Hindawi
Journal of Advanced Transportation
Volume 2019, Article ID 3958127, 12 pages
https://doi.org/10.1155/2019/3958127

Research Article

A Hybrid Approach Based on Variational Mode Decomposition for Analyzing and Predicting Urban Travel Speed

Eui-Jin Kim,[1] **Ho-Chul Park**,[2] **Seung-Young Kho**,[3] **and Dong-Kyu Kim**[3]

[1] *Department of Civil and Environmental Engineering, Seoul National University, Seoul 08826, Republic of Korea*
[2] *Department of Transportation Engineering, Myongji University, Yongin 17058, Republic of Korea*
[3] *Department of Civil and Environmental Engineering and Institute of Construction and Environmental Engineering, Seoul National University, Seoul 08826, Republic of Korea*

Correspondence should be addressed to Dong-Kyu Kim; dongkyukim@snu.ac.kr

Received 6 November 2018; Revised 13 August 2019; Accepted 17 September 2019; Published 12 December 2019

Academic Editor: Dongjoo Park

Predicting travel speeds on urban road networks is a challenging subject due to its uncertainty stemming from travel demand, geometric condition, traffic signals, and other exogenous factors. This uncertainty appears as nonlinearity, nonstationarity, and volatility in traffic data, and it also creates a spatiotemporal heterogeneity of link travel speed by interacting with neighbor links. In this study, we propose a hybrid model using variational mode decomposition (VMD) to investigate and mitigate the uncertainty of urban travel speeds. The VMD allows the travel speed data to be divided into orthogonal and oscillatory sub-signals, called modes. The regular components are extracted as the low-frequency modes, and the irregular components presenting uncertainty are transformed into a combination of modes, which is more predictable than the original uncertainty. For the prediction, the VMD decomposes the travel speed data into modes, and these modes are predicted and summed to represent the predicted travel speed. The evaluation results on urban road networks show that, the proposed hybrid model outperforms the benchmark models both in the congested and in the overall conditions. The improvement in performance increases significantly over specific link-days, which generally are hard to predict. To explain the significant variance of the prediction performance according to each link and each day, the correlation analysis between the properties of modes and the performance of the model are conducted. The results on correlation analysis show that the more variance of nondaily pattern is explained through the modes, the easier it was to predict the speed. Based on the results, discussions on the interpretation on the correlation analysis and future research are presented.

1. Introduction

Forecasting travel speed on urban road networks are the most critical component of intelligent transportation systems (ITSs). An accurate prediction of link travel time allows travelers to choose an alternative route to avoid congestion or traffic incidents, and it reduces congestion. Therefore, the reliable predictions of recurrent congestion and of nonrecurrent incidents are especially important, but they are challenging to provide due to the complex and elusive characteristics of traffic dynamics [1].

Travel speed on urban roads is more complicated than on freeways or arterial roads due to the intrinsic uncertainty associated with the roads. Many components in urban traffic dynamics, such as travel demand, traffic signals, and stochastic arrivals at intersections result in recurrent and nonrecurrent congestion [2]. These phenomena causes the nonlinearity, nonstationarity, and volatility, which are difficult to model for statistical approaches, such as Kalman filtering [3], autoregressive integrated moving average (ARIMA) [4], and its variants [5].

Some researchers have provided extensive reviews in which they compared various prediction schemes given on the different types of data and models [1, 6]. They generally have reported that a machine learning-based approach, such as artificial neural network (ANN) [7], the support vector machine (SVM) [8, 9], and the k-nearest neighbor method (KNN) [10–12], outperform the parametric statistical approach due to their ability to identify the nonlinear effects through flexible restructuring [13]. However, the existing learning-based approaches have difficulty dealing with the uncertainty of urban traffic conditions, and the difficulty is compounded in unstable congestion [14]. Another characteristic of urban travel speed is spatiotemporal heterogeneity. Some researchers have reported that the same prediction model shows significant differences in its prediction performance across the network and during different time periods [15–18]. This characteristic was observed between links, for different days of the week, for different traffic conditions on the same day. This phenomenon is due to the correlation of neighbor links [16] and to volatility in the data [17].

Recently, the effects of spatiotemporal variables on the prediction of travel speed have been discussed using multipoint data, and multivariate approaches, such as modified SVM [9], modified KNN [12, 13], deep neural networks [19], and Bayesian networks [20]. These researchers tried to capture the nonlinear spatiotemporal effects, and their critical processes were selecting the variables that reflected the dynamics of network traffic in various localities and determining the causal relationships between different road segments. Various techniques have been used to select the optimal variables, either before the training process [13] or during the training process [19, 20]. However, since the influences of links on their neighbors change with time and they are heterogeneous in space, only generally well-performed spatiotemporal variables (i.e., the number and location of links and the optimal time-lag) would be selected at a given point in time. These limitations cause poor performance, particularly in nonrecurrent conditions due to irrelevant variables [14, 16].

Alternatively, hybrid approaches using single-point data have been proposed in many studies to cope with the uncertainty of urban traffic conditions, and these are divided into two branches i.e., feature extraction and "divide and conquer". The feature extraction is the process of converting raw data into useful representations that enhance the efficiency and generality of models [21]. In the hybrid approaches, the features are mainly extracted in preprocessing with domain knowledge, such as known periodicity of the phenomena [22] and propagation of kinematic waves [23]. Therefore, the feature extraction requires the understanding of phenomena or extensive empirical research. However, traffic dynamics on urban networks remain elusive, and empirical calibration also is difficult due to the various patterns of each link and on each day. For practical application of the hybrid models, the concept of "divide and conquer" (DC) was proposed for accurate, data-adaptive, and easy-to-use prediction [24]. The principle of DC is that a complicated modeling task can be simplified by decomposing the data with multiple frequency components into orthogonal functions with local frequency components. The decomposed functions are respectively predicted and summed to represent the predicted time series. For the DC-based prediction, signal processing techniques, such as Fourier transform (FT) [25], wavelet transform (WT) [26], and empirical mode decomposition (EMD) [27, 28], have been used to decompose the freeway speed and passenger flow, and their improved performances were usually reported. Furthermore, analyzing spectral and statistical properties of decomposed functions can provide a better understanding of traffic dynamics on urban networks. However, this approach has not been applied to urban travel speed that has both prominent uncertainty, and prominent diversity.

In this study, variational model decomposition (VMD) is used to decompose the speed data for DC-based prediction. The VMD overcomes several limitations of conventional techniques, such as the stationary assumption in FT, basis wavelet selection in WT, and sensitivity to noise and sampling in EMD [29]. The VMD decomposes the data into orthogonal, and oscillatory sub-signals, called modes, with their respective center frequencies. The notable strength of VMD is the robustness to sampling frequency and noise, which can separate the useful oscillatory patterns from noisy and volatile speed data. The influential factors related to the uncertainty of urban traffic conditions can be decomposed as oscillatory modes, even in recurrent and nonrecurrent congestion, and the properties of the modes can be used to explain these uncertainties, which determines the possibility to predict the speed, i.e., forecastability. Therefore, the VMD can improve the prediction performance, and furthermore, it can provide new insights for the factors determining the difficulties of travel speed prediction.

