

Article

Estimating Freeway Level-of-Service Using Crowdsourced Data

Nima Hoseinzadeh ¹, Yangsong Gu ¹, Lee D. Han ^{1,*}, Candace Brakewood ¹ and Phillip B. Freeze ²

¹ Department of Civil and Environmental Engineering, University of Tennessee, Knoxville, TN 37996, USA; nhoseinz@vols.utk.edu (N.H.); ygu17@vols.utk.edu (Y.G.); cbrakewo@utk.edu (C.B.)

² Tennessee Department of Transportation, Nashville, TN 37243, USA; Phillip.b.freeze@tn.gov

* Correspondence: lhan@utk.edu

Abstract: In traffic operations, the aim of transportation agencies and researchers is typically to reduce congestion and improve safety. To attain these goals, agencies need continuous and accurate information about the traffic situation. Level-of-Service (LOS) is a beneficial index of traffic operations used to monitor freeways. The Highway Capacity Manual (HCM) provides analytical methods to assess LOS based on traffic density and highway characteristics. Generally, obtaining reliable density data on every road in large networks using traditional fixed location sensors and cameras is expensive and otherwise unrealistic. Traditional intelligent transportation system facilities are typically limited to major urban areas in different states. Crowdsourced data are an emerging, low-cost solution that can potentially improve safety and operations. This study incorporates crowdsourced data provided by Waze to propose an algorithm for LOS assessment on an hourly basis. The proposed algorithm exploits various features from big data (crowdsourced Waze user alerts and speed/travel time variation) to perform LOS classification using machine learning models. Three categories of model inputs are introduced: Basic statistical measures of speed; travel time reliability measures; and the number of hourly Waze alerts. Data collected from fixed location sensors were used to calculate ground truth LOS. The results reveal that using Waze crowdsourced alerts can improve the LOS estimation accuracy by about 10% (accuracy = 0.93, Kappa = 0.83). The proposed method was also tested and confirmed by using data from after coronavirus disease 2019 (COVID-19) with severe traffic breakdown due to a stay-at-home policy. The proposed method is extendible for freeways in other locations. The results of this research provide transportation agencies with a LOS method based on crowdsourced data on different freeway segments, regardless of the availability of traditional fixed location sensors.

Keywords: crowdsourced data; big data; Level-of-Service; traffic; machine learning



Citation: Hoseinzadeh, N.; Gu, Y.; Han, L.D.; Brakewood, C.; Freeze, P.B. Estimating Freeway Level-of-Service Using Crowdsourced Data. *Informatics* **2021**, *8*, 17. <https://doi.org/10.3390/informatics8010017>

Received: 8 February 2021

Accepted: 1 March 2021

Published: 5 March 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Intelligent transportation systems (ITS) are essential for assessing the state of traffic. ITS traffic measurements can be used in different applications, such as traffic operations, road work planning, assessing traffic queues, and congestion management. The United States Highway Capacity Manual (HCM) defines six Levels-of-Service (LOS) for estimating the traffic performance and state. The HCM provides analytical methods for assessing LOS from traffic density and highway characteristics [1]. The traffic density, speed, and flow are key components of LOS assessment [1–3]. The Department of Transportations (DOTs) and transportation agencies usually want real-time or historical hourly traffic status and LOS data for different freeway segments.

Traditionally, traffic data (speed, travel time, flow, and density) are collected by a variety of fixed location sensors, such as loop detectors, remote traffic microwave sensors (RTMS), magnetic sensors, laser sensors, video images, and License Plate Recognition (LPR) systems [4,5]. Fixed location data collection methods are typically expensive and have a limited network coverage. In recent years, data-driven ITS has led to multisource, high-performance, and powerful solutions in transportation systems [6]. Data collection

based on probe vehicles and floating cars has gained more attention. These approaches use new technologies such as smartphones, cellular networks, Bluetooth sensors, Wi-Fi, and connected vehicles (CVs) to provide traffic data [7,8]. These technologies not only generate useful data to be employed in different transportation analyses, such as traffic safety [9–12], public transit [13,14] and energy consumption and emissions analyses [15,16], but also provide new opportunities to collect crowdsourced data [15,16].

Crowdsourcing refers to obtaining data from a group of users who contribute their information via smartphones, social media, or the internet. The use of big data and crowdsourced data in transportation enables researchers to propose innovative ideas and solutions that were not studied in the past [17]. With an increased use of smartphone applications, road users can share information (e.g., speed, travel time, delay, incident, hazards, severe weather, and congestion) using navigation applications (e.g., Waze and Google Maps). Crowdsourced data are considered a promising alternative to traditional data collection methods [18–21]. Meanwhile, the advantages of crowdsourced data and probe vehicles over traditional fixed location data collection methods are the (i) expanded network coverage and resolution, (ii) low/no implementation and maintenance costs, (iii) improved real-time application, and (iv) ability to implement proactive applications [20,21]. Crowdsourcing enables transportation researchers to propose new methods and platforms that have various applications, such as incident detection [22,23], traffic condition analysis [23], traffic speed prediction [24], hotspot analysis [25], road anomaly detection [26], and road surface evaluation and indexing [27]. Crowdsourcing can also be helpful in the development of new applications for visually or mobility impaired road users in different modes of transportation [28].

Since leveraging crowdsourced data, many cities governments and DOTs have established partnerships with data providers companies such as Waze and INRIX [20,29]. They have utilized crowdsourced data in a variety of applications, such as performance measurement and incident detection. This paper focuses on Waze. Waze is a navigation app that provides crowdsourced data, such as the speed, travel time, and road user reports (incidents, traffic jams, and hazards), through the Waze for Cities (WFC) program. The Tennessee Department of Transportation (TDOT) is a Waze partner that uses crowdsourced data.

The acquisition of crowdsourced data generates an opportunity to propose a new methodology for assessing LOS based on the data's features and characteristics. This study proposes a new methodology that exploits features from crowdsourced data and speed/travel time deviation to assess LOS on freeways. This methodology can be used in developing new tools for LOS assessment and hourly traffic status data on freeways with no need for fixed location traffic volume sensors. The proposed approach can be considered a supplemental methodology for traditional HCM LOS calculation that relies on the traffic density and flow.

The remainder of this paper is organized as follows. The next section summarizes related literature about LOS assessment, traffic situation prediction, and crowdsourced data in transportation. The methodology section presents the traditional LOS calculation and the proposed method. Additionally, some data mining approaches are discussed in this section. Then, the data used in this study are discussed, followed by the results obtained by implementing the methodology. Finally, the paper is concluded, and areas for future work are provided.

2. Literature Review

This part reviews the most relevant literature pertaining to this study in the following order. First, traffic status and LOS assessment methods are summarized. Then, travel time reliability and alternative LOS methods are discussed. The following parts will review Waze data used in previous studies and discuss the research gaps.

2.1. Traffic Status and LOS Assessment Methods

Most transportation agencies and DOTs focus on the density or volume to capacity (V/C) ratio to assess LOS. HCM uses the density to define LOS for freeways and multilane highways. It also defines LOS for intersections using a metric called control delays [1,30]. Traditionally, studies use one or a combination [31–33] of parameters, such as the speed [34], flow [32,35], and density [36,37], to explain the traffic status and LOS. Previous studies have used different data sources, such as sensor data [38], probe vehicles [32,39], camera images and videos [40], CVs [2,41], and simulation [2,35,37,41]. In terms of the methodology, statistical modeling [37], Neural Networks [38,39], Kalman Filters [39], Image Processing [42], and Machine Learning [32,40] have been widely used.

2.2. Travel Time Reliability

Previous studies have captured useful information from speed and travel time variability and reliability to determine the traffic status and performance. These studies used statistical measures, such as the average, standard deviation, percentiles, and range [43–45]. Additionally, the relation between the speed deviation, travel time variability, planning time index (PTI), and buffer time index (BTI) with the V/C ratio has been explored in prior literature [43,46].

2.3. Alternative LOS Methods

The Strategic Highway Research Program 2 (SHRP2) Reliability Project L08 discussed supplemental methods for LOS measurement. This project used a density-based definition of LOS to form the distribution of LOS and presented a distribution instead of a single value for LOS [47]. This study proposed an innovative approach for LOS based on travel time reliability perspectives. The travel speed range, the most restrictive condition, and the travel time value were introduced in this project [47]. Travel time reliability and variability are measures of service quality [48].

