New Travel Time Reliability Methodology for Urban Arterial Corridors

Ernest Tufuor

Laurence Rilett

New Travel Time Reliability Methodology for Urban Arterial Corridors

Ernest O. A. Tufuor and Laurence R. Rilett

University of Nebraska-Lincoln, Lincoln, NE

Corresponding author — Ernest O. A. Tufuor, ernest.tufuor@huskers.unl.edu

Abstract

The need for reliable performance measures of urban arterial corridors is increasing because of the rise in traffic congestion and the high value of users' travel time. Consequently, travel time reliability (TTR), which attempts to capture the day-to-day variability in travel times, has recently received considerable research interest. The basis of all TTR metrics is the underlying travel time distribution (TTD) along the given link or corridor. Estimating and forecasting arterial corridor TTDs for TTR analysis is the focus of this paper. This paper proposes a TTR methodology that addresses some of the limitations of the current U.S. state-of-the-art methodology which was published in the 6th edition of the Highway Capacity Manual (HCM6). Specifically, HCM6 can only estimate average TTD and not the population TTD. However, the population TTD is needed for accurate trip decision-making by individual drivers and logistics companies. In addition, HCM6 cannot be used to analyze the effect of new technologies, such as connected and automated vehicles, nor can it be used easily for long corridors or networks. The proposed TTR methodology, which is traffic-microsimulation based, was applied on a 1.16-mi arterial testbed in Lincoln, Nebraska, U.S. It was shown that the proposed TTR methodology, when calibrated, could replicate the empirical population TTD at a 5% significance level. The population TTD could also be transformed into an average TTD that also replicated the corresponding empirical average TTD at a 5% significance level.

Published in *Transportation Research Record*, 2021
DOI: 10.1177/03611981211006104

Published April 7, 2021

The travel time to traverse a road corridor is, arguably, the most easily understood metric for measuring roadway performance by both road users and traffic managers (*1*, *2*). Providing accurate and timely travel time information to drivers can decrease downstream delays and congestion. Also, transportation agencies use travel times to measure how well their roadway systems are performing. With the recent increase in urban traffic congestion, there has been an increasing need for reliable travel time information. Consequently, travel time reliability (TTR), which attempts to capture the day-to-day variability in travel times, has recently received considerable research interest (*3*, *4*).

There has been significant progress in developing TTR metrics, and Pu provides a comprehensive review (*5*). In addition, there have been several studies on assessing the perceived value of TTR and incorporating TTR measures in traffic demand models. Carrion and Levinson provide a review of the studies on the value of TTR (*6*). An important, but less studied aspect of TTR is estimating and forecasting the travel time variability using mathematical formulations, Monte Carlo simulation, or both (*7*, *8*). The end goal is to better model the observed travel time distribution (TTD), key components of the TTD, or both, which are then used to forecast the corresponding TTR metrics.

Travel time differs across space (e.g., link, corridor, or system), time of day (e.g., AM peak, PM peak, off-peak), day of week (e.g., weekdays, weekends), and analysis period (e.g., summer). When these characteristics are defined, then travel time may be represented by a continuous distribution. Note that a TTD can be developed using average travel times (e.g., over 15 min) or by using individual travel times. In this paper, the former will be referred to as an ''average TTD'' and the latter will be referred to as a ''population TTD.'' Intuitively, for a given study period, the population TTD will have greater variance than the average TTD all else being equal. By definition, the two TTDs are related in that the average TTD can be derived directly from individual travel time distribution. Note that the corollary is not true in that the average travel times cannot be used to derive the individual travel times.

The Highway Capacity Manual (HCM) contains computational procedures for evaluating roadway performance (*8*). The latest and 6th edition of this manual, known as HCM6, included for the first time a methodology for estimating and forecasting average TTD. Specifically, the HCM6

TTR methodology considers the effect of inclement weather, traffic incidents, demand variations, work zones, and special events on the variability of travel times. The time period of analysis may be 15 min or 60 min. The average TTD from HCM6 can then be used to estimate and predict TTR metrics such as the TTI, planning time index (PTI), and the level of travel time reliability (*9*).

A recent study compared an average TTD calculated using the HCM6 procedure with a corresponding empirical average TTD on a 1.16-mi arterial testbed in Lincoln, Nebraska, U.S. It was found that the HCM6 TTR methodology underestimated the observed travel time variability by 67% (*4*). Additionally, a preliminary study on three other corridors in Lincoln, Nebraska, and two corridors in Omaha, Nebraska, showed similar results (*10*). A proposed HCM6 calibration methodology was applied to the same 1.16-mi arterial testbed and it was shown that a calibrated HCM6 TTR model was able to replicate the empirical TTD at a 5% significance level (*11*). While the proposed HCM6 calibration methodology is an important and significant contribution, there are some issues related to the HCM6 TTR methodology that need to be addressed.

First, there is a need to estimate and predict the population TTD, rather than simply the average TTD. The average TTD is typically used by traffic system operators to understand how a given link or corridor is operating, on average. In contrast, the population TTD is more useful to individuals and logistics firms as it defines the range of travel times individual drivers may encounter.

Second, the HCM6 TTR methodology is limited to estimating and forecasting the travel time distribution on a corridor that consists of, at most, eight segments. In addition, only seven work zone activities can be modeled for a given analysis. This may not be scalable to the needs of roadway managers who often need to analyze multicorridor networks.

Third, the HCM6 TTR methodology cannot be used to estimate or predict the effect of new traffic technologies on TTD and the corresponding TTR metrics. The recent advancements in intelligent transportation systems (ITS), computer technology, and the internet-of-things have increased the growth in the usage of these technologies. Examples of such technologies include the use of adaptive signal control systems, strategies for signal preemption at railroad crossings, and the adoption of connected and automated vehicles (CAV). It is important to note that the Highway Capacity and Quality of Service

Committee of the Transportation Research Board has funded several studies to develop capacity adjustment factors (CAF) for CAV applications (*12*). These CAF values will be based on microsimulation studies and will be included in a future HCM. To the authors' knowledge, there is no comparable approach for updating the HCM TTR methodology to account for CAV technologies.

To effectively estimate and predict the population TTD, it is important to capture both the endogenous variability (e.g., driver behavior) and the external or exogenous variability (e.g., inclement weather) that cause variability in arterial travel times. Traffic microsimulation models are ideal for capturing these effects because they can simulate the complex interactions of vehicles as they travel through a traffic network. When properly calibrated, they can readily model the stochastic and dynamic nature of traffic flow (*13*). Some of the advantages of this approach are:

1. Traffic microsimulation models have the potential of modeling longer corridors or even entire networks. They can also model an unlimited number of scenarios (e.g., work zones, weather);
2. Traffic microsimulation models of arterial corridors are becoming widespread and many transport agencies already have calibrated them to local conditions;
3. Calibrated microsimulation models are excellent at modeling key traffic parameters including travel times (*13*). For example, HCM6 has utilized microsimulation models for estimating capacity, capacity adjustment factors, and passenger car equivalents (e.g., HCM) (*8*); and
4. Microsimulation models can output travel time information at the individual vehicle level, which can be analyzed directly or aggregated to estimate the average TTD.