The objective of this study is to propose a hybrid model using VMD that contributes to understanding and improving the prediction of urban travel speed by investigating and mitigating its uncertainty. The key contents of this paper are as follows: (a) to propose a DC-based hybrid model using VMD for predicting travel speed in the urban networks; (b) to evaluate the proposed model on various links, days of the week, and for different traffic conditions for comparing with the benchmark models in terms of accuracy and robustness; and (c) to explain the intrinsic properties to forecast heterogeneous urban traffic dynamics using the spectral and statistical properties of the decomposed modes.

The remainder of this paper is organized as follows. First, we describe the studied site, and the collecting of the link-travel speed data. In the next section, the decomposition and prediction method are discussed in detail. Then, we present the performance evaluation with an analysis of the properties of the modes to interpret the results. Last, our concluding remarks and our potential future research are presented.

2. Study Site and Data Collection

The site of the study was the primary urban roadway network in the Daegu metropolitan area in South Korea. The complete data set included three months of individual vehicles' travel data collected by the roadside equipment (RSE) from April to June 2016. When the vehicles equipped with an onboard unit (OBU) device passed the RSE, travel data were collected using dedicated short-range communication (DSRC) which connect the OBU and RSE. The DSRC is more accurate than the data obtained by GPS-enabled vehicles (i.e., floating car data). In Korea, more than 60% of vehicles are equipped with OBU, which guarantees the reliability of the data. We aggregated the individual travel data into a 5-minutes time interval and converted it into link travel time data. Figure 1(a) shows the urban network that was studied, and it consists of 545 links equipped with DSRC roadside detectors, including expressways, arterials, and other roads. We studied the links that are

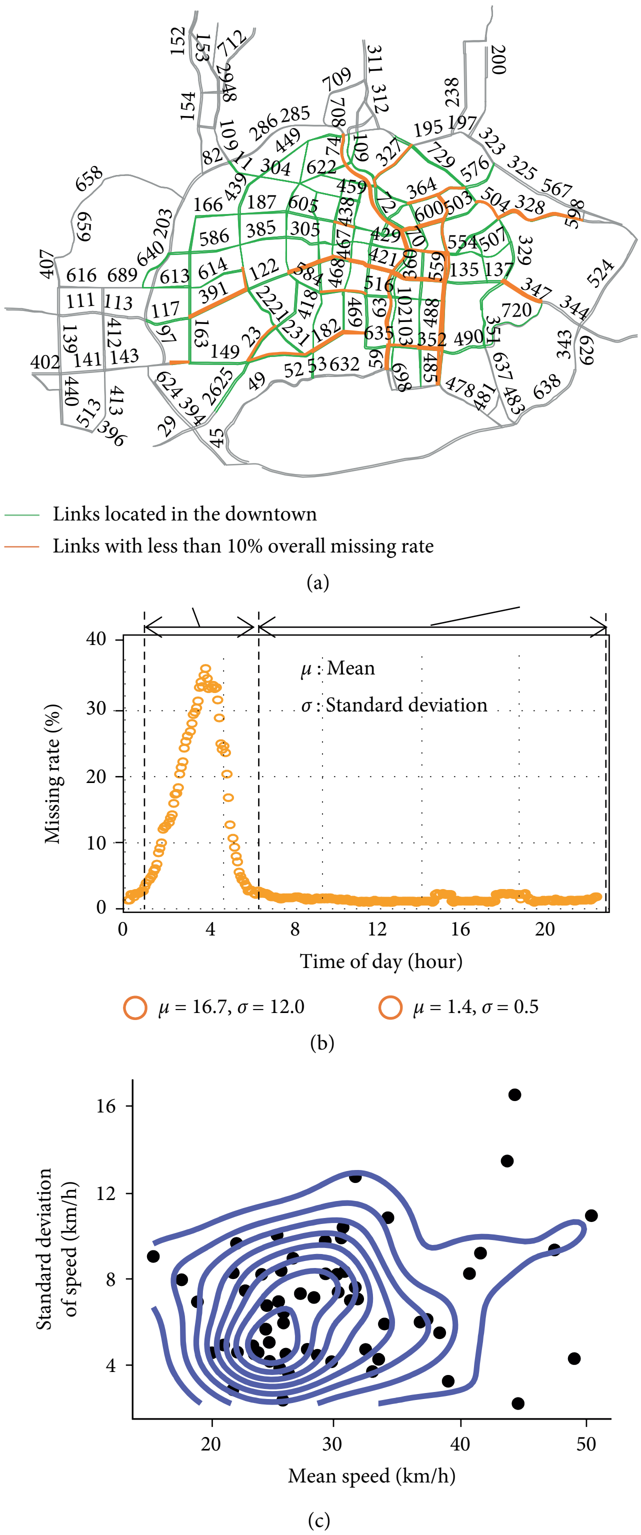


Figure 1: The urban road network that was studied: (a) spatial distribution of the studied site with missing rate of link travel speed data, (b) temporal distribution of missing rate of link travel speed data, and (c) mean and standard deviation of the travel speed on links that had a missing rate of less than 10%.

located in the downtown to ensure the reliability of DSRC data with sufficient traffic volume to minimize the missing rate of the data, and those links are highlighted in green lines in Figure 1(a). Missing data are represented by some links that remain unobserved during some time intervals. The missing rate of the data is calculated by the number of unobserved 5-minutes aggregated travel speeds divided by the number of unobserved and observed ones at each link. Among the links located in the downtown, the links with an overall missing rate of less than 10% are highlighted in orange lines in Figure 1(a). To evaluate the proposed method, it is necessary to conduct a consistent analysis of various locations of the urban network. However, the links show varying missing rate considering the whole time period as shown in Figure 1(b). Despite the high penetration rate of OBU, there exist some missing data especially at dawn when the volume of DSRC equipped vehicle was low.

In this study, we used those orange-lined 61 links data from 06:00 A.M. to 23:55 P.M., where the missing rate is less than 2%, to minimize the effect of the missing data, and the remaining missing data were imputed using a linearly-weighted moving average with 30-minutes windows. We also removed the outlier that are the observations below the 15th percentiles or above the 85th percentiles. The average length of the links that were used was 980 m. Distribution of characteristics of the travel speed in each link is shown in Figure 1(c). The mean travel speed and its standard deviation had bell-shaped distributions in the ranges of 20–40 km/h and 4–12 km/h, respectively, which reflected the heterogeneous nature of the urban links.

3. Methodology

3.1. Overview of Hybrid Model. The traffic characteristics of the urban roads can be treated as a combination of regular components, such as commuters' travel demand, geometric conditions, and irregular components, such as traffic incidents and other exogenous factors. Separating all of these individual components is impossible. However, typical patterns of regular components can be extracted directly, and the uncertainty of the irregular components can be transformed into a combination of orthogonal, and oscillatory modes, which is more predictable than the original components. Based on this principle, the VMD decomposed the travel speeds into modes, and they were predicted and summed up to represent the predicted travel speed. The summation of the modes can be predicted more accurately than the original travel speed data since some of the modes mitigate the uncertainty in the original data, such as nonlinearity, nonstationarity, and volatility. Figure 2 shows the flow diagram of the proposed hybrid model. In order to describe the processes, we first explain the VMD and then present travel time prediction models, i.e., ANN and SVM, for each mode. Next, the reconstruction process that can finally derive predicted travel speed from the mode that is presented.

