Component GARCH Models to Account for Seasonal Patterns and Uncertainties in Travel-Time Prediction

Yanru Zhang, Ali Haghani, and Xiaosi Zeng

Abstract—Uncertainty is often associated with travel-time prediction. Traditional point prediction methods only provide point values that are unable to offer enough information on the reliability of prediction results. The recent development of statistical volatility models has given us an effective way to capture uncertainties in data. Generalized autoregressive conditional heteroskedasticity (GARCH) models have been widely used in transportation systems as a way to account for this uncertainty by providing more accurate prediction intervals. However, a GARCH model argues does not consider the trend and seasonality in data. If there is a trend or seasonality, the performance of the GARCH model may be affected. In the context of travel-time prediction, this paper proposes two component GARCH models that are able to model trend and seasonal components through decomposition. The travel-time data obtained along a freeway corridor in Houston, TX, USA, were used to empirically test the performance of the proposed models. The study results indicate that the proposed models perform well when capturing uncertainties associated with travel-time prediction.

Index Terms—Component generalized autoregressive conditional heteroskedasticity (C-GARCH), prediction intervals (PIs), seasonality, time-series analysis, travel-time prediction, uncertainty.

I. INTRODUCTION

A

S an effective and widely acknowledged way to measure freeway traffic conditions, short-term travel-time forecasting is essential for the success of advanced traveler information systems and advanced traffic management systems. Accurate and timely forecasting can provide guidance for travelers’ decisions for departure time, route, and mode choice, and it can assist transportation agencies in proactive traffic control. Because of its importance, travel-time forecasting has generated great interest among researchers, and several forecasting algorithms have been proposed in literature, some of which include traditional statistical approaches [2], [3], neural networks [4]–[6], support vector machines [7], [8], and simulation-based models [9]. For more detailed information and extensive reviews of existing travel-time prediction models, see [10]. Within these different methods, autoregressive integrated moving average (ARIMA)-type models are one popular approach due to their well-defined theoretical foundation and their ease of estimation [11].

Although a wide range of technologies has been applied to this field and has shown promising abilities in predicting travel time, some are limited in their ability to capture the uncertainty and variability of travel time, as they only provide point values to represent future traffic conditions. A traffic condition is a complex phenomenon. It is often affected by interactions among different vehicles and by exogenous factors, such as traffic incidents, weather, and roadway conditions. Small changes in current traffic conditions may greatly affect future travel time. Due to the highly dynamic nature of traffic, predicting travel time is often associated with uncertainty, particularly during nonrecurrent congestion when incidents or bad weather occurs. A point value provides limited information regarding the uncertainty and unreliability of travel time. On the other hand, the prediction intervals (PIs) approach has the potential to provide more reliable forecasting results by providing a confidence band to indicate how reliable the forecasting results are. Khosravi et al. [12], [13] developed different approaches on the bases of a neural network to provide PIs to capture uncertainties in travel time. Their study indicates that there is always a mismatch between prediction and true values, but the constructed PIs capture most of the true values. Zeng and Zhang [5] employed an ensemble method using multiple instances of the same neural network model with different initial conditions to derive a prediction band. Fei et al. [3] proposed a Bayesian-inference-based dynamic linear model. Van Hinsbergen et al. [14] proposed an approach that combines neural networks in a committee using a Bayesian inference theory to predict travel time with confidence intervals. Li and Rose [15] developed a model that separately formulates the average travel time and the travel-time variability to incorporate uncertainty in travel-time prediction. All these studies indicate that travel-time prediction is a complex problem that is often associated with uncertainty.

Approaches based on PIs provide the potential to capture the dynamic changes in traffic.

In addition to the aforementioned methodologies, another way to capture the uncertainty and variations of data is the statistical volatility approach. Different from the neural-network-based approach, the statistical-volatility-based approach gives exact functions to model the evolution of the data series; therefore, it is easier to be interpreted. It relaxes the constant variance assumption and modeling time changing variances as
a function of its past values. As a result, the statistical volatility approach can capture the dynamic changes in travel time, and it provides more accurate PI's. The first volatility model, i.e., the autoregressive conditional heteroskedasticity (ARCH) model, was proposed by Engle in 1982 for financial analysis purposes [16]. Later, different variations of the ARCH model were formulated. The generalized ARCH (GARCH) [17] model has been one of the most widely used models. Driven by its successful applications in other areas [18], transportation professionals began to apply the GARCH family models to estimate traffic volatilities. Karlaftis and Vlahogianni [11] applied the ARFIMA–GARCH model by using the traffic-flow data in an urban network, ultimately demonstrating that the traffic-flow data displayed time-dependent volatilities. Later, Tsikeris and Statopoulos [20] incorporated fractionally integrated components in both the conditional mean and the conditional variance equations, and they proposed the autoregressive fractionally integrated moving average–fractionally integrated asymmetric power ARCH (ARFIMA–FIAPARCH) model. Similarly, Karlaftis and Vlahogianni [11] applied the ARFIMA–fractionally integrated GARCH (FIGARCH) model for traffic-flow prediction. Tsikeris and Statopoulos [21] predicted urban traffic variability through a stochastic volatility modeling approach. Other variations of a volatility model include a GARCH model with Kalman filtering [22], a GARCH model with different mean equations [23], a multivariate GARCH model [24], and asymmetric GARCH models [25], [26].