Pulugurtha and Imran (2020) and Kodupuganti and Pulugurtha (2019) modeled the LOS of freeways and urban links using travel time variability indicators such as the planning time index (PTI), buffer time index (BTI), average travel time, and 95th percentile. They suggested a threshold travel time reliability to assess LOS [49,50]. Singh et al. (2019) used Wi-Fi probe data to develop LOS thresholds based on travel time reliability and variability indices [30]. In a different approach, Altinatasi et al. (2016) used the average speed of Floating Car Data (FCD) to quantify LOS [34]. Moreover, Khan, Dey, and Chowdhury (2017) used simulation and artificial intelligence to assess LOS based on different CV penetration rates [2]. Table 1 presents the most relevant studies that have proposed alternative methods for LOS assessment.

Table 1. Summary of alternative Level-of-Service (LOS) assessment methods in the literature.

No.	Reference	Year	Data	Index Used	Method
1	Kittelson et al. [47]	2008	Sensor data (speed, density, travel time)	- Travel speed range - Most restrictive condition - Value of travel time	- Travel time reliability threshold
2	Altinatasi et al. [34].	2016	Floating Car Data (speed)	- Average speed	- Speed threshold
3	Khan, Dey, and Chowdhury [2]	2017	Simulation (speed, density)	- Average speed - CV penetration rate	- Artificial intelligence
4	Singh et al. [30]	2019	Wi-Fi probe vehicle (speed, travel time)	- Planning Time Index (PTI) - Buffer Time Index (BTI) - Travel Time Index (TTI)	- Travel time reliability threshold - Statistical regression

Table 1. Cont.

No.	Reference	Year	Data	Index Used	Method
5	Kodupuganti, and Pulugurtha [50]	2019	Travel time data provided by North Carolina DOT	- Planning Time Index - Buffer Time Index - Average travel time	- Travel time reliability threshold - Regression model
6	Pulugurtha and Imran [49]	2020	Simulation (travel time)	- Planning Time Index - Buffer Time Index	- Travel time reliability threshold - Statistical regression

2.4. Waze Data

Crowdsourcing enables researchers to use big data collected from road users, probe vehicles, bicycles, and pedestrians [21,51]. This research uses Waze crowdsourced alerts and speed/travel time data as the primary crowdsourced data source. The company Waze analyzes the app users' location to provide speed and travel time data. Waze also provides different event reports (congestion, incidents, severe weather, and road construction). Prior studies have explored and verified Waze incident report and travel time data to assess the reliability and coverage [20,29,52–54]. A recent study by Li et al. showed that Waze incident alerts are spatially correlated with police crash reports (PCR) and that Waze provides a broader coverage than PCR [25]. Several previous studies have also used Waze alerts data for applications such as accident clustering [55], safety hotspot detection [25], incident detection [22], and improving dynamic traffic lights [56]. A more recent study also verified the quality of Waze speed data on surface streets [21]. Table 2 summarizes prior studies using Waze data.

Table 2. Summary of studies using Waze data in the literature.

No.	Reference	Year	Country	Waze Data Used	Findings and Application
1	Santos et al. [52]	2016	Brazil	Waze accident alerts	- Showed acceptable reliability of Waze accident reports
2	Bahaweres et al. [53]	2017	Indonesia	Waze travel time	- Performed t-test to show the Waze travel time and ground truth are almost equal
3	Pack and Ivanov [20]	2017	United States	Waze alerts	- Explored the properties and possible benefits of Waze data
4	Amin-Naseri et al. [29]	2018	United States	Waze congestion and accident alerts	- Showed reasonable spatial and temporal accuracy of Waze data
5	Perez et al. [55]	2018	Mexico	Waze alerts	- Used Waze data for road accident clustering
6	Raul Sanchez et al. [56]	2019	Colombia	Waze jam and accident alerts	- Used Waze data to improve dynamic traffic lights and urban mobility
7	Turner et al. [54]	2020	United States	Waze alerts	- Found Waze data to be a valuable safety data source, especially for capturing unreported traffic incidents
8	Hoseinzadeh et al. [21]	2020	United States	Waze speed	- Assessed the quality of Waze speed on surface streets
9	Li et al. [25]	2020	United States	Waze alerts	- Showed Waze incident alerts are spatially correlated with PCR - They showed that Waze provides a broader coverage than PCR
10	Senarath et al. [22]	2020	United States	Waze alerts	- Proposed an incident detection platform

2.5. Gap in the Literature

As previously discussed, HCM density-based LOS were given the most attention in prior literature. Moreover, some studies used travel time and speed variability for determining LOS. Crowdsourced data have not been used to determine LOS. This study addresses a gap in integrating crowdsourced data (Waze incident report and speed data) for LOS assessment. This study's results can help agencies quantify LOS for different segments, without installing new fixed location equipment.

3. Data

This part describes the primary datasets used in this study. This study used Waze speed/travel time and Waze alert data for the LOS analysis. In the following, the Waze speed/travel time, Waze crowdsourced alerts, and fixed location data will be presented. Then, the study area will be introduced.

3.1. Waze Speed and Travel Time Data

The travel time and traffic speed for specific roadway segments are data sources that Waze shares with partners through the Waze for Cities (WFC) program. Waze obtains app users' kinematic information in specified segments to calculate and report the speed and travel time. If no user is passing in that time interval or segment, it reports historical speed and travel time data. Waze implements a tool called "traffic view" that allows transportation agencies and DOTs to specify a list of road segments. Authorized users can add links based on their priority or needs to the watch-list. Subsequently, real-time travel time/speed data for the predefined road segments are available at a one-minute level. Authorized users can use these data in real-time or archive them in a JavaScript Object Notation (JSON) format for further analysis. The archived JSON file for each time interval includes travel time, segment length, and geospatial information for all predefined segments in that time interval.

3.2. Waze Crowdsourced Alert Data

User report data are other valuable crowdsourced data that Waze provides to partners, which are referred to as alerts. Waze alerts can be used in different analyses, such as incident detection, hot spot clustering, and end of queue prediction. Waze users can report predefined incident types in the Waze App while traveling. These alerts include accidents (major or minor), traffic jams (heavy, moderate, standstill, or light), hazards (severe weather, stopped cars, or road potholes), construction, and closed roads. Users can also verify existing reports on the road. Waze shares all users' reports through the WFC program. Waze alerts data include the incident unique ID, time, spatial coordinates, direction, reliability, and confidence level of the reported alert. Waze partners can use real-time alerts or archive them in an Extensible Markup Language (XML) format for their analysis. This study had access to Waze alerts for Tennessee State.

3.3. Fixed Location Data

The Tennessee Department of Transportation (TDOT) uses Radio Data System (RDS) sensor data, which provide traffic information such as the traffic count, speed, and occupancy in 30-s time intervals. RDS stations are located on freeways close to four major cities, including Nashville, Memphis, Knoxville, and Chattanooga in Tennessee. This study uses the traffic volume (flow) from RDS data and traffic speed from Waze to calculate the density and LOS. The estimated LOS can be used as ground truth data. This will be elaborated on in the methodology section.

3.4. Study Time and Area

To quantify hourly data and assess the LOS for freeways, a study area was designated in Knoxville, Tennessee. A segment on the Interstate 40 (I-40) highway at westbound mile marker 385 was selected (Figure 1). The study segment length is about 1.5 miles (~2.4 km).

The speed limit for this segment is 65 mph (~105 km/h). This location was selected for two reasons: (1) The variability of traffic and LOS during the hours of the day, and (2) the availability of roadway sensor data (flow) to calculate different ground truth LOS.

