applicability of BT was examined at the Bosphorus Bridge in Istanbul, Turkiye, [39]. Results reflected challenges of travel time estimation with such limited data but established a linear relationship between speeds of different classes. Another study in Turkiye was conducted in Mersin to analyze travel times during peak hours and assess the travel time distribution between 07:30 and 09:30 for two weekdays [40]. The MAC-ID travel times that were captured in the study were filtered through statistical evaluation. The filtered travel times were then converted to speed for selected BTR pairings to estimate the Level of Service (LOS) by the Highway Capacity Manual (HCM). Careful filtering of the data yielded good estimations.

2.3 Studies Using Both FCD and BT Data

[41] compared travel times computed using BTRs to those obtained with floating cars. [42] proposed a method that uses a virtual probe car to estimate the amount of time that travelers spend in arterial traffic. A virtual probe is a simulated vehicle that travels from its point of origin to its final destination at a predetermined time. Instead of using data from actual probe vehicles, this method relied on loop detector data. [43] employed FCD as a reference to assess the reliability of the BT data and demonstrated that the travel times measured by BT sensors are not significantly different from the actual

travel times. The speed ranges were converted from time-to-speed measurements, and statistical analysis was conducted for four-speed ranges: 0–30 mph, 30–45 mph, 45–60 mph, and > 60 mph. [44] conducted a comprehensive FCD study on chosen segments to examine the travel time data supplied by the BT sensors. The study was performed on a 1.1-mile section of road in Maryland, USA, at a 5-min interval. Results revealed that the BT travel time data are of very high quality and are suitable for use as verifiable GT on freeways. [45] used BTRs installed in cars to get individual travel times. The unique MAC-ID of the BT device was recorded whenever a vehicle equipped with BT traveled through a detection point, whether it was upstream or downstream. The travel times were gathered through the process of matching the MAC-IDs. [46] developed an FCD-augmented approach to collecting spatiotemporal traffic data by measuring the BT travel times. [47] studied BT data processing technology and presented field-test results to demonstrate the applicability of the FCD-based method to gain link-based and network-wide travel times. [10] compared travel times gathered by BT sensors, INRIX, and TRANSMIT readers. The authors aimed to develop a method for comparing the reliability of FCD from various data sources, such as INRIX and TRANSMIT readers. It appears that the data acquired by BT sensors and the TRANSMIT readers are more in line with the GT data, and it even outperforms the data given by INRIX by a small margin. [48] presented a methodology to reduce congestion on a long section of a German freeway based on data obtained from three widely deployed sensor technologies: loop detectors, FCD, and low-frequency travel time data from BTRs.

[49] concentrated on field measurement, raw data analysis, and filtering the travel times using one of the fundamental statistical methods, such as Tukey's method or the interquartile range (IQR) method, utilizing RADAR as the GT data source. Following data filtering and processing, the results revealed adequate penetration rates and accurate travel time estimations. After filtering basic statistical methods using median and standard deviation, travel time estimation with BT detection systems (BDS) revealed more accurate and reliable results than commercial travel time estimations such as HERE, INRIX, and TomTom [6].

Although there are many studies in the literature on both FCD and BT data, there is no study on the use of commercial FCD with a data fusion method and also the evaluation of speed estimation performance by using BT data. For this reason, the importance of this study is not only the use of different data sources with the spatial fusion method but also the speed estimation in developing countries with limited data.

3 Methodology

The first step in managing the traffic problem in urban networks is to correctly identify the problem and understand how it is related to the underlying causes. Data sources are the most successful enablers for achieving that goal. Access to data, data gathering, penetration rate, and data analysis are crucial due to potential challenges, particularly in emerging economies like Turkiye. It is necessary to utilize available data and assess them to address transportation issues economically and effectively. Therefore, a comprehensive and fusion approach that may handle multiple data sources has been required. The study has introduced an adaptive comparison approach called the "Link-Based Generalized Data Fusion Method (LB-GDFM)" that utilizes commercial FCD and BT data. The proposed method provides a two-stage approach to fuse FCD and BT data. While FCD provides high-accuracy instantaneous traffic information, BT data provides large-scale estimates of travel time. Therefore, the main objective of the proposed method is to adapt these two datasets on a link-by-link basis to create a more accurate traffic performance model.

3.1 Preliminaries

3.1.1 FCD Description

In the study, the FCD type, which was commercially published by Be-Mobile, a Belgian company, was used. FCD in Turkiye is reported to come from more than 600,000 GPS-equipped vehicles. In the study of [50], FCD penetration rate in Turkiye was calculated as 3% and this corresponds to the total national population of 25.25 million registered vehicles. Considering the results of previous studies where the FCD penetration rate varied between 1.5% and 5%, it appears that it provides sufficient data for traffic speed estimation [51–53]. FCD has both static and dynamic data tables, including different attributes. The static data (see Table 1) includes basic design and geometric data for each segment ID, including segment length, road class, optimal travel time (tt), maximum (max) speed, as well as optimal speed (which is a company-specific measure not clearly stated), latitude and longitude coordinates of the center, and starting and ending of the segment. Segment lengths are shorter than 50 m (even the maximum length of a segment is 47 m, and the least one has almost 12,4 m on the corridor used in this study). The dynamic table (Table 2) comprised information that changed over time except for segment IDs such as travel time, time-based average speed, and timestamp, again marked by the same segment ID data, linking the static data to the dynamic one. The data come in 1-min intervals for each segment, but it is archived in 5-min intervals by the local provider due to storage problems. FCD availability in Turkiye was

observed mostly in and between major urban regions [20], which has higher reliability during daytime and working hours (7 a.m.–8 p.m.).

3.1.2 BT Data Description

A BTR scan (also called reading) detects any active BT device, including the capture time (discovery time), BTR identification number (BTR_ID), device MAC-ID, and type (Fig. 1). Discovery time reflects the EPOC capture time, which is a universal timestamp including both date and time information. MAC-ID is unique information for each device and has 18 characters. Device type indicates whether the captured device is a vehicle or not, with default brand information. It is necessary to characterize each MAC-ID by reviewing the data collected for it. In the study, the BTRs belong to the BT Class-1 type, covering a capture rate of 100 m diameter.

BTR-ID reveals information about the reading as the locations of the BTRs are fixed; follow-up of unique MAC-IDs across multiple BTR readings allows the estimation of speed, travel time, or movement for each BT device (hence the vehicles carrying them). A MAC-ID can belong to a vehicle or any other type of device, such as a cell phone, a television, and an internet receiver. The information obtained from the device type could lead to the classification of this data as a vehicle or not (non-vehicle). A thorough data preprocessing is required: each unique MAC-ID identified must be checked to confirm whether they are stationary BT devices in the network. For those non-stationary BT devices, travel time calculation by observing the same MAC-ID in consecutive BTR locations seems straightforward. However, there is no specific information about the exact location of the vehicle when the reading was done or whether the captured MAC-ID was from a vehicle or a device inside a vehicle. Some non-vehicle devices, such as cell phones or headphones, may be detected as "moving" simply because they are traveling inside a vehicle.

Deploying BTRs throughout the length of the roadway is one approach to achieving the goal. Each confirmation of receipt also includes a timestamp at the moment it was received. To calculate how long it takes to go somewhere, just subtract the timestamps from the two separate measurements. The mean speed could be determined by measuring the distance between BT sensors. Since BT sensors are sometimes positioned hundreds or even thousands of kilometers apart, their temporal resolution is substandard when compared to that of FCD.

3.2 Data Preparation and Fusion

In developing countries like Turkiye, due to limited data collection infrastructure, there is missing or incorrect information in the data sets. For this reason, before data fusion, a data cleaning and missing data completion process is applied to commercial FCD and BT data. Normalization and data cleaning processes are performed on the connection times in FCD and BT speed data. This process is necessary for the two data sets to become compatible.

In the second phase, FCD and BT data are analyzed and evaluated based on the links on a specific arterial. In this process, the speed profiles in FCD data are aggregated and thus fused according to BTR distances, and the new speed profiles obtained are compared with BT data. Thus, with this fusion process, the performance of existing data and the advantage of their combined use can be evaluated for the traffic situation on a specific link.