Nevertheless, these existing volatility models do not consider the possible cyclical patterns in the residual series, which are often referred to as a seasonal component. Conventional GARCH models are often criticized as unsatisfactorily modeling data series that show pronounced seasonal patterns [27]. Decomposition technologies provide the potential to deal with the trend and seasonal components in the data. Driven by the successful application of statistical volatility models in transportation analyses, this paper developed two different component-volatility-based travel-time prediction models to better characterize long-term and short-term volatility and cyclical patterns in travel-time data. Both models are similar to the structure of the GARCH model but include trend and seasonal elements. The proposed methods allow for a more versatile structure with the potential to provide more accurate traffic volatility forecasting along freeway corridors during peak hours. The rest of this paper is organized as follows. Section II provides the theoretical background of the mean equation, three different volatility models, the PI construction, and the comparison method. Section III describes the data and explains the prediction procedure. Section IV comprehensively evaluates the performance of the proposed methods. Section V summarizes and concludes this paper.

II. THEORY AND BACKGROUND

The observed travel time can be decomposed into a conditional mean \( u_t \) and a residual \( r_t \). The traditional time-series-based travel-time prediction methods only model the time variation of the data in the first-order moment and predict the mean part \( u_t \) while assuming a constant variance of the data \( r_t \) as constant across different time intervals. However, uncertainty often exists, particularly in the traffic field. Unexpected external factors, such as traffic incidents, work zones, weather, and special events, can dramatically affect the travel time. The reliability of the point prediction results would dramatically drop because of the presence of these unexpected factors. Modeling these uncertainties, e.g., see the conditional standard deviation (or residual \( r_t \)), would provide more reliable forecasting results. One prominent tool that characterizes the changing variance through time belongs to the statistical volatility model, which models the time variation of the data in second-order moments as

\[
x_t = u_t + r_t
\]

where \( x_t \) is the observed travel time at time \( t \), \( u_t \) represents the estimated conditional mean, and \( r_t \) is the residual part.

A. ARIMA Model

This paper chose the ARIMA model to represent the mean part \( u_t \) due to its ease of implementation and well-established theoretical foundation. The ARIMA \((p, d, q)\) model is comprised by three parts, i.e., the autoregressive (AR, \( p \)), integrated (I, \( d \)), and moving average (MA, \( q \)) parts. We define \( B \) as the backshift operator with \( B^k x_t = x_{t-k} \). An ARIMA \((p, d, q)\) model with \( p \) as the number of autoregressive terms, \( d \) as the number of differences, and \( q \) as the number of lagged forecast errors, for a given time series \( \{x_1, x_2, \ldots, x_n\} \), are of the following form:

\[
(1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p)(1 - B)^d x_t = (1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q) z_t
\]

where \( \phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_q \) are parameters, and \( z_t \) is the white noise process with zero mean and variance \( \sigma^2 \), \( \{z_t\} \sim WN(0, \sigma^2) \).

For the theoretical foundations of the ARIMA model, see [28] for details.

B. GARCH Type

The GARCH-type model aims at capturing the changes in the variance part \( r_t \). The first volatility model was proposed by Engle [16] in 1982 and was termed the ARCH model. In his paper, the discrete-time stochastic process \( r_{t,i} \) is expressed as [29]

\[
r_t = \sigma_t \epsilon_t
\]

\[
\sigma_t^2 = \text{Var}(r_t | F_{t-1}) = \text{Var}(x_t | F_{t-1})
\]

where \( \epsilon_t \) is an independent and identically distributed (i.i.d) process with a zero mean and one standard deviation, and \( F_{t-1} \) denotes the information available at time \( t-1 \). The aforementioned equation forms the foundation of the volatility model. Its various extensions are all based on this equation. Different ways of modeling \( \sigma_t \) lead to a wide variety of volatility models.
Engle suggested in his paper that $\sigma_t^2$ can be a linear function of the past squared values of process $r_t$ as

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{m} \alpha_i r_{t-i}^2$$

where $\alpha_0$ is the intercept term, $\alpha_0 > 0$, $\alpha_i$ represents the unknown coefficient of $r_{t-i}$ that satisfies $\alpha_i \geq 0$ to ensure that the conditional variance is positive, and $m$ denotes the number of lags selected for the model. This structure clearly captures the cluster of volatilities. The current magnitude of $\sigma_t^2$ is the transitory component that models short-term volatilities. The C-GARCH model considers both the long-term and short-term variations of the data series. It applies in situations when there is an obvious trend in the data.