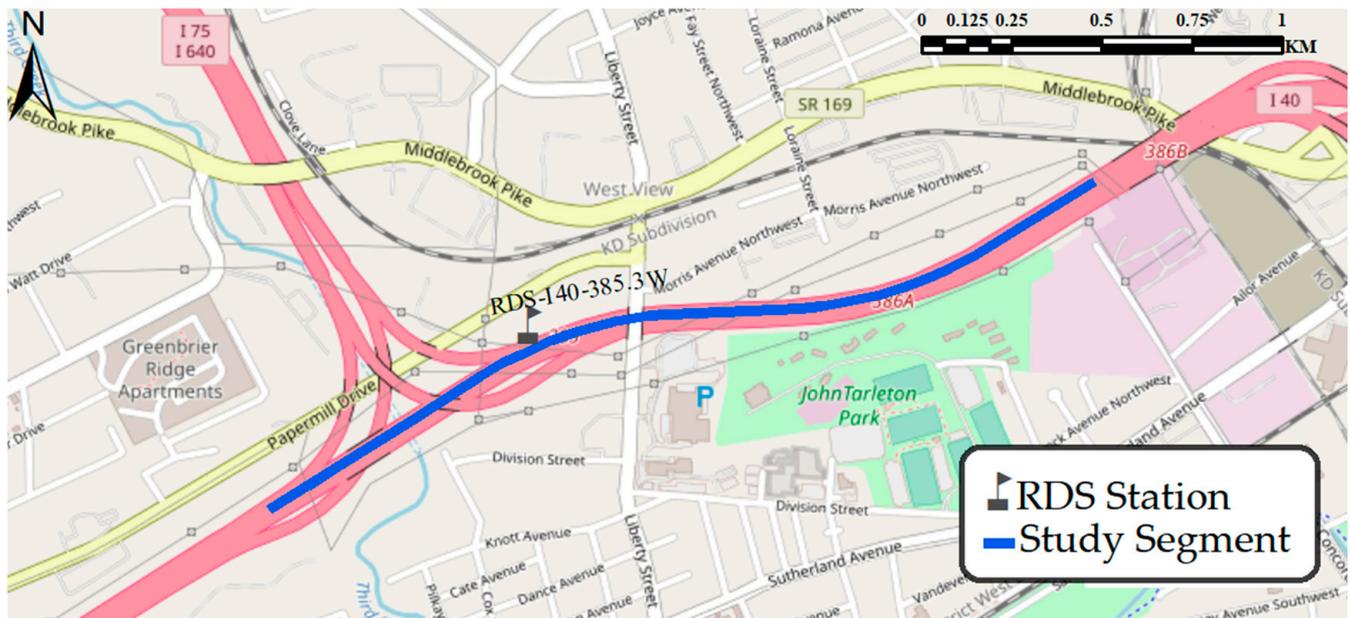


Figure 1. Framework of the proposed methodology.

One month of data, representing 1 October 2019 to 31 October 2019 (744 h), were selected to train the methodology. The world faced a significant challenge in 2020 from the coronavirus disease 2019 (COVID-19) pandemic [57]. Stay-at-home is known to be an effective policy [28] for preventing the spread of COVID-19 in the US and led to a major breakdown in mobility in March and April 2020. The traffic also recovered by about 90% in August 2020. Therefore, two months of data, consisting of 16 March 2020 to 15 April 2020 and 1 August 2020 to 3 August 2020 (overall 1488 h), were collected to test the final method. These two months were selected to test the method in both a normal and abnormal situation.

4. Methodology

The methodology used in this paper combines raw crowdsourced speed and user report data to obtain the hourly LOS-based traffic status. This methodology uses the speed variation, travel time reliability, and user alerts in the selected segment to define measures of traffic conditions. Here, the Waze speed/travel time and crowdsourced alerts are used as the primary data source in the study, which will be elaborated on in the following sections. Unlike some previous studies, this method does not solely depend on the average speed [34] or density.

This section provides more details about the proposed algorithm of this study. As shown in the framework of the study (Figure 2), the different steps of the proposed method are as follows:

- Step 1: Data collection, which includes archiving Waze data and traditional fixed location sensor data, as well as preprocessing and normalization;
- Step 2: Extract model inputs, which includes statistical measures, travel time performance measures, and crowdsourced Waze alerts;
- Step 3: Calculating ground truth LOS, using fixed location sensors, and labeling observations of Waze input data with the corresponding ground truth data;
- Step 4: LOS assessment, by performing different machine learning methods. This part includes feature selection, cross validation, and selecting the preferred method.

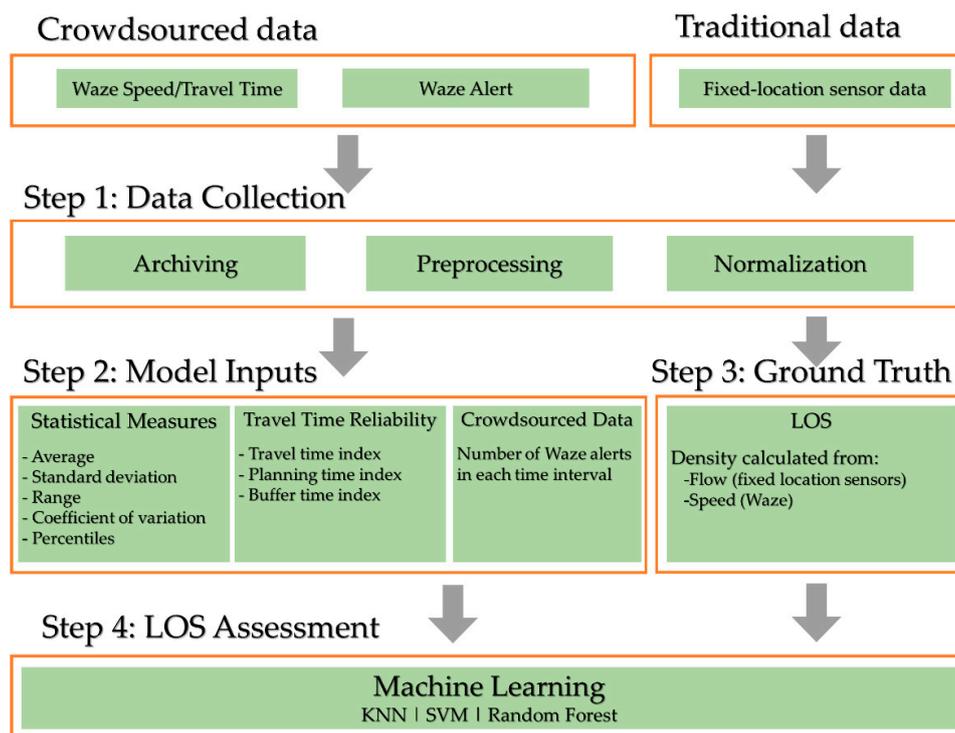


Figure 2. Framework of the proposed methodology.

4.1. Step 1: Data Collection

Waze continuously generates a massive amount of data. The first step in such a study is to archive Waze speed/travel time and Waze alert data. A Python code was implemented to capture crowdsourced alert, speed, and travel time data for 1-min time intervals. Employing real-world raw data can always present challenges, such as missing values or noise. In the next step, data were preprocessed by cleaning and removing possible errors. Possible missing values and outliers were removed/imputed. Next, RDS traffic volume data were collected to calculate the hourly traffic flow and LOS ground truth, which will be elaborated on in Step 3.

4.2. Step 2: Model Inputs

As explained, previous studies have explored the variation of speed/travel time to capture LOS. This study combined different speed and travel time variation indexes with crowdsourced data to assess LOS. Multiple indices were calculated as the inputs of the classification model. This paper divides these indicators into three categories, as follows. Each index will be elaborated on in the following paragraphs.

- Basic statistical measures, including the average speed, standard deviation, range, coefficient of variation, standard error, percentiles (25th, 50th, and 90th), and interquartile range.
- Travel time performance measures, including the Travel Time Index, Planning Time Index, and Buffer Time Index.
- Crowdsourced data, including the number of users' accident, jam, and hazard reports in the Waze alerts data.

4.2.1. Basic Statistical Measures

Pertaining to speed variation, different statistical measures were considered. As discussed, speed variation has been considered in prior studies [2,34,47,48]. All these measures were captured and measured during each period (in this study, hourly). Table 3 provides the different statistical measures used in this study.

Table 3. Model input equations.