Traffic microsimulation models contain several parameters that are used to describe driver behavior, traffic flow characteristics, and traffic controls. While the default values of these parameters are often provided, users are allowed to change the values to represent local conditions (*14*). The process of changing or adjusting the values of the parameters to replicate observed conditions is known as model calibration (*15*). Applying the default parameters or inappropriate calibration may result in misleading and erroneous models (*16*). It is hypothesized that properly calibrating the proposed TTR model will be a necessary step for estimating and predicting TTDs.

A new TTR methodology, which is based on a traffic microsimulation model, is proposed in this paper. The proposed approach explicitly considers the issue of calibrating the microsimulation model to empirical conditions. It is hypothesized that the new TTR methodology, once calibrated, can be used to estimate accurately both the population TTD, the average TTD, and their corresponding TTR metrics.

The remainder of this paper is organized as follows. The next section provides an overview of the HCM6 methodology and the proposed TTR methodology. Subsequently, the proposed TTR methodology is calibrated and used to estimate the population TTD for the same 1.16-mi testbed used in the previous studies (*11*). The estimated population TTD will be compared with the empirical population TTD using appropriate statistical techniques. Then the estimated population TTD is used to create an average TTD. This average TTD will be compared with the observed average TTD. Next, the proposed TTR methodology is used to estimate average TTDs when only empirical average travel time data is available. Lastly, the TTR metrics developed from the TTDs will be analyzed. Note that the focus of the analysis is combined conditions (i.e., inclement weather, work zone, and normal conditions). However, the methodology was also run for each of the individual scenarios including snow/rain conditions, work zone conditions, and normal conditions, and these results will be discussed.

The HCM6 TTR Methodology

A schematic diagram of the HCM6 TTR methodology is shown in **Figure 1**. The HCM6 TTR methodology is designed to model six different scenarios: work zones, snow/rain weather, special events, incidents, demand variation, and combined. The goal of the HCM6 approach is to account for the most important sources of variability when calculating the TTD. The analysis time period covered by the TTD is referred to as the reliability reporting period (e.g., 1 year). The travel time is calculated using the core HCM6 arterial travel time predictor model (e.g., from chapter 16 of the HCM) (*8*). A Monte Carlo simulation is used to model the

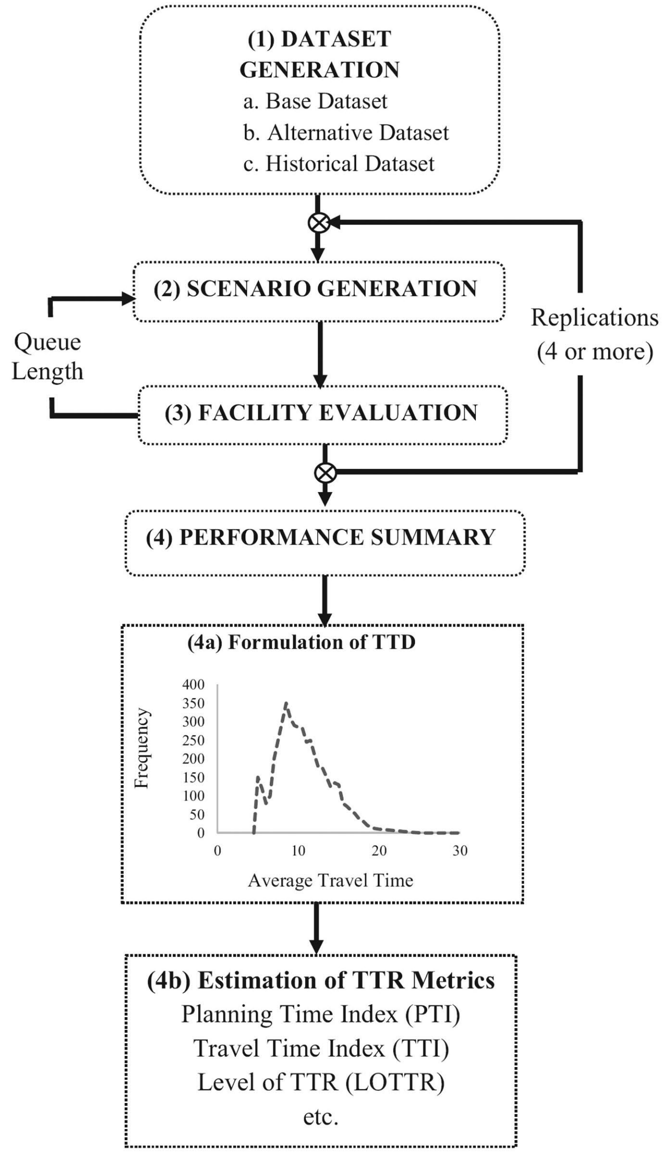

Figure 1. The Highway Capacity Manual (6th edition) (HCM6) travel time reliability (TTR) methodology general framework (8). Note: TTD = travel time distribution.

stochastic components (e.g., changes in weather, changes in incident). The user may choose to use a 15 min or 60 min aggregation interval for stimulating the average travel times. The output is the average TTD along the corridor. From this average TTD, a variety of TTR metrics are calculated. A detailed explanation of the HCM6 TTR methodology is beyond the scope of this paper, but detailed information on the underlying logic may be found elsewhere (*4*, *8*, *17*).