The multiplicative C-GARCH (MC-GARCH) model [27] assumes that variations increase with the level of data and decomposes the variance part into three multiplicative components, i.e., daily component $d_t$, deterministic diurnal pattern $s_t$, and stochastic intraday component $q_{t,i}$ as

$$x_{t,i} = u_{t,i} + r_{t,i}$$
$$r_{t,i} = \sqrt{d_t s_t q_{t,i} e_{t,i}}$$

where the travel time at time index $i$ in day $t$ is composed of the conditional mean $u_{t,i}$ and variance $r_{t,i}$, $e_{t,i}$ is the i.i.d $(0, 1)$ standardized innovation that can follow a normal Student’s $t$ distribution. The daily part $d_t$ models the variance of the data across different days. It can be derived from a multifactor risk model, a daily GARCH model, or a multiple indicator model [27]. The deterministic diurnal component for each time index is estimated as

$$\hat{s}_i = \frac{1}{T} \sum_{t=1}^{T} \frac{r_{t,i}^2}{d_t}$$

where $T$ is the total number of days, $t$ denotes the day, and $i$ denotes the regularly spaced time intervals.

As it is indicated in (11), the diurnal component at time index $i$ is the average of the variance at time index $i$ scaled by its corresponding variance for each day. Therefore, the diurnal component represents the regular intraday variations. After estimating the daily and deterministic intraday components, the rest of the component in the variance part is regarded as stochastic, which can be regarded as a GARCH $(p, q)$ process. The normalized residual is

$$z_{t,i} = r_{t,i} / \sqrt{d_t \hat{s}_i} = \sqrt{q_{t,i} e_{t,i}}$$

where the stochastic intraday component $q_{t,i}$ is assumed to follow the GARCH process as

$$q_{t,i} = \alpha_0 + \sum_{j=1}^{p} \alpha_j z_{t,i-j}^2 + \sum_{j=1}^{q} \beta_j q_{t,i-j}.$$  

From the perspective of travel-time prediction, the travel time exhibits both regular cyclical patterns (the seasonal component) and stochastic patterns. Daily cyclical patterns distinguish the travel time as peak-hour and nonpeak-hour traffic. Stochastic patterns are the results of unexpected influential events, such as bad weather conditions and traffic incidents. Capturing the time-varying features of traffic behavior is critical for travel-time forecasting. In addition, decomposing the data into cyclical and stochastic patterns provides a better understanding of the underlying structure of the data. The MC-GARCH is able to estimate the different structures of individual components and provides a more accurate estimation by treating these components separately. This model works best if there is seasonality in
the data, such as seasonality in the travel time observed during peak hours.

D. PI

PIs are one critical output of the volatility-based travel-time prediction model. This section briefly reviews the concept, calculation, and evaluation techniques for PIs.

A PI is an estimated range that captures the future observation, with a prescribed probability, given the current available observations. As illustrated in Fig. 1, a PI is comprised of upper and lower prediction limits that indicate the accuracy of the model output with respect to the observed value. Due to the uncertainty related with the data and the estimated model, there is often a mismatch between the model output and the observed value. PIs provide a range to predict how likely the travel time is during the next time interval. Therefore, PIs have the potential to capture the fluctuations and the stochastic traffic phenomena. Usually, it results in a wider PI associated with larger variations of the travel time. This indicates that there is a high likelihood that the predicted value may be further away from the true value.

The PIs for the volatility model are estimated based on the idea of PIs for regression models. To construct a PI with 100(1 − \( \alpha \))% confidence, we assume that the error follows a Gaussian distribution with a zero mean and a variance \( \sigma_t^2 \). The PIs can be calculated as

\[
(u_t - z_{\alpha/2} \sigma_t, u_t + z_{\alpha/2} \sigma_t)
\]

where \( u_t \) is the predicted mean, \( z_{\alpha/2} \) denotes the standard score that corresponds to the cumulative probability level of \( \alpha/2 \), and \( \sigma_t \) is the predicted variance from a volatility model.

As the concept of uncertainty or reliability is a relatively new area in traffic forecasting, there are few studies that provide criteria for PI assessment. One study by Khosravi et al. [12] suggested that two important aspects of PI assessment should be considered, i.e., the coverage probability and the length. The coverage probability measures the percentage of the targets that lies within the predicted PIs. It measures how effective the constructed PIs are. The mathematical representation of the PI coverage probability (PICP) is

\[
\text{PICP} = \frac{1}{n} \sum_{i=1}^{n} c_i
\]

where \( c_i = 1 \) if \( y_i \in [L(x_i), U(x_i)] \); otherwise, \( c_i = 0 \). \( L(x_i) \) and \( U(x_i) \) represent the upper and lower bounds of the PI of \( x_i \), and \( n \) is the total number of constructed PIs.

On the other hand, another criterion called the mean PI length (MPIL) measures the average length of the PIs. It measures how efficient the constructed PIs are. Assume that we have two models that provide PIs with the same coverage probability; the model that gives a narrower prediction band is more efficient. The following equation gives the definition of the MPIL:

\[
\text{MPIL} = \frac{1}{n} \sum_{i=1}^{n} (U(x_i) - L(x_i)).
\]

Therefore, both criteria should be considered when evaluating the volatility models.

III. CASE STUDY AND EXPERIMENT PROCEDURE

A. Data Description

The performance of the GARCH, C-GARCH, and MGARCH models are investigated here by using the data collected from automatic vehicle identification (AVI) stations located along U.S. Highway 290 (or U.S. 290) in Houston, TX, USA. The entire study corridor is about 5 mi long and covers the Northwest Freeway in the westbound direction between I-610 and the junction of Farm to Market Road 1960 (FM1960). The IDs of the selected AVI stations are 29, 30, 31, 32, 33, and 34. The travel times between each pair of consecutive detectors were collected and aggregated into 5-min time intervals.