Model Input	Measure	Equation	Eq. No.
Basic statistical measures of speed	Average Speed	$\bar{v} = \frac{\sum_{i=1}^n v_i}{n}$ where v_i is the speed and n is the number of observations in each time interval	(1)
	Standard Deviation (SD)	$\sigma = \sqrt{\frac{\sum_{i=1}^n (v_i - \bar{v})^2}{n}}$	(2)
	Range	$Range(v) = \max_{i=1,2,\dots,n} v_i - \min_{i=1,2,\dots,n} v_i$	(3)
	Coefficient of Variation (CoV)	$CoV = \frac{\sigma}{\bar{v}}$	(4)
	Standard Error (SE)	$SE = \frac{\sigma}{\sqrt{n}}$	(5)
	Percentiles (25th, 50th, 75th, 90th)	$kth\ percentile = rank\left(\frac{k}{100}(n+1)\right)$ where $k = 25, 50, 75, 90$ Here, rank is ordering the dataset from smallest to largest and finds the value with the $\frac{k}{100}(n+1)$ index	(6)
	Interquartile Range (IQR)	$IQR = Q_3 - Q_1$ where Q_3 is the 75th percentile and Q_1 is the 25th percentile of v_i	(7)
	Travel Time Index (TTI)	$TTI = \frac{TT_{Avg}}{TT_{free-flow}}$ where TT_{Avg} is the average travel time and $TT_{free-flow}$ is the free flow travel time	(8)
	Buffer Time Index (BTI)	$BTI = \frac{TT_{95th} - TT_{Avg}}{TT_{Avg}}$ where TT_{95th} is the 95th percentile of the travel time	(9)
	Planning Time Index (PTI)	$PTI = \frac{TT_{95th}}{TT_{free-flow}}$	(10)
Crowdsourced data	Hourly Number of Alerts	$Count(Waze\ Alert_s^t)$ where s is the study segment and t is the time interval (hour of day)	(11)

4.2.2. Travel Time Performance Measures

To analyze the travel time variability for each time period, the following well-known travel time performance measures were also calculated. It should be noted that all the travel time reliability indexes were derived based on a one-hour aggregation level. The travel time performance measures are as follows. Table 3 also presents the different travel time performance equations (Equations (8) to (10)).

- The Travel Time Index (TTI) captures the travel time variation by calculating the average travel time ratio to the free flow travel time in the segment. This index explores how the travel time deviates from the free flow travel time during the intended period, which is typically LOS A [30].
- The Buffer Time Index (BTI) represents the amount of extra time that the traveler needs to be on time [30,49,50].
- The Planning Time Index (PTI) calculates the ratio of the 95th percentile of travel time to the free flow travel time. A higher PTI value indicates a lower reliability and theoretically lower LOS [30,49,50].

4.2.3. Crowdsourced Data

This study incorporated crowdsourced data along with the speed and travel time variability. Here, the number of Waze user reports (alerts) in each period (one hour) for the study area was calculated. This number was then used as an input for the final model of LOS assessment.

4.3. Step 3: Ground Truth LOS

LOS is a widely used performance measure of the quality of service for a road segment. The HCM identified six LOS categories for freeways and highways based on density and road characteristics. HCM employs the traffic density as the primary measure of LOS for freeway segments [2,34,41,58,59]. Table 4 presents the density pertaining to each LOS [1]. In this study, traffic flow (from RDS sensors) and speed (from Waze) were used to calculate the hourly traffic density. The calculated density was used to obtain hourly LOS based on Table 4. The calculated LOS was used as the ground truth. The hourly input data were also labeled with ground truth values, which were used in the LOS model presented in the next section.

Table 4. Description of different LOS adapted from the Highway Capacity Manual (HCM) [1].

LOS	Density (Vehicle/Mile/Lane)	Description
A	≤11	Free flow
B	>11–18	Reasonably free flow
C	>18–26	Stable flow (acceptable delays)
D	>26–35	Speeds decline slightly with increasing flows
E	>35–45	Operation near or at capacity
F	>45	Forced or breakdown flow

4.4. Step 4: Machine Learning Methods

To accomplish the study objectives and estimate hourly LOS using crowdsourced data, machine learning classification methods were used. In this study, a variety of machine learning algorithms were tested. Among seven methods (Random Forest, Support Vector Machines, K-nearest Neighbor, Decision Tree, Boosted Tree, Naïve Bayes, and Multinomial Logistic Regression), the three methods with the highest accuracy were selected and are reported in this paper. These are as follows:

- Random Forest (RF): RF is an ensemble classification method that combines several random decision trees. In this method, all trees are built independently. Then, it classifies the data based on the majority of votes of all trees;
- Support Vector Machines (SVM): SVMs are well-known margin-based classification methods. For each class, the SVM algorithm finds the optimal support vector that provides the maximum distance to other classes. By calculating the optimal support vectors, the algorithm can identify the boundaries and classify the data;
- K-Nearest Neighbor (KNN): KNNs are non-parametric methods that are widely used for classification. All training data are considered in an n -dimensional feature space (n = number of input features) in this method. For each observation, the algorithm looks for the k (a predefined constant) nearest neighbors based on the Euclidean distance. Then, it assigns the category based on the most frequent label of the neighbors.

Since this study implemented different machine learning methods, they had to be compared to find the preferred model. The classification accuracy and Cohen's kappa coefficient were used to choose the preferred model and features. The accuracy captures the ratio of correctly classified predictions (LOS in this study) in comparison to ground truth data. Kappa is another classification performance measure that calculates how close the classified instances are to the labeled ground truth. Kappa eliminates the correct predictions occurring by chance. Kappa is useful when the data are unbalanced due to the number of observations in each category. It should be noted that the higher the accuracy and Kappa value, the better the performance of the method. The accuracy and Kappa can be calculated using the following equations:

$$Accuracy = \frac{\text{Number of correct predictions}}{\text{Total number of predictions}}, \quad (12)$$

$$Kappa = \frac{Pr(a) - Pr(e)}{1 - Pr(e)}, \tag{13}$$

where, $Pr(a)$ is the ratio of correct classification or accuracy (Equation (12)), and $Pr(e)$ represents the probability of success due to chance.

5. Results

This section first provides the descriptive statistics for all input variables. Then, the machine learning model results are presented. It should be noted that R programming language (version 4.0.0) was used for all analyses and visualization presented in this section. It should also be noted that missing values in the datasets represented less than 1% of the total population and were therefore removed from the dataset. Additionally, outlier values in the speed dataset represented less than 1% of the total; these were replaced with the median speed value.

5.1. Descriptive Statistics

Table 5 presents the descriptive statistics for the hourly input data. All input measures have a range of values during different hours. For example, the hourly average speed has a range between 31.1 and 119.4 km/h (19.3 and 74.2 mph). Furthermore, the number of Waze alerts has a range of 0–101 hourly alerts. This suggests that some of the measures require normalization to remove bias in the models. Therefore, some of the speed measures (average, maximum, minimum, and percentiles) were normalized to improve the dataset quality and prevent an imbalance bias of the dataset.

Table 5. Descriptive statistics of hourly data.

Model Input	Measure	Mean	Min.	Max.	Median	S.D.
Basic statistical measures on speed	Average speed (km/h)	100.6	31.1	119.4	110.7	21.7
	Speed standard deviation (km/h)	6.6	0.0	39.4	4.0	6.3
	Minimum speed (km/h)	88.4	18.0	118.4	105.4	25.7
	Maximum speed (km/h)	111.0	39.4	146.0	118.4	20.1
	Range of speed (km/h)	22.7	0.0	91.1	14.5	18.2
	CoV of speed	0.0	0.0	0.5	0.0	0.1
	SE of speed	0.5	0.0	3.1	0.3	0.5
	25th percentile (km/h)	96.9	25.9	114.4	109.8	24.5
	50th percentile (km/h)	100.9	29.1	118.4	110.7	22.9
	75th percentile (km/h)	105.1	36.5	121.0	111.5	21.1
	90th percentile (km/h)	107.7	39.4	126.5	118.4	20.3
Travel time performance	IQR	5.1	0.0	51.9	3.8	7.7
	TTI	1.2	1.0	3.9	1.0	0.5
	BTI	0.2	−0.4	3.4	0.0	0.4
	PTI	1.4	1.0	5.8	1.1	0.7
Crowdsourced data	Number of Waze alerts	9.0	0.0	101.0	4.0	20.0

Figure 3 presents a boxplot of the number of hourly crowdsourced alerts for each time of day. It shows that, typically, during the daytime and peak hours, there are a higher number of alerts than during night hours. Furthermore, Figure 4 shows a boxplot of the number of alerts, average speed, TTI, BTI, and PTI in each LOS category. This figure indicates that from LOS A to F, the range of the number of Waze alerts, TTI, BTI, and PTI for each LOS increases. On the other hand, the average speed decreases. Moreover, the range of measures in each LOS category is different. The results suggest that these features can be beneficial in describing the traffic status and LOS.