3.3 Link-Based Speed Estimation using Commercial FCD and BT

Examining speed distributions along corridor is necessary to identify recurrent patterns. The speed estimation using commercial FCD should start with the ordered FCD segmentation of a selected corridor part. Each FCD segment (x_i) along the road section is locally ordered in the direction of the traffic flow (Fig. 2). The road links can be defined as stretches between consecutive intersections. The segment numbers and geometry may change between road directions, especially on divided urban roads. For example, a study link may have 20 FCD segments in the east-to-west (E-W) direction, while there may be 22 segments in the opposite, west-to-east (W-E) direction. Additionally, the segment lengths may not be identical or an exact mirror image of the one in the opposite direction; thus, each direction should be treated as a separate link. To find the average link speed using commercial FCD speed, i) the amount of time spent traveling on each FCD segment should be estimated and added to calculate the "link travel time"; ii) the division of the link length by the "link travel time" will simply provide an average close to the "average link speed".

The concept of "Reoccurrence Time" simply stands for the time difference between two captures of the same MAC-ID by any BTR, while travel time is simply determined as the time difference of any MAC-ID between two different BTRs (called last-to-first in the literature). However, it is also necessary to consider the stay time duration within the capture zone of each BTR, for this purpose, half of the time spent in each BTR during the trip calculated within the scope of the study was also included in the link travel time calculations (mid-to-mid travel time).

If FCD-based speed is used against BT-based one, it is important to aggregate both datasets with the same time and space intervals and make them comparable. Several methodologies that have been examined for their accuracy in calculating travel times are based on statistical models; in a similar vein, the mean value of the FCD is averaged

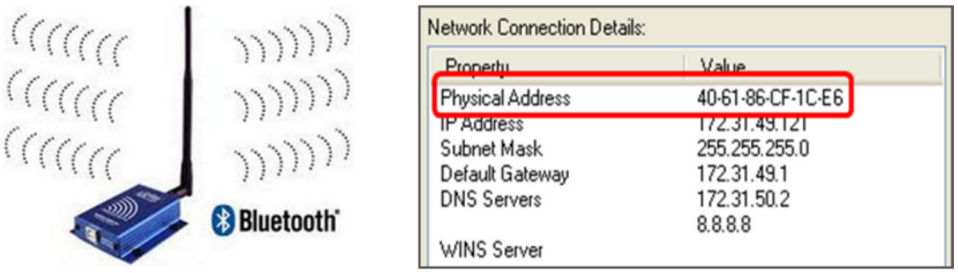

Fig. 1 BTR and BT data structure of any device

	BTRID	MacID	discoveryTime	deviceType
1	3367	x_04FE31ACFEF5	1562014803	0
2	3367	x_3021DA311DC4	1562014815	240408
3	3368	x_30C3D97E7424	1562014804	340408
4	3370	x_BC307EB3D3F1	1562014802	0
5	3370	x_AC7A4D5E98C0	1562014803	340408

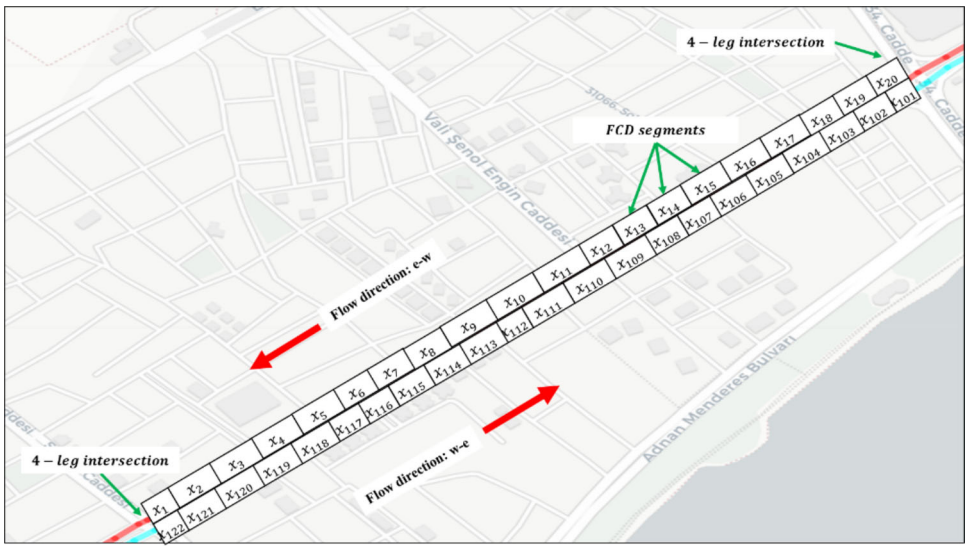

Fig. 2 FCD segmentation related to flow direction on a part of the road section

over several time spans [54]. Spatially, FCD segments can be aggregated between selected BTR locations; if there are multiple BTR locations along a road corridor, sub-corridors can be created for each consecutive BTR pair. Spatial processing is also required for time-wise aggregation; multiple FCD speed estimations for the segments (in 1-min or 5-min intervals) should be averaged over an aggregated duration and filter the BTR data to calculate the travel times of the BT devices detected in the same period. For example, to determine the 30-min mean speed at 7:00, the FCD travel times at 6 epoch times (7:00, 7:05, 7:10, 7:15, 7:20, and 7:25) are used. The speed profile of the link will be estimated for 30-min intervals through the total analysis time T, which could be taken as the daytime hours of 6:00–21:00. Speed profiles will be derived separately for both directions (E–W and W–E).

3.3.1 BT-based Link Speed Calculation

To begin the determination of fused FCD, the BT speeds should be calculated by deriving from travel times in a mathematical way. Readings of each unique MAC-ID, m (assumed to represent a vehicle), are ordered from the earliest to the latest at any BTR i and the next one, j. The notations for the speed calculation utilizing BT data are given in Table 3.

$\forall vehicle m$;

$$\begin{bmatrix} t_{i,m,1} \\ t_{i,m,2} \\ t_{i,m,3} \\ \cdots \\ t_{i,m,p} \end{bmatrix} \Rightarrow \delta t_{i,m} \text{ and } \begin{bmatrix} t_{j,m,1} \\ t_{j,m,2} \\ t_{j,m,3} \\ \cdots \\ t_{j,m,r} \end{bmatrix} \Rightarrow \delta t_{j,m}$$

Table 3 Notations for the fused commercial FCD and BT data components

x_i^{FCD}	Length of the FCD segment i
$L_{i,j}^{BT}$	Length of the link between BTR_i and BTR_j locations
$\Phi_{i,j,k}^{BT,FCD}$	Incidence matrix or association function of the $BTR_{i,j}$ and FCD segment k
$\omega_{k,\lambda}^{FCD}$	Travel time on the FCD segment k at epoch time λ
$\Omega_{i,j,\lambda}^{FCD}$	Link total travel time of FCD between BTR_i and BTR_j locations at epoch time λ
$\overline{\Omega}_{i,j,\tau}^{FCD}$	Average link total travel time of FCD at aggregated time interval τ between BTR_i and BTR_j locations
$u_{i,j,\lambda}^{FCD}$	Comparative average link fused FCD speed between BTR_i and BTR_j locations at epoch time λ
$\bar{u}_{i,j,\tau}^{FCD}$	Comparative average link fused FCD speed between BTR_i and BTR_j locations at aggregated time interval
$t_{i,m,p}$	p[th] reading time of vehicle m by BTR_i
$\delta t_{i,m}$	Stay time of vehicle m in the coverage area of BTR_i
$\Omega_{i,j,m}^{BT}$	Link travel time of vehicle m between BTR_i and BTR_j locations
$\overline{\Omega}_{i,j,\tau}^{BT}$	Average link travel time of cleaned-up vehicles between BTR_i and BTR_j locations at aggregated time τ
$\bar{u}_{i,j,\tau}^{BT}$	Average link BT speed between BTR_i and BTR_j locations at aggregated time τ
λ	Epoch time (5-min intervals in commercial FCD)
τ	Aggregated time interval (30-min intervals in commercial FCD between 6:00 and 21:00 h)
T	Analysis period (set of timestamps)

The time between readings is calculated as "reoccurrence time"

$$\Delta t = t_{i,m,n} - t_{i,m,(n-1)} \tag{1}$$

also allows the estimation of "stay time (at a BTR range)" at locations i and j as

$$\delta t_{i,m} = t_{i,m,p} - t_{i,m,1} \tag{2}$$

$$\delta t_{j,m} = t_{j,m,r} - t_{j,m,1} \tag{3}$$

Possible cross-BTR travel time is calculated as:

$$\Omega_{i,j,m}^{BT} = t_{j,m,1} - t_{i,m,p} + \frac{(\delta t_{i,m}) + (\delta t_{j,m})}{2} \tag{4}$$

which is also called the "mid-to-mid travel time" calculation method in the literature.

