TRR
JOURNAL OF THE TRANSPORTATION RESEARCH BOARD

Article

Transportation Research Record
2018, Vol. 2672(42) 137–147

DOI: 10.1177/0361198118792124
journals.sagepub.com/home/trr
SAGE

Resampling Methods for Estimating Travel Time Uncertainty: Application of the Gap Bootstrap

Bhaven Naik[1], Laurence R. Rilett[2], Justice Appiah[3], and Lubinda F. Walubita[4]

Abstract

To a large extent, methods of forecasting travel time have placed emphasis on the *quality* of the forecasted value—how close is the forecast point estimate of the mean travel time to its respective field value? However, understanding the reliability or uncertainty margin that exists around the forecasted point estimate is also important. Uncertainty about travel time is a fundamental factor as it leads end-users to change their routes and schedules even when the average travel time is low. Statistical resampling methods have been used previously for uncertainty modeling within the travel time prediction environment. This paper applies a recently developed nonparametric resampling method, the gap bootstrap, to the travel time uncertainty estimation problem, especially as it pertains to large (probe) data sets for which common resampling methods may not be practical because of the possible computational burden and complex patterns of inhomogeneity. The gap bootstrap partitions the original data into smaller groups of approximately uniform data sets and recombines individual group uncertainty estimates into a single estimate of uncertainty. Results of the gap bootstrap uncertainty estimates are compared with those of two popular resampling methods—the traditional bootstrap and the block bootstrap. The results suggest that, for the datasets used in this research, the gap bootstrap adequately captures the dependent structure when compared with the traditional and block bootstrap methods and may thus yield more credible estimates of uncertainty than either the block bootstrap method or the traditional bootstrap method.

In recent years, technological advances have enabled the collection and diffusion of real-time traffic information and this, combined with growing traffic volume and congestion, has triggered an increasing interest in traffic modeling (*1, 2*). The focus for researchers and users of predicted traveler information has been twofold: (i) traffic condition level analysis, and (ii) traffic condition uncertainty analysis. Research in the area of traffic condition level analysis (i.e., first moment of the traffic condition series) has been plentiful with primary focus on using the available traffic data to compute a prediction of the mean (speed, flow rate, or travel time) for a given time period and on ensuring that the difference between the predicted value and its real-time value for the same time period is as small as possible. In particular, this entails the ability to provide a forecasted point estimate of a traffic variable in close comparison with values encountered in real-time (*3*).

On the other hand, traffic condition uncertainty analysis refers to the second moment of the traffic condition series. It deals with any information on the quality (distributional properties) of the condition level value. This quality of confidence measure (or conditional variance) is often also referred to as the reliability of the forecast. A measure of the reliability is important because "drivers claim this influences their decisions that are based on the available travel time information" (*4, 5*) and it has been proven to have a notable impact on motorists' travel decisions, leading to a more efficient use of the traffic network (*6*). In effect, drivers are unlikely to follow the information if they do not believe in its validity (*7*). Therefore, the provision of a measure of reliability could not only increase user comfort by reducing the error risk associated with the traffic information, it could also (for the traffic

[1]Ohio University, Athens, OH
[2]University of Nebraska-Lincoln, Lincoln, NE
[3]Virginia Transportation Research Council, Charlottesville, VA
[4]Texas A&M Transportation Institute (TTI), The Texas A&M University System, College Station, TX

Corresponding Author:
Address correspondence to Bhaven Naik: naik@ohio.edu

engineer) be used to assess the predicted values for model selection. Intuitively, having a measure of the reliability of a forecast will make the job of the transportation professionals who operate the systems considerably easier. In statistical terms, this would mean quantifying the uncertainty (error bars or statistical intervals) associated with a given forecast.

Several researchers have shared insights into the uncertainty margin that exists around a predicted point (or condition level) estimate including Eisele et al. (*8*), Mazloumi et al. (*3*), and Van Lint (*9*). These research efforts assume that datasets are independent and identically distributed. However, this is generally not the case for Intelligent Transportation Systems' (ITS) data that are typically used in the prediction process. The ITS data, as will be described in later sections of this paper, have been observed to have nonstationary distribution and exhibit a general periodicity. Studies by Guo, Huang, and Williams have addressed this periodicity by adopting seasonal autoregressive integrated moving average (SARIMA) models at the condition level modeling stage to remove the autocorrelation structure from the vehicular traffic flow series (*10–12*). Next, the generalized autoregressive conditional heteroscedasticity (GARCH) approach is applied to model uncertainty in the univariate vehicular traffic flow series. Unlike previous studies on uncertainty that assume conditional variance is constant, the GARCH approach accounts for the heteroscedasticity observed in the traffic condition series. Other applications of the GARCH approach include modeling relative velocity (*13*), and also link travel time variability (*14*). These GARCH based traffic condition uncertainty models inherently assume that the second-order property of the traffic conditional series is time varying (or heteroscedastic in nature). It is worthwhile to note that research on traffic condition uncertainty is increasingly becoming popular because many prediction applications have issues with traditional methods (*9*) and is quickly gaining momentum because of its importance in many transportation applications (*5, 6*).

Based on the above background, a methodology that (i) will provide an accurate prediction of the mean travel time for a link, (ii) will give a measure of the reliability in the predicted estimate, and (iii) is based on nonparametric statistical theory is the focus of this paper. This approach will allow transportation agencies to better predict traffic parameters, and their associated reliability metrics, when the assumptions behind standard techniques do not apply.

Background and Problem Definition

The model prediction task for a given statistic $\hat{t}$ (such as the mean of a sample dataset) typically consists of two steps:

- *Step 1—Developing an initial model using a set of data*. This model (essentially a mean regression function) defines the relationship between inputs and the output. The model could be parametric (linear regression, autoregressive) or nonparametric (neural network) in nature.
- *Step 2—Estimating future values* (*predictions*) *of interest relative to current input*.

The predicted parameter, $\hat{t}$, is reflective of the "true" population parameter, t. However, it is unlikely to be exact. The value of the "true" population parameter is usually unknown and it is important that empirical measurements or estimates of the sample parameter are accompanied by a statement about their corresponding uncertainty (*15*). For travel time, this is traditionally expressed mathematically as

$$t = \hat{t} \pm k\sigma_t^2 \tag{1}$$

where
t = "true" travel time mean (s);
$\hat{t}$ = sample travel time mean (s);
σ_t^2 = variance (s^2); and
k = multiplier parameter (s^{-1}).