3.2. Variational Mode Decomposition (VMD). To look for the K number of modes that are compacted around a central frequency with limited bandwidth, the VMD algorithm solves the following minimization problem:

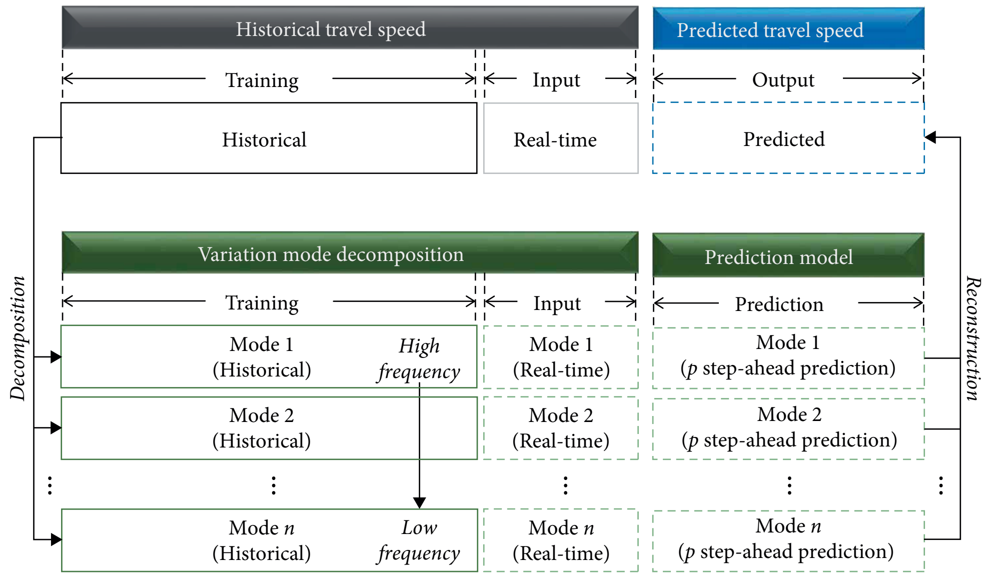


FIGURE 2: Flow diagram of a hybrid model using variational mode decomposition.

$$\min_{\{u_k\},\{\omega_k\}} \left\{ \sum_{k=1}^{K} \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \right\}$$
$$\text{s.t.} \sum_{k=1}^{K} u_k = f, \quad (1)$$

where u_k is the k^{th} decomposed mode, K is the number of predefined modes, f is the original time-series, ω_k is the central frequency of a mode k, δ is the Dirac distribution, $\|\cdot\|$ is the vector ℓ_2 norm, $j^2 = -1$, $*$ denotes convolutions, and ∂_t is partial derivative of time t. The reconstruction constraint indicates that summation over all modes equals to the original signal. The meaning of solving this problem is described as follows [29]: (a) the associated analytic signals, by means of the Hilbert transform to obtain a unilateral frequency spectrum are computed for each mode, u_k; (b) the frequency spectrum of the mode to "baseband" is shifted by mixing with an exponential tuned to the respective estimated central frequency; (c) the bandwidth is estimated through the Gaussian smoothness of the demodulated signal (the squared L2-norm of the gradient).

The solution to the problem is the saddle point of the augmented Lagrangian described with Lagrangian multipliers and quadratic penalty. The quadratic penalty is a way to encourage reconstruction fidelity, while the Lagrangian multipliers are a common way to enforce constraints strictly. By combining these two terms as in Equation (2), better convergence properties and strict enforcement of the constraint can be achieved [29].

$$\begin{aligned} L(\{u_k\},\{\omega_k\},\lambda) =& \alpha \sum_{k=1}^{K} \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \\ &+ \left\| x(t) - \sum_{k=1}^{K} u_k(t) \right\|_2^2 \\ &+ \left\langle \lambda(t), x(t) - \sum_{k=1}^{K} u_k(t) \right\rangle, \end{aligned} \quad (2)$$

where $\lambda(t)$ are Lagrangian multipliers, α is a balance parameter of the data fidelity constraint., $\left\| x(t) - \sum_{k=1}^{K} u_k(t) \right\|_2^2$ is a quadratic penalty term for accelerating speed for convergence, and $\left\langle \lambda(t), x(t) - \sum_{k=1}^{K} u_k(t) \right\rangle$ is the Lagranian term where $\langle \cdot, \cdot \rangle$ is the inner product. When α is large, the central frequency of a mode k, ω_k, might be precisely estimated, but it yields large violation of reconstruction constraint. The solution to Equation (2) is a sequence of iterative, sub-optimization algorithms, called the alternate direction method of multiplier (ADMM), and more details are presented in [29].

3.3. Artificial Neural Network (ANN). The ANN is a well-known learning-based approach for predicting time series. A multi-layer perceptron (MLP) is a conventional neural network that includes an input layer, one or more hidden layers, and an output layer. The MLP can capture the nonlinear relationship in time-series data by iteratively adjusting the weights and biases between the interactions of neurons in multiple layers. A standard backpropagation algorithm with a decay term [30] was used to train our MLP. The algorithm consists of two steps: firstly, the information of input neuron propagates forward to compute the output information, and then connection weights are modified by the difference between the computed and observed output information. For travel speed prediction, lagged values of speed, S_{t-j}, are used as an input data as in Equation (3) [7]

$$\hat{S}_t = \hat{h}(S_{t-m}, \ldots, S_{t-M}) = \sum_{i=1}^{I} W_i f\left(\sum_{j=m}^{M} w_{ij} S_{t-j} + \theta_i \right), \quad (3)$$

where t is the number of input data ($t = 1, \ldots, n$), m and M are minimum and maximum lagged time, $\hat{h}$ is an estimate of the nonlinear model, which is obtained by minimizing the $\sum_{t=1}^{n} (S_t - \hat{S}_t)^2$, W_i and w_{ij} are weights of interaction between neurons, and θ_i is the bias. A MLP has been used extensively for predicting short-term traffic parameters due to its ability to work with multi-dimensional data and its good predictive performance [31].