To capture the time-dependent changes during the day and to evaluate it for the same time periods with fused FCD, the collected data were evaluated by separating them into time intervals. For each time interval, data were sorted and outliers in the vehicle travel times were eliminated by defining interquartile range (IQR) (a simple 1.5 IQR limit below Q1 (Q1-1.5*IQR) and above Q3 (Q3 + 1.5*IQR) is taken as outliers) according to the basic statistical method. The filtered moving vehicle travel times for captured MAC-IDs are averaged.

$$\bar{\Omega}_{i,j,\tau}^{BT} = \mathrm{Avg}\left(\Omega_{i,j,m,\tau}^{BT}\right) \tag{5}$$

Also allowing the estimation of the average space mean speed between two consecutive BTRs during the aggregated time as:

$$\bar{u}_{i,j,\tau}^{BT} = \frac{L_{i,j}^{BT}}{\bar{\Omega}_{i,j,\tau}^{BT}} \tag{6}$$

3.3.2 Fused Commercial FCD-based Link Speed Calculation

A general joint formulation for fused FCD- and BT-based speed estimation is described below, showing the mathematical relations between the segment-based FCD and link-based BT data (Fig. 3). For a BT link consisting of k FCD segments, the link length is constructed by summing the FCD segment lengths (x_k^{FCD}) of all the individual segments positioned on the road, with the resulting sum equal to the distance ($L_{i,j}^{BT}$) between consecutive BTR nodes (BTR_i and BTR_j) as:

$$L_{i,j}^{BT} = \sum_k x_k^{FCD} \quad \Phi_{i,j,k}^{BT,FCD} \quad x_k^{FCD} \geq 0 \tag{7}$$

where

$$\Phi_{i,j,k}^{BT,FCD} = \begin{cases} 1 & \text{if FCD segment (k) is on the link between BTR pair i and j} \\ 0 & \text{otherwise} \end{cases}$$

This could therefore establish a spatial connection with the data from BT. It establishes a procedure for determining the average speed based on BT, as demonstrated by the framework that was produced (Fig. 4). See the notation table for indexing and labeling. The notations used for speed calculation of fused FCD are summarized in Table 3.

For a link between any two selected BTR locations, FCD travel time on every segment k ($\omega_{k,\lambda}^{FCD}$) for an epoch time (λ), total travel time spent traveling along the allocated link ($\Omega_{i,j,\lambda}^{FCD}$) is calculated by adding up the travel times spent on each individual segment ($\omega_{k,\lambda}^{FCD}$)

$$\Omega_{i,j,\lambda}^{FCD} = \sum_k \omega_{k,\lambda}^{FCD} \quad \Phi_{i,j,k}^{BT,FCD} \quad \lambda \, \varepsilon \, T \tag{8}$$

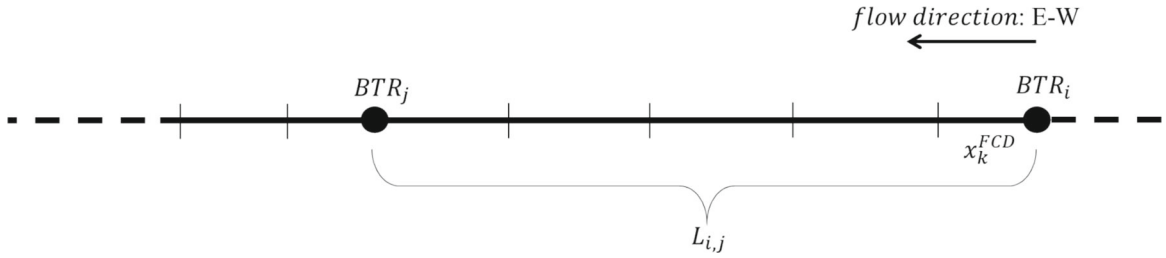

Fig. 3 Link-based spatial matching of FCD and BT (FCD segments between BTR_i and BTR_j)

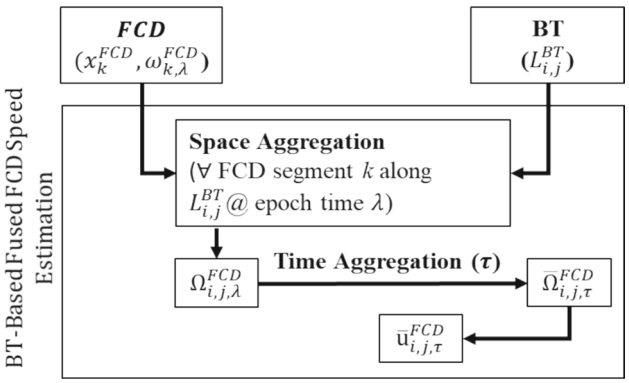

Fig. 4 Joint framework for the determination of the BT-based fused FCD speed

and corresponding link speeds can be found as:

$$u_{i,j,\lambda}^{FCD} = \frac{L_{i,j}^{BT}}{\Omega_{i,j,\lambda}^{FCD}} \tag{9}$$

The average total travel time of the link ($\bar{\Omega}_{i,j,\tau}^{FCD}$) may be calculated by averaging the travel times for each epoch period that falls within the aggregated duration (τ) as defined in [13, 38]

$$\bar{\Omega}_{i,j,\tau}^{FCD} = \frac{\sum_{\lambda}^{\tau} \Omega_{i,j,\lambda}^{FCD}}{\tau/\lambda} \qquad \lambda,\ \tau \varepsilon\ T \tag{10}$$

leading to the estimation of a link speed for the aggregated time

$$\bar{u}_{i,j,\tau}^{FCD} = \frac{L_{i,j}^{BT}}{\bar{\Omega}_{i,j,\tau}^{FCD}} \tag{11}$$

Considering space and time aggregation along the BTR link, the average link speed acts like a space mean speed, which will be easily compared to the space mean speed estimated from the BT data.

3.3.3 Performance Testing of Supposed Method and Data Sources

After determining the FCD and BT speeds, it is possible to test how compatible the estimated speeds are in two contexts: i) comparison of fused FCD and BT speed comparison data sets collected over the same time period; ii) testing effort could be made to interpret the effectiveness of the method with selected performance measures. The most common performance evaluation in urban arterial speed estimates is based on the following model metrics: i) mean absolute error ($\mathrm{MAE}_{FCD,BT}$), ii) mean absolute percentage error ($\mathrm{MAPE}_{FCD,BT}$), and iii) root-mean-square error ($\mathrm{RMSE}_{FCD,BT}$).

$$\mathrm{MAE}_{\mathrm{FCD,BT}} = \frac{1}{n}\sum_{i=1}^{n}\left|\bar{u}_{i,j,\tau}^{\mathrm{FCD}} - \bar{u}_{i,j,\tau}^{\mathrm{BT}}\right| \tag{12}$$

$$\mathrm{MAPE}_{\mathrm{FCD,BT}} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{\bar{u}_{i,j,\tau}^{\mathrm{FCD}} - \bar{u}_{i,j,\tau}^{\mathrm{BT}}}{\bar{u}_{i,j,\tau}^{\mathrm{FCD}}}\right| * 100 \tag{13}$$

$$\mathrm{RMSE}_{FCD,BT} = \left[\frac{1}{n}\sum_{i=1}^{n}\left(\left|\bar{u}_{i,j,\tau}^{FCD} - \bar{u}_{i,j,\tau}^{BT}\right|\right)^2\right]^{\frac{1}{2}} \tag{14}$$

Error $\mathrm{E}_{FCD,BT}$, is defined as the difference between average link speeds estimated from fused FCD and BT,

$$E_{FCD,BT} = \bar{u}_{i,j,\tau}^{FCD} - \bar{u}_{i,j,\tau}^{BT} \tag{15}$$

where

$\bar{u}_{i,j,\tau}^{FCD}$: the comparative average link fused FCD speed between BTR_i and BTR_j locations at the aggregated time period τ,

$\bar{u}_{i,j,\tau}^{BT}$: the average link BT speed between BTR_i and BTR_j locations at aggregated timeτ,

n: number of samples.

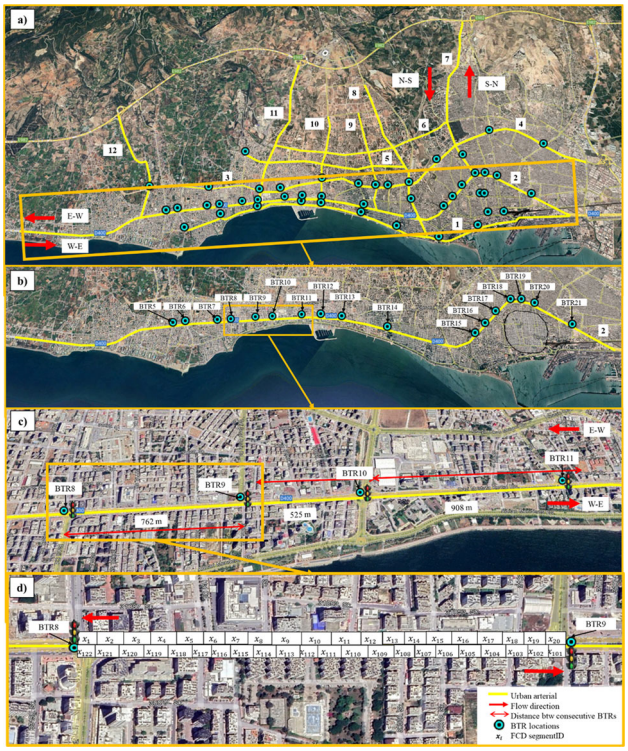

Fig. 5 **a** Simply map of the Mersin road network, including 12 major arterials and BTR locations, **b** GMK Boulevard with BTRs placed along it, **c** a part of the study corridor and selected BTRs, and **d** FCD segmentation on Link 1 (Google Earth, 2023)

4 Case Study and Analysis Results

4.1 Study Corridor in Mersin

The suggested method and speed evaluation are applied to urban arterials in the city of Mersin, Turkiye. Mersin, located on the Mediterranean coast of Turkiye, is a major metropolitan area and the 11th most populated city [55]. There are approximately 675,000 motorized vehicles registered in the city, of which close to half are cars [50]. The Mersin road network is mostly spread out along the coast and involves 12 arterials; half of them run in an E–W direction parallel to the coast, while the other half serves the north–south (N–S) direction as seen in Fig. 5a. The city center is located almost in the middle and close to the coast; thus, traffic moves mostly on the corridors parallel to the coast. Corridors are further designated as E–W or W–E, as well as N–S and south-to-north (S–N), to emphasize the traffic flow direction along an arterial. Also, BTR locations are located individually on each corridor.

For numerical evaluations, the longest arterial, Arterial 2, is selected in the study. Arterial 2 (Fig. 5b) is Gazi Mustafa Kemal (GMK) Boulevard, which has three lanes in each direction. It spans almost 19 km in the east–west direction with 29 junctions (one grade-separated). The shops are situated on both sides of it, so commercial activities are at a high level, which directly affects traffic conditions. It is a critical corridor connecting to the state highway D400. Roadside parking sometimes takes up one lane, mostly in the shopping districts along the road. A total of 21 BTR has been placed along the GMK, 17 of which can be seen in Fig. 5b. GMK has local speed characteristics and conditions: users of the road may have been traveling at the legal speed limit of 82 km/h; the operating speed limit might be assumed to be 90 km/h due to the 10% tolerance in penalties. Though some users on the route may violate the posted speed limit, commercial FCD values are truncated at 77 km/h, and the minimum speed is 13 km/h.

To simplify the BTR codes, a generic identification number creation is performed for this study; the scope of the analyses includes the 2.2-km-long sub-corridor between BTR08 and BTR11 (Fig. 5c), further divided into three parts (also called links): i) Link 1 is a 762-m-long part between BTR08 and BTR09; ii) Link 2 is a 525-m-long part between BTR09 and BTR10; and iii) Link 3 is a 908-m-long part between BTR10 and BTR11. Both the beginning and ending of these links overlap with the intersections. The details of study links are tabularized as in Table 4. For both directions, as the FCD segments are much smaller and defined mostly by road features, a link is matched with multiple FCD segments located on it. For example, Link 1 corresponds to 20 FCD segments on E–W, while it has 22 segments in the W–E direction (Fig. 5d).

Data were collected by both commercial FCD and BT sources for four days (6th, 7th, 13th, and 14th; two consecutive Tuesdays and Wednesdays) in September 2022. Since Monday and Friday were assessed to have different traffic patterns than usual weekdays, "regular weekdays" in this research comprise Tuesday and Wednesday. Two consecutive Tuesdays and Wednesdays were chosen so that the typical daily traffic pattern could be studied considering the beginning of the fall semester. The analysis period (T) was determined to encompass nearly the entire day, from 6:00 to 21:00. This period covers morning, lunch, and evening peak periods. The majority of the studied techniques approximate traffic conditions in daylight periods and peak hours [54, 56], each of which represents rush hours and time intervals in between. These time intervals are often taken from the literature and decided by experience.

4.2 Comparison of Link-Based Average Fused FCD and BT Speed Estimation

The average link FCD and BT speeds were calculated by utilizing link length and averaging the travel times of all segments on the specified corridor. The obtained FCD speed represents the comparative fused FCD. Each speed at epoch

Table 4 Key variables of commercial FCD segments on the corridor studied parts are limited by every BTR pair

Link	1		2		3	
BTR pairs	8–9		9–10		10–11	
FCD	E-W	W-E	E-W	W-E	E-W	W-E
minSL	13.01	12.39	13.78	13.01	6.12	17.05
maxSL	46.97	44.51	46.47	45.43	47.51	48.49
TotalL	764.97	757.12	539.78	540.05	900.97	903.02
NS	20	22	14	16	24	22

FD is flow direction, SL refers to FCD segment length, TL and SN represent total section part length between successive BTRs and number of FCD segment, respectively

times has been added to the link speed chart throughout the whole time of the analysis period, and thus, the average speed scattering around the mean speed point could be observed. The directions of flow were E–W and W–E.

The aggregated time (τ) was a 30-min interval with a date of 6th of September in 2022, and the graph below has been evaluated. Average link speed at even epoch time (λ) for 5-min intervals provided by both FCD and BT are used to create Fig. 6, which may be called average link speed profiles. Thereby, scattered speed values formed "a cloudy structure" generated by link average speeds belonging to timeframes of each day. The illustration shows a quantile–quantile (Q-Q) plot of fused FCD vs. BT for three links from BTR08–BTR11. Fused FCD and BT values have two different distributions, and outliers are visible. The outliers of average speeds are appended to the plotted diagrams. Most observations propagate around the mean speed for both directions, but data patterns typically have some higher extreme values (shown in red circles) than would be anticipated if the data were genuinely drawn from a normal distribution. The speed results for September 6 are spread around normal, whereas some outcomes on September 7, 13, and 14 indicate excessive ones.

The range of speed outliers is selected to be higher than 2.5 km/h and lower than 47 km/h. To create quantiles matching the same underlying probability, an interpolated quantile on the Q-Q plot estimate is required. Besides, the speed values in the corridor have no significance for values outside of this range. This condition can be observed across all links. These outliers appear around the early morning hour (up to 7:00 a.m.). The fused FCD speed results are close to the average link speed and indicate a more reliable outcome than BT data speed estimation. The average speed of both data sets is around 25 km/h. In this case, it may be considered that the analysis period (T) can be drawn between 7:00 and 21:00. Therefore, in Fig. 7, plots between the ranges of 7:00 and 21:00 are included to visualize the differences between the data. This condition explains the FCD and BT differences in terms of speed estimation in the early morning. Here, there are given results for September 6, 2022, at 30-min aggregated time intervals.

Figure 7 shows that Link 1 and Link 2 indicate similar speed patterns, but Link 3 has higher speed findings. The beginning and end of each link have intersections, which cause delays. Link 3 has a longer distance than the others, and vehicles can have enough time to accelerate over this distance. The scenario is different in Links 1 and 2; however, as acceleration times for halting vehicles at junctions indicate that they are often moving at slower speeds due to the proximity of the intersections. Even Link 2 has a lower speed than others. This situation demonstrates that analyses of urban arterials are more difficult than those on highways.

When examining FCD results, the traffic pattern reflected that of a free flow speed type up until 7:45. A significant period of dramatic deceleration ensued, lasting roughly half an hour during peak hour, from 7:45 to 8:15 a.m. Comparing the data, FCD speed profiles reveal variations for every link and direction. While reviewing the outcomes, it is clear that FCD speed results have varied depending on the link and the direction taken. Speed has been lowering/ed in the E–W than W–E. As a result, the overall link average speed for E–W is 27.6 km/h, whereas it is 30.2 km/h for W–E. In other words, traffic flow has been demonstrated lower in the W–E.

The peak hour period means that when demand is highest, there is a demonstrated drop in speed whose range is scattered within a small range of a very near value. This narrow speed interval has been included in the minimum speed values close to the mean speeds. 12 km/h was recorded as the minimum speed experienced on Link 2 and Link 1, which have 18.2 and 28.4 km/h, respectively, for E–W. It means that the minimum speed range is close to the queue threshold of 15 km/h and that an oversaturated traffic regime has dominated the corridor. Another thing that should be highlighted is that the queues within this range encourage congestion delays for a one hour in an unavoidable manner. The maximum speed is established to be 42.7 km/h on Link 3 concerning W–E. Given that [21] proposed a threshold of 42 km/h for the queue to start uninterrupted, it can be concluded that the LOS of the D400 state

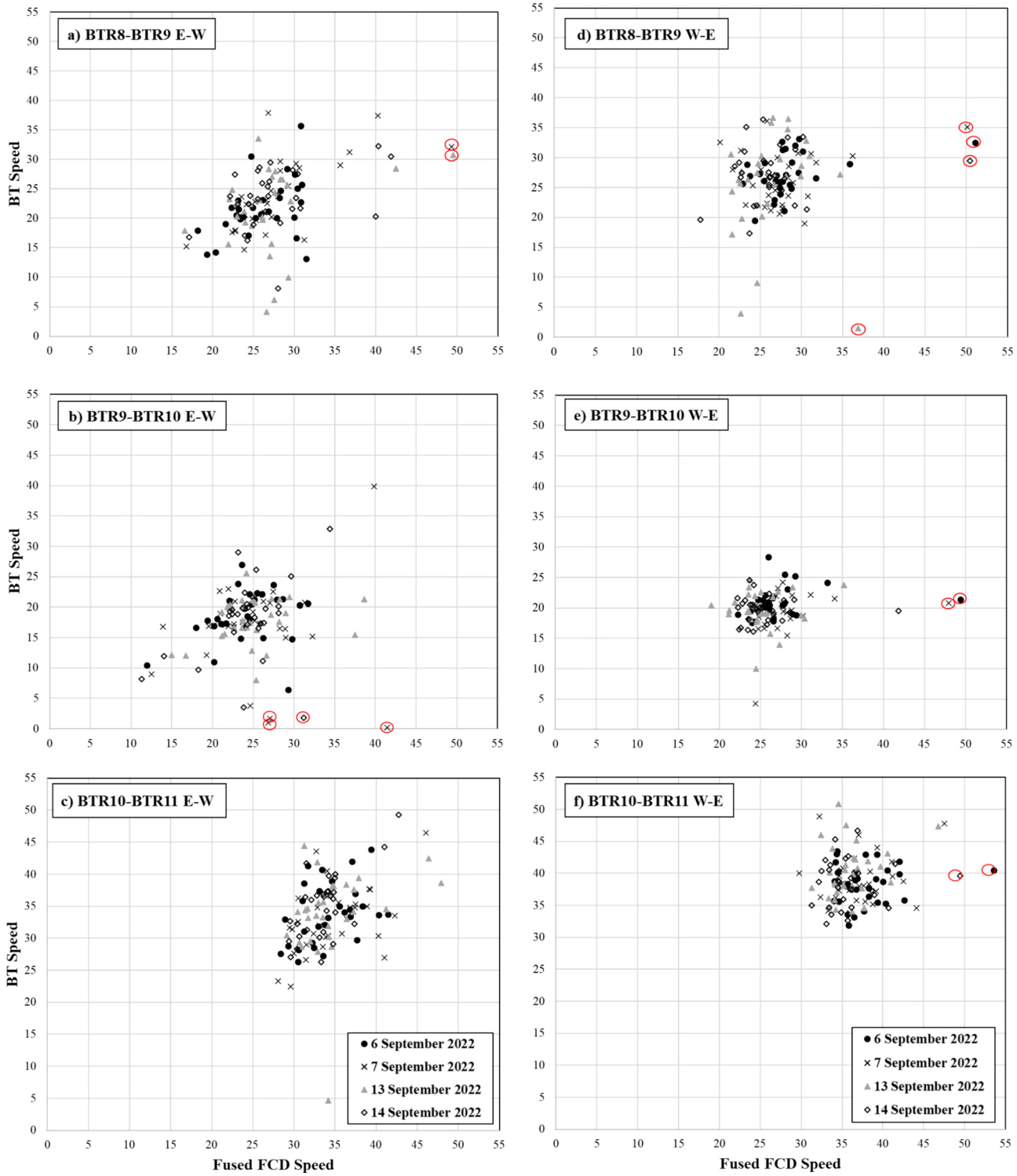

Fig. 6 Comparison of link-based average comparative fused FCD vs. BT speed distributions over aggregated time, $\tau = 30$ min, for both directions