Uncertainty in Travel Time Prediction

The sources of travel time related uncertainty can be categorized into two aspects—one related to input data quality and one related to model structure. Uncertainty related to the input data quality focuses largely on errors that emanate from failures in traffic monitoring equipment, communication between the field and traffic management center, and the traffic management archiving system (*16, 17*). Typically, uncertainties caused by input data quality are addressed by "cleaning" the data.

The uncertainty in the model structure relates to the noise in the parameter (or weight) selection process. For the reason that a model (with specified parameter) is based on a selected random sample of data, if the current input is outside of the domain, then the prediction would be uncertain.

Bootstrapping

Methods such as maximum likelihood estimation, approximate Bayesian, and the Delta method have previously been used for confidence estimation with nonlinear transportation models. However, these methods have been found to have potential sources of failure that may limit their applicability for transportation applications (*18–20*). Bootstrapping is a more commonly used data-driven simulation technique for computing statistical measures of accuracy for estimates. The idea behind bootstrapping is that the available dataset is nothing but a particular realization of some unknown probability distribution (*21, 22*). The method creates many replicates ("phantom samples") of the original data and then re-estimates the distribution of an estimate or test-statistic on each replicate sample to obtain an estimate of the standard error of the predicted values. This standard error is the basis for the reliability metric in many applications—including those related to transportation parameter forecasting. This paper demonstrates the application of a recently developed gap bootstrap method. A comparison against two existing methods (ordinary and block bootstrap) with specific application to link travel time prediction is then provided.

Ordinary Bootstrapping. The ordinary bootstrap is the simplest and most general bootstrapping technique (*23*). It is commonly adopted in most applications in which the data are (or are assumed to be) independent and identically distributed (i.i.d.). Given an original sample $x_n = \{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$, the bootstrap procedure implies that an unknown distribution F is estimated by constructing an empirical distribution $\hat{F}_n$ based on the n original observations. As illustrated in Figure 1*a*, the bootstrap samples $x_n^{*b} = \left\{\left(x_1^{*b}, y_1^{*b}\right), \left(x_2^{*b}, y_2^{*b}\right), \ldots, \left(x_n^{*b}, y_n^{*b}\right)\right\}$ are repeatedly drawn with replacement from the estimated empirical distribution until a total of B bootstrap samples are obtained (*23, 24*). Then, for each bootstrap sample, the bootstrap version $\hat{\theta}_B^*$ of an estimator $\hat{\theta}$ is calculated. The ordinary bootstrap uncertainty estimate (standard error) is then obtained (*25*) as

$$\text{standard error(se)} = \left\{\frac{1}{B-1}\sum_{b=1}^{B}\left[\hat{\theta}^{*b} - \bar{\theta}^{*}\right]\right\}^{1/2} \qquad (2)$$

where,

$$\text{bootstrap mean } \bar{\theta}^{*} = \frac{1}{B}\sum_{b=1}^{B}\hat{\theta}^{*b} \qquad (3)$$

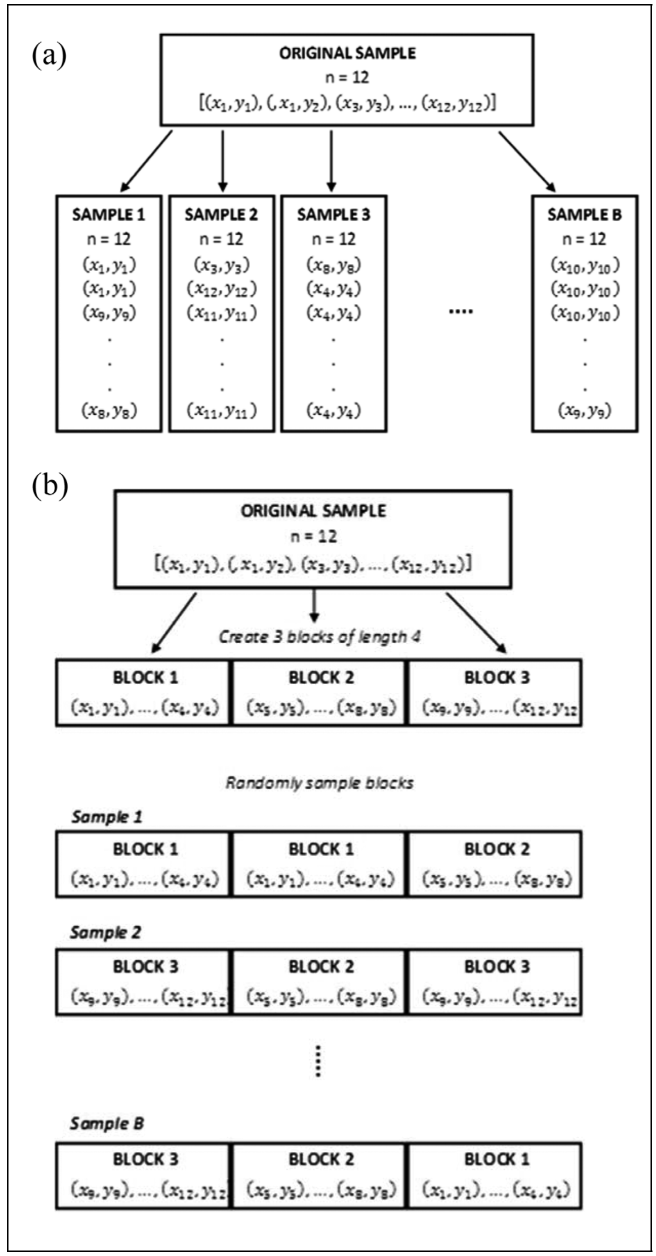

Figure 1. Illustration of bootstrap methods: (*a*) ordinary bootstrapping; and (*b*) block bootstrapping.

Block Bootstrapping. When implementing the ordinary bootstrapping method, each bootstrap sample is drawn from data that are "scrambled" in such a way that aspects of any dependence structure in the original dataset are lost. The ordinary bootstrap is therefore not appropriate for dealing with dependent data (such as link travel times) because the resampling does not capture the dependence structure. Block bootstrapping is currently "the best known method for implementing bootstrapping with dependent data" (*26*).