3.4. Support Vector Machine (SVM). The SVM have been used extensively for nonlinear regression. The SVM mapped the data from the input space into high-dimensional feature space and constructed the optimal decision function. Given the observed speed (S_t), and its lagged values $\left(\boldsymbol{S_m^M} = [S_{t-m}, \cdots, S_{t-M}]\right)$, the optimal decision function is shown in Equation (4):

$$\hat{S}_t = w^T \boldsymbol{\Phi}\left(\boldsymbol{S_m^M}\right) + b, \tag{4}$$

where $\boldsymbol{\Phi}\left(\boldsymbol{S_m^M}\right)$ is a nonlinear function that converts the data to feature space, w is a weight, and b is a bias. The optimal decision function is estimated by minimizing the regression risk [8]:

$$R_{\text{reg}}(f) = C \sum_{t=1}^{n} L\left(S_t, \hat{S}_t\right) + \frac{1}{2}\|w\|^2, \tag{5}$$

$$L\left(S_t, \hat{S}_t\right) = \begin{cases} 0 & \text{if} \left|y - \hat{S}_t\right| \le \varepsilon \\ \left|y - \hat{S}_t\right| - \varepsilon & \text{otherwise,} \end{cases} \tag{6}$$

where $L(\bullet)$ is a loss function, and C and ε are the regularization parameters. The w can be represented in terms of input data, and the optimal decision function can be written as Equation (8):

$$w = \sum_{t=1}^{n} \left(a_t - a_t^*\right) \boldsymbol{\Phi}\left(\boldsymbol{S_m^M}\right). \tag{7}$$

$$\begin{aligned} \hat{S}_t &= \sum_{t=1}^{n} \left(a_t - a_t^*\right) \boldsymbol{\Phi}\left(\boldsymbol{S_{m,t}^M}\right) \boldsymbol{\Phi}\left(\boldsymbol{S_m^M}\right) + b, \\ &= \sum_{t=1}^{n} \left(a_t - a_t^*\right) \boldsymbol{K}\left(\boldsymbol{S_{m,t}^M}, \boldsymbol{S_m^M}\right) + b, \end{aligned} \tag{8}$$

where $\boldsymbol{K}(\cdot)$ is the kernel function that coverts a nonlinear learning problem into a linear learning problem, and a_t, a_t^* are the Lagrangian multipliers used to estimate the optimal decision function. We used the radial basis function (RBF) kernel for nonlinear regression. The SVM shows the excellent generalized performance for predicting travel speed [9].

3.5. Reconstruction in Hybrid Model. The decomposed and predicted modes represent the regular and irregular components of travel speed, and each mode is summed up to reconstruct the predicted travel speed data. Based on two time-series prediction models, P, i.e., ANN and SVM, represented by the function, f_P, the hybrid models are formulated as Equation (9):

$$\widehat{S}_t^{VMD-P} = \sum_{k=1}^{K} \widehat{M}_{k,t}^P = \sum_{k=1}^{K} f_P\left(M_{k,t-m}, M_{k,t-m-1}, \ldots, M_{k,t-M}\right), \tag{9}$$

where $\widehat{S}_t^{VMD-P}$ is a predicted travel speed at time t, estimated by the hybrid prediction model, $VMD-P$. $\widehat{M}_{k,t}^P$ is a predicted value of the decomposed mode k at time t, estimated by prediction model, P. K is the number of decomposed modes by VMD. Since the VMD is conducted by constraining that the sum of the modes can be reconstructed to the original data as in Equation (2), the sum of the predicted modes can represent the predicted travel speed.

3.6. Identification of Traffic Congestion. To verify the robustness of the proposed method, which is crucial for the reliability of ITS, we tested our method in overall conditions and in congestion conditions. There is no consistent method to identify the traffic state. Many studies have used speed as a primary factor for identifying congestion where the duration time of speed under the threshold is sufficiently long [32, 33]. However, this measure cannot capture the transition from free-flow to congestion, which the traffic information systems should detect preemptively. In order to test the model just before, during, and after congestion in this study, we propose the congestion identification algorithm that is based on valley searching and consists of three steps, i.e., (a) the link travel speed is normalized to have a zero mean and unit variance for applying the algorithm irrespective of the link; (b) the valley that has a normalized speed less than −1.0 is detected; and (c) a peak-valley-peak sequence of longer than 20 min finally is identified as congestion.

Figure 3 shows the results of the proposed algorithm in eight different links on Fridays. Congestion in both long and short intervals was detected reasonably irrespective of various traffic patterns, and congestion was represented from the beginning to the end. The congestion detected by the algorithm accounted for 26% of the total data.

3.7. Tuning Parameters. Tuning parameters are important for reliable prediction of machine learning-based model. It should be conducted in a systematic and reproducible way to prevent underestimating or overestimating the performance of the proposed hybrid models, i.e., VMD-ANN and VMD-SVM, and the benchmark models, i.e., SVM and ANN. We consistently performed the five-fold cross-validation for a fair comparison of the proposed models with the benchmark models.

The common parameter for all of the models was the input lagged variable, and we calibrated them from *t-3* (prediction horizon in this study) to *t-p* where $p = 3, 4, \ldots, 13$, and determined whether the lagged variable presenting the day before was included or not. For the ANN model, we used the MLP with one hidden layer, and the number of hidden neurons and the weight decay in the learning algorithm were calibrated. For the SVM model, the RBF kernel was used with two regularization parameters, i.e., C and ε. For the VMD, the number of mode K which are the most critical parameters, were calibrated with the parameters of the prediction models for each of the IMFs. Among the K s from 8 to 12, the calibration of K was based on the accuracy of the prediction of the link-day unit. The balancing parameter, α, in Equation (2) was set as 5,000 so that the reconstruction error, which is the difference between original time-series and the sum of the decomposed modes, was less than 1%.

4. Results and Discussion

4.1. Evaluation Results. In this section, first, we evaluated the prediction performance of SVM and ANN as benchmark models and investigated the spatiotemporal patterns in

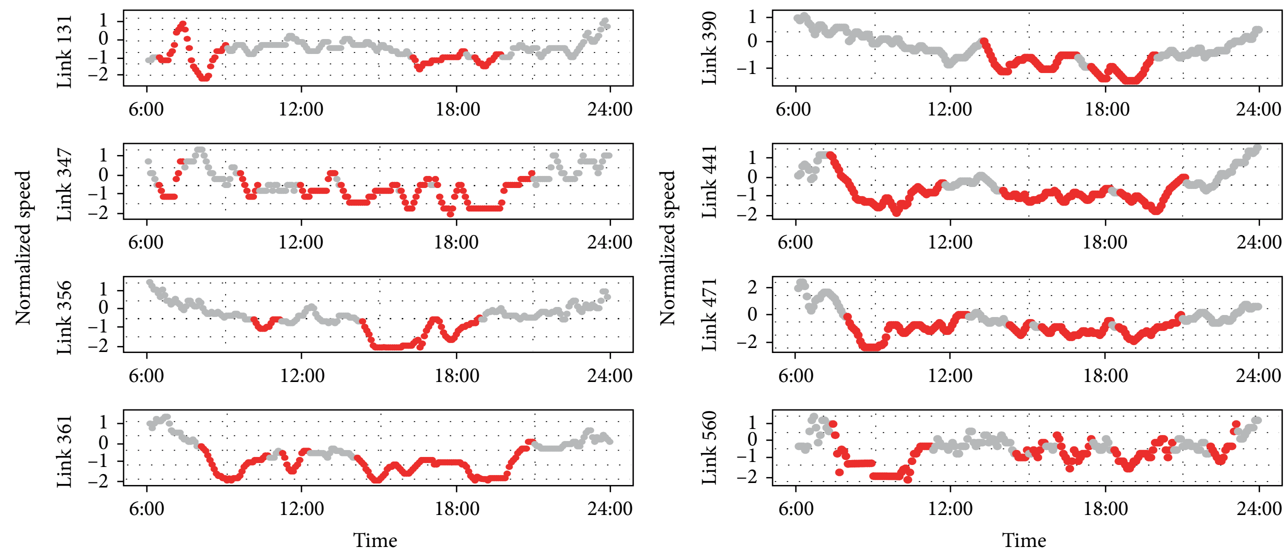


FIGURE 3: Traffic congestion on Friday identified by the valley searching-based algorithm.