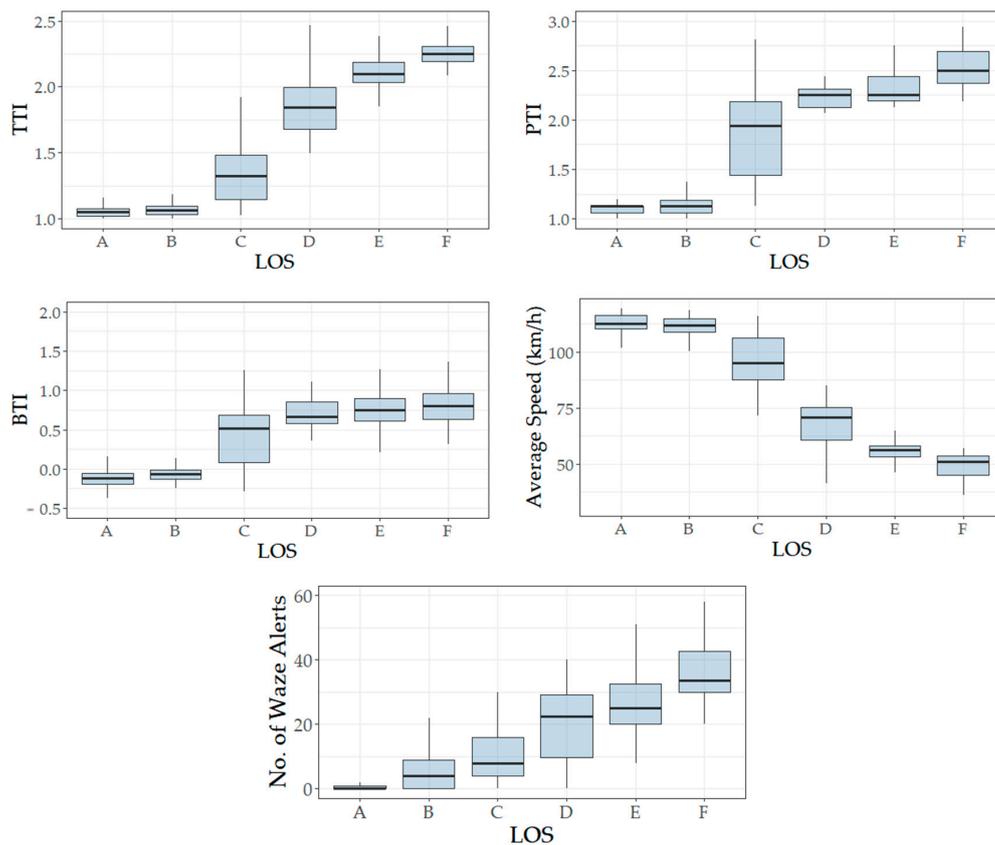


Figure 3. Boxplot of different measures in each LOS category.

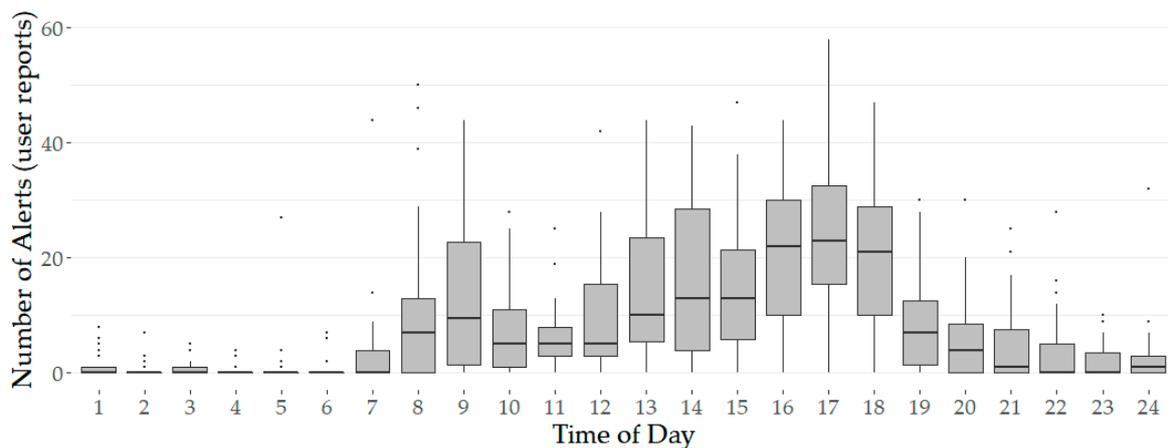


Figure 4. Boxplot of the number of alerts in each time period (hours).

5.2. LOS Classification Model Using Machine Learning

5.2.1. Model Training and Hyperparameter Tuning

To classify LOS, this study employed a variety of machine learning techniques. Among the tested methods, the highest accuracy methods are reported in this paper, which were SVM, RF, and KNN. As previously mentioned, one month of data (October 2019) was selected for the training and validation datasets. The stratified k-fold cross-validation technique was used for all three techniques (SVM, RF, and KNN) to remedy the overfitting problem, reduce the impact of unbalanced label frequencies, and maximize the use of data for both training and testing. In this cross-validation technique, the datasets were randomly divided into equal k-folds with approximately the same number of instances. One-fold

was used as the validation set, and the remaining k-1 folds were used for training. Each fold was used once as the validation dataset. Then, the final accuracy and Kappa value were calculated as the average of k validation results. The k-fold cross-validation technique enabled us to select tuning hyperparameters and increase the classification accuracy. For this purpose, different values of grids of hyperparameters were used in each method to tune hyperparameters and select the best model. To account for overfitting, this study limited the hyperparameters for each machine learning algorithm as follows:

- RF:
 - *Number of trees*: A higher number of trees typically avoids overfitting;
 - *Maximum number of features*: A smaller number of features basically reduces the chance of overfitting;
 - *Maximum tree depth*: The lower the tree depth, the less likely overfitting is;
- KNN:
 - *Number of neighbors (K)*: Increasing the number of neighbors can avoid overfitting;
- SVM:
 - *C*: Demonstrates a trade-off between a high and low accuracy, with a low C value resulting in a smoother decision surface and a lower chance of overfitting.
 - *Sigma*: A large gamma value can cause overfitting.

5.2.2. Model Selection

Here, three different models were estimated to elaborate on the impact of adding crowdsourced data in terms of the LOS assessment accuracy. By comparing these models using different machine learning methods, the preferred model could be selected. The proposed models are as follows:

- Model I uses only travel time performance measures as the model inputs and shows how accurately travel time performance measures can determine LOS;
- Model II uses travel time performance measures and basic statistical measures as the inputs;
- Model III incorporates crowdsourced Waze alerts and uses all three types of input. Model III captures the impact of the crowdsourced alerts in terms of improvement of the LOS classification.

Next, 10-fold cross validation was performed for the three models. Figure 5 displays the result of cross validation for the different machine learning techniques for each model. It can be inferred that, for most methods, adding a statistical measure (Model II) improves the accuracy and Kappa in comparison to Model I. Moreover, adding crowdsourced data to the input measures (Model III) increased the performance of all methods. Based on this step, it can be concluded that crowdsourced data improved the LOS classification performance. All three types of input measures were used in the final model.

In order to assess the sensitivity to the number of folds in cross validation, the values of k (3, 5, and 10) were selected. Table 6 compares the selected classification methods with 3, 5, and 10 cross validation folds for Model III (using all inputs).

Table 6. Summary of classification methods with 3-, 5-, 10-fold cross validation (train dataset).