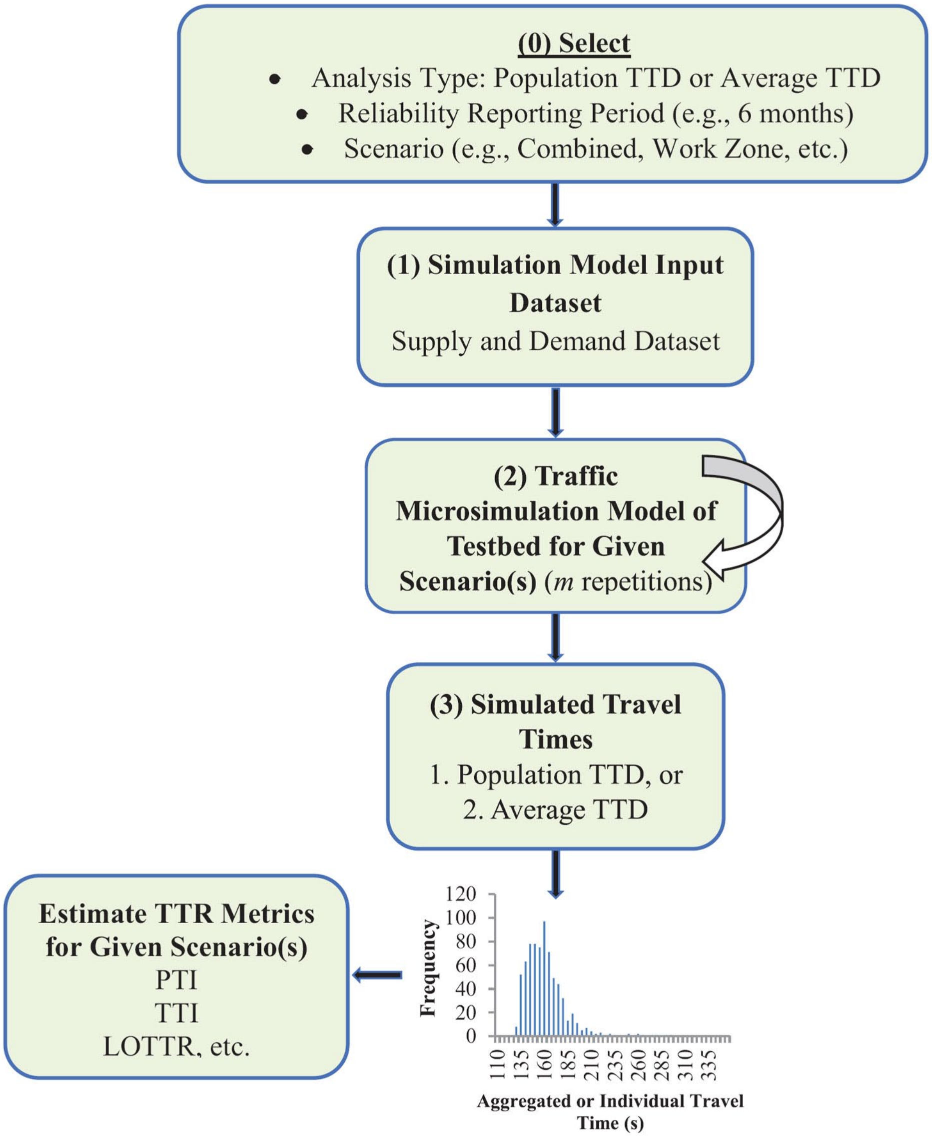

Figure 2. Proposed travel time reliability (TTR) methodology. Note: LOTTR = level of travel time reliability; PTI = planning time index; TTD = travel time distribution; TTI = travel time index.

The Proposed TTR Methodology

The proposed TTR methodology is shown in **Figure 2** and it follows a similar logic to the HCM6 TTR methodology. One main difference is that a traffic microsimulation model is used, instead of the HCM6 macroscopic model, to estimate travel times. Another difference is that the developers of HCM6 had to account for the effect of residual queues as the model transitioned from one time period to another. By definition, the queues are automatically accounted for in the microsimulation model. Because the microsimulation model, by definition, can model stochastic elements, the HCM6 Monte Carlo logic is no longer required when estimating travel times. In this situation, the number of weather events, incidents, type and extent of work zones, and so forth, would be known a priori and could be modeled explicitly. However, if the proposed approach is to be used for forecasting, then the exogenous variables (e.g., weather, incidents) will need to be forecast. In this case, a Monte Carlo approach, similar to HCM6, would be necessary. The key components of Figure 2 are described in the following sections.

Step 0. Select TTD Criteria

In Step 0, the user decides the type of TTD they wish to estimate. This could be either a population TTD (e.g., all vehicles) or an average TTD (e.g., aggregated at 15 min). Similar to HCM6, the user identifies the reliability reporting period (e.g., 6 months) and the scenario to be modeled (e.g., work zone, incident, combined).

Step 1. Collect Input Data

In this step, two datasets are obtained as shown in Step 1 of Figure 2. The supply dataset consists of the geometric features (e.g., segment lengths, number of lanes, road width), intersection control types and settings (signalized or un-signalized), and the road functional class (minor, major, or principal arterial). This information is required to model the physical road infrastructure and existing traffic controls for a given scenario. The demand dataset consists of the number of vehicles, by classification, that enter the arterial corridor at a given time at all origin nodes. The supply and demand datasets on the testbed to be analyzed in this paper were obtained from the City of Lincoln and by field observations. More details about the testbed may be found elsewhere (*4*, *11*). Note that the simulation model may have a different parameter set, demand dataset, and supply dataset for a given scenario.

Step 2. Run the Traffic Microsimulation Model for Given Scenario(s)

In Step 2, the interaction between the demand and supply is modeled using the traffic microsimulation model for a predefined simulation time (*18*). There are several available traffic microsimulation tools including TransModeler™, AIMSUM™, TRANSIMS™, PARAMICS™, CORSIM™, and VISSIM™. A detailed comparison of these tools is provided elsewhere (*13*).

In this paper, VISSIM™ (version 2020), a microscopic traffic simulation software package, is used to model the testbed. VISSIM was selected for four reasons. First, it can output detailed travel time information from each simulated vehicle (*19*). This information can be used to generate simulated population TTD. Secondly, many North American cities have calibrated VISSIM arterial corridor models and it is arguably the state of the practice in the U.S. Thirdly, and most importantly, VISSIM has been utilized in HCM6 for estimating capacity, capacity adjustment factors, and passenger-car equivalent factors. Lastly, VISSIM can be calibrated to a range of scenarios (e.g., weather) that are examined in this TTR methodology. While VISSIM was selected in this paper, the proposed TTR methodology can be used with any traffic microsimulation tool without loss of generality. Note that the parameter set used by VISSIM will depend on the scenario being examined. For example, it might be expected that the car following time headway parameter value will be larger for inclement weather, all else being equal.

Given that VISSIM is based on a simulation process, changing the random seed number will result in a different TTD. For each iteration, VISSIM is run *m* times, each with a new randomly generated seed number, to ensure the results are not biased by the inadvertent choice of a poor seed number. In this paper, *m* was set to 10 and the outputs from all 10 simulation runs were used to populate the TTD.

Note that most traffic microsimulation tools output similar performance measures including delay, travel time, queue length, and fuel consumption. These can all be examined from a reliability perspective. While travel time was the metric chosen in this paper, the new TTR methodology can be used with any reliability metric without loss of generality. All that would be required would be access to corresponding empirical data so the model could be calibrated.

Step 3. Compile Simulated Travel Time Data

In this step, shown as (3) in Figure 2, the simulated travel times that were output from VISSIM are used to develop the corridor TTD. The output travel times are converted, if necessary, to a desirable aggregation level. For example, if the scenario being examined is for 15 min average conditions then the output individual travel times must be aggregated appropriately. It should be noted that, unlike HCM6, the proposed TTR methodology can output simulated TTD for any spatial representation (e.g., links, segments, and corridors), at any travel time aggregation level preferred by the user, and for any direction of traffic.

The simulated corridor TTD that is output from VISSIM is used to estimate the corresponding travel time reliability metrics. The calculation of the TTR metrics will be discussed later in this paper.

Calibration of the Proposed TTR Methodology

The calibration process of the proposed TTR methodology is iterative, and it is shown in **Figure 3**. In this process, the empirical travel time distribution measured on the testbed over the reliability reporting period (e.g., 6 months, a year) is required as shown in Step 1b of Figure 3. In this paper, it is assumed that the TTD for the testbed for the given scenario over the reliability reporting period is available. Note that the model can still be calibrated with TTD metrics (e.g., mean, variance) but in all likelihood would not be accurate.

In this paper, Bluetooth (BT) travel time data, obtained as part of a long-term arterial testbed study conducted by the Nebraska Transportation Center (NTC), will be used. The PM peak (4:30 p.m. to 5:30 p.m.) for all weekdays in 2016 was used in the analysis as this period was shown, in a preliminary analysis, to have the greatest variability. Note that empirical TTDs were collected for every day in the analysis period and disaggregated by scenario (e.g., weather, work zone). A detailed description of the BT system and the data that was collected may be found elsewhere (*4*, *20*).

The empirical travel times may also need to be compiled into a user-defined format as shown in Step 3b. For example, the empirical travel time will vary across scenarios (e.g., snow/rain conditions, work zone

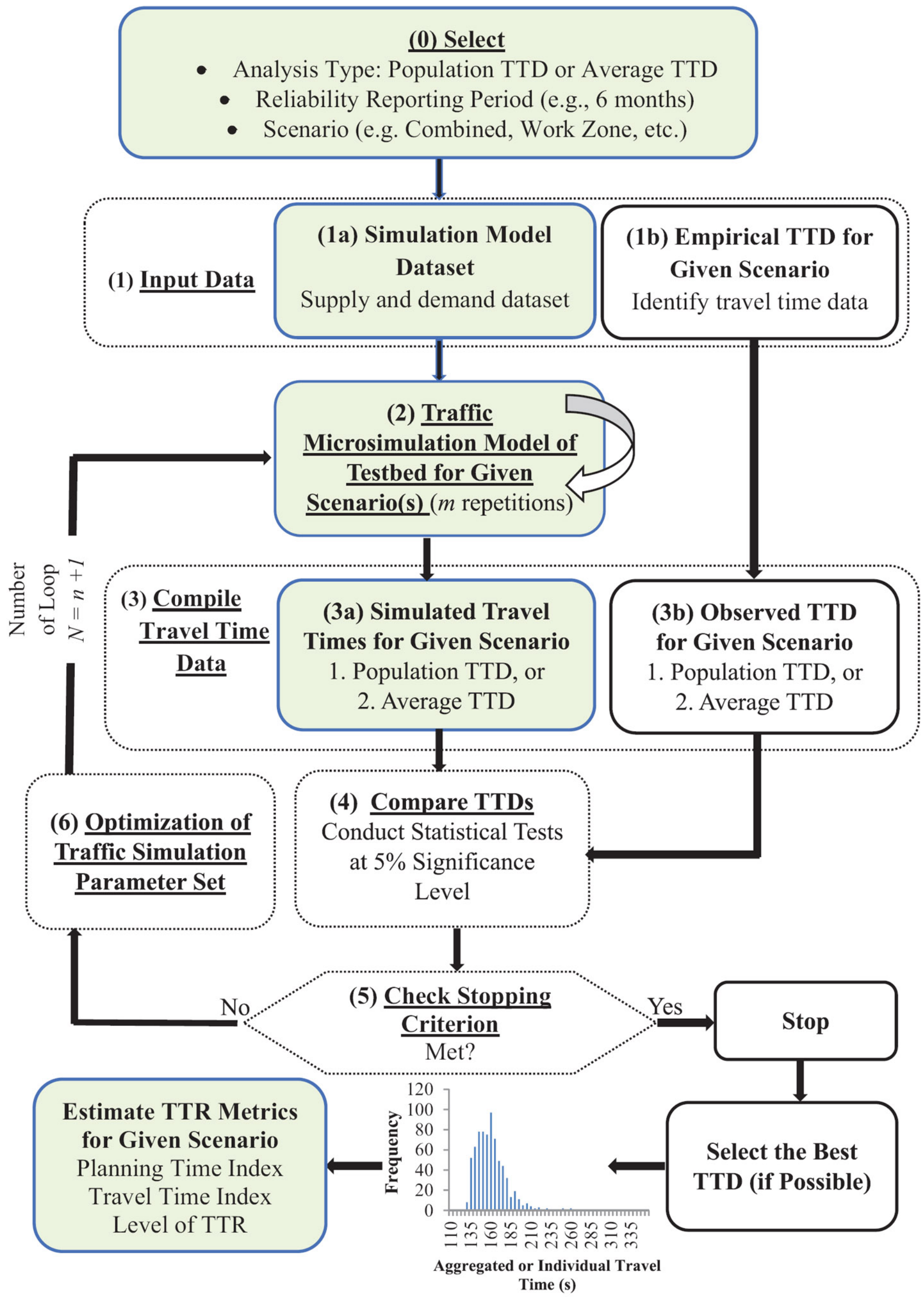

Figure 3. Proposed travel time reliability (TTR) methodology calibration flowchart. Note: TTD = travel time distribution.

conditions, normal conditions, and combined conditions) and the appropriate travel time aggregation must be identified (e.g., none, 15 min, 60 min). Intuitively, both the simulation TTD and the empirical TTD identified in this step must be comparable (e.g., same aggregation level, same scenario, same reliability reporting period). The key components of Figure 3 are described in the following sections.

Step 4. Compare the Simulated TTD with the Empirical TTD

In this step, the simulated TTD from Step 3a is compared statistically with the empirical TTD from Step 3b. The KS test was used in this paper to test the hypothesis that the simulated TTD and the empirical BT TTD are "similar." Let $t_1, t_2, ..., t_n$ be the empirical BT travel times with cumulative distribution function (CDF) F_{BT}, and let the F_S be the CDF of the simulated travel times. The KS test null hypothesis is as follows:

$$\text{The null hypothesis, } H_0 : F_{BT}(t) = F_S(t), \forall t$$

The maximum distance D between the CDFs is the KS test statistic, and it is defined as

$$D = \max\left|F_{BT}(t) - F_S(t)\right| \tag{1}$$

The H_0 is rejected if $D > C$, where C is the critical value in a KS table from any standard statistics textbook (*21*). The critical value corresponding to a 5% significance level was used in this paper.

Note that there are several non-parametric tests that can be used to test the differences between the CDFs as well as to test the differences between their expectations (e.g., mean, variance, median). The best comparison to use will be application-specific (*15*, *22*).

Step 5. Stopping Criteria

Step 5 is used to determine when to stop the calibration process. Because there is no guarantee of convergence, the algorithm is designed to have a maximum number of iterations. A preliminary study on this testbed has shown that the optimization tends to converge when the number of iteration loops (N) is set to 600 and this value was used in this paper. Note that, because VISSIM is run m times for a given iteration, the total number of VISSIM simulations analyzed will be the product of N=600 and m=10 or 6,000 for the example in this paper.

It took approximately 72 machine hours for a 64-bit Intel® Core™ i7 CPU 950@3.07GHz 12GB RAM desktop computer to complete the entire simulation process. Essentially, the level of corridor detail, the number of computers used, and the capability of each computer will determine the calibration time.

Step 6. Optimization of Traffic Simulation Parameters

Most traffic microsimulation models use psychophysical driver behavior algorithms that attempt to replicate human car-following behavior in vehicle traffic streams (*23*, *24*). For example, there are approximately 30 user-controlled VISSIM parameters associated with these models. These parameters can be grouped into car-following and lane-changing parameters and they control vehicle interactions, trajectories, and performance. These are the set of parameters that are changed during the calibration. The individual vehicle travel times output from the model are a direct result of these parameters.

The output of Step 6 is a new parameter set. The goal of the optimization step is that the resulting simulated travel times output from Step 2 will be ‘‘closer’’ to the empirical travel time data than any of the previous parameter sets. Based on past research, there are several excellent optimization algorithms that would be good candidates for use in Step 6. These include the genetic algorithm (GA), simulated annealing, and the simplex method (*11*, *15*, *18*, *25*). GA was chosen in this paper. The GA population size was equal to 20, the maximum generation was set to 30, the mutation probability was 1.75%, and the crossover rate was 70%. A detailed description of GA may be found elsewhere (*25*).

Step 7. Analysis of Results

Once the calibration procedure stops, there may be one acceptable parameter set, multiple acceptable parameter sets, or no acceptable parameter set. If there are multiple acceptable solutions, then it will be necessary to select secondary criteria to identify the ‘‘best’’ TTD. Candidate criteria include the comparison of root-mean-square errors, the sum of squared errors, and the mean average percentage errors. In this paper, the mean average error (MAE) defined by Equation 2 was selected as the secondary criteria. The acceptable TTD that also had the lowest MAE was chosen as the ‘‘best.’’

$$\text{MAE} = \frac{\sum_{i=1}^{n} \left| S_i - E_i \right|}{n} \tag{2}$$

where

S_i = the number of simulated vehicles that have corridor travel times corresponding to bin i
E_i = the number of observed vehicles that have travel times corresponding to bin i
n = the number of bins.

Note that for the examples in this paper the bins are 10 s wide and there are a total of 20 bins that have observed or simulated travel times. The first bin is from 100 to 110 s.

Note that if there are no acceptable solutions the user may choose to rerun the calibration procedure using a larger number of iterations (N) or adjust the type and number of calibration parameters in the parameter set that is being calibrated.

It is important to note that the Federal Highway Administration (FHWA) Traffic Analysis Toolbox: Volume III provides a calibration framework that is similar to the calibration process described in this paper and could potentially be used in its place (*26*). The FHWA process was designed to calibrate to the performance metrics (e.g., average travel time, variance) rather than the TTD. It also relies on parametric-based error bounds (e.g., 95% confidence intervals) to ascertain whether a given model is acceptable. In addition, the FHWA approach relies on empirical data from a representative day. Because this paper utilizes the entire TTD and not TTD metrics, and the empirical TTD is available over the entire analysis period, it was decided to use a nonparametric calibration procedure specifically designed to match travel time distributions (*15*, *22*). If a user only had a sampling of TTD metrics, then a parametric approach would be appropriate (*15*, *22*, *26*). However, with the recent advancements in travel time data collection systems, obtaining TTDs is relatively easy.

Estimating Test Corridor Population Travel Time Distribution

The proposed TTR methodology was used on the 1.16-mi Lincoln testbed. As a first step, the TTR model was calibrated to the observed population TTD for the combined scenario. In this scenario, all components of variability, including incidents, work zones, and weather, are modeled together. Only the PM peak period (4:30–5:30 p.m.) for all weekdays in 2016 was used in the analysis, as this period was shown, in a preliminary analysis, to have the greatest variability.

Figure 4 shows two CDFs of the population TTD. The solid blue line is the empirical population CDF which was obtained in Step 3b of Figure 3. The red dotted line shows the CDF of the population TTD when the VISSIM parameter set was based on the default parameters (e.g., no calibration). In this situation, as shown in Figure 2, the VISSIM model is run once using the default parameter set, the individual travel times are output, and the population TTD was developed from this output.

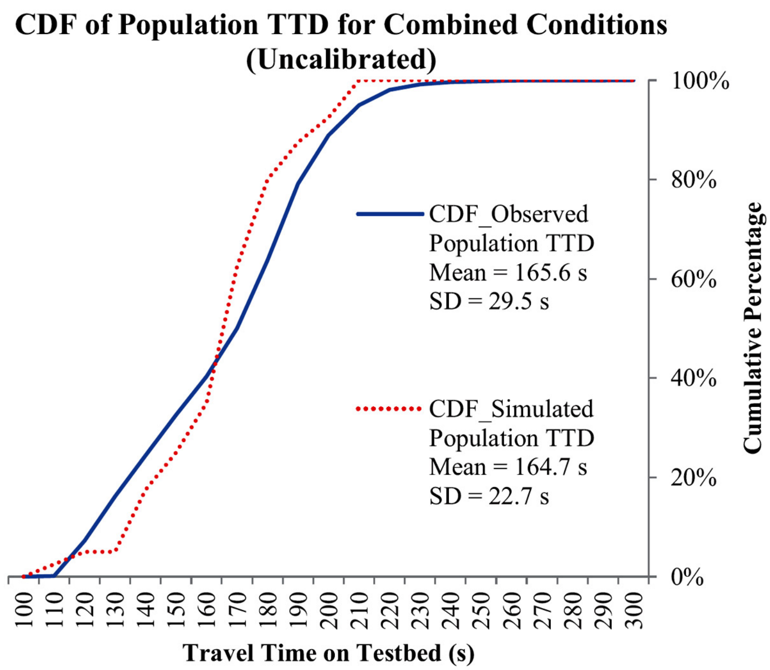

Figure 4. Cumulative distribution function (CDF) comparison of observed and uncalibrated simulated population travel time distributions (TTDs). Note: SD = standard deviation.

It may be seen that there are considerable differences between these two CDFs. Not surprisingly, the results from the KS test showed that there was a statistically significant difference between these two population TTDs at a 5% significance level. While the difference in the mean values was approximately 1 s, there was an approximately 23% difference in their standard deviations. In other words, if not properly calibrated, the proposed TTR model will indicate a more reliable testbed corridor than would have been observed in the field.

Subsequently, the new TTR methodology was calibrated following the logic shown in Figure 3. Of the 600 parameter sets examined, the number of acceptable VISSIM parameter sets was 25. In other words, the results of the KS test show that a total of 25 population TTDs were determined to be statistically the same as the empirical population TTD at a 5% significance level. The MAE for each of the 25 population TTDs was estimated by Equation 2. The population TTD with the lowest MAE value, which was 1.08, was selected as having the ''best'' calibrated parameter set.

Figure 5 shows the cumulative distribution functions of the ''best'' population TTD, shown by the red dotted line, and the empirical

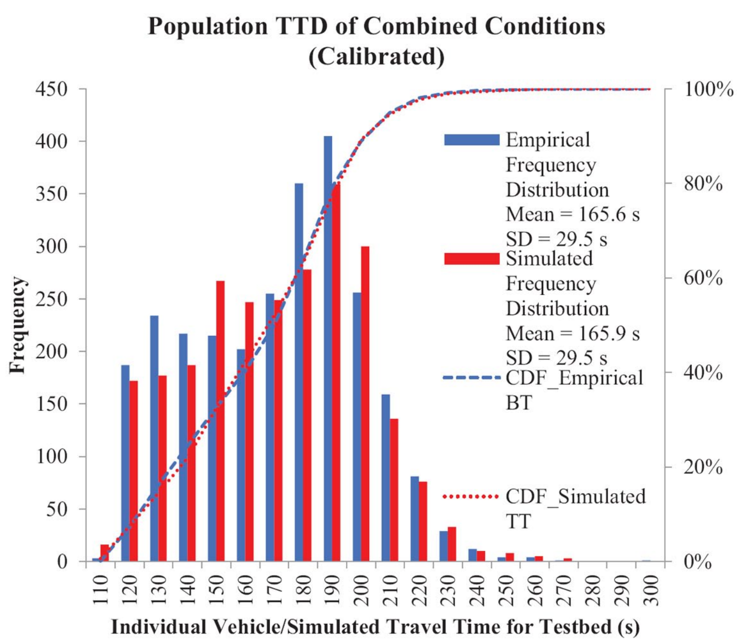

Figure 5. Observed and calibrated simulated population travel time distributions (TTDs) for combined conditions. BT = Bluetooth; CDF = cumulative distribution function; SD = standard deviation; TT = travel time..

population TTD shown by the blue dotted line. It may be seen that, visually, they are very close. In addition, the difference between the mean and standard deviation values of the two TTDs were 0.2% and 0.1%, respectively. Lastly, there were no statistically significant differences between the two TTDs, according to the KS test, at the 5% significance level. This implies that the new TTR methodology, when calibrated, can replicate the field population TTD for this test corridor.

It should be noted that the above process was repeated for each of the weather, work zone, incident, and normal conditions scenarios. Similar results were obtained in that the VISSIM model that was calibrated for a given scenario was able to replicate the empirical TTD for that scenario (e.g., the calibrated weather simulation model produced a TTD that is statistically similar to the empirical weather TTD). The statistical tests were acceptable at the 5% level. Because the conclusions across all four scenarios were essentially the same only the results for the combined scenario were provided in this paper.

Estimating Test Corridor Average TTD Based on Calibrated Population TTD

A natural question is whether the vehicle travel times, that were output from the calibrated VISSIM model and used to create the population TTD, can also be used to create an accurate average TTD. **Figure 6** shows various 15 min average travel time CDFs. The solid blue line represents the empirical average 15 min TTD obtained from the BT system. This is the target CDF that will be compared with the CDF derived from the VISSIM output.

The red dotted line in Figure 6 shows the 15 min average travel time CDF developed using the individual vehicle travel time from the population analysis. It may be seen that the CDFs are similar. The difference between the mean values and the standard deviation values of the two TTDs are only 1% and 4%, respectively. It was found that there were no statistically significant differences between the CDFs, mean values, and the median values of the 15 min aggregation of population TTD and the empirical 15 min average TTD at a 5% significance level. For this corridor, this implies that the user only needs to calibrate the proposed TTR model to the population TTD and then aggregate the travel time output to an appropriate aggregation level.

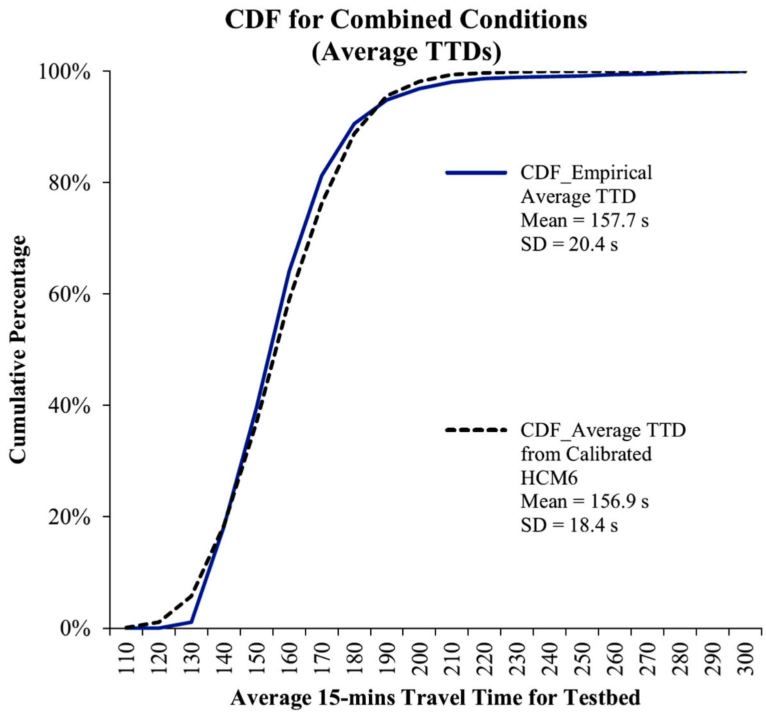

Figure 6. Cumulative distribution functions (CDFs) of simulated average travel time distributions (TTD) and the empirical averaged TTDs for the combined conditions. HCMC6 = Highway Capacity Manual (6th edition); SD = standard deviation.

Estimating Test Corridor Average Travel Time Distribution

Sometimes a user may only have access to average empirical travel times. For example, many private-sector travel time sources (e.g., INRIX) only provide average travel time values (*27*). In this situation, the methodology in Figure 3 can be run using the average TTD as the calibration goal. As before, an acceptable VISSIM parameter set is one in which the simulated average TTD is statistically the ''same'' at the 5% significance level as the empirical average TTD. The authors undertook a similar calibration process using the average TTD, and it was found that there was a 0.3% difference between the mean values of empirical average TTD and the simulated average TTD. However, the new TTR methodology underestimated the standard deviation value of the empirical average TTD by 4.2%. There were no statistically significant differences between the estimated average TTD from the new TTR methodology and the empirical average TTD based on a KS test at a 5% significance level. In addition, the calibrated TTR methodology and the calibrated HCM6 model performed equally well from a statistical point of view (*11*). As was illustrated previously, the advantage of the new TTR methodology is that it can also estimate the population TTD.

The above process was repeated for the work zone, incident, and normal conditions scenario using the 15 min average travel time. Similar conclusions were obtained in that the calibrated VISSIM model was able to replicate the empirical TTD for each of these scenarios. All statistical tests were acceptable at the 5% level. Because the findings across all four scenarios were essentially the same, only the results for the combined scenario were provided in this paper.

Estimate the Associated Travel Time Reliability Metrics

Obtaining an acceptable TTD is the first, and arguably the most important, step in the TTR analysis. Once this is done, then TTR metrics may be calculated. Note that there is no commonly accepted definition of TTR. Consequently, apart from standard statistical measures (e.g., mean, median, variance, coefficient of variance), several TTR metrics have been developed (*5*). The commonly used TTR metrics are defined by Equations 3 to 5.

It may be seen that the TTI, as shown in Equation 3, is the ratio of the mean value of the TTD to the free-flow travel time. The free-flow travel time is typically measured during off-peak conditions. In this paper, the methodology presented in the latest edition of HCM was used to estimate the free-flow travel time (*8*). TTI may be used as a measure of congestion and it is based on the assumption that a more reliable travel time is one that is closer to the free-flow travel time, all else being equal (*28*).

$$\text{Travel Time Index (TTI)} = \mu_T / \mu_F \tag{3}$$

$$\text{Planning Time Index (PTI)} = T_{95} / \mu_F \tag{4}$$

$$\text{Level of Travel Time Reliability (LOTTR)} = T_{80} / T_{50} \tag{5}$$

where

μ_T =mean corridor travel time of TTD under consideration (e.g., simulated or observed) (seconds);

μ_F =mean of corridor free-flow travel time (seconds);
T_{95}=95th percentile corridor travel time (seconds);
T_{80}=80th percentile corridor travel time (seconds);
T_{50}=50th percentile corridor travel time (seconds).

PTI, shown in Equation 4, compares the ''near worst'' travel time condition with the free-flow travel time condition. PTI is intended to provide information on how much additional time is required for a traveler to arrive on time for at least 95% of their trips (*9*).

The level of travel time reliability, shown in Equation 5, measures the range of travel time defined by the median value and the 80th percentile travel times. It was found by Rilett et al. to be highly correlated to the standard measures of dispersion (e.g., standard deviation and interquartile range) and fairly poor at identifying changes in TTR brought about by the COVID-19 pandemic (*27*). The level of travel time reliability (LOTTR) is recommended by both the Fixing America's Surface Transportation (FAST) Act and the Moving Ahead of Progress in the 21st Century (MAP-21) as a performance measure for the U.S. National Highway System (*29*).

Table 1 shows the estimated TTR metrics calculated using the simulated travel time distribution (ST), the empirical Bluetooth travel time distribution (BT), and the differences between them. These are provided for the population TTD and every TTD for all five scenarios examined.

It may be seen from Table 1 that the new TTR methodology developed in this paper, when calibrated, results in TTR metric estimates that are very close to the empirical TTR metrics. It may be seen that this was true for all scenarios (normal, snow/rain, work zone, and combined conditions) as evidenced by the differences with the empirical TTR metrics (TTI, PTI, and LOTTR) being all below 3.5%. It is not surprising all the TTR metrics have similar characteristics because the simulated TTD is statistically the ''same'' as the observed TTD for all comparisons.

The TTR metrics for the scenario where the calibrated population TTD was aggregated to create the average TTD also performed well as evidenced by the differences between the empirical TTI, PTI, and LOTTR metrics being 0.6%, 2.6%, and 3.6%, respectively. Interestingly, the calibrated average TTD performed marginally better than when the average TTD was created from the calibrated population TTD. However, the differences in the TTR metrics are minimal and the two TTDs were statistically similar.

Table 1. Travel Time Reliability (TTR) Performance Metrics (Unitless) Scenarios

	TTR performance metrics								
	Travel time index (TTI)			Planning time index (PTI)			Level of travel time reliability (LOTTR)		
	ST	BT	Diff %	ST	BT	Diff %	ST	BT	Diff %
Population TTD									
Normal scenario	1.64	1.65	–0.6	2.00	2.02	–1.0	1.16	1.17	–0.9
Snow and rain	1.70	1.66	2.4	2.46	2.38	3.4	1.25	1.22	2.5
Work zone	1.60	1.60	0.0	1.87	1.86	0.5	1.13	1.12	0.9
Combined	1.64	1.64	0.0	2.09	2.08	–0.5	1.14	1.12	–1.8
Average TTD									
Normal scenario	1.56	1.57	–0.7	1.90	1.92	–1.0	1.10	1.11	–0.9
Snow and rain	1.62	1.58	2.5	2.34	2.27	3.1	1.19	1.16	2.6
Work zone	1.52	1.52	0.0	1.78	1.77	0.6	1.08	1.07	0.9
Combined	1.56	1.56	0.0	1.92	1.89	1.6	1.10	1.10	0.0
Average TTD from calibrated population TTD for combined conditions	1.57	1.56	0.6	1.94	1.89	2.6	1.14	1.10	3.6

BT = empirical Bluetooth travel time distribution; ST = simulated travel time distribution; Diff% = percentage difference between BT and ST; TTD = travel time distribution; base free-flow travel time, $T_{freeflow}$ = 101 s.

Concluding Remarks

This paper proposes a new TTR estimation and forecasting methodology that addresses some of the limitations of the current U.S. state-of-the-art methodology which was published in HCM6. For example, HCM6 can only estimate average TTD and not the population TTD. However, the population TTD is needed for accurate trip decision-making by individual drivers. In addition, HCM6 cannot be used to analyze the effect of new technologies such as CAVs on corridor reliability.

The proposed TTR methodology follows a similar logic to the HCM6 TTR methodology. The main differences are that a traffic microsimulation model is used instead of the HCM macroscopic model. Unlike HCM6, the new TTR methodology can be used to estimate and forecast both an average TTD and a population TTD. The new TTR methodology was illustrated by applying a VISSIM microsimulation tool to model the field data from a 1.16-mi principal arterial in Lincoln, Nebraska.

It was shown that the proposed TTR methodology can be used to estimate the population TTD. The estimated population TTD can then be used to create average TTD at any user-defined aggregation level. An average 15 min TTD created from the calibration of the population TTD was found to replicate the empirical average 15 min TTD at the 5% significance level.

If the analyst only has empirical average TTD, it was shown how the proposed TTR methodology can still be used. The calibrated proposed TTR model was shown to accurately estimate the average TTD for the snow/rain conditions, work zone conditions, normal conditions, and combined conditions.

It was also shown that the TTR metrics derived from the proposed methodology were acceptable as evidenced by the difference between the empirical TTI, PTI, and LOTTR metrics being only 0.6%, 2.6%, and 3.6%, respectively. While the results are promising, the approach was applied to a single arterial corridor in Lincoln, Nebraska for only the PM peak period for 1 year. Repeating the analysis on other arterial corridors across the U.S. for other time periods is recommended. In addition, further study is recommended to test the temporal and spatial transferability of the calibrated TTR model.

Lastly, unlike the current HCM6 approach, the proposed approach in this paper can be used to estimate and forecast TTR for large corridors and networks in a single run. This contrasts with the current HCM6 that can only model corridors with a maximum size of eight segments. For these reasons, the authors believe that the proposed TTR methodology may be a viable alternative for inclusion in the next HCM update.

Acknowledgment The authors appreciate the assistance of Andy Jenkins of the City of Lincoln in making local data readily available. The contents of this paper reflect the views of the authors, who are responsible for the facts and accuracy of the information presented here, and are not necessarily representative of U.S. DOT or the City of Lincoln.

Author Contributions The authors confirm contribution to the paper as follows: study conception and design: E. Tufuor, L. Rilett; data collection: E. Tufuor, L. Rilett; analysis and interpretation of results: E. Tufuor, L. Rilett; draft manuscript preparation: E. Tufuor, L. Rilett. Both authors reviewed the findings and approved the final version of the manuscript.

Conflicting Interests The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding This research was made possible through the funding of the U.S. Department of Transportation and the Nebraska Department of Transportation (DPS-STWP-118, and DPSSTWP-118[a]), as well as the Nebraska Transportation Center.

Data Availability Some or all data, models, or codes used during the study were provided by a third party (testbed demand and supply data). Direct requests for these materials may be made to the provider as indicated in the Acknowledgment.

References

1. Lomax, T. J., and D. L. Schrank. *Using Travel Time Measures to Estimate Mobility and Reliability in Urban Areas*. No. FHWA/TX-02/1511-3. Federal Highway Administration, Washington, D.C., 2002.
2. Yang, S., and P. Cooke. How Accurate is your Travel Time Reliability?—Measuring Accuracy Using Bootstrapping and Lognormal Mixture Models. *Journal of Intelligent Transportation Systems*, Vol. 22, No. 6, 2018, pp. 463–477.
3. Taylor, M. A. P. Travel through Time: The Story of Research on Travel Time Reliability. *Transportmetrica B: Transport Dynamics*, Vol. 1, No. 3, 2013, pp. 174–194.
4. Tufuor, E. O. A., and L. R. Rilett. Analysis of Component Errors in the Highway Capacity Manual Travel Time Reliability Estimations for Urban Streets. *Transportation Research Record: Journal of the Transportation Research Board*, 2020. 2674: 85–97.
5. Pu, W. Analytic Relationships between Travel Time Reliability Measures. *Transportation Research Record: Journal of the Transportation Research Board*, 2011. 2254: 122–130.
6. Carrion, C., and D. Levinson. Value of Travel Time Reliability: A Review of Current Evidence. *Transportation Research Part A: Policy and Practice*, Vol. 46, No. 4, 2012, pp. 720–741.
7. Clark, S., and D. Watling. Modelling Network Travel Time Reliability under Stochastic Demand. *Transportation Research Part B: Methodological*, Vol. 39, No. 2, 2005, pp. 119–140.
8. Highway Capacity Manual (HCM). *A Guide for Multimodal Mobility Analysis*. Transportation Research Board, Washington, D.C., 2016.
9. FHWA Office of Operations, U.S. Department of Transportation. Travel Time Reliability: Making it there on Time, All the Time. 2017. https://ops.fhwa.dot.gov/publications/tt_reliability/TTR_Report.htm#Whatmeasures Accessed December 3, 2019.
10. Murphy, S., L. Rilett, and E. Tufuor. A Comparison of Empirical Data to the Highway Capacity Manual Estimated Travel Time Distribution on Urban Streets. Presented at the Annual Meeting of the Transportation Research Board, Washington, D.C., 2021.

11. Tufuor, E. O., L. R. Rilett, and L. Zhao. Calibrating the Highway Capacity Manual Arterial Travel Time Reliability Model. *Journal of Transportation Engineering, Part A: Systems*, Vol. 146, No. 12, 2020, p. 04020131.
12. Adebisi, A., Y. Liu, B. Schroeder, J. Ma, B. Cesme, A. Jia, and A. Morgan. Developing Highway Capacity Manual Capacity Adjustment Factors for Connected and Automated Traffic on Freeway Segments. *Transportation Research Record: Journal of the Transportation Research Board*, 2020. 2674: 401–415.
13. Federal Highway Administration, Office of Safety Research and Development. Surrogate Safety Measures from Traffic Simulation Models. 2016. https://www.fhwa.dot.gov\publications\research\safety\03050\04.cfm Accessed February 24, 2020.
14. Park, B., and J. D. Schneeberger. Microscopic Simulation Model Calibration and Validation: Case Study of VISSIM Simulation Model for a Coordinated Actuated Signal System. *Transportation Research Record: Journal of the Transportation Research Board*, 2003. 1856: 185–192.
15. Spiegelman, C., E. S. Park, and L. R. Rilett. *Transportation Statistics and Microsimulation*. CRC Press, Boca Raton, FL, 2010.
16. Park, B., and H. Qi. Development and Evaluation of a Procedure for the Calibration of Simulation Models. *Transportation Research Record: Journal of the Transportation Research Board*, 2005. 1934: 208–217.
17. Zegeer, J., J. A. Bonneson, R. G. Dowling, P. Ryus, M. Vandehey, W. Kittelson, N. M. Rouphail, B. J. Schroeder, A. Hajbabaie, B. Aghdashi, and T. Chase. *Incorporating Travel Time Reliability into the Highway Capacity Manual*. No. SHRP 2 Report S2-L08-RW-1. National Research Council, Transportation Research Board, Washington, D.C., 2014.
18. Rilett, L. R. Using Simulation to Estimate and Forecast Transportation Metrics: Lessons Learned. In *CIGOS 2019, Innovation for Sustainable Infrastructure* (C. Ha-Minh, D. Dao, F. Benboudjema, S. Derrible, D. Huynh, and A. Tang, eds.), Springer, Singapore, 2020, pp. 23–33.
19. PTV AG. *PTV VISSIM User Manual*. Planung Transport Verkehr AG, Karlsruhe, Germany, 2020.
20. Tufuor, E. O. A., and L. R. Rilett. Validation of the Highway Capacity Manual Urban Street Travel Time Reliability Methodology Using Empirical Data. *Transportation Research Record: Journal of the Transportation Research Board*, 2019. 2673: 415–426.
21. Wackerly, D., W. Mendenhall, and R. L. Scheaffer. *Mathematical Statistics with Applications*. Cengage Learning, Belmont, CA, 2014.
22. Kim, S.-J., W. Kim, and L. R. Rilett. Calibration of Microsimulation Models Using Nonparametric Statistical Techniques. *Transportation Research Record: Journal of the Transportation Research Board*, 2005. 1935: 111–119.
23. Wilson, R. E., and J. A. Ward. Car-Following Models: Fifty Years of Linear Stability Analysis–A Mathematical Perspective. *Transportation Planning and Technology*, Vol. 34, No. 1, 2011, pp. 3–18.

24. Brackstone, M., and M. McDonald. Car-Following: A Historical Review. *Transportation Research Part F: Traffic Psychology and Behaviour*, Vol. 2, No. 4, 1999, pp. 181–196.

25. Kramer, O. *Genetic Algorithm Essentials*, Vol. 679. Springer, Cham, 2017, pp. 11–19.

26. Wunderlich, K. E., M. Vasudevan, and P. Wang. *TAT Volume III: Guidelines for Applying Traffic Microsimulation Modeling Software 2019 Update to the 2004 Version*. No. FHWA-HOP-18-036. United States Federal Highway Administration, Washington, D.C., 2019.

27. Rilett, L., E. Tufuor, and S. Murphy. Arterial Roadway Travel Time Reliability and COVID-19 Pandemic. *Journal of Transportation Engineering, Part A: Systems*, 2020.

28. Schrank, D., T. Lomax, and B. Eisele. *2019 Urban Mobility Report – Appendix A*. Texas Transportation Institute, 2019. http://mobility.tamu.edu/ums/report

29. Federal Register. FHWA National Performance Management Measures. https://www.federalregister.gov/documents/2017/01/18/2017-00681/national-performance-management-measures-assessing-performance-of-the-national-highway-system Accessed August 23, 2020.