The individual segment travel time at free-flow conditions is less than 4 min. The total period of the sample was the entire year of 2008 with the missing data replaced by the annual medians of the missing intervals. Since the travel-time patterns during weekdays and weekends are quite different, the weekend data were excluded from the sample. As a result, 262 weekdays of the travel-time data that contain 75 456 5-min observations are selected for this paper.

B. Modeling Conditional Mean

The first step of the modeling stage was to estimate the mean of the data. In the literature of traffic-parameter forecasting, different mean equation models have been tested. For example, Kamarianakis et al. [19] applied the ARIMA model as the mean equation of the volatility model. Karlaftis and Vlahogianni [11] proposed using the ARFIMA model to capture the long memory in the conditional mean. Yang et al. [23] adopted three different mean equations for volatility models, i.e., the seasonal ARIMA, artificial neural network, and historical average methods. Among these existing methods that have been proposed in literature, the ARIMA-type model becomes one of the most widely used methods due to its ease of implementation and its well-known ability in traffic-parameter modeling and forecasting [11]. Therefore, this paper applies the ARIMA model as the mean equation. However, it is worth noting that a proper mean equation model should not be restricted to the ARIMA-type model. Properly choosing a mean equation regarding the structure of the data can lead to better model performance [23].
Order selection and parameter estimation are two major steps for the ARIMA-model forecasting. The order selection process of the ARIMA model is often considered subjective and difficult to apply. This paper utilizes the method proposed by Hyndman and Khandakar [32] to select the orders of the appropriate ARIMA model automatically. The parameters of the ARIMA model are estimated based on the maximum likelihood method.

C. Characteristics of the Residuals Series

Traffic data often show periodic patterns. The travel time increases and significantly varies during peak hours compared with the travel time during nonpeak hours. It is difficult to precisely predict traffic when congestion occurs. Point prediction methods are often unable to capture the traffic variation in congested situations, therefore providing less reliable or less accurate prediction. As the performance of the point prediction methods often decreases when congestion occurs, it is expected that the residual series (after removing the predicted mean by the ARIMA model) show higher variations during peak hours. Fig. 2 provides a box plot of absolute deviation from the predicted mean for each 20-min time interval (outliers have been removed from this plot). Each box statistically represents the day-to-day variations of the residual series. The green line indicates the mean of the residual, and the lower and upper boundaries of each box are the 25th and 75th percentiles of the data for the corresponding time intervals. The entire plot gives interval-to-interval variations. As observed in this plot, the statistics of each interval are different from each other. Both the mean and percentiles of the data are different at different time intervals. This indicates that the residual series vary over time, and the constant variance assumption of the traditional time-series models is violated. This further proves that a volatility model, which released the constant variance assumption, is necessary. In addition, there is a pronounced increased variation at the beginning of 15:00; subsequently, the variation decreases at 19:00. Comparatively, the variations during other time periods (nonpeak hours) are less significant. The other four studied segments also detected this diurnal pattern. It is clear that seasonal components exist in the residual time series. In addition, the mean of the absolute deviations during nonpeak hours is close to zero, which means that the ARIMA model provides adequate prediction during nonpeak hours. On the other hand, the prediction performance of the ARIMA model decreases during peak hours, as both the mean and the 75% statistics increases.

D. Testing the ARCH Effect

The basic assumption of the GARCH-type model is that the square values of the residuals are correlated. Therefore, before applying GARCH-type models to the data, there is a need to test if the data meet this assumption. Two tests are available, i.e., the Ljung–Box statistics and the Lagrange multiplier test [33]. This paper chose the Ljung–Box statistics to test if the first lags of the squared residuals are uncorrelated. The Ljung–Box test is

$$H_0 : \rho_1 = \rho_2 = \cdots = \rho_m = 0$$

$$Q(m) = N(N + 2) \sum_{h=1}^{m} \frac{\rho_h^2}{N - h}$$

where $N$ is the number of data points under study, $\rho_h$ is the sample autocorrelation at lag $h$, and $m$ is the number of lags being tested.

The critical region for rejecting the null hypothesis at significance level $\alpha$ is

$$Q > \chi^2_{1-\alpha, m}.$$  

In terms of $p$ values, the null hypothesis will be rejected if the $p$ value is less than $\alpha$. In our paper, the Ljung–Box test is applied to the residual data of all five segments. The $p$ values of the test for all studied segments are significantly less than 0.01. Therefore, the null hypothesis is rejected at the significance level of 0.01. That is, correlations exist between the squared values of residuals. The GARCH-type models are necessary.

E. Estimating the Volatility Model

Similar to the ARIMA-type model, estimating the volatility model also involves order selection and parameter estimation. Several studies indicate that the GARCH family model with an order of $(1, 1)$ was found adequate in representing the volatility dynamics [11], [18], [20]. Therefore, the GARCH, C-GARCH, and MC-GARCH models with an order of $(1, 1)$ were adopted for ease of implementation and comparison by using the R package “rugarch” [34].

Since the MC-GARCH model decomposes the data into a daily component, a deterministic diurnal pattern, and a stochastic intraday component, the first step is to model the daily component. This paper terms the average volatility for each day as the daily component. The daily component is estimated through the standard GARCH $(1, 1)$ process. There are 262 daily data in total; the first 242 data points are used as the training data. After removing the daily component, the deterministic diurnal part is estimated as the annual average of the residual data at each time interval. The normalized residuals (12) are...
Fig. 3. MC-GARCH forecasting results: the decomposition of the volatility into its various components (32–33).

Fig. 4. Predicted mean and PI for the MC-GARCH model.

then used to produce the stochastic intraday component. The volatility components estimated by the MC-GARCH model are displayed in five panels, as shown in Fig. 3. The top panel shows the true values of the residual data series. The second panel gives the estimated conditional variance, being the product of the following three components, i.e., the deterministic intraday (panel three), daily (panel four), and stochastic intraday (panel five) components. As indicated in this figure, the MC-GARCH model is able to model the trend, seasonal, and stochastic components of the data. This feature provides a better understanding of the basic structure of the data and is easy to interpret. For example, the Intraday [Deterministic] components indicate the regular cyclical patterns of travel-time volatility. The uncertainties associated with travel-time prediction are higher during peak hours. Therefore, there are a small peak and a large peak of intraday volatility each day. The Intraday [Stochastic] components specify the daily variations due to the demand variation, an incident, or other abnormal traffic phenomena. As indicated in the lower panel in Fig. 3, the stochastic intraday components are more daily specific because traffic conditions can be different each day. It is evident that each component of the data series has its unique characteristics and structure. Separately modeling them potentially improves the prediction accuracy.

F. Construction of the Mean and PIs

The final output of the proposed volatility models includes two measures, i.e., the predicted mean and the predicted PIs. The predicted mean part generally tells the expected value of the travel time in the future, whereas the PIs tell how likely the true value will lie within a certain range. In other words, wider PIs often indicate unreliable travel time and prediction. Thus, based on the combined information of the predicted mean and PIs, travelers and operators would have a better sense of future traffic conditions. In this paper, the ARIMA model provides the mean values, and the PIs are constructed according to the method aforementioned in Section II. Fig. 4 plots some sample prediction results of the MC-GARCH model. The blue dot stands for the true travel time, and the red triangle stands for the predicted mean. The green lines represent the PIs constructed by the MC-GARCH model. It is obvious in this figure that there is always a mismatch between the predicted mean and the true value. This partly results from the dynamic nature of traffic, i.e., the travel time varies from time to time. The PIs, on the other hand, are able to adequately capture this variation by covering most of the true values. Therefore, this model provides an effective and efficient way to measure the uncertainty associated with the future travel time.

IV. RESULTS AND DISCUSSION

This section comprehensively evaluates the effectiveness and efficiency of the constructed PIs. The PIs with different confidence levels are constructed through the GARCH, C-GARCH, and MC-GARCH models for the five studied segments. Our experiment involves two steps. For the first step, we estimate all three models and evaluate the effectiveness of each model based on the criteria of coverage probability and PI length. For each of the five studied segments, 30 days of 5-min travel-time data with 8640 observations are used as the training data set; ten days travel-time data with 2880 observations are used as the comparison (testing) data set. We estimate the individual model for each of the five segments. PIs are constructed with 95%, 90%, and 85% confidence levels, respectively.

Table I provides the average MPIL and PICP values of the three models with different confidence levels during peak hours, nonpeak hours, and all day. Both the MPIL and PICP values are calculated based on (15) and (16). The PICP criterion measures the real coverage for the constructed PIs. The 95%, 90%, and 85% confidence levels are the theoretical coverage for
the constructed PIs. Therefore, the higher the PICP value, or the closer the PICP to the confidence level, the better. It is expected that if we expand the PIs to make them wide enough to cover all the true values, the constructed PIs will not provide any useful information. Therefore, we use another criterion, i.e., the MPIL, that measures the length of the contracted PIs. A good PI should have a lower value of MPIL and a higher value of PICP. As shown in this table, during peak hours, the PI coverage rates of the MC-GARCH model are the highest compared with those of the C-GARCH and GARCH models. As the confidence level decreases, the advantage of the MC-GARCH model becomes obvious, proving that the MC-GARCH model is able to capture the volatility of traffic data during peak hours. In terms of the MPIL, the C-GARCH is the smallest. However, on average, the C-GARCH model only reduced the length by 0.75 compared with the GARCH model. During nonpeak hours, the MC-GARCH model also provides the highest coverage. However, the advantage of the MC-GARCH model is not very obvious compared with the GARCH and C-GARCH models in terms of either the MPIL or the PICP during nonpeak hours. This is expected as the travel time is relatively stable with small variations and as the trend and seasonal patterns are not obvious during this period. Therefore, the performance of these three models should be similar during nonpeak hours. Investigating the all-day performance of these three models indicates that the MC-GARCH model provides the highest PICP value, whereas both the C-GARCH and GARCH models give lower MPIL values compared with the MC-GARCH model. In general, based on the estimation results, we can conclude that the MC-GARCH model tends to cover more targets compared with the C-GARCH and GARCH models, particularly during peak hours. The C-GARCH and GARCH models give a lower prediction band compared with the MC-GARCH model with the compromise of a lower coverage rate. Since the coverage rates of the C-GARCH and GARCH models are much lower than the corresponding confidence level during peak hours (e.g., the PIs of both the GARCH and C-GARCH models cover around 78% of the targets for the 85% confidence level), the MC-GARCH model generates more effective PIs, although a little bit wider than others.

To check the consistency of each model’s performance, Figs. 5 and 6 compare the PICP values of the GARCH, C-GARCH, and MC-GARCH models at different confidence levels for individual segments. The orange, blue, and green columns represent the GARCH, C-GARCH, and MC-GARCH models, respectively.
TABLE II

<table>
<thead>
<tr>
<th>Segments</th>
<th>Model</th>
<th>95%</th>
<th>PICP</th>
<th>90%</th>
<th>PICP</th>
<th>85%</th>
<th>PICP</th>
</tr>
</thead>
<tbody>
<tr>
<td>29-30</td>
<td>GARCH</td>
<td>80.33</td>
<td>87.76%</td>
<td>67.22</td>
<td>84.08%</td>
<td>59.02</td>
<td>80.20%</td>
</tr>
<tr>
<td></td>
<td>C-GARCH</td>
<td>82.84</td>
<td>88.78%</td>
<td>69.32</td>
<td>85.10%</td>
<td>60.86</td>
<td>81.63%</td>
</tr>
<tr>
<td></td>
<td>MC-GARCH</td>
<td>98.83</td>
<td>90.41%</td>
<td>82.69</td>
<td>87.76%</td>
<td>72.61</td>
<td>85.31%</td>
</tr>
<tr>
<td>30-31</td>
<td>GARCH</td>
<td>146.75</td>
<td>87.96%</td>
<td>122.79</td>
<td>82.86%</td>
<td>107.82</td>
<td>78.16%</td>
</tr>
<tr>
<td></td>
<td>C-GARCH</td>
<td>144.55</td>
<td>87.96%</td>
<td>120.95</td>
<td>82.65%</td>
<td>106.2</td>
<td>79.39%</td>
</tr>
<tr>
<td></td>
<td>MC-GARCH</td>
<td>181.57</td>
<td>91.22%</td>
<td>151.93</td>
<td>89.18%</td>
<td>133.4</td>
<td>85.10%</td>
</tr>
<tr>
<td>31-32</td>
<td>GARCH</td>
<td>141.71</td>
<td>88.37%</td>
<td>118.57</td>
<td>83.88%</td>
<td>104.11</td>
<td>78.57%</td>
</tr>
<tr>
<td></td>
<td>C-GARCH</td>
<td>136.14</td>
<td>88.57%</td>
<td>113.91</td>
<td>82.45%</td>
<td>100.02</td>
<td>76.33%</td>
</tr>
<tr>
<td></td>
<td>MC-GARCH</td>
<td>238.59</td>
<td>96.94%</td>
<td>199.64</td>
<td>94.29%</td>
<td>175.29</td>
<td>92.45%</td>
</tr>
<tr>
<td>32-33</td>
<td>GARCH</td>
<td>94.87</td>
<td>86.53%</td>
<td>79.38</td>
<td>81.22%</td>
<td>69.7</td>
<td>77.96%</td>
</tr>
<tr>
<td></td>
<td>C-GARCH</td>
<td>94.27</td>
<td>86.33%</td>
<td>78.88</td>
<td>81.43%</td>
<td>69.26</td>
<td>76.53%</td>
</tr>
<tr>
<td></td>
<td>MC-GARCH</td>
<td>113.79</td>
<td>92.06%</td>
<td>95.21</td>
<td>87.96%</td>
<td>83.6</td>
<td>84.29%</td>
</tr>
<tr>
<td>33-34</td>
<td>GARCH</td>
<td>118.41</td>
<td>90.00%</td>
<td>99.08</td>
<td>85.31%</td>
<td>86.99</td>
<td>80.61%</td>
</tr>
<tr>
<td></td>
<td>C-GARCH</td>
<td>117.68</td>
<td>89.59%</td>
<td>98.47</td>
<td>85.10%</td>
<td>86.46</td>
<td>79.80%</td>
</tr>
<tr>
<td></td>
<td>MC-GARCH</td>
<td>174.24</td>
<td>96.73%</td>
<td>145.79</td>
<td>92.86%</td>
<td>128.01</td>
<td>90.00%</td>
</tr>
<tr>
<td>Average</td>
<td>GARCH</td>
<td>116.41</td>
<td>88.12%</td>
<td>97.41</td>
<td>83.47%</td>
<td>85.53</td>
<td>79.10%</td>
</tr>
<tr>
<td></td>
<td>C-GARCH</td>
<td>115.1</td>
<td>88.24%</td>
<td>96.31</td>
<td>83.35%</td>
<td>84.56</td>
<td>78.73%</td>
</tr>
<tr>
<td></td>
<td>MC-GARCH</td>
<td>161.41</td>
<td>93.67%</td>
<td>135.05</td>
<td>90.41%</td>
<td>118.58</td>
<td>87.43%</td>
</tr>
</tbody>
</table>

The advantage of the MC-GARCH model is very significant, with the highest difference of 6.94% compared with the PIs (at the 85% confidence level) provided by the C-GARCH model for segment 33–34. On the other hand, the PICP values of the GARCH and C-GARCH models are very similar. The largest difference of the PICP values between the GARCH and C-GARCH models is 1.63%. During nonpeak hours (see Fig. 6), all three models provide a high coverage rate at corresponding confidence levels. The differences among individual models during nonpeak hours are not as significant as those during peak hours. Comparing all three models’ performance between peak and nonpeak hours suggests that the peak-hour coverage is relatively low, as the traffic variations increase during peak hours. In addition, the PI lengths during peak hours are also longer than the PI lengths during nonpeak hours. This is because the uncertainties during peak hours are more evident compared with those during nonpeak hours.

Fig. 7 provides an intuitive comparison of one-day peak-hour PIs at the 95% confidence level constructed by the MC-GARCH, C-GARCH, and GARCH models. The green dash lines stand for the PIs of the MC-GARCH model, the yellow
dash lines stand for the PIs of the C-GARCH model, and the pink dash lines stand for the PIs of the GARCH model. As shown in this figure, the PIs constructed by the C-GARCH and GARCH models almost overlap. It has been also depicted in Table I that there is no significant difference between the MPIL and PICP values of the C-GARCH and GARCH models. It seems that the effect of the long-term component in the C-GARCH model is limited in this case. On the other hand, the PIs of the MC-GARCH model are different from both the C-GARCH and GARCH models. The MC-GARCH model tends to cover more targets by increasing the width of the PIs at certain time intervals (the points identified by blue arrows). The increase in the width of the PIs at those time intervals indicates a higher uncertainty associated with travel-time prediction. Overall, the C-GARCH and GARCH models create PIs that are similar to each other. Compared with the C-GARCH and GARCH models, the MC-GARCH model tends to cover more targets by increasing the length of its PIs during certain time intervals.

We now turn to one-step ahead (5 min ahead) rolling forecasting method. The rolling forecasting method automatically updates the training data set and parameters of the model when new observations become available. Therefore, this forecasting method is able to quickly respond to the external changes in traffic conditions. To guarantee both the speed and accuracy of the model, we re-fit the model for every 1-h interval. The ten days travel-time data are used as a training data set, and the next ten days travel-time data are used as the evaluating data set in this paper. Tables II and III summarize the PI forecasting results of the GARCH, C-GARCH, and MC-GARCH models at different confidence levels during peak and nonpeak hours. As shown in Table II, the MC-GARCH model provides the highest coverage in all cases during peak hours. In terms of the MPIL, the GARCH and C-GARCH models construct PIs that have a shorter length with the compromise of providing a lower coverage rate. During nonpeak hours, all three models have similar performance. In general, all three models provide a good estimate during nonpeak hours, as their actual coverage rates, on average, are all higher than the corresponding confidence level. Both the estimating and forecasting results highlighted the strength of the MC-GARCH model for generating PIs with a higher coverage rate during peak hours. The high coverage rate of the MC-GARCH model is a benefit from the consideration of the seasonal effect during the model building process. Therefore, the MC-GARCH model shows an advantage in adjusting its PI values during peak and nonpeak periods. Comparisons among all individual segments show the consistent performance of the MC-GARCH model.

To statistically determine if there are significant differences in the PI performance of the MC-GARCH, C-GARCH, and GARCH models, we applied the two-sample Wilcoxon test. The two-sample Wilcoxon test is a nonparametric statistical test to determine if two populations are the same without assuming them to follow the normal distribution. The null hypothesis of the test is that the performance measures for each pair of models (I: MC-GARCH versus GARCH, II: MC-GARCH versus C-GARCH, and III: GARCH versus C-GARCH) are identical or similar populations. We set 0.05 as the significance level for
the test. If the $p$ value is less than 0.05, we can conclude that the null hypothesis is violated at the 0.05 significance level. Table IV provides the $p$ values of the test for each pair of models. As indicated in this table, the $p$ values of the statistical test for the PICP values of the MC-GARCH versus GARCH and the MC-GARCH versus C-GARCH are less than 0.05 during peak hours. This indicates that the PICP values of the MC-GARCH model are statistically different from those of the C-GARCH and GARCH models during peak hours. On the other hand, the MPIL values of the MC-GARCH are similar with those of the C-GARCH and GARCH models, as the $p$ values are higher than 0.05. This result confirmed the fact that the PIs constructed by the C-GARCH model provide a higher coverage (see Table II) with a statistically nondifferent mean prediction length compared with the GARCH and C-GARCH models. In addition, according to this test, we also find that the performance of the C-GARCH model and that of the GARCH model are statistically identical at the 0.05 significance level.

### Table IV

<table>
<thead>
<tr>
<th>Test</th>
<th>Period</th>
<th>PICP $%$</th>
<th>MPIL $%$</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Peak</td>
<td>0.01 0.01</td>
<td>0.01 0.22</td>
</tr>
<tr>
<td></td>
<td>Non-peak</td>
<td>0.55 0.53</td>
<td>0.83 1.00</td>
</tr>
<tr>
<td>II</td>
<td>Peak</td>
<td>0.01 0.01</td>
<td>0.01 0.22</td>
</tr>
<tr>
<td></td>
<td>Non-peak</td>
<td>0.17 0.14</td>
<td>0.25 1.00</td>
</tr>
<tr>
<td>III</td>
<td>Peak</td>
<td>0.75 0.83</td>
<td>0.84 0.84</td>
</tr>
<tr>
<td></td>
<td>Non-peak</td>
<td>0.42 0.4</td>
<td>0.42 1.00</td>
</tr>
</tbody>
</table>

V. CONCLUSION

As the uncertainty associated with travel-time prediction becomes an important topic for implementing an intelligent transportation system, statistical volatility models provide a promising way to generate more accurate PIs that account for variability in travel-time prediction. The traditional GARCH model is argued to be inadequate when modeling the data that show pronounced seasonal patterns. This paper has developed the C-GARCH and MC-GARCH models in travel-time prediction. To empirically evaluate the performance of the proposed models, this paper has tested the GARCH, C-GARCH, and MC-GARCH models by using the freeway travel-time data collected from AVI stations located along U.S. Highway 290 (or U.S. 290) in Houston, TX, USA. The forecasting results of the proposed models are attractive, particularly during peak hours. The main contributions of this paper include a proposal of two novel volatility models in predicting the uncertainty associated with travel time, the comparison and evaluation of the performance of different volatility models under congested and noncongested situations, and a detailed insight into the unique characteristic of each model. The findings of this paper include the following.

1) The proposed MC-GARCH model outperforms the GARCH and C-GARCH models during peak-hour prediction. A case study of the five selected segments highlighted the strength of the MC-GARCH model in providing more effective PIs in terms of the coverage rate. The idea of decomposing the travel-time volatility is promising when the data show cyclic patterns. By decomposing the travel-time volatility into daily, diurnal, and stochastic components, the MC-GARCH model is able to capture the uniqueness of each component and captures the seasonal effect of the data.

2) The C-GARCH model treats travel time as a long-term and temporary component. It works best if there is a trend. Based on the case study, the performance of the PIs constructed by the C-GARCH model and the GARCH model are similar to each other. The effect of the long-term volatility component in the C-GARCH model is not significant in this case.

3) During nonpeak hours, there is no obvious advantage of all three models in terms of the MPIL and the PICP. This is partly due to the fact that the travel time during nonpeak hours is relatively stable, with small variations around the mean. The trend and seasonal patterns are not obvious during this period.

The C-GARCH model decomposes the travel-time data into long-term, short-term, and cyclical components. If there are cyclic components in the data, the MC-GARCH model has the potential to better capture the uncertainties associated with the travel time. In addition, the MC-GARCH model decomposes the traffic volatility into several different components that can be easily interpreted and estimated. In our paper, the daily component and the normalized residuals have been modeled as a simple GARCH model. In addition to the GARCH model, the daily component can be also estimated through a multifactor risk model, as suggested by Engle and Sokalska [27]. It is also worth trying different variations of GARCH models to estimate the normalized residuals. In addition, our paper has treated the intraday component as an average term. Further study could also explore different ways in defining the intraday component.

### References


A. Ghalanos, Rugarch: Univariate GARCH Models 2013, in R package version 1.2-7 ed.

Yanru Zhang received the B.S. degree in transportation engineering from Beijing Jiaotong University, Beijing, China, and the M.S. degree in civil engineering from Texas A&M University, College Station, TX, USA, in 2009 and 2011, respectively. She is currently working toward the Ph.D. degree in civil engineering at the University of Maryland, College Park, MD, USA.

She is a Research Assistant with the Department of Civil and Environmental Engineering, A. James Clark School of Engineering, University of Maryland. Her research interests include data mining applications in the transportation field, intelligent transportation systems, transportation modeling and simulation, and predictive modeling/forecasting.

Ali Haghani received the B.S. degree from Shiraz University, Shiraz, Iran, and the M.S. and Ph.D. degrees from Northwestern University, Evanston, IL, USA, in 1976, 1982, and 1986, respectively, all in civil engineering.

He is a Professor with the Department of Civil and Environmental Engineering, A. James Clark School of Engineering, University of Maryland, College Park, MD, USA, where he is also the Director of the Center for Integrated Transportation Systems Management. His research interests include transportation network modeling, freight transportation and logistics, emergency response, dynamic fleet management, container transportation, and mass transit operations.

Xiaosi Zeng was born in Guangzhou, China, in 1983. He received the B.S. degree in transportation engineering from South China University of Technology, Guangzhou, China, and the M.S. degree in civil engineering in 2009 from Texas A&M University, College Station, TX, USA, where he is currently working toward the Ph.D. degree in civil engineering.

From 2007 to 2013 he was a Research Assistant with the Texas A&M Transportation Institute.