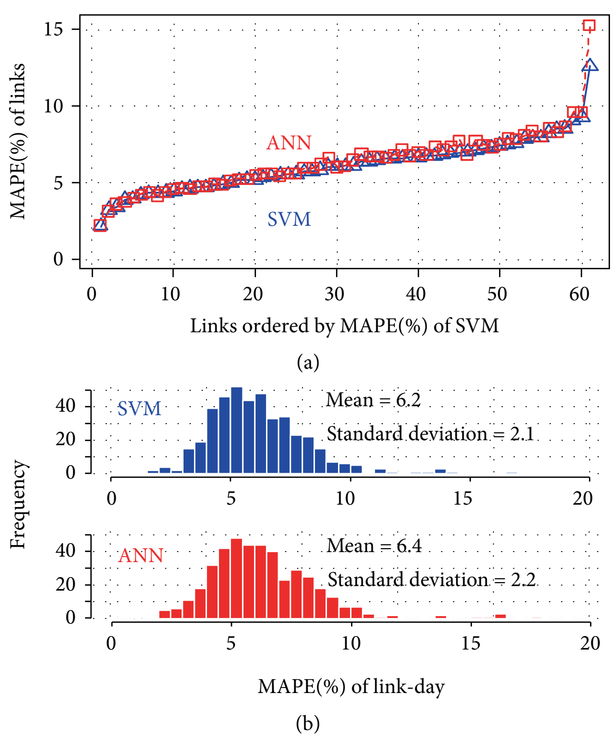


FIGURE 4: Comparision of prediction performance of benchmark models, ANN and SVM: (a) MAPE(%) of link unit and (b) MAPE(%) of link-day unit.

urban traffic parameter predictions. Next, we compared the prediction performance of two hybrid models, i.e., VMD-ANN and VMD-SVM, with the benchmark models. We used the travel speed data from 91 days to train and test the model. The 84 days of data were used for training and the remaining 7 days for testing the model. The five-fold cross-validation was conducted using the training data to calibrate the parameters of models for each link. The prediction performance was evaluated by the results of 15-minute-ahead speed prediction, i.e., three-step ahead prediction of 5 minutes aggregated data.

Figure 4 shows the prediction performance using two benchmark models of which the mean absolute percentage errors (MAPE) were calculated using 61 links with 7 days (427 link-days). With the evaluation in link units in Figure 4(a), the SVM showed slightly better performance than ANN in each link, and the difference became larger on the hard-to-predict link. With the evaluation in link-day units in Figure 4(b), the SVM also showed more accurate and robust performance by producing the lower average and the lower standard deviation of MAPEs, respectively, compared to the ANN. Also, the results showed that the prediction performance had a considerably broad range according to links and days. This indicated that the prediction performances were affected by different spatiotemporal patterns of the traffic state, which vary depending on the day of the week and the geometric and environmental condition of the links. As shown in Figure 4(b), the MAPE of both models showed skewed, bell-shaped distributions, which presented the heterogeneity of urban traffic dynamics with the existence of harsh conditions, such as traffic incidents and special events.

Table 1 compares the performances of the proposed models with benchmark models in the overall condition and in the congested condition. In the detailed comparison and analyses, we used 16 randomly-selected links from Friday to Tuesday (80 Link-days). The results indicated that the hybrid models had better performances than moving average (MA) with 15 min window, ANN, and SVM. In combination with the VMD, the VMD-ANN outperformed the VMD-SVM in both of the conditions. This indicated that, unlike raw travel speed, the decomposed modes, which are the oscillatory patterns mitigating the uncertainty of the original data, were trained more properly in the ANN than in the SVM.

The comparisons of the models were focused on the best benchmark and hybrid models, i.e., SVM and VMD-ANN. The VMD-ANN showed an overall better performance than SVM, and its MAPE value was slightly greater in congestion.

Table 1: Comparison of performances of benchmark models and proposed models.

Statistic	MAPE (%)		MAE (km/h)	
	Overall	Congestion	Overall	Congestion
MA	7.45	15.73	2.13	3.21
ANN	7.11	9.34	1.77	1.92
SVM	6.76	8.84	1.70	1.83
VMD-ANN	6.37	8.08	1.63	1.74
VMD-SVM	6.66	8.48	1.66	1.78
Percent difference (VMD-ANN, SVM)	−6%	−9%	−4%	−5%

Note. Percent difference (A, B) = {measure (A)−measure (B)} / mean (measure (A), measure (B)).

In particular, we found that VMD-ANN and SVM had different characteristics in the prediction models according to traffic state, which is supported by gray-shaded region in Figures 5(a) and 5(b). Figure 5(a) presents the congestion on the weekend, and it shows that the VMD-ANN stably tracked the periodic patterns of transition to congestion, while the SVM sensitively reacted to local speed variation before the heavy congestion. However, in the case of stable and recurrent congestion, such as a peak-hour or the congestion on a weekday, the predictions of SVM were a little better than those of VMD-SVM as shown in Figure 5(b).

Figures 5(c) and 5(d) show the advantage of the VMD-ANN by plotting the percent differences of each link-day. The percent differences of MAPE ranged from −72% to 18%. In many cases, both models may have similar performance, but the VMD-ANN showed a significant improvement in performance over specific link-days, while the performance of SVM was relatively low. The specific link-days may have large variations, such as transitions of traffic state and abrupt changes in speed.

4.2. Relationship between the Properties of the Modes and the Accuracy of the Prediction. In this section, we examined how the properties of the modes affected the performances of the prediction models. We computed the spectral and statistical properties of modes and investigated the relationship between those properties and the prediction performance of SVM and VMD-ANN. We conducted the VMD upto 84 days of training data, and Figures 6(a) and 6(b) show the speed data for the last three days and their corresponding modes from link 523. Except for mode 1, which had an approximately monotonic trend, each mode represented the oscillatory patterns of the travel speed with a central period, i.e., the inverse of frequency.

Table 2 shows the central periods of modes according to the best K from the link-day data. The modes obtained by the different numbers of K had different bandwidths, but some of the modes with a similar period usually were included. To understand the effect of the characteristics of the mode on prediction performance irrespective of K, we reclassified the modes based on their periods. The bandwidths of the reclassified modes were determined to minimize the coefficient of variance (CV) of each mode's period. These modes could be considered as the traffic patterns in an urban network, i.e., daily travel demand, commuting, stochastic fluctuation, and overlapped effects of traffic dynamics. In subsequent analyses, we investigated the properties of these eight reclassified modes.

To examine the effects of the modes on the prediction performance, we conducted a correlation analysis on the performance of two models and the explanatory power (EP) of each mode based on the data. The EP was defined by the percent of the variance of each mode, c_i, as in Equation (9). Because the modes are approximately orthogonal and collectively reproduced the original signal, S_K, the variance of the original data approximates to the summation of the variance of each mode, as in Equation (10) [29].

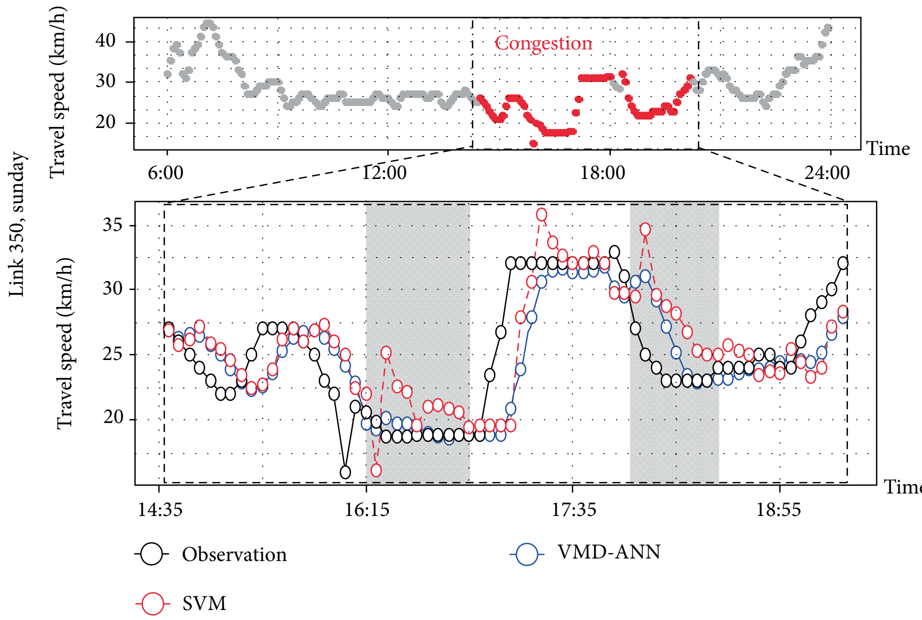
$$EP_i = \frac{\mathrm{Var}(c_i)}{\sum_{i=1}^{K}\mathrm{Var}(c_i)}, \tag{10}$$

$$\mathrm{Var}(S_K) = \sum_{i=1}^{K}\mathrm{Var}(c_i) + 2\sum_{i\neq j}\mathrm{cov}(c_i c_j) \approx \sum_{i=1}^{K}\mathrm{Var}(c_i). \tag{11}$$

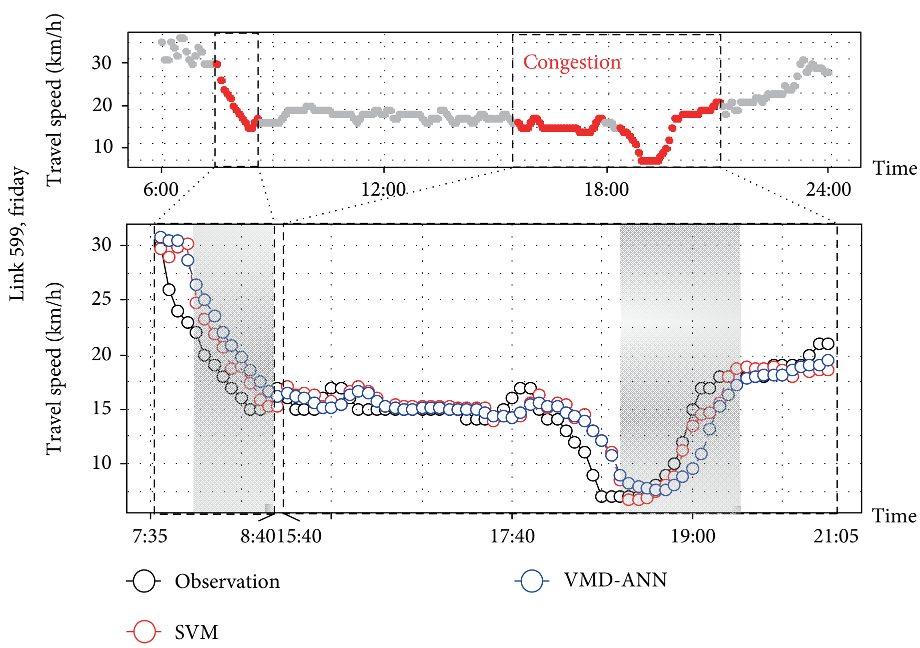
Table 3 shows the correlation between the MAPE of the models and the EP of each mode in overall and congested conditions. The explained variance in the summation of the modes ranged from 92% to 98% in our data. Only the training data were used for the correlation analysis to enable pre-emptive diagnosis for practical use.

For the correlation between performance measures, MAPE of the SVM in the overall conditions and the congested conditions have significant negative correlations with the percent difference between the VMD-ANN and SVM. This result shows that the VMD-ANN provided a better prediction in congested and complex condition, whereas the SVM had trouble predicting. Such an improvement in the model's performance, where reliable prediction needs, is a significant advantage of our model.

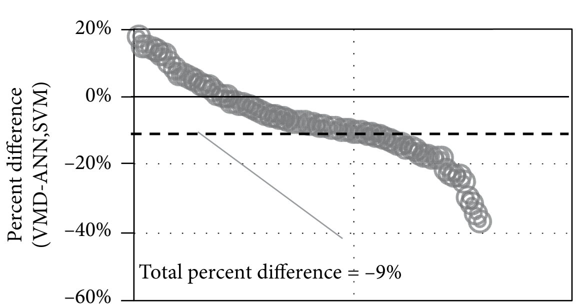
For the EP on the overall patterns of the link, it should be noted that modes with short periods (modes 5–8) showed negative results that were statistically significant in all situations. In particular, in the EP of the training data in congested traffic condition, the only modes with short periods were statistically significant. This indicated that the larger the proportion of the modes with short periods becomes, the more accurate the forecast will be in congested traffic condition. In other words, the stochastic fluctuations in congested traffic condition, which may be explained by the modes with short periods, were transformed into combinations of more predictable modes than the original components. This was supported



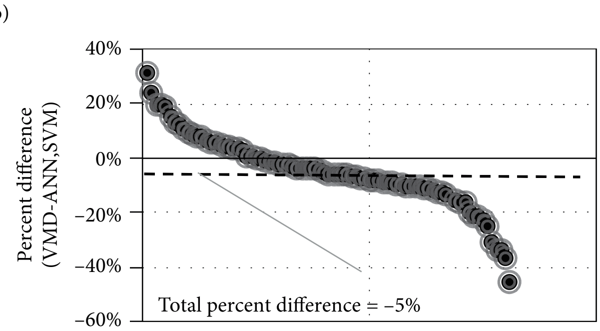
(a)



(b)



(c)



(d)

Figure 5: Detailed comparison of SVM's and VMD-ANN's predicted values in congestion: (a) VMD-ANN's predicted value is better than SVM's, (b) VMD-ANN's predicted value is worse than SVM's. Variation of the percent difference between SVM and VMD-ANN by different link-day, (c) MAPE, and (d) MAE.

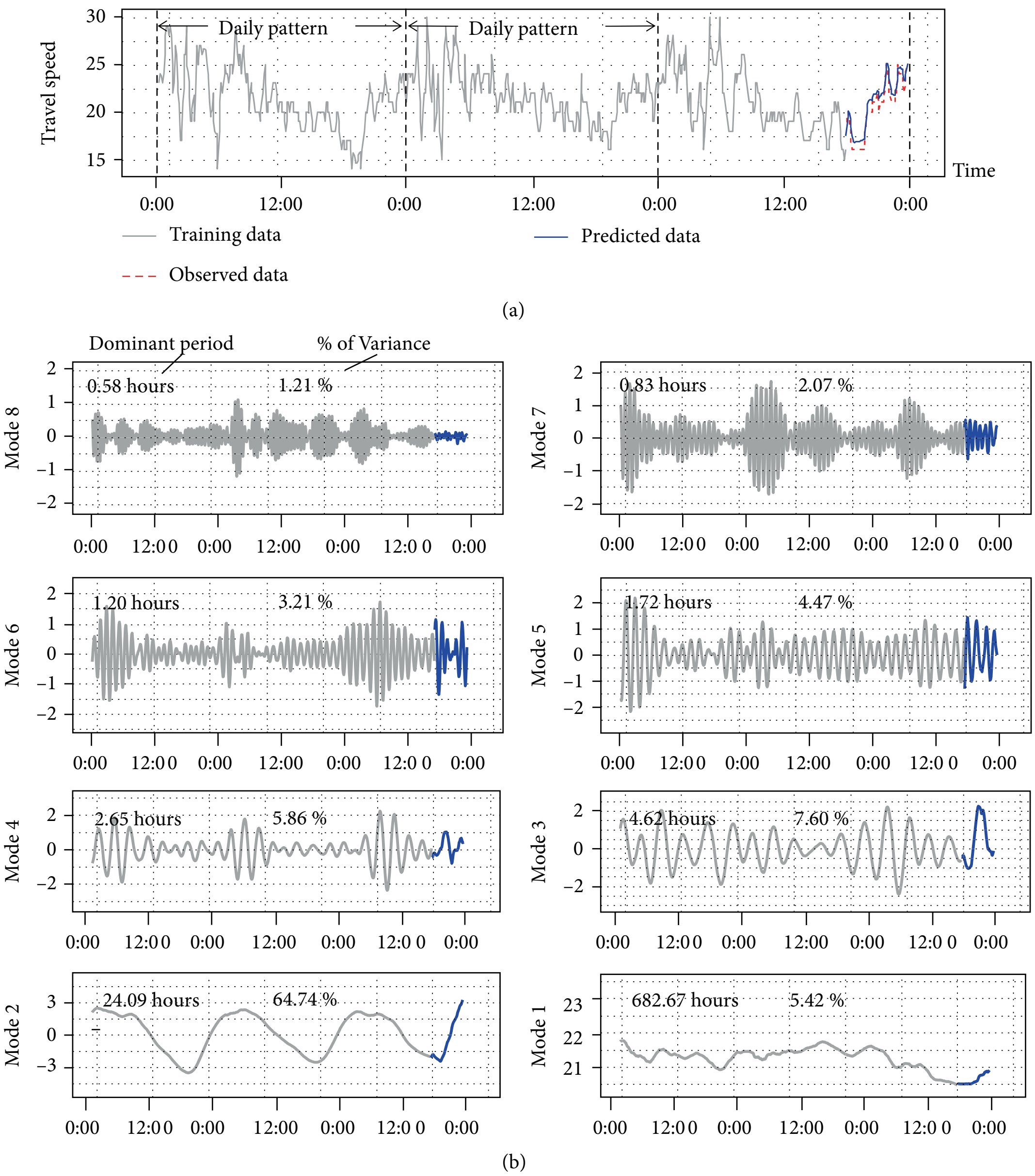


Figure 6: (a) Three days of travel speed data from Link 523; (b) Its corresponding decomposed modes with central frequencies.

by the fact that mode 3 with the commuting trend was statistically insignificant in all situations of congested traffic condition. The results showed how the proposed model in the study could achieve better performance in congested traffic condition.

5. Discussion and Conclusions

The purpose of this study was to develop a hybrid model to predict and analyze the urban travel speed using VMD that decomposes the data into multi-scale oscillatory modes. Travel speed data were predicted by the "divide and conquer" strategy, which respectively predicts the modes and summed up to represent the predicted travel speed. The modes were predicted more accurately than their original speed since they were mitigating the complexity of the original data by extracting the oscillatory pattern. In the performance evaluation, the proposed hybrid method, VMD-ANN, outperformed the existing machine learning models, SVM and ANN, and this improvement in performance was greater in congested traffic condition where it was difficult for the existing method to predict.

We also analyzed the correlation between the performance of the model, and the spectral and statistical properties of the decomposed modes to provide an understanding of the inherent ability to forecast travel speed in the urban networks. Our analyses showed that the more the variance of nondaily patterns were explained through the multiscale modes, the easier

Table 2: Central period of modes of different K from speed data and reclassified modes with their spectral and statistical properties.

	Best K for link-day units			Reclassified modes				
	8	10	12					
	Period (hours) (min – max)	Period (hours) (min – max)	Period (hours) (min – max)		N	Period (hours) (min – max)	CV	% of variance (min – max)
Mode 1	–	–	–	Mode 1	80	–	–	1–15
Mode 2	11.9–28.9	12.3–29.2	12.0–29.2	Mode 2	70	23.0–30.0	0.07	44–86
Mode 3	3.9–12.7	4.0–12.7	5.9–11.4	Mode 3	38	10.5–13.0	0.05	11–56
Mode 4	2.6–4.8	2.7–6.0	3.4–5.8	Mode 4	32	7.0–9.0	0.04	11–15
Mode 5	1.7–2.7	1.8–3.5	2.6–3.4	Mode 5	52	4.5–6.5	0.10	2–16
Mode 6	1.2–1.9	1.5–2.0	1.8–2.2	Mode 6	120	3.5–4.5	0.21	1–19
Mode 7	0.7–1.2	1.1–1.4	1.4–1.7	Mode 7	156	2.5–3.5	0.19	1–11
Mode 8	0.5–0.8	0.8–1.1	1.1–1.2	Mode 8	196	0–1.5	0.25	0–3
Mode 9	–	0.6–0.8	0.8–0.9					
Mode 10	–	0.3–0.6	0.6–0.7					
Mode 11	–	–	0.4–0.6					
Mode 12	–	–	0.3–0.5					

Table 3: Results of correlation analysis for properties of modes and predicted performances.

		Performance (MAPE)					
		VMD_ANN		SVM		Percent Difference	
		Overall	Congestion	Overall	Congestion	Overall	Congestion
Performance (MAPE)	SVM_Overall	0.959**	0.795**	1	0.835**	−0.178	−0.164
	SVM_Cong	0.738**	0.858**	0.835**	1	−0.289**	−0.292**
EP in overall conditions	Mode 1	−0.494**	−0.350**	−0.487**	−0.321**	−0.011	−0.002
	Mode 2	0.454**	0.304*	0.464**	0.287*	-0.165	-0.009
	Mode 3	−0.518**	−0.277	−0.466**	−0.264	−0.238	0.086
	Mode 4	0.251	0.155	0.280	0.174	-0.092	0.008
	Mode 5	−0.623**	−0.395**	−0.651**	−0.442**	0.127	0.012
	Mode 6	−0.503**	−0.283**	−0.461**	−0.228*	−0.120	0.026
	Mode 7	−0.462**	−0.287**	−0.426**	−0.233**	−0.125	−0.040
	Mode 8	−0.545**	−0.377**	−0.520**	−0.317**	−0.035	−0.028
EP in congested conditions	Mode 1	−0.442**	−0.271*	−0.441**	−0.267*	0.020	0.045
	Mode 2	0.255*	0.167	0.284*	0.173	−0.223	−0.107
	Mode 3	0.038	−0.092	0.078	−0.001	−0.204	−0.081
	Mode 4	0.306	0.394*	0.307	0.339*	−0.228	0.173
	Mode 5	−0.436**	−0.286*	−0.463**	−0.349*	0.090	0.058
	Mode 6	−0.467**	−0.239**	−0.425**	−0.213*	−0.148	0.066
	Mode 7	−0.451**	−0.245**	−0.423**	−0.223**	−0.126	0.007
	Mode 8	−0.446**	−0.257**	−0.416**	−0.238**	−0.118	0.049

Note: $^{*}p < 0.05$; $^{**}p < 0.01$.

it was to predict the speed. This tendency also was significant even in unstable congestion. The possibility to forecast each link-day, therefore, can be measured by the explained variance of modes with a nondaily pattern and used to evaluate the robustness of models for heterogeneous urban traffic dynamics. Also, it can be used in the studies of influence factors that affect the complexity of urban traffic conditions, such as the geometric condition, traffic signals, and the spatiotemporal correlation.

Further studies should be able to extend both the performance enhancement of the hybrid model and the detailed analysis of urban traffic dynamics. In order to emphasize the practical uses of the model, we simplified the prediction process into only three stages. i.e., decomposition, modes prediction, and summation. However, a sophisticated model to predict the modes [24], and to optimize the reconstruction [28] can improve the performance of the hybrid models. Although our model predicts the travel speed only using the travel speed data from the target link, the neighbor links surrounding it can affect the forecastability of the target link which is explained by the statistical and spectral properties of the modes. Spatiotemporal effects from the neighbor links and its impact on multiscale modes need to be discussed for a better understanding of traffic dynamics. The physical meaning of the modes other than the typical daily and commuting patterns cannot be explained in this study since, we identified the modes based only on their central periods. Additional research will be required to complement the interpretability of modes by introducing a statistical method, such as the independent components analysis [34], and by analyzing microscopic trajectory data associated with the travel speed data [35].

Data Availability

The data used in this research are provided by the Trlab research programme conducted at the Seoul National University, Seoul, Republic of Korea. The data are available when readers ask the authors for academic purposes.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the publication of this paper.

Acknowledgments

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT & Future Planning (2019R1H1A1080045).

References

[1] E. I. Vlahogianni, M. G. Karlaftis, and J. C. Golias, "Short-term traffic forecasting: where we are and where we're going," *Transportation Research Part C: Emerging Technologies*, vol. 43, pp. 3–19, 2014.

[2] F. Zheng and H. Van Zuylen, "Urban link travel time estimation based on sparse probe vehicle data," *Transportation Research Part C: Emerging Technologies*, vol. 111, pp. 145–157, 2013.

[3] Y. Wang and M. Papageorgiou, "Real-time freeway traffic state estimation based on extended kalman filter: a general approach," *Transportation Research Part B: Methodological*, vol. 39, pp. 141–167, 2005.

[4] D. Billings and Y. Jiann-Shiou, "Application of the ARIMA models to urban roadway travel time prediction-a case study," in *IEEE International Conference on Systems, Man and Cybernetics*, IEEE, pp. 2529–2534, Taipei, Taiwan, 2006.

[5] M. C. Tan, S. C. Wong, J. M. Xu, Z. R. Guan, and P. Zhang, "An aggregation approach to short-term traffic flow prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 10, pp. 60–69, 2009.

[6] S. Oh, Y. J. Byon, K. Jang, and H. Yeo, "Short-term travel-time prediction on highway: a review on model-based approach," *KSCE Journal of Civil Engineering*, vol. 22, pp. 298–310, 2018.

[7] E. Vlahogianni and M. Karlaftis, "Testing and comparing neural network and statistical approaches for predicting transportation time series," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2399, pp. 9–22, 2013.

[8] C. H. Wu, J. M. Ho, and D. T. Lee, "Travel-time prediction with support vector regression," *IEEE Transactions on Intelligent Transportation Systems*, vol. 5, pp. 276–281, 2004.

[9] B. Yao, C. Chen, Q. Cao et al., "Short-term traffic speed prediction for an urban corridor," *Computer-Aided Civil and Infrastructure Engineering*, vol. 32, pp. 154–169, 2017.

[10] J. Myung, D.-K. Kim, S.-Y. Kho, and C.-H. Park, "Travel time prediction using k-nearest neighbor method with combined data from vehicle detector system and automatic toll collection system," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2256, pp. 51–59, 2011.

[11] H. H. Chang and B. J. Yoon, "High-speed data-driven methodology for real-time traffic flow predictions: practical applications of ITS," *Journal of Advanced Transportation*, vol. 2018, pp. 1–11, 2018.

[12] J. Zhao, Y. Gao, J. Tang, L. Zhu, and J. Ma, "Highway travel time prediction using sparse tensor completion tactics and k-nearest neighbor pattern matching method," *Journal of Advanced Transportation*, vol. 20181, no. 16, 2018.

[13] P. Cai, Y. Wang, G. Lu, P. Chen, C. Ding, and J. Sun, "A spatiotemporal correlative k-nearest neighbor model for short-term traffic multistep forecasting," *Transportation Research Part C: Emerging Technologies*, vol. 62, pp. 21–34, 2016.

[14] G. Fusco, C. Colombaroni, and N. Isaenko, "Short-term speed predictions exploiting big data on large urban road networks," *Transportation Research Part C: Emerging Technologies*, vol. 73, pp. 183–201, 2016.

[15] M. T. Asif, J. Dauwels, C. Y. Goh et al., "Spatiotemporal patterns in large-scale traffic speed prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, pp. 794–804, 2014.

[16] T. Cheng, J. Haworth, and J. Wang, "Spatio-temporal autocorrelation of road network data," *Journal of Geographical Systems*, vol. 14, pp. 389–413, 2012.

[17] C. