Scientific African 15 (2022) e01105

Contents lists available at ScienceDirect

Scientific African

journal homepage: www.elsevier.com/locate/sciaf

Modelling of segment level travel time on urban roadway arterials using floating vehicle and GPS probe data

Kwame Kwakwa Osei[a,d,*], Charles Anum Adams[a], Ramaswamy Sivanandan[b], Williams Ackaah[c]

[a] *Department of Civil Engineering, Regional Transport Research and Education Centre Kumasi, Kwame Nkrumah University of Science and Technology, PMB, Kumasi, Ghana*
[b] *Department of Civil Engineering, Indian Institute of Technology Madras, Chennai 600036, India*
[c] *CSIR-Building and Road Research Institute, P.O. Box 40, Kumasi, Ghana*
[d] *Kumasi Technical University, P.O. Box 854, Kumasi, Ghana*

ARTICLE INFO

Article history:
Received 20 November 2020
Revised 30 June 2021
Accepted 25 January 2022

Editor: DR B Gyampoh

Keywords:
Traffic congestion
Travel time
Global position system (GPS)
Heterogeneous traffic
Urban roadway arterials
Multiple linear regression

ABSTRACT

With the increasing traffic congestion levels on urban arterials, an essential step to tackling this challenge is to effectively quantify it and understand how it relates to its contributing factors. Although researchers have proposed models that relate travel time with several traffic and roadway factors for quantifying congestion especially in the Western world, most of these models lack the predictive accuracy for arterials in low-income countries due to differences in roadway and roadside interference factors, heterogeneous traffic flow, and others. A number of them do not incorporate delays from factors including on-street parking activities. Additionally, some existing models are complex in structure and require several parameters which may not be available in many low-income countries such as Ghana or may be expensive to collect. This study aimed at exploring a simple model for predicting travel time which will capture the contributing factors of congestion typical of low-income country arterial road environment and flow characteristics using a multiple linear regression model. Using moving observer method, traffic and roadway data were collected from eight arterial roadways in the Greater Kumasi Metropolitan Area for modeling travel time at the segment level. The fitted model that captures the impact of factors including on-street parking, access density, traffic density, and segment downstream bounding conditions on travel time will aid decision making by transport planners on the factors to consider to mitigate congestion. The model demonstrates how inadequate enforcement of on-street parking restrictions on arterial roadways exacerbates congestion during periods of high demand.

* Corresponding author at: Department of Civil Engineering, Regional Transport Research and Education Centre Kumasi, Kwame Nkrumah University of Science and Technology, PMB, Kumasi, Ghana.

E-mail addresses: oseikwakwa@gmail.com (K.K. Osei), carladams1702@yahoo.com (C.A. Adams), rsiva@iitm.ac.in (R. Sivanandan), ackaahwillie@yahoo.com (W. Ackaah).

https://doi.org/10.1016/j.sciaf.2022.e01105

Introduction

Traffic congestion has continuously been a major problem that contributes to higher travel time delays in most countries. According to Fleuren [1], investigations across four advanced countries (i.e., France, Germany, UK and US) conducted by INRIX & CEBR [2] show that the total economic costs associated with congestion from 2013 to 2030 are predicted to be 3.35 trillion euros. Similarly, Takyi et al. [3], in their study in the Kumasi Metropolis, reported a loss of 21.9% and 14.3% of the expected daily income of trotro (paratransit) and taxi drivers respectively due to congestion. Billions of cedis worth productive hours are lost to congestion, costing Ghana about 8.21% of its GDP [4]. Congestion has therefore gained much attention from transport engineers and planners, all in the quest of mitigating its impacts and largely providing sustainable transport. In order to suggest mitigation strategies to tackling the problem of congestion, there is the need to understand congestion from all perspectives. The first and essential step requires the effective quantification of the current performance of the road network [5] and assessing how it interacts with the local roadway and traffic conditions. Among the performance measures that are useful to quantifying congestion, travel-time-based measures are of special interest to both travelers and traffic engineers and planners because these measures relate directly to the traveler perspective [6]. Moreover, travel time is used for traffic assignment in various transportation planning models. Consequently, much attention has been given to developing travel time models over the years, for quantifying congestion on arterials.

It is clear in the literature that several models have been developed for estimating travel time with parameters including volume, capacity, segment length, free-flow speed and signal timing as predictors. With the quest of improving the accuracy of travel time estimation, the models have evolved from simple BPR function which related travel time with free-flow travel time, link volume and link capacity to complex ones which relates travel time to several parameters including signal timing. Although analytical models such as the Akcelik model, Skabardonis-Dowling Model, Highway Capacity Manual (HCM) [7] models have been widely accepted for operational analysis, their application for long range planning is difficult. Due to their complex structure, they require large scope of input data of parameters such as signal timing configurations (e.g., cycle length, splits, phases, and progression quality) which may not be readily available for base year and future years analysis. Although having a lot of input parameters in these models may improve the accuracy of prediction, it may however increase cost in terms of the collection of input data for predictive and model maintenance purposes.

Further, most of these models such as the HCM [7] methods were developed by fitting data from homogenous traffic and roadway conditions and may therefore not be appropriate for estimating travel time on arterials with heterogeneous traffic conditions. Such models do not explicitly incorporate frictional factors such as indiscriminate on-street parking activities at midblock sections, access density, and segment downstream conditions such as lane drops. These factors significantly impact traffic flow along urban arterials and failure to capture their impact in the estimation method of travel time may result in inaccurate predictions. Moreover, the application of data-driven models that appear to be promising in the context of travel time estimation and prediction may not be practically possible on arterials that lack extensive traffic sensors along the urban arterials for collecting “big data” for the modeling process. Several major arterials in Ghana lack permanent counters and sensors to acquire large and comprehensive data to serve as inputs for data-driven travel time models.

Undoubtedly, traffic congestion in sub-Saharan Africa has gained some considerable attention of governments and policy makers transport planners and engineers. However, most of the congestion studies in sub-Saharan Africa, for example, Musah et al. [8], Agyapong and Ojo [9] and Olagunju [10] have been mainly qualitative assessment of the congestion in terms of identifying causes; and the health, economic, social and environmental impacts. There is a marginal body of works in sub-Saharan Africa that seek to quantify traffic congestion on urban arterials from operational perspective. These studies, for example, Kwakwa and Adams [11], Tuffour et al. [12] and Dumba et al. [13] have mostly focused on evaluating congestion on urban arterials at the intersection level with little attention to the midblock traffic conditions (e.g., indiscriminate on-street parking activities, lane drops/additions, and driveways) and temporal factors that are expected to affect travel time. Generally, studies that model travel time of urban roadways at the arterial level under the constraints of the heterogeneous nature of the local traffic and roadway conditions are lacking in sub-Saharan Africa.

Against this backdrop, this study aims at exploring the joint impact of contributing factors of congestion by developing and validating a simple alternative travel time model for quantifying congestion on urban arterials through multiple linear regression modeling. The study explores factors including on-street parking, access density and downstream bounding conditions (e.g., lane drops/additions and roundabouts) which are not considered in previous congestion prediction models that are based on data from developed countries. Alternative models that incoporate heterogenous traffic and roadway conditions are developed and the best one selected and further tested with a new independent data set.

Previous studies on travel time prediction models

The volume delay functions (VDFs) are one of the widely used models for estimation travel time. VDFs have been extensively employed for standard highway assignments problems in travel demand modeling software. These VDFs operate with link-specific, fixed-variable values for free-flow speed and capacity typically defined by road type [14]. The most common type of the VDFs for travel demand modeling is the Bureau of Public Roads (BPR) function which was developed in the 1950s. Due to the inherent limitations of the BPR function of underestimating travel time in oversaturated conditions [15], the conical volume – delay function was developed by Spiess [16]. Following the same spirit, Akcelik [17] built on Davidson function [18] and developed a time-dependent travel time model to overcome the inaccurate prediction of oversaturated conditions. This model unlike the BPR and the conical delay function incorporates explicitly the impact of signal timing [17,14]. Skabardonis and Dowling [19] also modified the BPR function by introducing traffic signal component based on Highway Capacity Manual into their function. Further, Highway Capacity Manual (HCM) [7] developed a series of models as part of the method of estimating arterial travel time to quantify the running time and control delay over a link to estimate travel time on signalized arterials. Since the HCM [7] method does not incorporate the impact of roundabout control on travel time, Bugg et al. [20] developed a series of regression models based on geometric and operational data to predict arterial travel time for a corridor with roundabouts using data from seven existing roundabout corridors. The authors developed and calibrated sub-models for free-flow speed, roundabout influence area, geometric delay, impeded delay, and average travel speed. Fitzpatrick and Das [21] employed regression analysis to investigate the relationship between suburban vehicle operating speed and roadway characteristics including the presence of bicyclists. Recently, Kukkapalli and Pulugurtha [22] employed generalized linear modeling approach to model the effect of a freeway road construction project on link-level travel times. The authors found that subject link, upstream and downstream link characteristics have a significant effect on freeway and connnecting arterial street link level travel times. Considering temperature and rainfall spares GPS data in Wuhan, Zhu (2018) used artificial neural network approach to model link travel time.

Additionally, a summary of other recent studies on travel time in terms of location, method of modeling, application, dataset and results has been provided in Table 1. It is obvious from the literature that studies on travel time modeling at the arterial level have not received much attention in sub-Saharan African countries. This study therefore attempts to contribute to the body of knowledge in the field of congestion and/or travel time prediction modeling at the arterial level from the sub-Sahara African context.

Techniques for travel time modeling

In the literature, the techniques for developing travel time prediction and/or estimation models can be grouped into six categories: Simulation modeling, Regression modeling, Kalman Filtering modeling, Historical Average modelling, Time Series modelling and Machine Learning modelling.

Historical average models make use of data from the historical bus travel time of previous journeys to give the current and future travel time with the assumption that current traffic condition remains stationary. Therefore, a model of this kind is reliable only when the traffic pattern in the area of interest is relatively stable or where congestion is minimal (e.g., rural areas) [23].

Simulation models have gained popularity among transport engineers and planners for predicting the operational performance of highway facilities. They are computer based and analyze traffic at either macroscopic level or, microscopic or mesoscopic levels depending on the level of analysis required and the complexity of the conditions.

Regression models predict and explain a dependent variable with a function formed by a set of independent variables. One advantage of these models is that they are able to work satisfactorily under unstable traffic conditions, unlike historical databased prediction models. Moreover, regression models usually measure the simultaneous effects of various factors (e.g., density, and offset) that are independent between one another; affecting the dependent variable. They can reveal which variables are less or more important for estimating travel times. Also, it does not require big data for modeling contrary to other techniques such as machine learning [5].

Kalman filtering models have been used extensively for predicting bus arrival time. These models have sophisticated mathematical representations and the potential to adequately accommodate traffic fluctuations with time-dependent parameters (e.g., Kalman gain). Their basic function is to provide estimates of the current state of the system, but they also serve with an advantage of improving estimates of variables at earlier times; they have the capacity to filter noise [24].

With the advancement of technology for harnessing and processing traffic data, data-driven models for travel time prediction have been receiving much attention in the literature. These include Machine Learning techniques which unlike traditional statistical techniques, do not require detailed mathematical functional forms and assumptions on error distributions [25]. They are therefore able to capture complex underlying relationships among different variables even when their relationships are not easily apparent [25]. Support Vector Machines (SVM) and Artificial Neural Network (ANN) models can estimate and capture the linkage of very complex traffic flows relationships even under rapidly changing conditions [26]. Studies including Jabamony and Shanmugavel [27]; Mane and Pulugurtha [28]; Qiao et al. [29]; and Bharti et al. [30] have previously employed machine learning technologies for travel time prediction.

Time series models construct the time series relationship of travel time or traffic state, and then current and/or past traffic data are used in the constructed models to predict travel times in the next time interval [31]. Commonly used time series techniques includes Moving average (MA), Auto-Regressive Moving Average (ARMA), Auto-Regressive Integrated Moving Average (ARIMA), and Auto-Regressive Fractionally Integrated Moving Average (ARFIMA).

When choosing a method for estimating or predicting travel time, one of the most important factors to consider is the availablility data. In spite of the different number of estimation techniques and their advantages and disadvantages, it would be impossible to use any of them if the required and related sources of data are lacking. This study could not employ techniques which require relatively large data such as Machine Learning modeling, Time Series modeling, Kalman Filtering modeling, Historical Average modeling and due to the lack of sensors and permanent count stations on the Ghanaian arterials for obtaining the required traffic data size. Regression modeling approach which require relatively smaller data size was therefore employed in this study.

Table 1
Summary of previous studies on travel time modeling.

Study / Location	Method	Application	Dataset (Conditions)	Results
Kumar et al. [53] / India	• Time-space discretization	• Bus travel time prediction	GPS data	• The proposed method was able to perform better than historical average, regression, and ANN methods
Wang et al. [62] / China	• Regression model	• Real-time travel time estimation of different traffic streams	Video image processing	• A regression model was built and integrated into the travel time estimation model which reduced prediction errors
Rahmani et al. [54] / Sweden	• Consistent path inference (fixed point approach)	• Accurate path finding and route optimization	FCD from a GPS-equipped taxi	• Fixed point algorithm improves shortest path finding. • The solution is robust under different initial travel times assumptions and data sizes.
Tang et al. [55] / China	• Data-driven approach (A tensor-based context-aware approach)	• Advance Traveler Information System (ATIS)	Sparse and large-scale taxi GPS trajectories	• Proposed model provides an effective and robust approach for citywide personalized travel time estimation, and outperforms the competing methods
Qi et al. [56] / China	• Discrete and continuous combined analysis • Gaussian mixture regression	• Policy-making processes in traffic planning and management • improve reliability of trip schedules. • Navigation information systems	Trajectory data from commercial vehicles, temporal factors, weather conditions, driver differences, speed factors, travel distance	• The methodology is able to reduce the long-term travel time prediction error between 14% and 43% compared with the traditional average speed method and other baseline methods
Strauss and Miranda-Moreno [57] / Island of Montreal	• Linear regression model	• Identify factors to promote cycling route navigation • Identify most popular route, fastest route, shortest route	GPS cyclist trip data, from the Mon RésoVélo Smartphone application; Montreal	• Grade of segments as well as geometric design and built environment characteristics affect speed • Morning peak and segments which do not have signalized intersections at either end increase speed
Mil and Piantanakulchai [58] / South Korea	• Modified Bayesian datafusion approach combined with the Gaussian mixture model	• Transportation system management, • For an Advance Traveler Information System (ATIS)	Data from a microscopic simulation model, loop detector data, GPS data, historical data, and ground truth data.	• Show significantimprovement in the accuracy of travel time estimation in terms of mean absolute percentage errors(MAPE) in the range of 3.46% to 16.3%.
Li et al. [59] / China	• Coupled application of deep learning model and quantile regression	• Advanced transportation information system (ATIS)	Real-time travel time data via the float car technique	• The proposed hybrid model outperforms other models in all cases for probabilistic prediction. • deep learning models keep stable as the number of probabilistic prediction steps is increased

(*continued on next page*)

Table 1 (*continued*)

Study / Location	Method	Application	Dataset (Conditions)	Results
Lee et al. [63] / Taiwan	• Vehicular ad-hoc network (VANET) technology	• Navigation information systems	Real data collected from a GPS-based taxi dispatching system	• The TTE of the suggested paths has the possibility of being 82.2% better than the paths traveled by taxi, and the travel time is thus reduced by 15.9% on average over a year.
Kukkapalli and Pulugurtha [22] / North Carolina	• Generalized linear modeling	• Provides insights into factors that influence the average travel time on the freeway during construction • Traffic control and management during construction on freeways	Link travel time per time-of-the-day and day of-the-week, distance of a link from the work zone, link characteristics of subject link, upstream link and downstream link	• Subject link, upstream link and downstream link characteristics have a significant effect on the freeway and connecting arterial street link level travel times
Akandwanaho & Nakamura [64] / Japan	• Microsimulation	• Strategies for traffic management and control on arterials	Video data in Nagoya City	• Increased density of intersections reduces travel speed • Function of the intersection significantly impacted travel speed
Zhu et al. [65] / Wuhan, China	• Artificial Neural Network modeling	• Predict link travel time using sparse data	Meteorological factors: Temperature and rainfall sparse GPS data	• ANN model based on feature relationships between target links and adjacent links and big historical data produce better predictions when data are sparse • Day of week, the 30 min interval of the day, and the expected speed of adjacent links had higher influence on link travel time prediction than other factors
Liu et al. [60] / China	• Bayesian fusion	• Advance Traveler Information System (ATIS)	Loop detector and probe vehicle data from simulation model	• Proposed method outperforms probe vehicle data based method, loop detector based method and single Bayesian fusion
Kumar and Sivanandan [5] / India	• Multiple Linear Regression Modeling	• Real-time display of congestion status in ATIS applications in a cheap and effective manner.	Travel time data from GPS-fitted buses, cars, 3-wheelers and 2 wheelers, carriageway width, presence/abscense of signalized intersection	• A wide variation in congestion index (CI) between classes of vehicles was observed • Increase in carriageway width reduces CI • Presence of signalized intersection increases congestion
Bugg, et al. [20] / USA	• Regression modeling	• A step-by-step framework for calculation of travel time on a roundabout corridor in a manner consistent with the HCM [7], Chapter 17, methodology.	Roundabout geometry and capacity, subsegment length, presence of raised median, driveway density, grade, posted speed limit, and circulating speed, volume	• A series of simple linear models could be used successfully to describe the relationship between the components of travel time on a roundabout corridor and the geometric and operational elements of a roundabout corridor
Fitzpatrick and Das [21] / USA	• Regression modeling	• Development of strategies that improve traffic operations on urban arterials.	On-road tube speed data, crowdsourced data, and roadway characteristics	• Traffic flow, access density, and the posted speed limit, presence of bus stop significantly influenced speed.

(*continued on next page*)

Table 1 (*continued*)

Study / Location	Method	Application	Dataset (Conditions)	Results
Jain and Jain [61] / India	• Regression modeling	• Forcasting and prioritising arterial segments in terms of congestion levels	Traffic and roadway factors including travel time data from moving car method, volume and density, number of intersections	• A multiple regression model of travel time gave substantially accurate predictions • Congestion index were estimated for segment for prioritization in terms of congestion levels.
Qiao et al. [29] / USA	• Comparative evaluation of Historical average, (ARIMA), KNN, KNN-T travel time prediction models	• Short-time travel time prediction for stochastic freeway applications	Real time Bluetooth travel time data	• KNN-T outperformed the rest. It captures the time-varying trend, which leads to a more precise future traffic conditions.
Saw et al. [50] / India	• Multiple Linear Regression (MLR) approach	• Development of mitigation strategies to improve travel time	Classified traffic volume, speed and travel time data from GPS, road side friction severerity rating, average segment speed, dropdown speed at intersections	• Sensitivity analysis of the travel time model indicates that mixed traffic flow, road side friction and presence of intersections have significant impact on travel time
Mane et al. [28] / USA	• Comparative evaluation of neural network model, nonlinear autoregressive with external inputs (NARX) model, and nonlinear autoregressive (NAR) model	• Advance Traveler Information System (ATMS) applications	Historical travel time data from a private data source TOD, DOW, weekday / weekend, and cumulative time-stamp (seconds) variables	• The results obtained indicate that the NARX model outperformed the two-layer feedforward neural network model and NAR model. • NAR model performed better than the traditional neural network model
Bharti et al. [30] / India	• Comparative evaluation of Stochastic Response Surface Method (SRSM), Artificial Neural Network (ANN) and Linear regression	• Identifying sources of travel time variation uninterrupted traffic flow facility • Application in traveler route and departure time decisions	Classified traffic volume data from video cameras, Travel time from License plate matching technique	• SRSM and ANN models are more efficient for stochastic modeling than regression model. • Traffic volume and composition of car significantly affect travel time of cars

Table 2
Geometric characteristics of the study arterials in both directions.

S/N	Arterial	# of Carriageways	# of through lanes	# of access points	# of signals	# of segments	Length (km)
A1	Suame Rbt/Buoho road	Varies -1 / 2*	1-2	24 (25)	2	5 (5)	9.97
A2	Sofoline/Abuakwa road	Varies -1 / 2*	1-2	20 (27)	0	6 (6)	8.80
A3	Santasi/Kotwi road	Varies -1 / 2*	1-2	35 (37)	0	5 (5)	9.06
A4	Atonsu/Esereso road	Varies -1* / 2	1-2	21	0	5 (5)	6.63
A5	Kumasi/Ejisu road	2	2	19 (31)	3	5 (5)	14.74
A6	Manhyia/Buokrom road	1	1	22	0	4 (4)	4.45
A7	Tafo/Mampong road	Varies -1* / 2	1-2	28	0	4 (4)	9.70
A8	Ring road	Varies -1 / 2*	1-2	42 (47)	9	8 (8)	18.69

xx (xx) – Direction towards CBD (Direction away from CBD).
* Dominating in length relative to the other.

Materials and methods

The following sections describe the approach used to model the urban arterial travel time, including study arterials, data collection and extraction, and model development. Fig. 1 presents the flow chart for conducting this study.

Study arterials

For this research, eight major urban arterials were selected for data collection in the Greater Kumasi Metropolitan area. The selected arterials include one ring road and seven radial roads of length ranging from 4.5 to 18.7 km. The arterials are designated as A1 through to A8 as shown in Fig. 2 and described in Table 2. All study arterials are mixed controlled facilities with signalized intersections, roundabouts and other unsignalized intersections. Except for the Manhyia / Buokrom road which has a single carriageway throughout, all others have some segments as two lane dual carriageways and the rest are one lane single carriageway. The average carriageway width ranges between 5 and 13 m. The arterials are of high importance since most of the trips towards the central business district (CBD) in the morning and away from the CBD in the evening are made on these arterials.

Data collection

Roadway characteristics

Roadway characteristics and features including geometry were obtained from Google Earth and were later confirmed on the field. The selected study arterials have two or more intersection control types including roundabouts, signalized intersections and unsignalized intersections. They are characterized by side frictions, particularly from access points (driveways) and on-street parking activities by public transport minibuses and taxis. For the purpose of this study, the arterials have been divided into segments using major intersection as boundaries or a point where there is a major change in roadway characteristics such as lane drops. A typical segment begins from the yield/stop line of the upstream intersection to the yield/stop line of the downstream intersection. The segment was defined to be consistent with the Highway Capacity Manual (HCM) urban streets procedure making sure that segments exhibit relatively constant traffic conditions and geometric features (HCM, [7]). The length of the study arterials varies with a minimum of about 4.5 km for the Manhyia / Buokrom road and a maximum of about 18.7 km for the ring road. Table 2 shows a summary of the geometric characteristics of the study arterials.

Traffic data collection

This study collected travel time and segment flow data on the selected major arterials using the moving observer method. The floating vehicle was equipped with an android mobile application called My Tracks. Observers within the floating vehicles entered waymarkers as the vehicle got to the yield or stop lines which served as the segment boundaries. GPS files were obtained for all runs and uploaded into Microsoft Excel for the extraction of travel time data. Using the timestamps corresponding to the waymarkers, the travel time, *tij*, in seconds, between each pair of consecutive waymarkers at locations *i* and *j* was computed as shown in Eq. (1).

$$tij = (3600 \times hj + 60 \times mj + sj) - (3600 \times hi + 60 \times mi + si) \tag{1}$$

The moving observer method involves a series of runs by the test vehicle made traveling 'with' and 'against' a traffic stream whiles recording the following for each segment of length *l*:

- number of opposing vehicles met (m_a)
- number of vehicles overtaken by the test vehicle (m_p)
- number of vehicles that overtook the test vehicle (m_o) and
- travel time over each segment for a trip with traffic stream (t_w) and a trip against traffic stream (t_a)

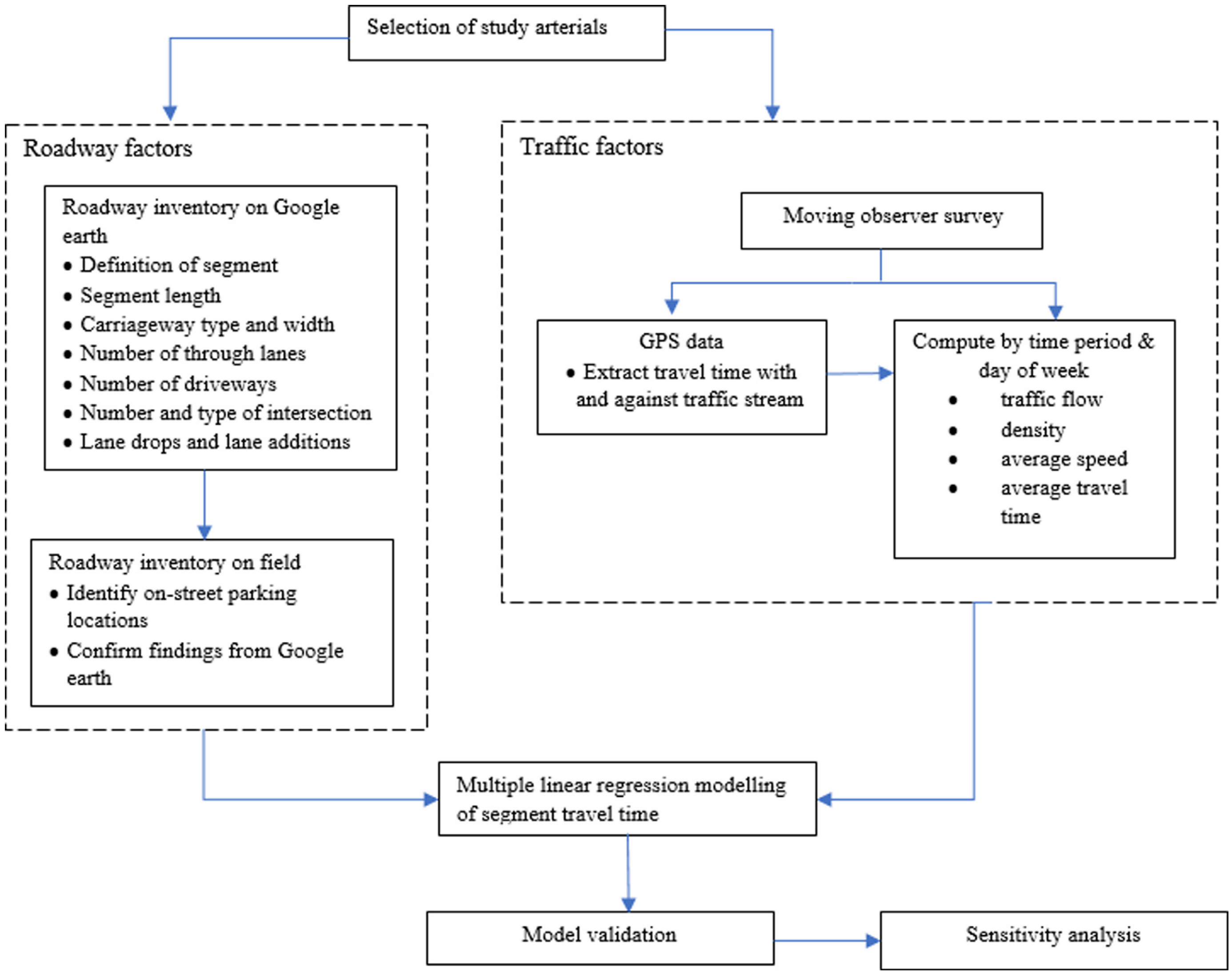

Fig. 1. Flow chart for the travel time modeling.

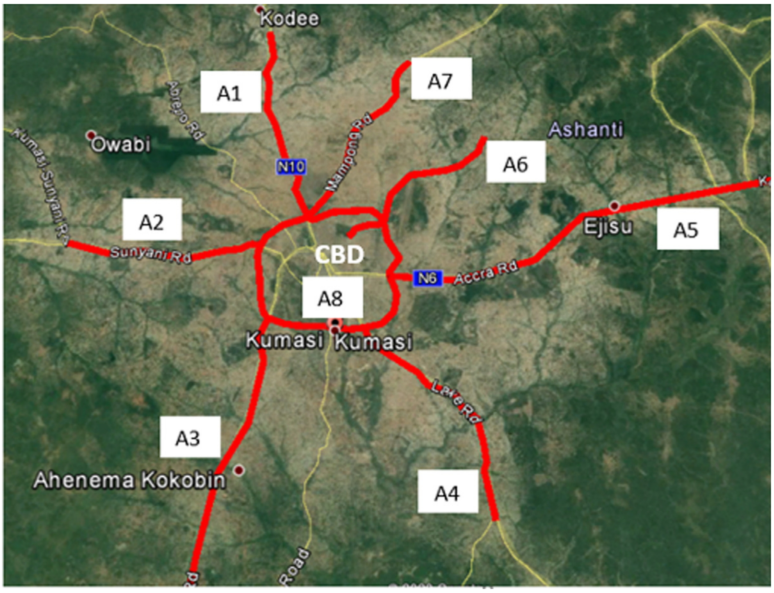

Fig. 2. Map showing the study arterials in Kumasi.

Traffic stream parameters flow (q), speed (v_s), density (k) and average travel time (t_{avg}) were estimated for each segment per direction of trip using Eq. (2) through to (5) [32,33].

$$q = \frac{m_o - m_p + m_a}{t_w + t_a} \tag{2}$$

$$t_{avg} = t_w - \frac{m_w}{q} \tag{3}$$

$$v_s = \frac{l}{t_w - \frac{m_w}{q}} \tag{4}$$

$$k = \frac{q}{v_s} \tag{5}$$

Additionally, another phone was set up in the test vehicle for videotaping the trips. This method made it possible to record the roadway environment to identify some roadway factors contributing to congestion. The data collection was conducted over four days (i.e., two weekdays and two weekends) for the AM (6:00 am–10:00 am) and PM (2:00 pm–6:00 pm) time periods. All put together, GPS data from a total number of 157 trips across 84 segments were used for this study. Considering the number of segments per direction for each arterial and the number of trips made per direction over the four days, a total of data set of 803 cases were available for modeling purposes.

Variable description

This study sought to develop a statistical model that captures the impact of traffic factors, roadway features, roadside frictions and temporal factors for estimating travel time of through traffic on arterials. An initial inventory undertaken across the arterials revealed that locations with the presence of indiscriminate on-street parking activities, lane drops, signalized intersections and driveways (access points) were characterized by slower movement and stop-and-go phenomena. Traffic factors considered for modeling were traffic flow, traffic density and free-flow travel time. The roadway features considered include segment length, number of through lanes, carriageway width, carriageway type and the segment downstream bounding condition. The downstream bounding condition was considered under two cases: Case 1 which considers it as a categorical variable with five levels, which are roundabout, traffic signal, two-way stop control (TWSC), lane drop and lane addition and Case 2 considered downstream bounding condition as a binary variable as either intersection control or other conditions (lane drop/lane addition).

On-street parking activities and access points are sources of friction on the arterials. The on-street parking for this study represents the illegal parking activities of vehicles (e.g. paratransits) on the travel lanes which disrupt the normal flow of traffic. The on-street parking areas are characterized by bus bunching conditions as a result of several public transport minibuses and taxis illegally parking in the travel lanes in search of passengers. It is worth mentioning that, the access point density does not include the first and last intersections bounding the segment.Time periods and day of week were the temporal factors considered in the model because the traffic condition is expected to vary considerably from time to

Table 3
Key variables considered for travel time modeling.

Variable	Symbol	Descriptions
Segment length (km)	L	Length of segment; Continuous variable
Free flow travel time (sec)	T_o	L/ Free flow speed; Continuous variable
Number of lanes	NL	Number of lanes in travel direction; Discrete variable
Carriageway type	MDN	Dummy variable; = 1 if Divided road, 0 otherwise
Time of day	TOD	Dummy variable; = 1 if PM peak, 0 otherwise
Day of week	DOW	Dummy variable; = 1 if weekend, 0 otherwise
Travel direction	DIR	Dummy variable; = 1 if travelling in non-peak direction, 0 otherwise
Access point density (access points/km)	ACD	Number of access points per km; Discrete variable
Presence of on-street parking	PAK	Dummy variable; = 1 if present, 0 otherwise
Traffic density (veh/km)	K	Number of vehicles per km; Continuous variable
Carriageway width (m)	CW	Average width of carriageway; Continuous variable
Downstream bounding condition (Case 1)		
Signalized	SG	Downstream condition dummy variable; = 1 if present, 0 otherwise
Roundabout	RBT	Downstream condition dummy variable; = 1 if present, 0 otherwise
Lane drop	LDP	Downstream condition dummy variable; = 1 if present, 0 otherwise
Lane addition	LAD	Reference category
Two Way Stop Control (TWSC)	STP	Downstream condition dummy variable; = 1 if present, 0 otherwise
Downstream bounding condition (Case 2)		
Intersection controlled	IC	Downstream condition dummy variable; = 1 if present, 0 otherwise
Other conditions (lane drops/lane additions)	OC	Reference category

time and from day to day. The time period was modeled at two levels as AM peak and PM peak periods. Also, day of week was categorized into weekday and weekend. Moreover, the direction of travel for a specific time period was expected to influence the travel time along an arterial. This factor was modeled as peak direction and non-peak direction. The traffic density and free-flow travel time on each segment were considered as the traffic operational factors. The free-flow travel time is the average travel time that can be experienced on a roadway segment under zero traffic density or relatively low traffic volumes conditions. The free-flow travel time was estimated as the segment length divided by the free-flow speed. In free-flow conditions, vehicles are assumed to travel at a free-flow speed approximately equal to the posted speed limit of 50 km/h within the urban context. Table 3 gives a summary of all factors considered for modeling average travel time.

Model development

Model formulation

Travel time has been reported to increase exponentially with the increasing demand intensity and its interactions with other factors over a fixed capacity base, and directly with free-flow travel time [34,35]. It was therefore assumed that the proposed model will take the exponential functional form shown in Eq. (6). To estimate the parameters of the travel time model, the exponential functional form was log-transformed into a linear form as shown in Eq. (7).

$$\mathrm{TT} = a_0 L^{a_1} T_0^{a_2} \exp^{(\beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \ldots + \beta_n X_n)} \tag{6}$$

$$\ln(\mathrm{TT}) = \ln(a_0) + a_1 \ln(L) + a_2 \ln(T_0) + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \ldots + \beta_n X_n \tag{7}$$

where

- L is segment length
- T_o is free - flow travel time (in seconds) estimated as segment length divided by average free flow speed
- $\beta_{1,}$ β_2, $\beta_{3,}$ β_n are the estimated regression coefficients of the other independent variables
- X_1, X_2, X_3, X_n are additional independent variables to L and T_o.

It is worth mentioning that, two alternative models were explored with this functional form: Model 1 considers the downstream bounding condition as a variable with five (5) categories (Case 1) and Model 2 considers the downstream bounding condition as a binary variable (Case 2).

Modeling procedure

The study employed multiple linear regression (MLR) approach to quantify the joint impact of roadway and traffic factors on travel time as a measure of congestion. Multiple linear regression was chosen because it does not need large data points for modeling relative to other methods such as Machine learning. It allows for the effect of predictors to be quantified, unlike methods such as ANN which are more like a black box. After removing outliers, a total of 791 data set was available to build the model. The Stepwise AIC backward regression function in R programming was employed for the modeling effort. This function builds a regression model from a set of candidate predictor variables by removing predictors based on

Akaike information criterion, in a stepwise manner until removing any of the remaining variables does not improve the model fit. Further, the final models from the backward stepwise process were compared based on the Akaike information criterion (AIC), Bayesian information criterion (BIC) and adjusted R squared metrics to select the best model. The AIC and BIC criterion shown in Eqs. (8) and (9), respectively were included in the performance metrics because they penalize for the number of predictors included in the model, unlike the R square. Models with lesser AIC and BIC values are considered to be the better ones. Although the model selection criterion was based on the aforementioned metrics, it was ensured that all predictor variables included in the selected model were intuitively logical in sign. Also, the level of significance of the predictor variables was taken into consideration.

$$\mathrm{AIC} = -2\log(\mathrm{maximum\,likelihood}) + 2\mathrm{k} \tag{8}$$

$$\mathrm{BIC} = -2\log(\mathrm{maximum\,likelihood}) + \mathrm{k}\log(\mathrm{n}) \tag{9}$$

Where k is the number of model parameters and n is the sample size.

Model performance testing

After calibrating the models with the training data and selecting the best model, the next step was to test its performance through a cross-validation process. The cross-validation is an essential process for assessing how the predictive model will generalize to an independent data set. For this purpose, a data set of 466 cases obtained from a moving observer survey were used as the testing data. The testing data were collected on different days independent of when the training data were collected. The performance of the model was evaluated in two contexts, which are, point predictions and prediction intervals (PIs).

The model's performance evaluation in point estimates of travel time was based on the following model performance metrics: root mean squared error (RMSE), mean absolute error (MAE), and mean absolute percentage error, (MAPE) expressed mathematically as shown in Eqs. (10), (11) and (12), respectively.

$$RMSE = \sqrt{\frac{1}{n}\sum_{j=1}^{n}\left(y_j - \hat{y}_j\right)^2} \tag{10}$$

$$MAE = \frac{1}{n}\sum_{j=1}^{n}\left|y_j - \hat{y}_j\right| \tag{11}$$

$$MAPE = \frac{100}{n}\sum_{j=1}^{n}\frac{y_j - \widehat{y}_j}{y_j} \tag{12}$$

Where

- n is number of observations
- y_j is observed value
- $\hat{y}_j$ is predicted value

On the other hand, two main metrics are of interest when evaluating prediction intervals [36,37]. The first one is called the Prediction Interval Coverage Probability (PICP). This metric essentially represents the percentage of test instances for which the target value (i.e., actual travel time) falls within the PIs. PICP assesses the validity of the prediction intervals. Intervals should on average enclose the actual values in (1-α)% of all cases [38]. It is expressed mathematically as:

$$PICP = \frac{1}{n}\sum_{j=1}^{n}c_j \tag{13}$$

where n is the number of samples. If the actual observations of the *jth* travel time sample fall into the predicted interval, then c_j =1; otherwise, c_j = 0. When the PICP is significantly lower than the predetermined confidence level, the PI is unreliable. The second interval prediction metric is the normalized mean prediction interval width (NMPIW) which assesses the optimality; where narrower intervals are more desirable [37,38]. NMPIW is written mathematically as shown in Eq. (14).

$$NMPIW = \frac{1}{n}\sum_{j=1}^{n}\frac{U\left(x_j\right) - L\left(x_j\right)}{t_{max} - t_{min}}. \tag{14}$$

Where $U(x_j)$ and $L(x_j)$ are the upper and lower interval boundaries of the jth sample, respectively; $t_{\max}$ and $t_{\min}$ denote the maximum and minimum values of the actual travel time, respectively. Generally, a prediction interval with a high PICP

Table 4
Constant values and conditions of variables for sensitivity analysis.

Variable	Constant values and conditions
Free flow travel time	60 seconds
Number of lanes	1
On-street parking	Yes
Day of week	Weekday
Time of day	AM peak
Access point density	4 access points per km
Downstream bounding condition	Signalized intersection
Direction of travel	Peak direction

and low NMPIW is considered high quality. Both metrics, however, assess the PIs from only one aspect. Therefore, a combined metric, that is, coverage width-based criterion (CWC) modified is considered as a comprehensive evaluation metric of the quality of PIs. The CWC is expressed as shown in (15).

$$CWC = NMPIW + \gamma(PICP)e^{-\delta|PICP-\mu|} \tag{15}$$

where μ is the predetermined confidence level, δ is the penalty parameter (usually a large number), and

$$\gamma(PICP) = \begin{cases} 1, & PICP < \mu \\ 0, & PICP \geq \mu \end{cases} \tag{16}$$

From Eq. (16), the CWC is equal to NMPIW when PICP is greater than μ. Both PICP and NMPIW, however determine CWC when PICP is lower than μ. A high CWC means poor quality PIs, and vice versa. It is worthy of note that the interval prediction evaluation of the selected model was performed under 95% confidence level.

Sensitivity analysis

Sensitivity analysis is a method of measuring the changes to which fluctuations in variables of a system or mathematical model have on the performance of the system. This technique was employed to evaluate the robustness of the model, identify the factors that cause significant impact on travel time and gain deeper insights into the complex interactions among the contributing factors of traffic congestion. In this study, the impact of on-street parking, day of week, number of through lanes, travel direction, time of day and access density under varying traffic demand (density) were explored. For each case, the variable under investigation and traffic density were varied while all other variables were kept constant. For the case for access density, impact on travel time was investigated for access density range of 1–6 acess points/km at an interval of 1, under varying traffic density, ranging from 20 to 80 veh/km at 20 veh/km interval. For all other cases, traffic density was varied from 0 to 110 veh/km at an interval of 10 veh/km. Table 4 shows the values and conditions used for keeping the variables constant when required.

Results and discussions

Bivariate correlation analysis

A bivariate correlation analysis was first performed to identify significant variables and how the contributing factors relate individually to travel time and other variables at 95% confidence interval. Only correlation coefficients between potential predictor variables and travel time have been shown in Table 5. The results indicate that all predictors except carriageway width, time of day and roundabout significantly correlated with travel time. However, the correlation results are not conclusive to make causal inferences concerning the data. As expected, segment length, access point density, on-street parking activities, free-flow travel time and traffic density positively correlated with travel time. Similarly, the negative correlation exhibited by number of through lanes, carriageway width, carriageway type, travel direction and day of week were logical. Interestingly, flow correlated negatively with travel time. A possible explanation to the negative correlation between traffic flow and travel time is that oversaturated traffic conditions are characterized by low traffic flows associated with low vehicular speeds which tend to increase the density over a section of roadway and consequently increase travel time. The opposite is true for unsaturated conditions. This is an inherent problem of using traffic flow as a lone predictor of congestion because for any given flow there are two possible traffic states. As agued by Kachroo and Sastry [39], “A vehicle travels with a speed that is consistent with the traffic density, and hence its travel time depends on that density. Since flow gives two different densities and hence two different travel times, flow cannot determine a unique travel time on a link.” Due to this inherent problem with traffic flow, traffic density was used as a predictor instead. It is worth mentioning that, the correlation checks among the predictor variables revealed multicollinearity between carriageway type and number of lanes as well as free-flow travel time and segment length.

Table 5
Correlation coefficient between independent variables and travel time.

Variable	Travel time
Segment length	0.588**
Carriageway width	-0.038
Number of through lanes	-0.230**
On-street parking	0.223**
Time of day	0.034
Day of week	-0.107**
Travel direction	-0.119**
Traffic flow	-0.190**
Access density	0.129**
Traffic density	0.334**
Free flow travel time	0.588**
SignalizedSignalized intersection	0.110**
Roundabout	0.011
Two way stop control	-0.015
Lane drop	-0.122**
Lane addition	0.082*
Carriageway type	-0.141**

** . Correlation is significant at the 0.01 level (2-tailed).
* . Correlation is significant at the 0.05 level (2-tailed).

Table 6
Estimates of the alternative travel time models.

	Model 1			Model 2		
Predictors	Beta	Std. Error	Sig	Beta	Std. Error	Sig
(Intercept)	1.367	0.135	0.000	1.345	0.126	0.000
Log(Free flow travel time)	0.867	0.022	0.000	0.873	0.022	0.000
Number of Lanes	-0.263	0.026	0.000	-0.242	0.022	0.000
On street parking	0.092	0.037	0.014	0.113	0.036	0.002
Day of week (Weekend)	-0.137	0.032	0.000	-0.139	0.032	0.000
Time of day (PM peak period)	0.063	0.032	0.050	0.064	0.032	0.048
Travel direction (Non peak direction)	-0.067	0.033	0.041	-0.070	0.032	0.030
Access point density	0.045	0.01	0.000	0.040	0.010	0.000
Traffic density	0.011	0.000	0.000	0.011	0.000	0.000
Downstream bounding condition (Intersection control)	na			-0.075	0.040	0.053
Downstream bounding condition	na			na		
Lane addition ***(reference category)***				na		
	na					
Roundabout	-0.071	0.078	0.363	na		
SignalizedSignalized intersection	0.082	0.084	0.331	na		
Two way stop control	-0.066	0.071	0.356	na		
Lane drop	0.064	0.085	0.451	na		
R-sq	0.75			0.75		
Adj R-Sq	0.75			0.75		
AIC	982.93			984.65		
BIC	1048.36			1036.06		

Note: na = not applicable.

Alternative travel time regression models

This section presents the estimates of the two models: Model 1, which considers the downstream bounding condition as a five categorical variable and Model 2, which considers the downstream bounding condition as a binary variable. Only the final outputs of the stepwise process of the two models are shown in Table 6. It is worth noting that, segment length and carriageway type were removed from the final model during the step-wise process due to multicollinearity conditions. Both models, Model 1 and Model 2 appear to be a good fit of the data with $F\,(12,\,778) = 198.62$, and $F\,(9,\,781) = 262.4$, respectively, $p < 0.0005$ at 95% confidence interval.

The results in terms of variable signs and significance are very similar in both models. Free-flow travel time, number of lanes, on-street parking, day of week, travel direction, access density, traffic density and time of day are all significant predictors of travel time in both models at 95% confidence level. Surprisingly, all categories of the downstream bounding conditions in Model 1 appear to be non-significant contributing factors to travel time. However, when the downstream bounding condition is considered as a binary variable (i.e., either intersection control or other conditions) in Model 2, it contributes significantly to travel time. Relative to intersection controls, other conditions at the downstream of segments (e.g., lane drops) increase travel time.

Further, all variables in both models have logical coefficient signs. All other variables held constant, travel time is seen to increase with increasing free-flow travel time, presence of on-street parking activities, access density and traffic density; hence their positive coefficient. As expected, number of lanes, weekends and traveling in the non-peak direction have negative coefficients. The addition of travel lanes intuitively increases the physical roadway capacity which reduces congestion and consequently reduces travel time [7,40]. The negative coefficient of the weekend variable is also logical because traffic demand on weekends is generally low compared to weekdays. Therefore, trips on weekends are expected to be associated with lower travel time relative to weekday trips. Similarly, traveling in the non – peak direction with less traffic demand is expected to reduce travel time. Interestingly, both models show that PM peak period trips are associated with higher travel times relative to AM peak period trips.

With lane addition as reference condition in Model 1, signalized intersections and lane drops are observed to be associated with higher travel times relative to roundabout and two-way stop-controlled intersection at the segment downstream. The positive effect of signalized intersections is consistent with the congestion index model developed by Kumar and Sivanandan [5]. This finding is also in line with the body of literature [41–44] that reported signalized intersections to be associated with higher delays relative to roundabouts. The negative coefficients of roundabout could be attributed to the fact that vehicles do not to come to a complete stop at roundabouts but yield to the circulating traffic. Another reason for the negative coefficient for the presence of roundabouts is that the driver behavior under high traffic demand conditions sometimes changes the control scheme into priority reversal or limited priority. Priority reversal occurs due to forced entry by the entering vehicles. Moreover, limited priority is said to exist when vehicles in the circulatory roadway make co-operative adjustments to accommodate the merging and crossing behavior of the entering vehicles [45,46].

Similarly, mainline vehicles at a two-way stop-controlled (TWSC) intersection are not mandated to stop because they have the right of way. This tends to reduce travel time on segments with TWSC at the downsteam compared to signalized intersection and lane drops. Signalized intersections and lane drops are however characterized by vehicular stops; hence their positive effect on travel time relative to roundabouts and TWSC intersections. Moreover, Model 2 suggests that having lane drops rather than intersection control at downstream of the segment is associated with higher travel times. Lane drops are characterized by distruptive driver behaviors which further reduce capacity and therefore increase travel time. Lane straddling, forced and late merges and queue jumping are typical driver behaviors that increase congestion and travel time at lane drops. For instance, queue jumping usually evoke aggressive behavior by other drivers to lane straddle [47]. Drivers who lane straddle attempt to prevent late merging by vehicles behind them or prevent queue jumpers from moving ahead of them. The entire queue behind drivers who lane straddle end up beign slowed down and consequently experience longer travel time [48].

Model selection and predictor quantification

Both models have most of the predictor variables to be significant and consistent in terms of signs of coefficient, their performances are not the same. Interestingly, both models also explain 75% of the variability of travel time. However, Model 1 yields AIC and BIC values of 982.93 and 1048.36, respectively whiles Model 2 yields AIC and BIC values of 984.65 and 1036.06. Although the AIC of Model 1 is slightly lower than that of Model 2, Model 2 is considered a better model of the two since it has the least BIC due to its least number of predictor variables. Eq. (13) shows the equivalent exponential form of Model 2.

$$TT = 3.84xT_o{}^{0.873}x\exp^{(-0.242xNL+0.113xPAK-0.139xDOW+0.064xTOD--0.070xDIR+0.040xACD+0.011xK--0.075xIC)} \quad (13)$$

Where all variables are as described in Table 3.

One advantage of MLR is the possibility of quantifying the effect of change in a predictor on the respondent variable with the assumption that all other variables are held constant. Focusing on the best model (Model 2), the presence of on-street parking on a segment is associated with about 11% increase in travel time, all other variables held constant. This is consistent with the findings of studies such as [49,50] which mentioned that parking activities affect the quality of stream flow and thus increases travel time. Similarly, shifting from an AM peak period trip to PM peak period trip is expected to result in 6.4% increase in travel time. Further, increasing access point density by one access point/km is associated with 4% increase in travel time similar to the findings by Eisele et al. [6]. As mentioned in the HCM [7], some drivers naturally slow down when approaching an access point to allow for potential conflicts and this contributes to delays. On the other hand, an additional travel lane to the existing is expected to reduce travel time by 24.2%, all other predictors held constant. Similarly, switching from a weekday trip to a weekend trip is associated with 14% reduction in travel time. Moreover, the presence of lane drops at the downstream of segments are expected to result in travel time about 7.5% more than having intersection controls at the downstream of segments.

Model prediction performance evaluation

Point predictions

The results of the cross-validation of Model 2, performed with an independent data set as test data and the training data are presented here. The point evaluation results indicated very low RMSE, MAPE and MAE of 0.44 s, 6.8% and 0.35 s,

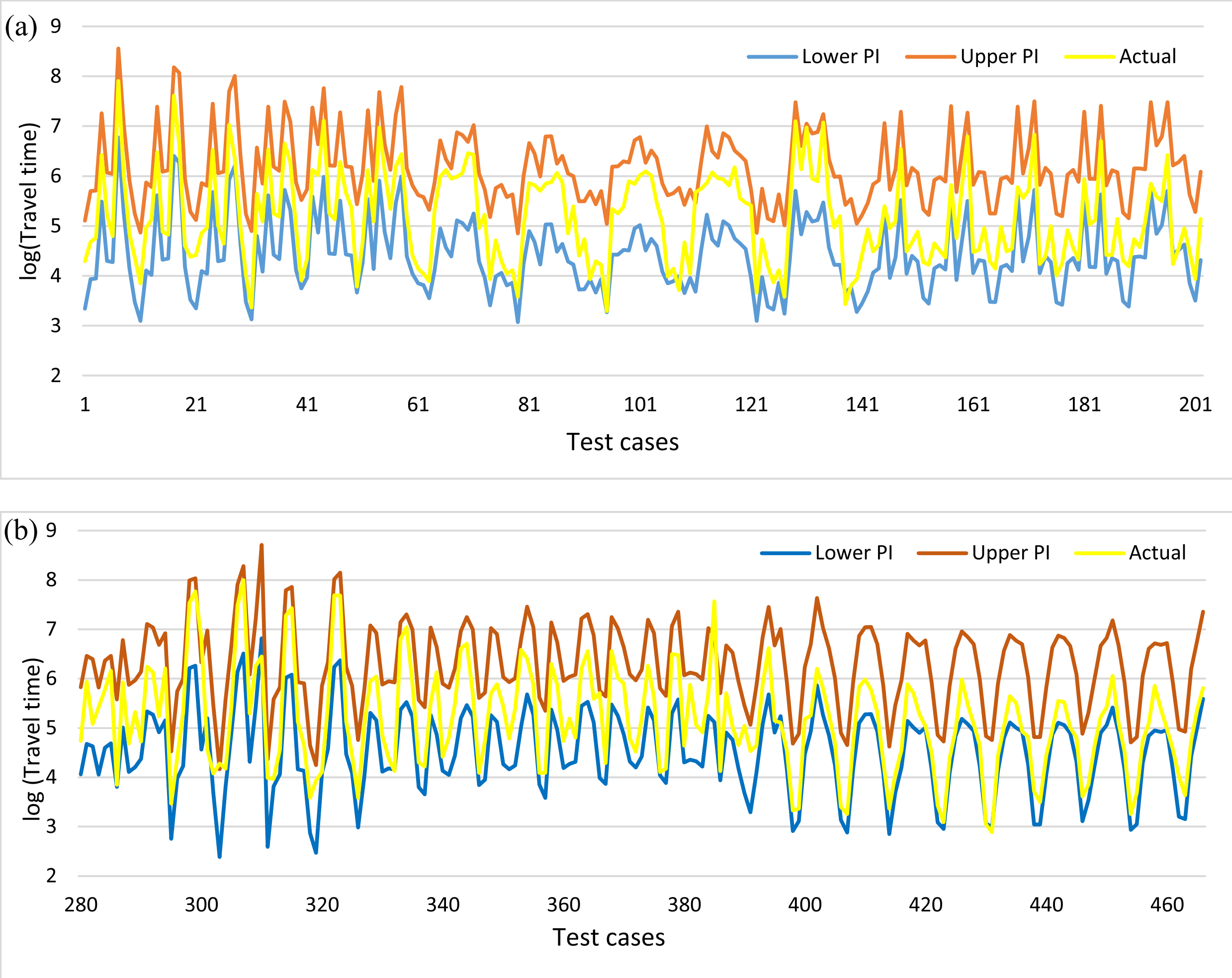

Fig. 3. (a). Prediction intervals with 95% confidence level: Cases 1–201 (b) Prediction intervals with 95% confidence level: Cases 280–460.

respectively when used for predicting training data. Similarly, RMSE, MAPE and MAE of 0.47 s, 7.2% and 0.37 s, respectively were recorded when used for predicting the test data. The performance metrics of the two data sets (i.e., training and test data) do not differ substantially; an indication of the high predictive power of the model. A high training and high testing error indicates an underfit model while extremely low training error but extremely high testing error indicates an overfit model [51]. Based on the point predictions from the cross-validation, it can be said that the model has no overfitting and underfitting problem. The model is therefore considered to be effective in terms of point prediction with a new independent data set.

Interval predictions

Predictions and their corresponding prediction intervals (PIs) were carried out from the test data under 95% confidence level to assess the validity and optimality of the model. Fig. 3a and b show the constructed PIs with a 95% confidence level for cases 1–201 and 280–460, respectively. These figures highlights the fact that, in most cases, the actual travel time fall within the PIs. Moreover, it is observed that the trends of the upper and lower PIs are basically consistent with the actual observations. However, there are few cases (e.g., cases 385 and 431 in Fig. 3b) where the actual travel times in the test data fall outside the constructed PIs. This could be attributed to the fact that the model is trained to construct PIs with 95% confidence level, not 100%. The PICP of the predicted intervals, calculated to be 97.2% exceeds the predetermined confidence level. This indicates that the predicted interval generated by the model is effective and able to satisfactorily cover the actual travel time. Moreover, the NMPIW value was found to be 34.6%. The NMPIW value is equal to the CWC since the calculated PICP was greater than the predetermined confidene level (95%). In summary, the above results indicate that the model provide satisfactory PIs and can adequately evaluate the uncertainties corresponding to the travel time prediction.

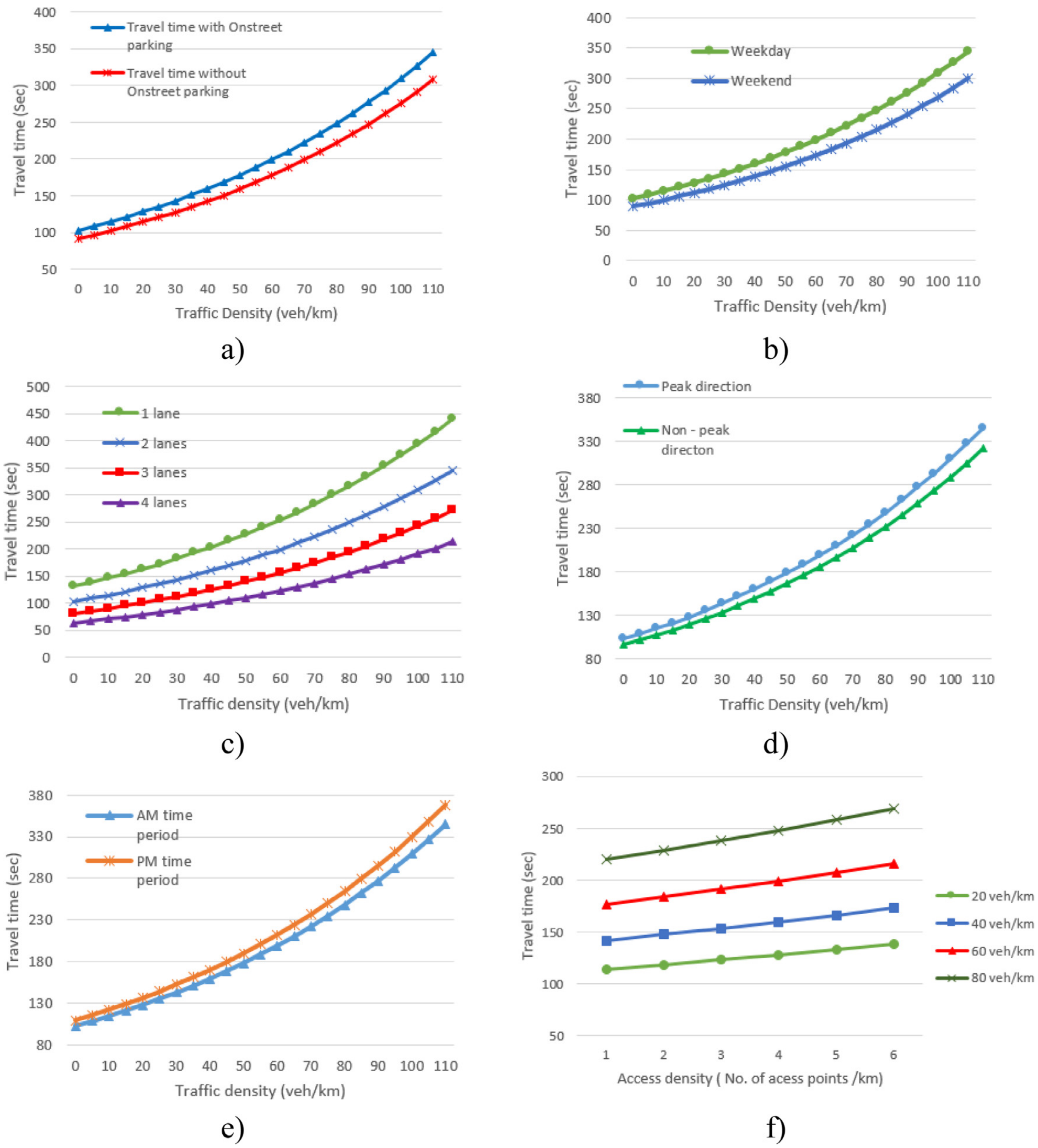

Fig. 4. Sensitivity analysis plots of (a)Travel time against traffic density and on street parking condition; (b)Travel time against traffic density and day of week; (c)Travel time against traffic density and number of lanes; (d)Travel time against traffic density and travel direction; (e)Travel time against traffic density and time of day; and (f)Travel time against access density and traffic density.

Sensitivity analysis results

The sensitivity analysis reveals the responsiveness of the travel time model under different traffic and roadway conditions. Fig. 4a–f highlights the impact of predictor variables on travel time under varying traffic density. From Fig. 4a–e, travel time is seen to increase exponentially with increasing traffic density (K). The model is seen to satisfy the properties of i) Monotonicity ($K_1 \geq K_2 \Rightarrow T(K_1) \geq T(K_2)$) and ii) Positivity ($\forall K \geq 0,\ T(K) \geq 0$), as part of the properties all travel time functions are expected to have [16,34]. Fig. 4a shows that segments with presence of on-street parking activities will have higher travel times relative to segments without on-street parking activities. However, this effect appears to increase with increasing traffic density. Similarly, weekdays and PM time periods depicted by Fig. 4b and 4e, respectively are associated with higher travel time relative to weekends and AM time periods. But this effect also increases with increasing traffic density. The low weekend travel times relative to weekdays could be attributed to less vehicular movements on weekends [52], consequently reducing the impact of on-street parking and access point density on travel time.

Moreover, the exponential effect of density on travel time is seen to be significantly reduced with increasing number of lanes as shown in Fig. 4c. The reduction in travel time is however seen to increase with increasing traffic density. This indicates that higher benefits could be achieved on segments with relatively high traffic density when travel lanes are added as a mitigation strategy of congestion. Further, traveling in the peak direction is associated with longer travel time relative to the non-peak direction for a given density as shown in Fig. 4d. It is interesting to note from Fig. 4f that the rate of impact of access point density on a segment travel time depends on the level of traffic density as all four lines appear to divert from each other as access density increases.

These results highlight the possible interactions of the contributing factors of congestion. Since travel time increases with traffic intensity on a segment with a fixed capacity, for both on-street parking and without on-street parking conditions; and reduces with increasing number of lanes, increasing the number of lanes (physical capacity) could reduce the impact of on-street parking activities on travel time. The physical capacity could also be increased and consequently improve travel time by other strategies including enforcement of traffic and roadway regulations to curb indiscriminate parking activities on travel lanes; and hard shoulder usage in the peak direction. The relationship of travel time with access point density indicates that checks should be put into place regarding the provision of access points during the design and management of urban arterials.

Conclusions, contributions and recommendations

Conclusions

The study developed a model for estimating segment travel time as a measure for quantifying congestion on urban arterials for a city. To achieve this, multiple linear regression technique was employed to analyze the impact of roadway features, roadside friction, temporal and traffic factors on travel time. Two alternative models were explored and the better one was selected. The selected model was found to explain about 75% of the variability in the data. Further, the model prediction performance was conducted by point prediction and interval prediction evaluations with an independent data set. The prediction performance evaluation results indicated that the model provides satisfactory point predictions and interval predictions, and can adequately evaluate the uncertainties corresponding to the travel time prediction.

Presence of on-street parking, access point density, traffic density, and weekday trips are among the variables found to significantly increase travel time. On the contrary, number of through travel lanes was found to reduce travel time significantly. Interestingly, there was no significant difference in the extent to which intersection control types individually affect travel time on the segments, although their effect directions were logical. However, having intersection control at the segment downstream was found to be significantly associated with lower travel times relative to having lane drops at the segment downstream. Moreover, the sensitivity analysis threw more light on the impact and the interdependencies of the contributing factors of congestion. The extent of the impact of factors such as presence of on-street parking and access point density and number of lanes on travel time were seen to be dependent on the level of traffic density. Specifically, their impacts on travel time increase with increasing traffic density.

Contributions

This study contributes to the body of literature by bridging the gap of lack of travel time models developed from heterogenous traffic and roadway conditions of sub-Saharan Africa to quantify congestion at the arterial level. Factors which otherwise have not been incorporated in previous travel time models developed from homogenous traffic and roadway conditions have been considered in the travel time models developed in this study. The model is therefore not applicable on only signalized arterials compared to other previously developed models. As such the travel time models developed are applicable for planning purposes in a typical urban arterial with prevalent indiscriminate on-street parking activities, lane drops/additions, aggressive usage of unsignalized intersections, and fixed time signalized intersections. This study provides valuable information to transport planners and engineers concerning factors that significantly affect urban arterial travel time. This informs their decisions when prioritizing measures to improve travel time and mitigate congestion at large. The empirical relation between the segment travel time and the contributing factors (e.g., on-street parking, access density and downstream bounding conditions) can be possibly applied to evaluate traffic control and management strategies for mobility improvement purposes.

Recommendations

Although this model shows a good fit for prediction purposes, it incorporates just the impact of the presence of on-street parking activities which does not explicitly capture the characteristics of the parking activities. Future work can therefore look further into the joint impact of parking maneuvers, parking duration, and the proportion of paratransits within the traffic stream on travel time. It will also be interesting to investigate how the intensity of on-street parking (e.g., length of on-street parking area relative to segment length) impact on segment travel time.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] S.T.G. Fleuren, Optimizing Pre-Timed Control at Isolated Intersections, TechnischeUniversiteit Eindhoven, Eindhoven, 2017.
[2] INRIX, I. & CEBRThe Future Economic and Environmental Costs of Gridlock in 2030, Centre for Economics and Business Research, 2014.
[3] H. Takyi, K. Poku, E.K. Anin, An assessment of traffic congestion and its effect on productivity in urban Ghana, Int. J. Bus. Soc. Sci. 4 (3) (2013) 225–234.
[4] J.T. Mordy, 2017. Bad roads, traffic congestion cost Ghana billions yearly. Myjoyonline.com. [Online] Available at: https://www.myjoyonline.com/news/2017/may-19th/bad-roads-traffic-congestion-cost-ghana-billions-yearly.php [Accessed Saturday February 2020].
[5] S.V. Kumar, R. Sivanandan, Traffic congestion quantification for urban heterogeneous traffic using public transit buses as probes, Period. Polytech. Transp. Eng. 47 (4) (2019) 257–267.
[6] W.L. Eisele, et al., Developing and applying models for estimating arterial corridor travel time index for transportation planning in small to medium–sized communities, Transp. Res. Rec. J. Transp. Res. Board 2244 (2011) 81–90 No.p. pp..
[7] Highway Capacity Manual, Fith Edition. National Research Council (U.S.). Transportation Research Board. Washington, D.C.: Transportation Research Board, ©2010.
[8] B.I. Musah, L. Peng, Y. Xu, Urban congestion and pollution: a quest for cogent solutions for Accra city, in: Proceedings of the IOP Conference Series: Earth and Environmental Science, 435, IOP Publishing, 2020, pp. 1–8. 012026.
[9] F. Agyapong, T.K. Ojo, Managing traffic congestion in the Accra central market, Ghana, J. Urban Manag. 7 (2) (2018) 85–96.
[10] K. Olagunju, Evaluating traffic congestion in developing countries. A case study of Nigeria, J. Chart. Inst. Logist. Transp. Niger. 2 (3) (2015) 23–26.
[11] O.K. Kwakwa, C.A. Adams, Assessment of turbo and multilane roundabout alternatives to improve capacity and delay at a single lane roundabout using microsimulation model vissim: a case study in Ghana, Am. J. Civ. Eng. Archit. 4 (4) (2016) 106–116.
[12] Y.A. Tuffour, A.A. Obiri-Yeboah, M. Salifu, P.C. Acquah, Adjusting for the effect of bus blockage on saturation flows, Int. J. Sci. Res. 3 (6) (2014) 749–753.
[13] S. Dumba, L.D. Vassileva, T. Gumbo, Methodological issues in modelling signalized intersection capacity under informal public transport operations: case study, in: Proceedings of the Transportation Research Procedia, World Conference on Transport Research- WCTR 2016 Shanghai, Harare, Zimbabwe, 2017, pp. 4891–4915. 10-15 July 2016, Volume 25.
[14] R. Neuhold, M. Fellendorf, Volume delay functions based on stochastic capacity, Transp. Res. Rec. J. Transp. Res. Board 2421 (2014) 93–102.
[15] P. Tisato, Suggestions for an improved davidson travel time function, Aust. Road Res. 21 (2) (1991) 85–100.
[16] H. Spiess, Conical volume–delay functions, Transp. Sci. 24 (2) (1990) 153–158.
[17] R. Akçelik, Travel time functions for transport planning purposes: Davidson's function, its time-dependent form and an alternative travel time function, Aust. Road Res. 21 (3) (1991) 49–59 (Minor revisions: December 2000).
[18] K.B. Davidson, A flow travel time relationship for use in transportation planning. In Australian Road Research Board (ARRB) Conference, 3rd, 1966, Sydney, 1966, Vol. 3, No. 1. pp. 183-194.
[19] A. Skabardonis, R. Dowling, Improved speed-flow relationship for planning applications, Transp. Res. Rec. J. Transp. Res. Board 1 (1572) (1997) 18–23.
[20] Z. Bugg, et al., A methodology to compute roundabout corridor travel time, Transp. Res. Rec. J. Transp. Res. Board 2483 (1) (2015) 20–29 No. 2483.
[21] K. Fitzpatrick, S. Das, Vehicle operating speed on urban arterial roadways.Safe-D National UTC, Texas A&M Transportation Institute Report No. TTI-01-04. (January, 2019)
[22] V.M. Kukkapalli, S.S. Pulugurtha, Modeling the effect of a freeway road construction project on link-level travel times, J. Traffic Transp. Eng. 8 (2) (2021) 267–281 (English Edition).
[23] Z.K. Gurmu, W.D. Fan, Artificial neural network travel time prediction model for buses using only GPS data, J. Public Transp. 17 (No. 2) (2014) 45–65.
[24] R.E. Kalman, a new approach to linear filtering and prediction problems, Trans. ASME J. Basic Eng. 82 (1960) 35–45.
[25] Y. Hou, P. Edara, Network scale travel time prediction using deep learning, Transp. Res. Rec. 2672 (45) (2018) 115–123.
[26] A. Csikós, Z.J. Viharos, K.B. Kis, T. Tettamanti, I. Varga, Traffic speed prediction method for urban networks—an ANN approach, in: *2015 International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS)*, IEEE, 2015, June, pp. 102–108.
[27] J. Jabamony, G.R. Shanmugavel, IoT based bus arrival time prediction using Artificial Neural Network (ANN) for Smart Public Transport System (SPTS), Int. J. Intell. Eng. Syst. 13 (1) (2020) 312–323.
[28] A.S. Mane, S.S. Pulugurtha, Link-level travel time prediction using artificial neural network models, in: Proceedings of the 21st International Conference on Intelligent Transportation Systems (ITSC), Maui, Hawaii, USA, 2018.
[29] W. Qiao, A. Haghani, M. Hamedi, A nonparametric model for short-term travel time prediction using bluetooth data, J. Intell. Transp. Syst. 17 (2) (2013) 165–175.
[30] A.K. Bharti, C.R. Sekhar, S. Chandra, Modeling travel time of car with varying demand on an urban midblock section, J. Intell. Transp. Syst. 22 (2) (2018) 99–105, doi:10.1080/15472450.2017.1408013.
[31] H. Chen, H.A. Rakha, C.C. McGhee, Dynamic travel time prediction using pattern recognition, in: Proceedings of the 20th World Congress on Intelligent Transportation Systems, TU Delft, 2013.
[32] S. Barua, J. Hossain, Estimation of traffic density to compare speed-density models with moving observer data, Int. J. Res. Eng. Technol. 4 (8) (2015) 471–474.
[33] M. Guerrieri, G. Parla, R. Mauro, Traffic flow variables estimation: an automated procedure based on moving observer method. potential application for autonomous vehicles, Transp. Telecommun. 20 (3) (2019) 205–214.
[34] P. Kachroo, S. Sastry, Travel time dynamics for transportation systems: theory and applications, IEEE Trans. Intell. Transp. Syst. 17 (2) (2016) 385–394.
[35] Cambridge Systematics, Traffic congestion and reliability: trends and advanced strategies for congestion mitigation, Federal Highway Administration, Washington, DC, 2005. https://rosap.ntl.bts.gov/view/dot/20656 (Accessed 18 April 2021) Report Number: FHWA-HOP-05-064.
[36] D. Pevec, I. Kononenko, Prediction intervals in supervised learning for model evaluation and discrimination, Appl. Intell. 42 (4) (2015) 790–804.
[37] L. Zhang, L. Xie, Q. Han, Z. Wang, C. Huang, Probability density forecasting of wind speed based on quantile regression and kernel density estimation, Energies 13 (22) (2020) 6125.
[38] E. Häglund, Estimating prediction intervals with machine learning and Monte Carlo methods in online advertising, Master Thesis, KTH Royal Institute of Technology, Stockholm, Sweden, 2020. https://www.diva-portal.org/smash/get/diva2:1472065/FULLTEXT01.pdf (Accesseddate:18April2021).
[39] P. Kachroo, S. Sastry, Traffic assignment using a density-based travel-time function for intelligent transportation systems, IEEE Trans. Intell. Transp. Syst. 17 (5) (2016) 1438–1447.
[40] A. Kondyli, et al., Evaluating the operational effect of narrow lanes and shoulders for the highway capacity manual, Transp. Res. Rec. Transp. Res. Board 2673 (10) (2019) 558–570.

[41] H.G. Demir, Y.K. Demir, A comparison of traffic flow performance of roundabouts and signalized intersections: a case study in Nigde, Open Transp. J. 14 (2020) 120–132.
[42] M.A. LeBas, 2015. A comparative analysis of roundabouts and traffic signals through a corridor. Louisiana State University and Agricultural and Mechanical College. Master's Theses 3753. https://digitalcommons.lsu.edu/gradschool_theses/3753 (Accessed 16 August 2020).
[43] R.T. Luttinen, 2004. Capacity and Level of Service at Finnish Unsignalized Intersections. (pp. 210). Finnish Road Administration, Traffic Engineering, S 12 Solutions to improve main roads. (Finnra Reports; No. 1/2004).
[44] I.I. Otkoviæ, I. Dadic, Comparison of delays at signal-controlled intersection and roundabout, Sci. Traffic Transp. 21 (3) (2009) 157–165.
[45] W. Brilon, N. Wu, Capacity at unsignalized intersections derived by conflict technique, Transp. Res. Rec. 1776 (1) (2001) 82–90.
[46] Z. Qu, Y. Duan, H. Hu, X. Song, Capacity and delay estimation for roundabouts using conflict theory, The Scientific World Journal 2014 (2014) (Article ID 710938), p. 12 pages, doi:10.1155/2014/710938.
[47] Hallmark, et al., Behavior Study of Merge Practices for Drivers at Work Zone Closures, Institute for Transportation, Iowa State University, Ames, IA, 2011.
[48] S. Hallmark, N. Oneyear, 2015. Modeling merging behavior at lane drops: smart work zone deployment initiative. Center for Transportation Research and Education. Institute for Transportation at Iowa State University. Report No: InTrans Projects 13-466 and 06-277.
[49] H. Naghawi, M.A. Shattal, W. Idewu, Application of AIMSUN microscopic simulation model in evaluating side friction impacts on traffic stream performance, Int. J. Transp. Veh. Eng. 13 (1) (2019) 10–15.
[50] K. Saw, A.K. Das, B.K. Katti, G.J. Joshi, Travel time estimation modelling unde heterogeneous traffic: a case study of urban traffic corridor in Surat, India, Period. Polytech. Transp. Eng. 47 (4) (2019) 302–308.
[51] W. Koehrsen, 2018. Overfitting vs. underfitting: a complete example. Towards Data Science [Online] Available at: https://towardsdatascience.com/overfitting-vs-underfitting-a-complete-example-d05dd7e19765 (Accessed 24 March 2020).
[52] T.D. Wemegah, S. Zhu, C. Atombo, Modeling the effect of days and road type on peak period travels using structural equation modeling and big data from radio frequency identification for private cars and taxis, Eur. Transp. Res. Rev. 10 (39) (2018) 1–14.
[53] B.A. Kumar, L. Vanajakshi, S.C. Subramanian, Bus travel time prediction using a time-space discretization approach, Transp. Res. Part C Emerg. Technol. 79 (2017) 308–332.
[54] M. Rahmani, H.N. Koutsopoulos, E. Jenelius, Travel time estimation from sparse floating car data with consistent path inference: a fixed point approach, Transp. Res. Part C Emerg. Technol. 85 (2017) 628–643.
[55] K. Tang, S. Chen, A.J. Khattak, Personalized travel time estimation for urban road networks: a tensor-based context-aware approach, Expert Syst. Appl. 103 (2018) 118–132.
[56] G. Qi, A.A. Ceder, Z. Zhang, W. Guan, D. Liu, New method for predicting long-term travel time of commercial vehicles to improve policy-making processes, Transp. Res. Part A Policy Pract. 145 (2021) 132–152.
[57] J. Strauss, L.F. Miranda-Moreno, Speed, travel time and delay for intersections and road segments in the montreal network using cyclist smartphone GPS data, Transp. Res. Part D Transp. Environ. 57 (2017) 155–171.
[58] S. Mil, M. Piantanakulchai, Modified Bayesian data fusion model for travel time estimation considering spurious data and traffic conditions, Appl. Soft Comput. 72 (2018) 65–78.
[59] L. Li, B. Ran, J. Zhu, B. Du, Coupled application of deep learning model and quantile regression for travel time and its interval estimation using data in different dimensions, Appl. Soft Comput. 93 (2020) 106387.
[60] K. Liu, M.Y. Cui, P. Cao, J.B. Wang, Iterative Bayesian estimation of travel times on urban arterials: fusing loop detector and probe vehicle data, PLoS One 11 (6) (2016) e0158123.
[61] S. Jain, S.S. Jain, 2017. A methodology for modelling urban traffic congestion based on ITS.
[62] D. Wang, F. Fu, X. Luo, S. Jin, D. Ma, Travel time estimation method for urban road based on traffic stream directions, Transportmetrica A: Transport Science 12 (6) (2016) 479–503.
[63] W.H. Lee, K.P. Hwang, W.B. Wu, An intersection-to-intersection travel time estimation and route suggestion approach using vehicular ad-hoc network, Ad Hoc Netw. 43 (2016) 71–81.
[64] E. Akandwanaho, H. Nakamura, Analysis of travel speed on 4-lane signalized arterials, Transp. Res. Procedia 48 (2020) 890–903.
[65] X. Zhu, Y. Fan, F. Zhang, X. Ye, C. Chen, H. Yue, Multiple-factor based sparse urban travel time prediction, Appl. Sci. 8 (2) (2018) 279.