Classifier	3-Fold Cross Validation		5-Fold Cross Validation		10-Fold Cross Validation	
	Accuracy	Kappa	Accuracy	Kappa	Accuracy	Kappa
SVM	0.91	0.81	0.91	0.81	0.90	0.79
RF	0.91	0.82	0.93	0.83	0.92	0.83
KNN	0.88	0.77	0.89	0.79	0.88	0.76

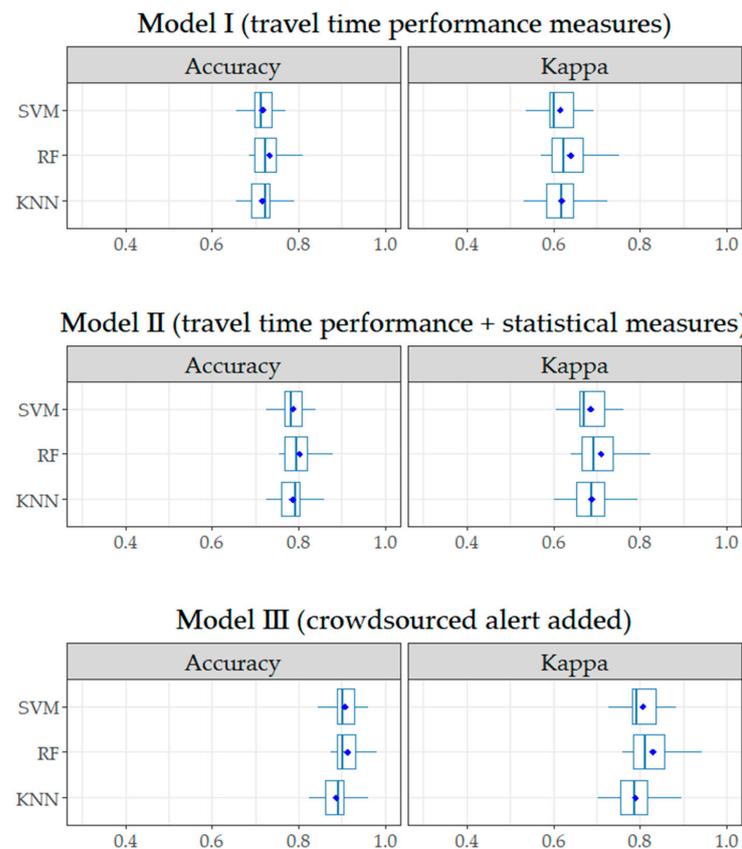


Figure 5. Different model accuracy and Kappa values.

The result of this study shows that machine learning techniques are capable of determining LOS. The RF accuracy of the method with 3, 5, and 10 cross validation was 0.91, 0.93, and 0.92, respectively. Additionally, all Kappa values for RF were above 0.8, which is acceptable for a classification with six different categories. Among the selected machine learning techniques, the RF performed the best result using model III inputs. Here, LOS calculated from RDS data was used as the ground truth. The best RF model (with the highest accuracy and kappa values) was selected to be evaluated with the test dataset. The hyperparameters for the best model (selected RF) included a number of trees of 250, maximum number of features of 2, and maximum tree depth of 3.

5.2.3. Test Result

As mentioned earlier, two months of data in 2020 were collected to test the methodology. The first month (March 16 to April 15) is known to have exhibited a major breakdown in traffic and mobility due to the stay-at-home policy regarding the COVID-19 outbreak. The second month (August) was selected since the traffic breakdown was slightly recovered (about 90%) to normal traffic. The selected RF model was evaluated by using the test data. Table 7 shows that the test result is close to training datasets. The proposed method was also applied to and tested in other segments of I-40 in the Knoxville area, and the result showed a similar accuracy. It shows that the proposed method is extendible to other locations.

Table 7. Summary of testing classification methods.

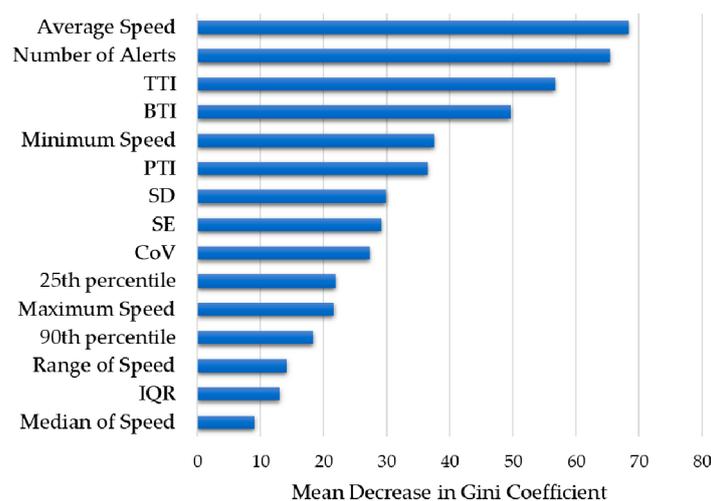
Date	Random Forest Test Result	
	Accuracy	Kappa
03/15/2020 to 04/15/2020	0.95	0.86
08/01/2020 to 08/31/2020	0.92	0.83

5.3. Sensitivity Analysis

The sensitivity of the preferred model (Model III using RF) to different hours of the day was investigated. In the RF model, the corresponding accuracy of each hour of the day was calculated. The accuracy values displayed a range from 0.92 to 0.94 during 24 h of the day. The result highlights that the LOS estimation model is not dependent on the time of day. This method can be used for both peak hours and non-peak hours to estimate the traffic state and LOS. Additionally, the classification accuracy for each ground truth LOS was calculated based on the confusion matrix. Similar to the hourly analysis, each LOS category's accuracy did not deviate from the total accuracy (0.93). This suggests that the proposed method results are not biased due to the frequency of each LOS category.

5.4. Variable Importance

This study suggests that crowdsourced data can improve the LOS classification accuracy. Accordingly, variable importance analysis was performed based on the preferred RF model (Figure 6). The mean decrease in the Gini index was computed from the RF model. A higher value of this index indicates a higher variable importance. The average speed, number of crowdsourced alerts, TTI, BTI, minimum speed, PTI, and standard deviation (SD) of speed are the most important variables in determining the LOS, respectively. This result is consistent with previous studies in the literature that employ speed and travel time reliability measures when determining LOS thresholds. However, the number of crowdsourced alerts seems to impact the LOS prediction accuracy significantly more than TTI, BTI, and TTI.

**Figure 6.** Variable importance plot.

6. Limitations and Future Work

This study proposed a new methodology for estimating LOS using crowdsourced data and machine learning algorithms. However, there were some limitations to this study. This study did not consider the variability and sensitivity of the methodology regarding weather conditions. Additionally, the proposed method used crowdsourced data, travel time variability, and speed statistics measures to estimate LOS. The travel time reliability and speed statistics measures captured the temporal variability of speed

and travel time; however, the spatial variation in speed was not considered. In future research, spatial variation can be used as an input variable for LOS assessment. The speed deviation from upstream and downstream segments can also be addressed in LOS estimation. To this end, more complex methods such as deep learning can be deployed. Using deep neural networks such as convolutional neural network (CNN) and recurrent neural network (RNN) could enable future research to simultaneously capture spatial and temporal variation. Furthermore, in this study, Waze speed and alert data were used as the primary data source. In the future, other crowdsourcing methods (e.g., social media) can be examined when estimating LOS. Finally, this study used Waze alert counts regardless of the event type (jam, accident, and hazards). The impact of each type of event on traffic conditions and LOS should be evaluated in the future.

7. Conclusions

Crowdsourced data availability is increasing rapidly, and machine learning offers the opportunity to analyze it. This study proposed a new methodology to incorporate crowdsourced data in LOS assessment. The method was applied to a 1.5-mile (~2.4 km) segment of freeway on I-40 in Knoxville, Tennessee. Crowdsourced data from Waze were collected, and three categories of input measures (basic statistical measures, travel time reliability, and Waze crowdsourced alerts) were calculated. Machine learning techniques were performed to classify LOS on an hourly basis. Additionally, data collected from fixed location RDS sensors were used to calculate the traffic density and estimate the LOS ground truth using HCM density thresholds.

The results of this study highlight that crowdsourced data and machine learning techniques can be used to estimate LOS. The results revealed that using crowdsourced alerts as an input can significantly improve the model accuracy (about 10%). Moreover, the RF method showed the highest performance among other classification methods in training datasets (accuracy = 0.93 and Kappa = 0.83). Evaluating and testing the trained method also confirmed the classification accuracy. In this method, the LOS estimation accuracy value was relatively consistent among different times of day and LOS categories. Sensitivity analysis confirmed that the accuracy of this methodology does not deviate in traffic peak-hours or non-peak hours. The results also suggest that the average speed, number of alerts, TTI, and BTI are the most important variables in determining LOS.

This method helps to explore the traffic status of freeways without relying on fixed location sensors, times of day, or days of the week. The proposed method has the potential to be applied to different freeway segments to assess LOS. This method does not need fixed location sensors, potentially resulting in lower implementation and maintenance costs. Transportation agencies and DOTs can utilize this method for traffic operation purposes. This method can also analyze freeway traffic in locations outside of urban areas with no fixed location sensors. It benefits from crowdsourced data and can be applied for different time periods, such as hourly, daily, and traffic peak hours.

Author Contributions: Conceptualization, N.H., Y.G., L.D.H., C.B., and P.B.F.; data curation, N.H., Y.G., and L.D.H.; formal analysis, N.H., Y.G., L.D.H., and C.B.; funding acquisition, L.D.H., C.B., and P.B.F.; investigation, N.H. and L.D.H.; methodology, N.H., Y.G., L.D.H., and C.B.; project administration, L.D.H., C.B., and P.B.F.; resources, N.H., Y.G., L.D.H., C.B., and P.B.F.; software, N.H. and Y.G.; supervision, L.D.H. and C.B.; validation, N.H., Y.G., L.D.H., and C.B.; visualization, N.H. and Y.G.; writing—original draft, N.H., Y.G., and C.B.; writing—review and editing, N.H., L.D.H., and C.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Tennessee Department of Transportation (TDOT), contract number: TDOT RES2019-11.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Restrictions apply to the availability of these data. Waze crowdsourced data was obtained from Waze with the permission of TDOT. RDS data was obtained from TDOT.

Acknowledgments: Special thanks to TDOT and the Waze for Cities (WFC) program for providing data used in this study.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Board, T.R. *Highway Capacity Manual 6th Edition: A Guide for Multimodal Mobility Analysis*; The National Academies Press: Washington, DC, USA, 2016; p. 6.
2. Khan, S.M.; Dey, K.C.; Chowdhury, M. Real-time traffic state estimation with connected vehicles. *Ieee Trans. Intell. Transp. Syst.* **2017**, *18*, 1687–1699. [[CrossRef](#)]
3. Hernandez, S.; Tok, A.; Ritchie, S.G. *Density Estimation Using Inductive Loop Signature Based Vehicle Re-Identification and Classification*; Institute of Transportation Studies University of California: Irvine, CA, USA, 2013.
4. Hargrove, S.R.; Lim, H.; Han, L.D.; Freeze, P.B. Empirical evaluation of the accuracy of technologies for measuring average speed in real time. *Transp. Res. Rec.* **2016**, *2594*, 73–82. [[CrossRef](#)]
5. Underwood, S.E. *A Review and Classification of Sensors for Intelligent Vehicle-Highway Systems*; University of Michigan Transportation Research Institute: Ann Arbor, MI, USA, 1990.
6. Zhang, J.; Wang, F.-Y.; Wang, K.; Lin, W.-H.; Xu, X.; Chen, C. Data-driven intelligent transportation systems: A survey. *IEEE Trans. Intell. Transp. Syst.* **2011**, *12*, 1624–1639. [[CrossRef](#)]
7. Herrera, J.C.; Work, D.B.; Herring, R.; Ban, X.J.; Jacobson, Q.; Bayen, A.M. Evaluation of traffic data obtained via GPS-enabled mobile phones: The Mobile Century field experiment. *Transp. Res. Part C Emerg. Technol.* **2010**, *18*, 568–583. [[CrossRef](#)]
8. Hoseinzadeh, N.; Arvin, R.; Khattak, A.J.; Han, L.D. Integrating safety and mobility for pathfinding using big data generated by connected vehicles. *J. Intell. Transp. Syst.* **2020**, *24*, 404–420. [[CrossRef](#)]
9. Khattak, A.J.; Mahdinia, I.; Mohammadi, S.; Mohammadnazar, A.; Wali, B. Big Data Generated by Connected and Automated Vehicles for Safety Monitoring, Assessment and Improvement, Final Report (Year 3). *arXiv* **2021**, arXiv:06106.
10. Mohammadnazar, A.; Arvin, R.; Khattak, A.J. Classifying travelers' driving style using basic safety messages generated by connected vehicles: Application of unsupervised machine learning. *Transp. Res. Part C Emerg. Technol.* **2021**, *122*, 102917. [[CrossRef](#)]
11. Mousavi, S.M.; Lord, D.; Dadashova, B.; Mousavi, S.R. Can Autonomous Vehicles Enhance Traffic Safety at Unsignalized Intersections? In Proceedings of the International Conference on Transportation and Development, Seattle, WA, USA, 26–29 May 2020; American Society of Civil Engineers: Reston, VA, USA, 2020; pp. 194–206.
12. Mousavi, S.M.; Osman, O.A.; Lord, D.; Dixon, K.K.; Dadashova, B. Investigating the safety and operational benefits of mixed traffic environments with different automated vehicle market penetration rates in the proximity of a driveway on an urban arterial. *Accid. Anal. Prev.* **2021**, *152*, 105982. [[CrossRef](#)] [[PubMed](#)]
13. Azad, M.; Hoseinzadeh, N.; Brakewood, C.; Cherry, C.R.; Han, L.D. Fully autonomous buses: A literature review and future research directions. *J. Adv. Transp.* **2019**, *2019*, 4603548. [[CrossRef](#)]
14. Azad, M.; Hoseinzadeh, N.; Brakewood, C.; Cherry, C.R.; Han, L.D. A literature review on fully autonomous buses. In Proceedings of the Transportation Research Board 98th Annual 2019, Washington, DC, USA, 13–17 January 2019.
15. Mahdinia, I.; Mohammadnazar, A.; Arvin, R.; Khattak, A.J. Integration of automated vehicles in mixed traffic: Evaluating changes in performance of following human-driven vehicles. *Accid. Anal. Prev.* **2021**, *152*, 106006. [[CrossRef](#)]
16. Mahdinia, I.; Arvin, R.; Khattak, A.J.; Ghiasi, A. Safety, energy, and emissions impacts of adaptive cruise control and cooperative adaptive cruise control. *Transp. Res. Rec.* **2020**, *2674*, 253–267. [[CrossRef](#)]
17. Chatzimilioudis, G.; Konstantinidis, A.; Laoudias, C.; Zeinalipour-Yazti, D. Crowdsourcing with smartphones. *IEEE Internet Comput.* **2012**, *16*, 36–44. [[CrossRef](#)]
18. Kanhere, S.S. Participatory sensing: Crowdsourcing data from mobile smartphones in urban spaces. In Proceedings of the 2011 IEEE 12th International Conference on Mobile Data Management, Luleå, Sweden, 6–9 June 2011; Volume 2, pp. 3–6.
19. Nair, D.J.; Gilles, F.; Chand, S.; Saxena, N.; Dixit, V. Characterizing multicity urban traffic conditions using crowdsourced data. *PLoS ONE* **2019**, *14*, e0212845.
20. Pack, M.; Ivanov, N. Are You Gonna Go My WAZE? *Inst. Transp. Eng. ITE J.* **2017**, *87*, 28.
21. Hoseinzadeh, N.; Liu, Y.; Han, L.D.; Brakewood, C.; Mohammadnazar, A. Quality of location-based crowdsourced speed data on surface streets: A case study of Waze and Bluetooth speed data in Sevierville, TN. *Comput. Environ. Urban. Syst.* **2020**, *83*, 101518. [[CrossRef](#)]
22. Senarath, Y.; Nannapaneni, S.; Purohit, H.; Dubey, A. Emergency Incident Detection from Crowdsourced Waze Data using Bayesian Information Fusion. *arXiv* **2020**, arXiv:05440.
23. Ali, F.; Ali, A.; Imran, M.; Naqvi, R.A.; Siddiqi, M.H.; Kwak, K.-S. Traffic accident detection and condition analysis based on social networking data. *Accid. Anal. Prev.* **2021**, *151*, 105973. [[CrossRef](#)]

24. Farajiparvar, P.; Hoseinzadeh, N.; Han, L.D.; Hedayatipour, A. Deep Learning Techniques for Traffic Speed Forecasting with Side Information. In Proceedings of the 2020 IEEE Green Energy and Smart Systems Conference (IGESSC), Long Beach, CA, USA, 2–3 November 2020; pp. 1–5.
25. Li, X.; Dadashova, B.; Yu, S.; Zhang, Z. Rethinking Highway Safety Analysis by Leveraging Crowdsourced Waze Data. *Sustainability* **2020**, *12*, 10127. [[CrossRef](#)]
26. Chen, Y.; Zhou, M.; Zheng, Z.; Huo, M. Toward practical crowdsourcing-based road anomaly detection with scale-invariant feature. *IEEE Access* **2019**, *7*, 67666–67678. [[CrossRef](#)]
27. Daraghmi, Y.-A.; Wu, T.-H. Crowdsourcing-Based Road Surface Evaluation and Indexing. *IEEE Trans. Intell. Transp. Syst.* **2020**. [[CrossRef](#)]
28. Martinez, M.; Yang, K.; Constantinescu, A.; Stiefelwagen, R.J.S. Helping the Blind to Get through COVID-19: Social Distancing Assistant Using Real-Time Semantic Segmentation on RGB-D Video. *Sensors* **2020**, *20*, 5202. [[CrossRef](#)] [[PubMed](#)]
29. Amin-Naseri, M.; Chakraborty, P.; Sharma, A.; Gilbert, S.B.; Hong, M. Evaluating the reliability, coverage, and added value of crowdsourced traffic incident reports from Waze. *Transp. Res. Rec.* **2018**, *2672*, 34–43. [[CrossRef](#)]
30. Singh, V.; Gore, N.; Chepuri, A.; Arkatkar, S.; Joshi, G.; Pulugurtha, S. Examining Travel Time Variability and Reliability on an Urban Arterial Road Using Wi-Fi Detections—A Case Study. *J. East. Asia Soc. Transp. Stud.* **2019**, *13*, 2390–2411.
31. Celikoglu, H.B. An approach to dynamic classification of traffic flow patterns. *Comput. Aided Civ. Infrastruct. Eng.* **2013**, *28*, 273–288. [[CrossRef](#)]
32. Sekula, P.; Marković, N.; Laan, Z.V.; Sadabadi, K.F. Estimating historical hourly traffic volumes via machine learning and vehicle probe data: A Maryland case study. *Transp. Res. Part C Emerg. Technol.* **2018**, *97*, 147–158. [[CrossRef](#)]
33. Marwah, B.; Singh, B. Level of service classification for urban heterogeneous traffic: A case study of Kanpur metropolis. In Proceedings of the Fourth International Symposium on Highway Capacity, Maui, HI, USA, 27 June–1 July 2000.
34. Altintasi, O.; Tuydes-Yaman, H.; Tuncay, K. Detection of urban traffic patterns from Floating Car Data (FCD). *Transp. Res. Procedia* **2017**, *22*, 382–391. [[CrossRef](#)]
35. Geistefeldt, J.; Giuliani, S.; Vortisch, P.; Leyn, U.; Trapp, R.; Busch, F.; Rascher, A.; Celikkaya, N. Assessment of level of service on freeways by microscopic traffic simulation. *Transp. Res. Rec.* **2014**, *2461*, 41–49. [[CrossRef](#)]
36. Wu, N.; Lemke, K. A new model for level of service of freeway merge, diverge, and weaving segments. *Procedia-Soc. Behav. Sci.* **2011**, *16*, 151–161. [[CrossRef](#)]
37. Jolovic, D.; Stevanovic, A.; Sajjadi, S.; Martin, P.T. Assessment of Level-Of-Service for Freeway Segments Using HCM and Microsimulation Methods. *Transp. Res. Procedia* **2016**, *15*, 403–416. [[CrossRef](#)]
38. Celikoglu, H.B.; Silgu, M.A. Extension of traffic flow pattern dynamic classification by a macroscopic model using multivariate clustering. *Transp. Sci.* **2016**, *50*, 966–981. [[CrossRef](#)]
39. Aljamal, M.A.; Abdelghaffar, H.M.; Rakha, H.A. Developing a Neural–Kalman Filtering Approach for Estimating Traffic Stream Density Using Probe Vehicle Data. *Sensors* **2019**, *19*, 4325. [[CrossRef](#)] [[PubMed](#)]
40. Wassantachat, T.; Li, Z.; Chen, J.; Wang, Y.; Tan, E. Traffic density estimation with on-line SVM classifier. In Proceedings of the 2009 Sixth IEEE International Conference on Advanced Video and Signal Based Surveillance, Genoa, Italy, 2–4 September 2009; pp. 13–18.
41. Khan, S.M. Real-Time Traffic Condition Assessment with Connected Vehicles. Master’s Thesis, Clemson University, Clemson, SC, USA, 2015.
42. Kurniawan, F.; Sajati, H.; Dinaryanto, O. Image Processing Technique for Traffic Density Estimation. *Int. J. Eng. Technol.* **2017**, *9*, 1496–1503. [[CrossRef](#)]
43. Lomax, T.; Margiotta, R. *Selecting Travel Reliability Measures*; Citeseer: University Park, TX, USA, 2003.
44. Schrank, D.; Eisele, B.; Lomax, T. *TTI’s 2012 Urban Mobility Report*; Texas A&M Transportation Institute, The Texas A&M University System: College Station, TX, USA, 2012; Volume 4.
45. Martchouk, M.; Mannering, F.L.; Singh, L. *Travel time reliability in Indiana*; Publication FHWA/IN/JTRP-2010/08; Joint Transportation Research Program, Indiana Department of Transportation and Purdue University: West Lafayette, IN, USA, 2010. [[CrossRef](#)]
46. Chepuri, A.; Wagh, A.; Arkatkar, S.S.; Joshi, G. Study of travel time variability using two-wheeler probe data—an Indian experience. In Proceedings of the Institution of Civil Engineers-Transport, London, UK, 4 August 2018; Thomas Telford Ltd.: London, UK, 2018; Volume 171, pp. 190–206. [[CrossRef](#)]
47. Kittelson, W.; Vandehey, M. *Incorporation of Travel Time Reliability into the HCM*; Transportation Research Board: Washington, DC, USA, 2013.
48. Chen, C.; Skabardonis, A.; Varaiya, P. Travel-time reliability as a measure of service. *Transp. Res. Rec.* **2003**, *1855*, 74–79. [[CrossRef](#)]
49. Pulugurtha, S.S.; Imran, M.S. Modeling basic freeway section level-of-service based on travel time and reliability. *Case Stud. Transp. Policy* **2020**, *8*, 127–134. [[CrossRef](#)]
50. Kodupuganti, S.R.; Pulugurtha, S.S. Link-level travel time measures-based level of service thresholds by the posted speed limit. *Transport. Res. Interdiscipl. Perspect.* **2019**, *3*, 100068. [[CrossRef](#)]
51. Misra, A.; Gooze, A.; Watkins, K.; Asad, M.; Le Dantec, C.A. Crowdsourcing and its application to transportation data collection and management. *Transp. Res. Rec.* **2014**, *2414*, 1–8. [[CrossRef](#)]

52. Santos, S.R.d.; Davis, C.A., Jr.; Smarzaro, R. Integration of data sources on traffic accidents. In Proceedings of the GeoInfo, Campos do Jordão, Brazil, 27–30 November 2016; pp. 192–203.
53. Bahaweres, R.B.; Akbar, M.R. Analysis of travel time computation accuracy from Crowdsourced data of hospitality application in South of Tangerang City with estimated travel time method. In Proceedings of the 2017 5th International Conference on Cyber and IT Service Management (CITSM), Bali, Indonesia, 8–10 August 2017; pp. 1–5.
54. Turner, S.; Martin, M.; Griffin, G.; Le, M.; Das, S.; Wang, R.; Dadashova, B.; Li, X. *Exploring Crowdsourced Monitoring Data for Safety*; Bureau of Transportation Statistics U.S. Department of Transportation: Washington, DC, USA, 2020.
55. Perez, G.V.A.; Lopez, J.C.; Cabello, A.L.R.; Grajales, E.B.; Espinosa, A.P.; Fabian, J.L.Q. Road Traffic Accidents Analysis in Mexico City through Crowdsourcing Data and Data Mining Techniques. *Int. J. Comput. Inf. Eng.* **2018**, *12*, 604–608.
56. Sanchez, R.; Martinez, D.; Mitnik, O.A.; Yanez-Pagans, P.; Lanzalot, M.L.; Stucchi, R.; Sanguino, L. *Dynamic Traffic Lights and Urban Mobility: An Application of Waze Data to the City of Medellin*; The Research Institute for Development, Growth and Economics (RIDGE): Montevideo, Uruguay, 2019.
57. Naming the Coronavirus Disease (COVID-19) and the Virus that Causes it. 2020. Available online: [https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/naming-the-coronavirus-disease-\(covid-2019\)-and-the-virus-that-causes-it/](https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/naming-the-coronavirus-disease-(covid-2019)-and-the-virus-that-causes-it/) (accessed on 8 February 2021).
58. Al-Sobky, A.-S.A.; Mousab, R.M. Traffic density determination and its applications using smartphone. *Alex. Eng. J.* **2016**, *55*, 513–523. [[CrossRef](#)]
59. Ensley, J.O. Application of Highway Capacity Manual 2010 Level-of-Service Methodologies for Planning Deficiency Analysis. Master's Thesis, University of Tennessee, Knoxville, TN, USA, 2012.