The basic concepts of block bootstrapping are similar to those of ordinary bootstrapping discussed earlier. However, with block bootstrapping, a bootstrap sample is drawn by dividing the data into contiguous blocks that are randomly sampled (as opposed to sampling individual observations in the ordinary bootstrap method), and laying these blocks end-to-end in the order sampled. This method is illustrated in Figure 1*b*. The blocks could be created as nonoverlapping or overlapping with fixed or varying lengths. The four common block bootstrapping methods available are the moving block bootstrap (*27*), the nonoverlapping block bootstrap (*28*), the circular block bootstrap (*29*), and the stationary block bootstrap (*30*). Synopses of these techniques can be found in the publication by Lahiri (*31*).

Gap Bootstrapping. Gap bootstrapping is a procedure that is appropriate for datasets that can be partitioned into approximately exchangeable partitions (*32*). The only previous applications of the gap bootstrap evaluate uncertainties in point estimates of origin–destination split proportions (*32, 33*).

To illustrate the gap bootstrap technique for travel time estimation, denote the time series of mean travel times (for a given link) by the $p \times q$ matrix X and denote the estimator based on the entire data by $\hat{\theta}$. The rows of X are then partitioned into m groups $X_i, i = 1, 2, 3, \ldots, m$ such that there are $k = T/m$ observations in each column of the partition matrices. Let $\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3, \ldots, \hat{\theta}_m$ be the mean travel times predicted from the submatrices $X_1, X_2, X_3, \ldots, X_m$ respectively. The main result of the gap bootstrap is that the usual covariance matrix estimator obtained from $\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3, \ldots, \hat{\theta}_m$ provides a consistent estimate of the covariance matrix of $\hat{\theta}$ and the rate of convergence is $\sqrt{T}$. Further, the estimator

$$\text{bootstrap mean } \bar{\theta} = \sum_{i=1}^{m} \hat{\theta}_i / m \tag{4}$$

is an asymptotically efficient estimator of θ. The component-wise variance of the m groups is calculated by performing ordinary bootstrapping within each subset. Assuming a constant covariance among the subsets, the final gap bootstrap uncertainty estimate of $\bar{\theta}$ is given by

$$\text{Var}(\bar{\theta}) = \left(\sum_{i=1}^{m} \text{Var}(\hat{\theta}_i) + \sum_{i \neq j} \text{Cov}(\hat{\theta}_i, \hat{\theta}_j) \right) / m^2 \tag{5}$$

where

$\text{Var}(\hat{\theta}_i)$ = the ordinary bootstrap estimator for the variance of $\hat{\theta}_i$,

$\text{Cov}(\hat{\theta}_i, \hat{\theta}_j)$ = the variance estimate of $(\hat{\theta}_i - \hat{\theta}_j)$ given by,

$$\sum_{i \neq j} \left(\hat{\boldsymbol{\theta}}_i - \bar{\boldsymbol{\theta}} \right) \left(\hat{\boldsymbol{\theta}}_j - \bar{\boldsymbol{\theta}} \right)^t / \boldsymbol{m}(\boldsymbol{m} - 1) \tag{6}$$

Study Methodology

A conceptualization of the methodology to calculate the predicted travel time values and additional computation of an estimate of the standard error (or confidence interval) discussed above is presented in Figure 2. It may be seen that the methodology consists of five steps.

Step 1: Initial sample selection—A sample dataset $x_n = \{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$ of size n is selected from the population $x_N = \{(x_1, y_1), (x_2, y_2), \ldots, (x_N, y_N)\}$ of size N.

Step 2: Bootstrapping—"Phantom" samples of the sample travel time dataset (now assumed as the "population") are created. A "bootstrap pairs" approach that considers each training case as a sampling unit and samples with replacement from the training set cases to create a bootstrap sample is adopted (*25*).

Step 3: Fitting ensemble of models—An ensemble of models is fitted on each bootstrap dataset. This study used neural network models. However, it should be noted that the approach is generic and other model types could be used. Essentially, for a given bootstrap sample $\{(x_1^{*b}, y_1^{*b}), (x_2^{*b}, y_2^{*b}), \ldots, (x_n^{*b}, y_n^{*b})\}$, a model is fitted by

$$\text{minimizing} \sum_{i=1}^{n} \left[y_i^{*b} - y(x_i^{*b}; w) \right]^2 \tag{7}$$

subject to $0 < w \leq 1$

where

w = set of model parameters (or weights),

$y(x_i; w)$ = predicted value for input x_i and model parameter w; and

b = bootstrap sample ($b = 1, \ldots, B$).

A multi-layer feed forward neural network with five input nodes and a hidden layer with 15 neurons and 15 output nodes was adopted. The Levenberg-Marquardt algorithm was used to train the neural networks. The neuron activation functions used are the log-sigmoid and the linear function for the hidden and output layers, respectively. Specific details on the design of the neural network model are presented elsewhere (*34*).

Step 4: Calculate predicted values—Models trained in Step 3 are each used here to predict the travel times for a given test dataset. Therefore, if B ensemble models were fitted in Step 3, then for a given set of values there would be B sets of predictions available (i.e., B predictions for period Δt). Note that the neural network in this study was trained on 127 days of data records and tested on 3 days of data (test data).

Step 5: Calculate the mean and standard error - A mean for a given time period Δt, and for all the B predictions from Step 4 is computed in this step. The standard deviation of the sampling distribution of means (i.e., standard error) is also computed.

Demonstration Case Study

The data used in this study were from a corridor located northeast of downtown San Antonio, Texas, along Interstate 35 (IH-35) as indicated in the dashed box shown in Figure 3*a*. This corridor section is part of one of

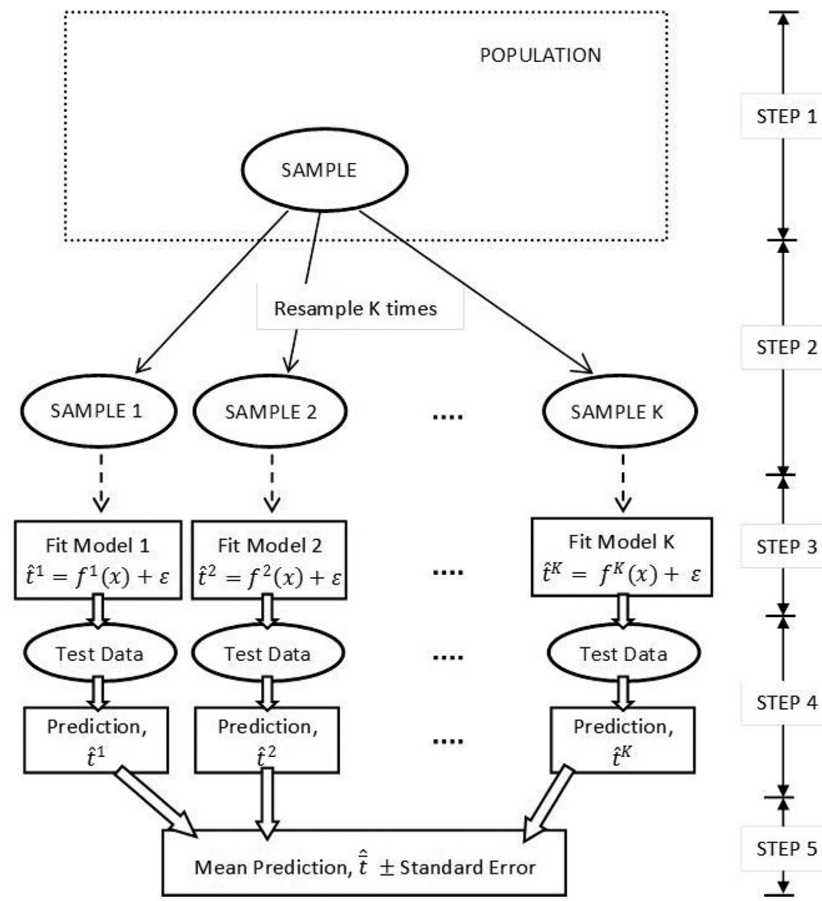

Figure 2. Schematic of the proposed gap bootstrap methodology.

the busiest interstates in Texas. The IH-35 merges into the IH-10 (north leg of the San Antonio downtown loop) and connects the cities of San Antonio and Austin, Texas.

Figure 3*b* shows the specific corridor, which is approximately 1.39 miles and has three lanes with three links. There were four main lane dual loop detectors (stations 159.500 to 160.892) including two on-ramps and one off-ramp on the test bed. Speed, volume, and occupancy data from the loop detectors for the morning peak period (7:00–9:00) for all weekdays between April 1, 2007 and September 30, 2007 were obtained. The data were aggregated at 2-min intervals and, therefore, 7,930 observations were available per link. The link travel times were estimated as a function of the speed, volume, and occupancy using a methodology developed by Vanajakshi et al. (*35*).

Preliminary Data Analyses

Point data such as that from loop detectors have been shown to have a nonstationary distribution and exhibit a general periodicity (*36*). For example, a given link may exhibit a repeating pattern such as a large peak at 8:00 a.m., a smaller peak at around 4:00 p.m., and an absolute minimum around 2:00 p.m. Additionally, the data are also weakly dependent. That is, the dependence diminishes as the aggregation interval becomes large (i.e., traffic characteristics on Monday at 7:30 a.m. influence the traffic characteristics on Monday at 7:35 a.m., but do not influence traffic characteristics on the next Monday at 7:35 a.m.). It was necessary to explore the characteristics of the data—temporally and spatially—to select the most appropriate estimation and prediction models for both the mean and the confidence interval measures.

The preliminary analyses were performed to explore the characteristics of the observed speed, volume, occupancy, and travel time data. Note that, although each data type was explored as an example, only results for the volume data and travel times are presented here.

Periodicity

A periodic behavior (repeating pattern) can be observed when each day's data are plotted side-by-side as shown

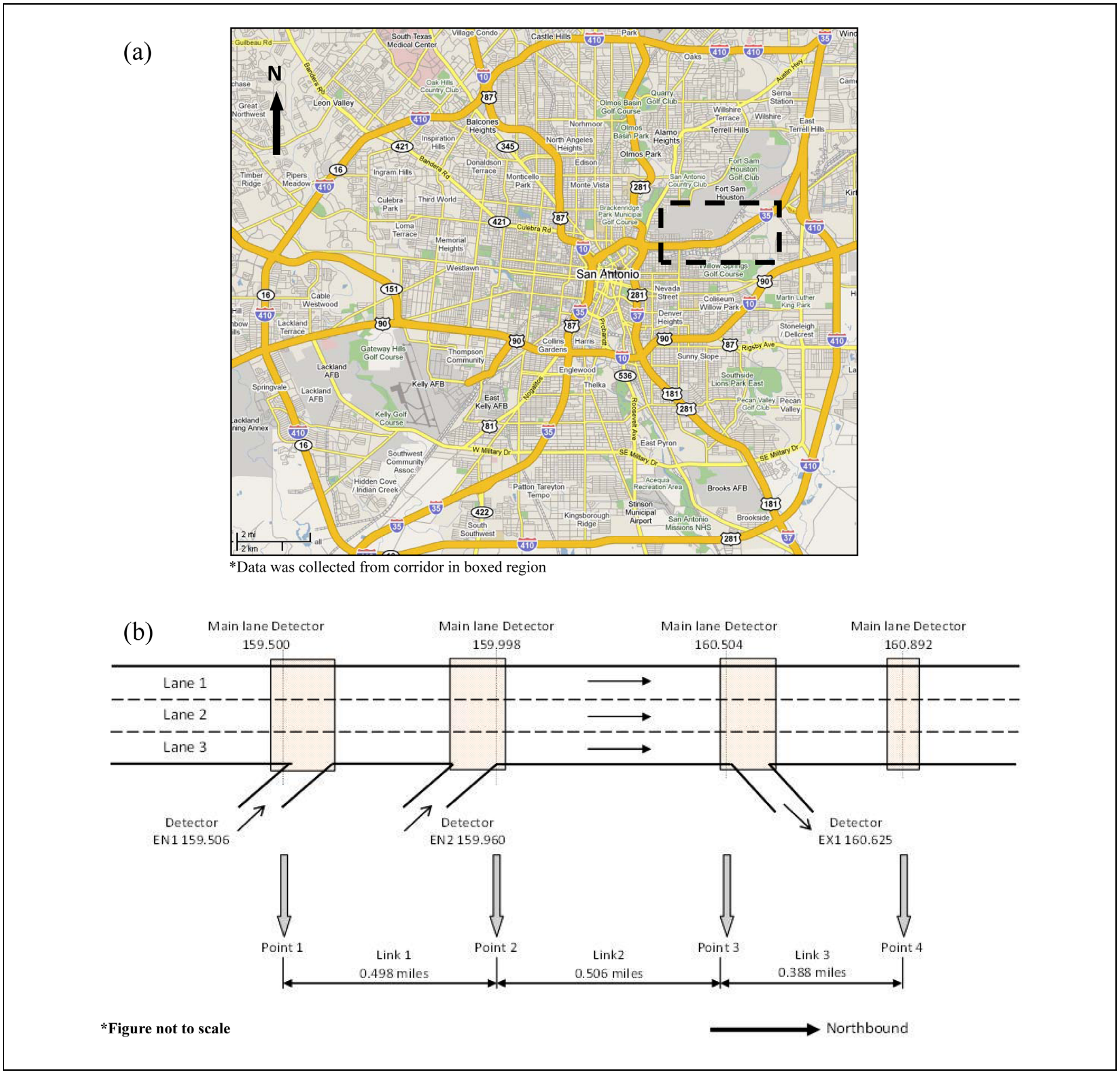

Figure 3. Data collection test bed, showing (*a*) the location on a regional map of San Antonio, Texas, and (*b*) a schematic diagram of the test bed at San Antonio, Texas.

in Figure 4. Figure 4, *a* and *b*, show the 2-min aggregated data for the morning peak (7:00–9:00 a.m.) as a function of the time of the day. A noticeable pattern can be observed. It should be noted that it was difficult to view the entire data (130 days) when plotted at once, so only data for 5 days (305 data points) were plotted as a demonstration example.

Dependence and Nonstationarity

An autocorrelation function, or a correlogram, is a plot of the autocorrelation coefficients against the time lags and these were created for the volume and travel time data as shown in Figure 4, *c* and *d*, respectively. The autocorrelation coefficients measure the correlation between observations at different lags or distances apart (*37*). The autocorrelation coefficient r_k at time lag k is computed (*37*) as

$$r_k = \frac{\sum_{t=1}^{N-k}(x_t - \bar{x})(x_{t+k} - \bar{x})}{\sum_{t=1}^{N}(x_t - \bar{x})^2} \tag{8}$$

where

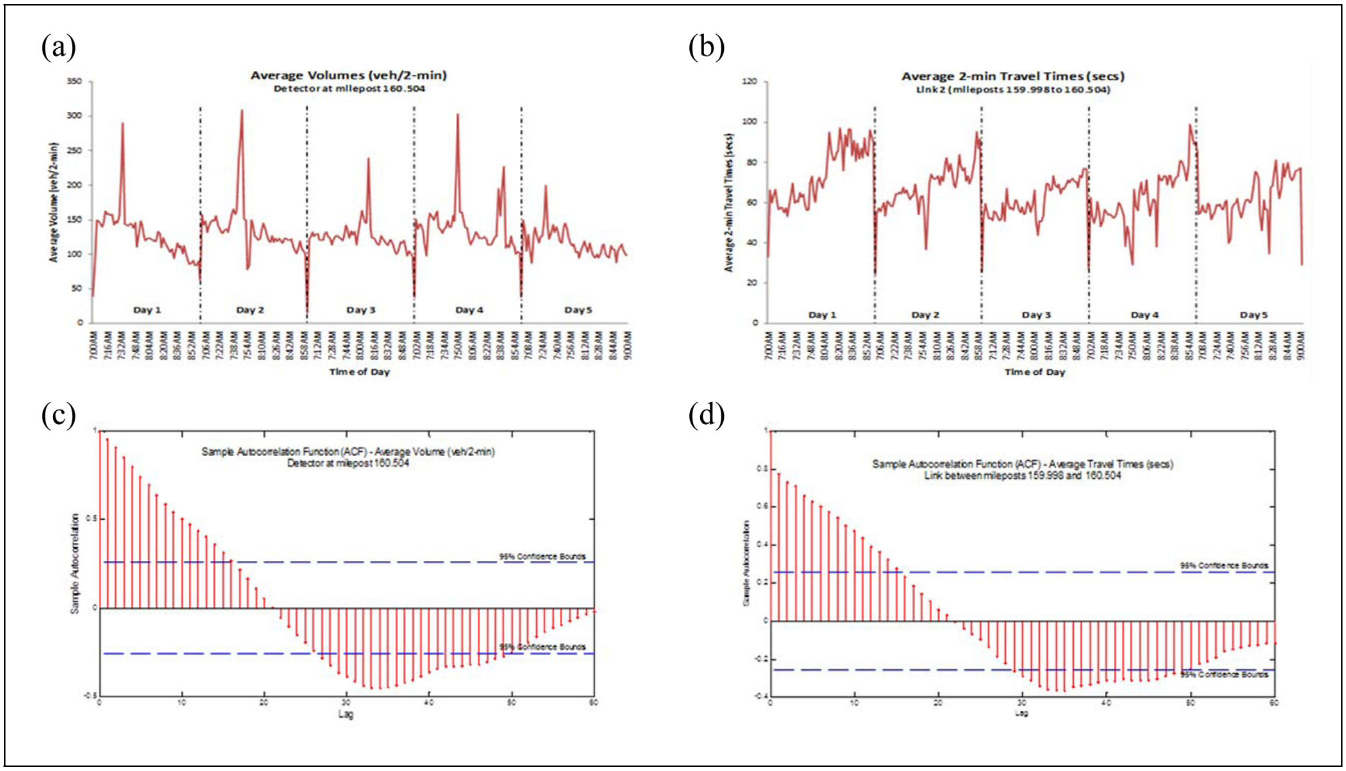

Figure 4. Results of exploratory data analyses, showing (*a*) peak hour traffic volume data, (*b*) peak hour travel time data, (*c*) correlogram of traffic volume data, and (*d*) correlogram of travel time data.

$$\bar{x} = \frac{\sum_{t=1}^{N} x_t}{N} = \text{mean};$$

r_k = correlation coefficient at lag k;
x_t = observation at current time;
x_{t+k} = observation at one lag ahead; and
N = size of series.

In Figure 4, *c* and *d*, the dashed horizontal lines are the confidence bounds: $\pm 1.96\sqrt{n}$ (1.96 is the .975 quantile of the standard normal distribution). The 60-lags represent the 07:00–09:00 a.m. time period with each lag denoting a 2-min interval. The correlogram illustrates that more than 95% of the autocorrelation coefficients are outside the confidence bounds. This indicates that there is time dependence in the data (*38*). That is, data from the current time step are correlated with the data from the previous time step. As an example, travel time at 7:12 a.m. is correlated with travel time at 7:10 a.m. Additionally, the correlation coefficients tend to decay slowly as the time lag increases and are indicative of the data being nonstationary. That is, the statistical characteristics (mean and variance) of the data change over time.

In conclusion, the exploratory analyses suggest that the data (point and interval) exhibit a periodic behavior, are time dependent, and have a nonstationary distribution. Standard historical statistical methods do not apply and any modeling with this data must account for this characteristic behavior. For such data sets, Lahiri et al. found that the gap bootstrap method outperformed other commonly used resampling methods in closeness to the true standard errors based on Monte-Carlo simulation (*32*).

Implementation

Mean travel times were forecasted for fifteen 2-min time periods (30 min ahead) using the methodology shown in Figure 2. An overall mean travel time (at each 2-min interval) as well as its corresponding bootstrap estimate of uncertainty (standard error) was then computed. The overall means and standard error values were computed for the three test days.

Ordinary bootstrap estimates were computed using 250 realizations of the original travel time data that were randomly sampled. The block bootstrap estimates were calculated by dividing the original data into 25 blocks, each consisting of 305 rows (i.e., travel time values for 5 days). Each bootstrap sample was then formed by randomly drawing the blocks with replacement and laying them end-to-end. The overall mean and standard error of the predicted travel times (block bootstrap estimates) were computed as calculated from 250 such samples.

The gap bootstrap estimates were obtained by dividing the original data into 42 subsets of independent data, in which the elements of the *i*th subset are the *i*th 2-min travel times for each of the 125 days in the training dataset. That is, the first subset consists of 7:00 a.m. travel times, the second subset consists of 7:02 a.m. travel times, the third subset consists of 7:04 a.m. travel times, and so on. The predicted overall mean travel times for each 2-min period were obtained by averaging the 2-min predictions from the 42 independent subsets. Component-wise variance estimators of the 42 subsets were calculated by performing ordinary bootstrapping within each subset. These were calculated using 250 realizations of each subset by resampling across the rows of the subsets. Based on the preliminary analysis, it was assumed there is constant covariance among the subsets and the final gap bootstrap uncertainty estimates were computed by combining the component-wise variance estimates with the usual covariance matrix estimator obtained from the subsets.

Results and Analysis

The predicted overall mean travel times and corresponding estimates of the standard error are presented in Table 1. It can be observed that the mean travel time values computed using the three bootstrapping techniques were within 7% of each other. This suggests that the choice of the resampling technique will not affect the predicted point estimate.

However, the estimated standard errors are clearly different. In general, the uncertainty estimates are lower for the ordinary bootstrap than those of the block and gap bootstrap. In addition, the gap bootstrap estimates are the largest. To illustrate these results, the standard error values computed on test day 1 are shown in Figure 5. The block bootstrap uncertainty estimates were on average only 1.2 times larger than the ordinary bootstrap estimates. As expected, the gap bootstrap provided substantially larger estimates of the standard error for the predicted travel times than the ordinary bootstrap (on average 2.7 times) and the block bootstrap (on average 2.3 times). These results are not surprising given the heteroscedastic nature of the data. In this situation, standard techniques would underestimate the confidence interval of the estimate.

It is hypothesized that the more conservative confidence interval would be beneficial to both system operator and system users because it more accurately captures the nature of the data and associated prediction values.

Summary and Conclusions

To a large extent, methods of forecasting travel time have placed emphasis on the *quality* of the forecasted value.

Table 1. Predicted Means and Standard Errors

	Prediction interval (min)														
Method	2	4	6	8	10	12	14	16	18	20	22	24	26	28	30
Test day 1															
Ord	26 (0.616)	26 (0.617)	26 (0.618)	26 (0.615)	26 (0.616)	26 (0.616)	26 (0.615)	26 (0.617)	26 (0.617)	26 (0.616)	26 (0.615)	26 (0.616)	26 (0.615)	26 (0.614)	26 (0.617)
Blk	26 (0.725)	26 (0.754)	26 (0.748)	26 (0.756)	26 (0.727)	26 (0.706)	26 (0.702)	26 (0.701)	26 (0.721)	26 (0.730)	26 (0.714)	26 (0.716)	26 (0.710)	26 (0.728)	26 (0.738)
Gap	28 (2.120)	28 (2.745)	28 (1.614)	27 (2.425)	28 (1.564)	28 (1.314)	28 (1.548)	26 (1.998)	27 (1.461)	27 (1.540)	28 (1.484)	28 (1.690)	27 (1.353)	28 (1.098)	27 (1.258)
Test day 2															
Ord	27 (0.823)	27 (0.824)	27 (0.819)	27 (0.824)	27 (0.824)	27 (0.821)	27 (0.824)	27 (0.823)	27 (0.822)	27 (0.823)	27 (0.822)	27 (0.824)	27 (0.824)	27 (0.825)	27 (0.824)
Blk	27 (0.850)	27 (0.891)	27 (0.870)	27 (0.879)	27 (0.858)	27 (0.846)	27 (0.843)	27 (0.843)	27 (0.848)	27 (0.846)	27 (0.831)	27 (0.833)	27 (0.834)	27 (0.838)	27 (0.843)
Gap	29 (1.753)	28 (1.852)	28 (2.304)	27 (2.245)	27 (1.547)	28 (1.882)	27 (1.783)	26 (1.298)	27 (1.161)	27 (1.704)	27 (1.911)	27 (1.427)	27 (1.489)	27 (1.817)	27 (1.491)
Test day 3															
Ord	25 (1.586)	25 (1.590)	25 (1.590)	25 (1.587)	25 (1.584)	25 (1.590)	25 (1.585)	25 (1.587)	25 (1.584)	25 (1.586)	25 (1.585)	25 (1.588)	25 (1.587)	25 (1.591)	25 (1.587)
Blk	25 (1.550)	25 (1.366)	25 (1.335)	25 (1.338)	25 (1.332)	25 (1.310)	25 (1.305)	25 (1.287)	25 (1.310)	25 (1.307)	25 (1.281)	25 (1.275)	25 (1.264)	25 (1.275)	25 (1.255)
Gap	27 (3.232)	26 (4.723)	28 (4.217)	27 (2.409)	27 (2.726)	26 (2.850)	27 (2.566)	27 (2.829)	27 (2.201)	26 (2.860)	27 (2.720)	27 (3.065)	27 (3.649)	27 (2.812)	27 (4.375)

Note: Bootstrap method: Ord = Ordinary; Blk = Block; and Gap = Gap. Predicted travel times are in seconds. Value in parenthesis is the estimated uncertainty (standard error).

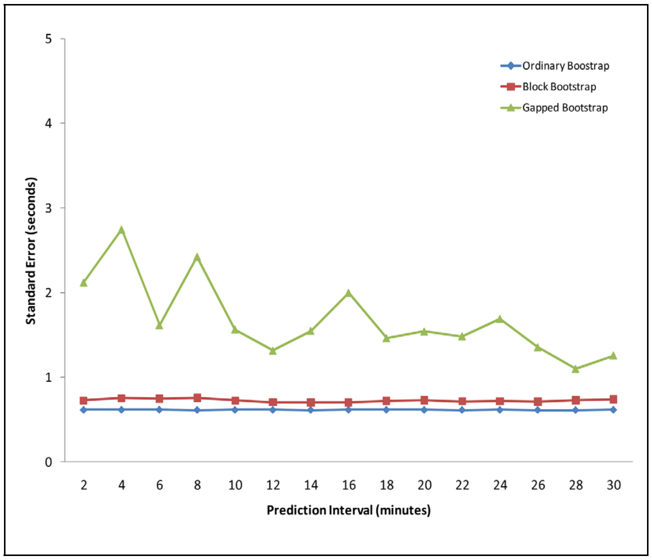

Figure 5. Estimated standard errors for predicted travel time.

However, understanding the reliability or uncertainty margin that exists around the forecasted point estimate is also important, especially understanding that it has been shown to lead end-users to change their routes and schedules even when the average travel time is low (*6*). This paper demonstrates the application of the recently developed gap bootstrap method and compares it against two existing methods (ordinary and block bootstrap) with specific application to link travel time prediction. The method is unique in that, unlike the techniques used in past research, it accounts for datasets that are not independent and identically distributed. These characteristics are common within ITS data which are typically used in the travel time prediction process.

From the results and analysis presented in this paper, it was concluded that uncertainty estimates from the ordinary and block bootstrap methods are similar. In comparison with the ordinary bootstrap, the block bootstrap does not appear to adequately capture the dependence structure (identified in the exploratory analyses) in a dataset. Therefore, although the ordinary bootstrap is not suitable for dependent data, the block bootstrap does not seem to provide much advantage when applied to the dataset used in this study. The estimates from the gap bootstrap are larger than the ordinary and block bootstrap estimates. It is hypothesized that the difference between the gap bootstrap estimate and those from the other two resampling methods are a result of the gap bootstrap explicitly accounting for the dependent structure evident in the dataset. Moreover, unlike the ordinary bootstrap, which is suitable only for data that are independent, it appears the gap bootstrap can adequately address uncertainties for both independent and dependent structured datasets.

Because the true uncertainty estimate of travel times is unknown for the test network, further work using simulation data is required to validate the findings. Such research will allow for comparing uncertainty estimates obtained with the various resampling methods as well as other time series based methods with the true values of the uncertainty based, for example, on Monte-Carlo simulation.

This paper makes a strong case for the gap bootstrap uncertainty estimator given that the existing methods, including the ordinary and block bootstrap, might not be appropriate. The empirical example in this paper, in which the exchangeability of partitions might have been only approximate, led to fairly conservative estimates of uncertainty. In many applications, having a slightly conservative estimate of uncertainties is not a major disadvantage and, in transportation engineering in particular, it is more desirable than the converse.

As indicated earlier, the GARCH approach has also been applied to model uncertainty in the univariate vehicular traffic flow series (*10–12*). The proposed gap bootstrap approach accounts for datasets that are not independent and identically distributed (which is common for ITS data) whereas the GARCH approach accounts for the heteroscedasticity observed in the traffic condition series. The authors suggest further research that would involve a direct comparison of the performance of the proposed gap bootstrap method and the existing GARCH based methods. The comparison should be made under different conditions such as locations, weather, and traffic patterns. An investigation of this kind will provide insights on which is a better method under what set of prevailing conditions.

One potential application of the proposed methodology is in advanced traffic management systems (ATMS) in which, given the number of travel time prediction models available, engineers are faced with the challenge of selecting and implementing the most effective modeling scheme. As a general guide to selecting one of the numerous models for implementation, it would be instructive to have a means of assessing the reliability of these models by being able to accurately quantify the uncertainties in the predictions in terms of variances or confidence intervals but also accounting for the structure of the ITS data which has a dependent structure.

Acknowledgments

The authors would like to thank Dr. Clifford Spiegelman and Dr. Soumen Lahiri for providing background information and technical expertise on the technique used in this research. The authors are also grateful to the Texas Department of Transportation for the use of the Transguide data.

Author Contributions

The authors confirm contribution to the paper as follows: study conception and design: Naik 50%, Rilett 50%; data collection: Naik 50%, Appiah 50%; analysis and interpretation of results: Naik 25%, Rilett 25%, Appiah 25%, Walubita 25%; draft manuscript preparation: Naik 25%, Rilett 25%, Appiah 25%, Walubita 25%. All authors reviewed the results and approved the final version of the manuscript.

References

1. Mori, U., A. Mendiburu, M. Álvarez, and J. A. Lozano. A Review of Travel Time Estimation and Forecasting for Advanced Traveller Information Systems. *Transportmetrica A: Transport Science*, Vol. 11, No. 2, 2015, pp. 119–157.
2. Du, L., S. Peeta, and Y. H Kim. An Adaptive Information Fusion Model to Predict the Short-Term Link Travel Time Distribution in Dynamic Traffic Networks. *Transportation Research Part B*, Vol. 46, No. 1, 2012, pp. 235–252.
3. Mazloumi, E., G. Currie, G. Rose, and M. Sarvi. Statistical Confidence Estimation Measures for Artificial Neural Networks: Application in Bus Travel Time Prediction. Presented at 89th Annual Meeting of the Transportation Research Board, Washington, D.C., 2010.
4. Mahmassani, H. S., and Y. H. Liu. Dynamics of Commuting Decision Behavior Under Advanced Traveler Information Systems. *Transportation Research Part C*, Vol. 7C, 1999, pp. 91–107.
5. Van Berkum, E. C., and P. H. J. Van Der Mede. *The Impact of Traffic Information: Dynamics in Route and Departure Time Choice*. Delft University Press, Delft, the Netherlands, 1993.
6. Wang, X., and A. Khattak. Role of Travel Information in Supporting Travel Decision Adaption: Exploring Spatial Patterns. *Transportmetrica A: Transport Science*, Vol. 9, No. 4, 2013, pp. 316–334.
7. Yang, H. *An Analysis of Route Choice Behavior under Advanced Traveler Information Systems: A Neural Network Approach*. ITS, UC Davis, 1992.
8. Eisele, W. L., B. Naik, and L. R. Rilett. Estimating Route Travel Time Reliability from Simultaneously Collected Link and Route Vehicle Probe Data and Roadway Sensor Data. *International Journal of Urban Sciences*, Vol. 19, No. 3, 2015, pp. 286–304.
9. Van Lint, J. W. C. *Reliable Travel Time Prediction for Freeways*. PhD dissertation. The Netherlands TRAIL Research School, 2004.
10. Guo, J., B. M. Williams, and B. L. Smith. Data Collection Time Intervals for Stochastic Short-Term Traffic Flow Forecasting. *Transportation Research Record: Journal of the Transportation Research Board*, 2008. 2024: 18–26.
11. Guo, J., and B. M. Williams. Real Time Short Term Traffic Speed Level Forecasting and Uncertainty Quantification Using Layered Kalman Filters. *Transportation Research Record: Journal of the Transportation Research Board*, 2010. 2175: 28–37.
12. Guo, J., W. Huang, and B. M. Williams. Real Time Traffic Flow Outlier Detection Using Short-Term Traffic Conditional Variance Prediction. *Transportation Research Part C*, Vol. 50, 2015, pp. 160–172.
13. Kamarianakis, Y., A. Kanas, and P. Prastacos. Modeling Traffic Volatility Dynamics in an Urban Network. *Transportation Research Record: Journal of the Transportation Research Board*, 2005. 1923: 18–27.
14. Sohn, K., and D. Kim. 2009. Statistical Model for Forecasting Link Travel Time Variability. *Journal of Transportation Engineering*, Vol. 135, No. 7, pp. 440–453.
15. Ellison, S. L. R., M. Rosslein, and B. Williams. *Quantifying Uncertainty in Analytical Measurement*, 2nd ed. EURACHEM/CITAC Guide, 2000.
16. Robinson, S., and J. W. Polak. ILD Data Cleaning Treatments and their Effect on the Performance of Urban Link Travel Time Models. Presented at 85th Annual Meeting of the Transportation Research Board, Washington, D.C., 2006.
17. Vanajakshi, L. D., and L. R. Rilett. Loop Detector Data Diagnostics Based on Conservation-of-Vehicles Principle. *Transportation Research Record: Journal of the Transportation Research Board*, 2004. 1870: 162–169.
18. Zapranis, A., and E. Livanis. Prediction Intervals for Neural Network Models. *Proc., WSEAS International Conference on Engineering Education*, Greece, 2005.
19. Bishop, C. M. *Neural Networks for Pattern Recognition*. Clarendon Press, Oxford, 1995.
20. Nix, D. A., and A. S. Weigend. Learning Local Error Bars for Nonlinear Regression. *Proc., Advances in Neural Information Processing Systems (NIPS 7)*, MIT Press, Cambridge, 1995.
21. Singh, K., and M. Xie. *Bootstrap: A Statistical Method*. Unpublished manuscript, Rutgers University, USA, 2008. http://www.stat.rutgers.edu/home/mxie/RCPapers/bootstrap.pdf.
22. Heskes, T. Practical Confidence and Prediction Intervals. In *Advances in Neural Information Processing Systems* (Mozer, M., M Jordan, and T. Petsche, eds.), MIT Press, Cambridge, Vol. 9, 1997, pp. 176–182.
23. Chernick, M. R. *Bootstrap Methods: A Practitioner's Guide*. John Wiley & Sons Inc., New York, NY, 1999.
24. Davison, A. C., and D. V. Hinkley. *Bootstrap Methods and their Applications*. Cambridge University Press, New York, NY, 1997.
25. Tibshirani, R. A Comparison of Some Error Estimates for Neural Network Models. *Neural Computation*, Vol. 8, 1996, pp. 152–163.
26. Hardle, W., J. Horowitz, and J. Kriess. Bootstrap Methods for Time Series. *International Statistics Review*, Vol. 71, No. 2, 2003, pp. 435–459.
27. Kunsch, H. R. The Jackknife and the Bootstrap for General Stationary Observations. *Annals of Statistics*, Vol. 17, 1989, pp. 1217–1261.
28. Carlstein, E. The Use of Subseries Methods for Estimating the Variance of a General Statistic From a Stationary Time Series. *Annals of Statistics*, Vol. 14, 1986, pp. 1171–1179.
29. Politis, D., and J. P. Romano. A Circular Block Resampling Procedure for Stationary Data. In *Exploring the*

Limits of Bootstrap (Lepage, R., and L. Billard, eds.), Wiley, New York, 1992, pp. 263–270.
30. Politis, D., and J. P. Romano. The Stationary Bootstrap. *Journal of the American Statistical Association*, Vol. 89, 1994, pp. 1303–1313.
31. Lahiri, S. N. Theoretical Comparison of Block Bootstrap Methods. *Annals of Statistics*, Vol. 27, 1999, pp. 386–404.
32. Lahiri, S. N., C. Spiegelman, J. Appiah, and L. Rilett. Gap Bootstrap Methods for Massive Data Sets with an Application to Transportation Engineering. *Annals of Statistics*, Vol. 6, No. 4, 2012, pp. 1552–1587.
33. Appiah, J. *Quantifying Uncertainties in Synthetic Origin-Destination Trip Matrix Estimates*. PhD dissertation. University of Nebraska-Lincoln, 2009.
34. Naik, B. *A Generalized Nonparametric Approach for Uncertainty Evaluation in Travel Time Prediction Models*. PhD dissertation. University of Nebraska-Lincoln, 2010.
35. Vanajakshi, L. D., B. M. Williams, and L. R. Rilett. Improved Flow-Based Travel Time Estimation Method from Point Detector Data for Freeways. *Journal of Transportation Engineering*, Vol. 135, No. 1, 2009, pp. 26–36.
36. May, A. D. *Traffic Flow Fundamentals*. Prentice Hall, Upper Saddle River, NJ, 1990.
37. Chatfield, C. *The Analysis of Time Series: An Introduction*. Chapman and Hall/CRC, Boca Raton, FL, 2003.
38. Brockwell, P. J., and R. A. Davis. *Introduction to Time Series and Forecasting*, 2nd ed. Springer-Verlag, NY Inc., 2002.

The Standing Committee on Urban Transportation Data and Information Systems (ABJ30) peer-reviewed this paper (18-03453).

The contents of this paper reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented here.