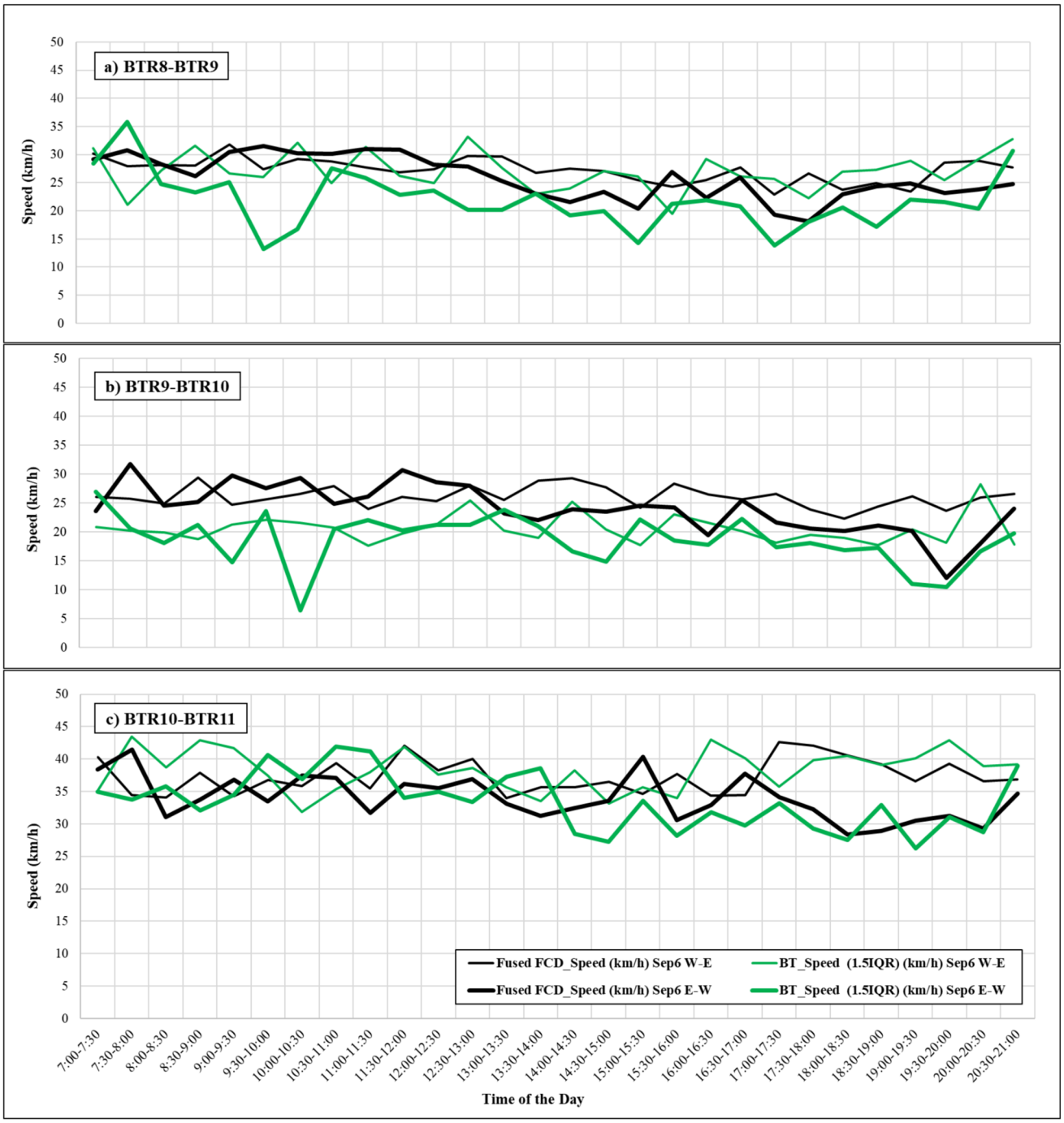

Fig. 7 Link-based average comparative fused FCD vs. BT speed over aggregated time, $\tau = 30$ min, for both directions on September 6, 2022

highway may not meet this requirement. However, it could be fulfilled for a major city arterial to serve as intended if the serviceability is selected as least comfortable in terms of congestion delay. Moreover, it can be inferred that the vehicles floated in the corridor with an almost constant mean speed of close to 30 km/h corresponding to the morning peak. Similarly, lunchtime speed has changed in the same pattern, but there has been a gradual decline up to almost 10 km/h. This fluctuating trend in speed continued until the evening peak. Additionally, all links experienced minimum speed levels and extended durations during the evening peak. Approximately 10 km/h speed changes were figured out after 16:00. The next timeframe on Link 2 revealed a speed reduction of up to 15 km/h. After 17:00, the significant speed decrease persisted until 21:00. The speeds of the remaining aggregated time intervals until 21:00 exhibited a smooth average speed

Table 5 Error comparison of each link on average fused FCD vs. BT speeds over aggregated time, $\tau = 30$ min, on E–W

Day	Link	$\text{MAE}_{FCD,BT}$ (km/h)	$RMSE_{FCD,BT}$ (km/h)	$\text{MAPE}_{FCD,BT}$ (%)
Sep 6	BTR08-BTR09	4.82	6.14	17.96
	BTR09-BTR10	5.75	7.44	22.59
	BTR10-BTR11	3.81	4.61	11.16
Sep 7	BTR08-BTR09	4.55	5.62	16.95
	BTR09-BTR10	8.16	10.58	32.19
	BTR10-BTR11	3.88	4.71	11.39
Sep 13	BTR08-BTR09	5.67	8.40	21.47
	BTR09-BTR10	6.41	7.49	26.02
	BTR10-BTR11	4.26	6.99	12.65
Sep 14	BTR08-BTR09	4.72	6.85	16.78
	BTR09-BTR10	6.06	7.31	25.59
	BTR10-BTR11	3.26	3.93	9.84

Table 6 Error comparison of each link on average fused FCD vs. BT speeds over aggregated time, $\tau = 30$ min, on W-E

Day	Link	$\text{MAE}_{FCD,BT}$ (km/h)	$RMSE_{FCD,BT}$ (km/h)	$\text{MAPE}_{FCD,BT}$ (%)
Sep 6	BTR08-BTR09	2.96	3.40	10.96
	BTR09-BTR10	5.69	6.03	21.73
	BTR10-BTR11	3.37	4.17	9.21
Sep 7	BTR08-BTR09	4.21	5.36	16.35
	BTR09-BTR10	7.73	8.50	28.69
	BTR10-BTR11	3.83	5.27	10.77
Sep 13	BTR08-BTR09	4.99	6.55	20.07
	BTR09-BTR10	5.53	6.69	21.49
	BTR10-BTR11	4.59	5.88	13.14

trend, consistently hovering around 20–25 km/h. Therefore, it can be safely assumed that, except for 6:00–7:00 h, the average speed of the vehicles cruising the corridor has remained lower at 25–30 km/h throughout the analysis period, resulting in congestion delays throughout the day for the corridor in question.

It should be noted that aggregating speeds near the free flow along the corridor presented no difficulty when assessing traffic conditions throughout the duration; however, the average speed status seen around congestion limitations resulted in considerable losses due to averaging the travel times and/or speeds pertaining to each segment embedded on the links.

4.3 Performance Evaluation of Link-Based Average Fused FCD and BT Speed Estimation

To evaluate the performance of data sources, different metrics have been chosen. Both directions and all days are considered, and the analysis results are given as follows: Table 5 indicates the $\text{MAE}_{FCD,BT}$, $\text{RMSE}_{FCD,BT}$, and $\text{MAPE}_{FCD,BT}$ of E–W in comparison with speed estimation errors by the comparative fused FCD and BT data, whereas Table 6 displays results in the opposite direction. The calculation bases the outcomes on the absolute errors of FCD and BT. The results show that the estimated approach effectively captures the daily trend in all cases, particularly for the $\text{MAE}_{FCD,BT}$ and $\text{RMSE}_{FCD,BT}$ values. The results of each day reveal the promising findings, but especially longest Link 3 gives more reliable ones. $\text{MAPE}_{FCD,BT}$ also performs better in those circumstances but worse in Link 2, and Link 1 shows acceptable success. So, the conclusion results suggest that the comparable fusion of BT-based FCD increases the reliability of the estimation on longer links. The flow direction has a direct impact on the link-based speed estimation. Direction with lower speed displays fewer errors.

5 Discussion

The increasing levels of traffic congestion, particularly on urban arterials, present significant challenges for city planners, traffic engineers, and policymakers. Understanding the dynamics of traffic flow and its relationship with congestion patterns requires accurate and cost-effective data collection

methods. The present study introduced a comparative fusion model that effectively integrates two powerful traffic data sources—FCD and BT—to estimate average speeds along an urban corridor. The methodology explored in this research not only highlights the strengths and weaknesses of each data source but also demonstrates how their complementary attributes can be used to derive more reliable traffic estimates.

A key observation from the analysis is the efficacy of FCD in providing consistent speed updates due to its high spatial and temporal coverage. Its ability to gather data at 5-min intervals across all links on the corridor provides a fine-grained representation of traffic dynamics. However, the use of BT data, although lower in frequency and subject to greater variation for short distances, offers a low-cost alternative that is particularly useful for travel time and OD analysis. The combination of these two data sources offers a powerful approach to monitoring traffic behavior, as evidenced by the generally consistent results between the two during peak hour periods.

Another important finding in this study was the identification of outliers, particularly in the early morning hours. These outliers represent instances where the FCD and BT data showed marked discrepancies, primarily in the range of 2.5 to 47 km/h. These variations, especially during off-peak times, highlight the need for further refinement in understanding the specific factors contributing to speed fluctuations, such as varying traffic demands, vehicle types, and signal timing patterns at intersections.

The performance of the fusion model was assessed through multiple performance metrics, including the $\mathrm{MAE}_{FCD,BT}$, $\mathrm{MAPE}_{FCD,BT}$, and $\mathrm{RMSE}_{FCD,BT}$. The results reveal that the model performs well in capturing general traffic trends, particularly in the E–W direction where speeds were slightly lower and error metrics were more favorable. Longer links, such as Link 3, exhibited more reliable estimates, likely due to the greater time available for vehicles to accelerate between intersections. Conversely, shorter links, such as Links 1 and 2, demonstrated more significant speed variation, as the proximity of intersections caused frequent halts, resulting in higher estimation errors.

The study also revealed that congestion patterns varied significantly across the day, with the morning peak showing the greatest volatility in speed. This is particularly evident between 7:45 and 8:15 a.m., where dramatic decelerations were observed. The evening peak presented a more prolonged congestion period, especially after 17:00, as speeds dropped below 25 km/h and did not recover until after 21:00. This insight is critical for traffic management strategies, as it indicates that managing congestion at specific times of the day requires targeted interventions that address both recurrent and non-recurrent traffic patterns.

6 Conclusions

The study successfully demonstrated the utility of a generalized adaptable comparative model using FCD and BT data to estimate traffic speeds and analyze congestion patterns along urban arterial sections. By leveraging the high-resolution FCD and the cost-effective BT technology, this model provides a robust and scalable approach to traffic monitoring and analysis, particularly in emerging economies where resource constraints often limit the deployment of more expensive traffic management systems. Several key conclusions can be drawn from the research findings:

- Commercial FCD provides a highly reliable dataset for estimating average speeds across various timeframes. Its ability to update every five minutes ensures that it captures the dynamic nature of traffic flow, especially during peak hours.
- While BT has limitations in terms of coverage and precision over shorter distances, its low cost and ability to provide accurate travel time and speed estimates make it a valuable complement to FCD. The comparative fusion of BT-based FCD significantly improves the reliability of speed estimates for longer links.
- The length of the link and the direction of traffic flow have a direct impact on speed estimation accuracy. Longer links allow for greater acceleration, while intersections on shorter links result in more frequent stops and delays, making them more difficult to model accurately.
- The study identified specific periods (early morning and peak hours) where speed estimates showed significant deviations from normal traffic conditions. These outliers, particularly in the early morning, suggest the need for further investigation into external factors influencing traffic speeds, such as road geometry, signal timing, or driver behavior.
- Both morning and evening peak periods exhibit significant congestion delays, with speeds dropping as low as 12 km/h in some cases. The consistent decrease in speed during these times highlights the need for targeted traffic management interventions to alleviate bottlenecks and improve flow efficiency.

This research provides a solid foundation for future work aimed at integrating additional data sources and refining the fusion model to better account for varying congestion patterns. Expanding the model to include real-time traffic management applications and other data types, such as video-based traffic surveillance or mobile phone location data, could further enhance its predictive capabilities. Moreover, by applying this methodology to other urban corridors, policymakers and traffic managers can develop more effective

strategies to reduce congestion, improve travel time reliability, and enhance overall traffic efficiency in metropolitan areas.

In summary, this study contributes to the growing body of knowledge on traffic data fusion methods and highlights the potential for using advanced data analytics to address the complex challenges of urban traffic management. The findings offer practical implications for transportation engineers, city planners, and decision-makers seeking to optimize traffic flow and mitigate congestion in cost-effective and scalable ways.

Acknowledgements The authors would like to thank Integrated Systems and Systems Design (ISSD) Company, Turkiye, for providing them with information about the FCD, corridor segments, and BT data.

Funding Open access funding provided by the Scientific and Technological Research Council of Türkiye (TÜBİTAK).

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

1. Budimir, D.; Jelušić, N.; Perić, M.: Floating car data Technol. Pomorstvo **33**(1), 22–32 (2019)
2. Robinson, S.; Polak, J.: Overtaking rule method for the cleaning of matched license-plate data. J. Transp. Eng. **132**(8), 609–617 (2006)
3. Zhu, L.; Guo, F.; Polak, J.W.; Krishnan, R.: Urban link travel time estimation using traffic states-based data fusion. IET Intel. Transp. Syst. **12**(7), 651–663 (2018)
4. Chen, Y.; Chen, C.; Wu, Q.; Ma, J.; Zhang, G.; Milton, J.: Spatial-temporal traffic congestion identification and correlation extraction using floating car data. J. Intell. Transp. Syst. **25**(3), 263–280 (2021)
5. Cotten, D.; Codjoe, J.; Loker, M.: Evaluating advancements in Bluetooth technology for travel time and segment speed studies. Transp. Res. Rec. **2674**(4), 193–204 (2020)
6. Gong, Y.; Abdel-Aty, M.; Park, J.: Evaluation and augmentation of traffic data including Bluetooth detection system on arterials. J. Intell. Transp. Syst. **25**(6), 561–573 (2021)
7. Osei, K.K.; Adams, C.A.; Sivanandan, R.; Ackaah, W.: Modelling of segment level travel time on urban roadway arterials using floating vehicle and GPS probe data. Scientific African **15**, e01105 (2022)
8. Jayan, A.; Anusha, S.P.: Travel time prediction under mixed traffic conditions using RFID and bluetooth sensors. Period. Polytech. Transp. Eng. **48**(3), 276–289 (2020)
9. Kessler, L.; Karl, B.; Bogenberger, K.: Spatiotemporal traffic speed reconstruction from travel time measurements using bluetooth detection. In 2019 IEEE Intelligent Transportation Systems Conference (ITSC) (pp. 4275–4280). IEEE. (2019)
10. Liu, X.; Chien, S.; Kim, K.: Evaluation of floating car technologies for travel time estimation. J. Mod. Transp. **20**(1), 49–56 (2012)
11. Zhou, H.; Benz, R. J.; Voigt, A.; Mao, A. C.: Bluetooth travel time/speed data analysis for houston-galveston regional transportation study (No. 16–3553). (2016)
12. Day, C. M.; Li, H.; Lavrenz, S. M.; Remias, S. M.; Bullock, D. M.: Enhanced Dynamic Travel Time Estimation from Probe Vehicle Segment Speeds. Transportation Research Board 97th Annual Meeting, (No. 18–00114). (2018
13. Chase, R.T.; Williams, B.M.; Rouphail, N.M.; Kim, S.: Comparative evaluation of reported speeds from corresponding fixed-point and probe-based detection systems. J. Transp. Res. Record (2012). https://doi.org/10.3141/2308-12
14. Kim, S.; Coifman, B.: Comparing INRIX speed data against concurrent loop detector stations over several months. Transp. Res. Part C: Emerging Technol. **49**, 59–72 (2014)
15. Yuan, Q.; Liu, Z.; Li, J.; Zhang, J.; Yang, F.: A traffic congestion detection and information dissemination scheme for urban expressways using vehicular networks. Transportation Research Part C: Emerging Technologies, 47, Part 2, 114–127. (2014)
16. Cohen, S.; Christoforou, Z.: Travel time estimation between loop detectors and FCD: A compatibility study on the Lille network. France. Transp. Res. Procedia **10**, 245–255 (2015)
17. Rempe, F.; Franeck, P.; Fastenrath, U.; Bogenberger, K.: A phase-based smoothing method for accurate traffic speed estimation with floating car data. Transp. Res. Part C: Emerging Technol. **85**, 644–663 (2017)
18. Kessler, L.; Huber, G.; Kesting, A.; Bogenberger, K.: Comparing speed data from stationary detectors against floating-car data. IFAC-PapersOnLine **51**(9), 299–304 (2018)
19. Shahrbabaki, M.R.; Safavi, A.A.; Papageorgiou, M.; Papamichail, I.: A data fusion approach for real-time traffic state estimation in urban signalized links. Transp. Res. part C: Emerging Technol. **92**, 525–548 (2018)
20. Altintasi, O.; Tuydes-Yaman, H.; Tuncay, K.: Monitoring urban traffic from floating car data (FCD): using speed or a los-based state measure. In: Directions of development of transport networks and traffic engineering: 15th scientific and technical conference, Transport systems. theory and practice 2018, Katowice, Poland, September 17–19, 2018, Selected Papers. Springer, Cham. (2019)
21. Altintasi, O.: Detecting queue length on urban arterials using floating car data (FCD). Doctoral dissertation, Middle East Technical University, Ankara Turkiye. (2018)
22. Bachmann, C.: Multi-sensor data fusion for traffic speed and travel time estimation. Doctoral dissertation, University of Toronto. (2011)
23. Ambuhl, L.; Menendez, M.: Data fusion algorithm for macroscopic fundamental diagram estimation. Transp. Res. Part C: Emerging Technol. **71**, 184–197 (2016)
24. Ji, Y.; Xu, M.; Li, J.; Van Zuylen, H.J.: Determining the macroscopic fundamental diagram from mixed and partial traffic data. Promet-Traffic & Transp. **30**(3), 267–279 (2018)
25. Jiang, Y.; Song, G.; Zhang, Z.; Zhai, Z.; Yu, L.: Estimation of hourly traffic flows from floating car data for vehicle emission estimation. J. Adv. Transp. **2021**, 1–11 (2021)
26. Sun, C.; Zhang, H.; Chen, X.: Road traffic operation assessment based on multi-source floating car data fusion. J. Tongji Univ. (Natural Sci.) **46**(1), 46–52 (2018)
27. Schmidt, D.; Friedrich, B.; Weichenmeier, F.: Data fusion techniques for traffic state estimation–DINO within DMotion. In 15th World Congress on Intelligent Transport Systems (pp. 15–20). (2008)

28. Bluetooth.com (Bluetooth®).: Bluetooth®, Learn about Bluetooth and Specifications. Accessed on February of 2023 and retrieved from https://www.bluetooth.com/ (2023)
29. Wasson, J.S.; Sturdevant, J.R.; Bullock, D.M.: Real-time travel time estimates using media access control address matching. Inst. Transp. Eng. (ITE) J. **78**(6), 20–23 (2008)
30. Haseman, R.J.; Wasson, J.S.; Bullock, D.M.: Real-time measurement of travel time delay in work zones and evaluation metrics using bluetooth probe tracking. Transp. Res. Record **2169**, 40–53 (2010). https://doi.org/10.3141/2169-05
31. Bachmann, C.; Roorda, M.J.; Abdulhai, B.; Moshiri, B.: Fusing a Bluetooth traffic monitoring system with loop detector data for improved freeway traffic speed estimation. J. Intell. Transp. Syst. **17**(2), 152–164 (2013). https://doi.org/10.1080/15472450.2012.696449
32. Bhaskar, A.; Chung, E.: Fundamental understanding on the use of Bluetooth scanner as a complementary transport data. Transp. Res. Part C: Emerging Technol. **37**, 42–72 (2013). https://doi.org/10.1016/j.trc.2013.09.013
33. Quayle, S.M.; Koonce, P.; DePencier, D.; Bullock, D.M.: Arterial performance measures with media access control readers-portland, oregon, pilot study. Transp. Res. Record **2192**, 185–193 (2010). https://doi.org/10.3141/2192-18
34. Saeedi, A.; Park, S.; Kim, D.S.; Porter, D.: Improving accuracy and precision of travel time samples collected at signalized arterial roads with bluetooth sensors. Transp. Res. Record **2380**, 90–98 (2013). https://doi.org/10.3141/2380-10
35. Araghi, B.N.; Pedersen, K.S.; Christensen, L.T.; Krishnan, R.; Lahrmann, H.: Accuracy of travel time estimation using Bluetooth technology: Case study limfjord tunnel Aalborg. Proc., 19th Intelligent Transport Systems World Congress, Vienna, Austria. (2012)
36. Araghi, B.N.; Skoven Pedersen, K.; Tørholm Christensen, L.; Krishnan, R.; Lahrmann, H.: Accuracy of travel time estimation using Bluetooth technology: Case study Limfjord tunnel Aalborg. Int. J. Intell. Transp. Syst. Res. **13**(3), 166–191 (2015)
37. Moghaddam, S.S.; Hellinga, B.: Real-Time Prediction of Arterial Roadway Travel Times Using Data Collected by Bluetooth Detectors. Transp. Res. Record **2442**, 117–128 (2014). https://doi.org/10.3141/2442-13
38. Yucel, S.: Use of Bluetooth Technology for Traffic Analysis in Urban Road Networks. MSc Thesis, Middle East Technical University, Ankara, Turkiye. (2015)
39. Erkan, I.; Hastemoglu, H.: Bluetooth as a traffic sensor for stream travel time estimation under Bogazici Bosporus conditions in Turkiye. J. Mod. Transp. **24**(3), 207–214 (2016). https://doi.org/10.1007/s40534-016-0101-y
40. Altintasi, O.; Balci, M.E.; Ozen, M.: Urban arterial travel time evaluation using Bluetooth data. Politeknik Dergisi (2022). https://doi.org/10.2339/politeknik.784652
41. Stevanovic, A.Z.; Martin, P.T.: Assessment of the suitability of microsimulation as a tool for the evaluation of macroscopically optimized traffic signal timings. J. Transp. Eng. **134**(2), 59–67 (2008)
42. Liu, H.X.; Ma, W.: A virtual vehicle probe model for time-dependent travel time estimation on signalized arterials. Transp. Res. Part C: Emerging Technol. **17**(1), 11–26 (2009)
43. Haghani, A.; Hamedi, M.; Sadabadi, K.F.; Young, S.; Tarnoff, P.: Data collection of freeway travel time ground truth with Bluetooth sensors. Transp. Res. Rec. **2160**(1), 60–68 (2010)
44. Sharifi, E., Hamedi, M., Haghani, A., & Sadrsadat, H.: Analysis of vehicle detection rate for Bluetooth traffic sensors: A case study in maryland and delaware. In 18th World Congress on on Intelligent Transport Systems. (2011)
45. Hainen, A.M.; Wasson, J.S.; Hubbard, S.M.L.; Remias, S.M.; Farnsworth, G.D.; Bullock, D.M.: Estimating route choice and travel time reliability with field observations of bluetooth probe vehicles. Transp. Res. Record **2256**, 43–50 (2011). https://doi.org/10.3141/2256-06
46. Ruppe, S.; Junghans, M.; Haberjahn, M.; Troppenz, C.: Augmenting the floating car data approach by dynamic indirect traffic detection. Procedia Soc. Behav. Sci. **48**, 1525–1534 (2012)
47. Gurczik, G.; Touko Tcheumadjeu, L. C.; Luber, A.; Ruppe, S.; Junghans, M.: Floating Observer Information Processing on the Basis of Mobile Bluetooth Data. Eng. Appl. (2014)
48. Kessler, L.; Rempe, F.; Bogenberger, K.: Multi-sensor data fusion for accurate traffic speed and travel time reconstruction. Front. Future Transp. **2**, 766951 (2021)
49. Cvetek, D.; Boji´c, V.; Jeluši´c, N.; Muštra, M.: Initial Bluetooth probe vehicle penetration rate analysis: A case study in the city of Zagreb. In Proceedings of the International Scientific Conference: Science and Traffic Development (ZIRP 2019), Opatija, Croatia. (2019)
50. Turkish Statistical Institute (TUIK).: "Vehicles registered in each province" and "Number of motorized vehicles by year". Accessed on 12.12.2022 and retrieved from https://data.tuik.gov.tr/Bulten/Index?p=Motorlu-Kara-Tasitlari-Aralik-2021-45703 (2022a)
51. Kerner, B.S.; Demir, C.; Herrtwich, R.G.; Klenov, S. L.; Rehborn, H.; Aleksic, M., & Haug. A.: Traffic State Detection with Floating Car Data in Road Networks. Proceeding of the 8th International IEEE conference on Intelligent Transportation Systems, Vienna, Austria. September 13–16. (2005)
52. Hong, J.; Zhang, X.; Wei, Z.; Li, L.; Ren, Y.: Spatial and temporal analysis of probe vehicle-based sampling for real-time traffic information system. Proc. 2007 IEEE Intelligent Vehicles Symposium. Istanbul, Turkey, pp. 1234–1239. https://doi.org/10.1061/(ASCE)TE.1943-5436.0000815. (2007)
53. Sunderrajan, A.; Viswanathan, V.; Cai, W.; Knoll, A.: Traffic State Estimation Using Floating Car Data. Procedia Comput. Sci. **80**, 2008–2018 (2016)
54. Erdelić, T.; Carić, T.; Erdelić, M.; Tišljarić, L.; Turković, A.; Jelušić, N.: Estimating congestion zones and travel time indexes based on the floating car data. Comput. Environ. Urban Syst. **87**, 101604 (2021)
55. Turkish Statistical Institute (TUIK).: Population of Provinces by Years. Accessed on 12.12.2022 and retrieved from https://data.tuik.gov.tr/Bulten/Index?p=Adrese-Dayal%C4%B1-N%C3%BCfus-Kay%C4%B1t-Sistemi-Sonu%C3%A7lar%C4%B1-2021-45500&dil=1 (2022b)
56. Song, J.; Zhao, C.; Zhong, S.; Nielsen, T.; Prishchepov, A.: Mapping spatiotemporal patterns and detecting the factors of traffic congestion with multi-source data fusion and mining techniques. Comput. Environ. Urban Syst. (2019). https://doi.org/10.1016/j.compenvurbsys.2019.101364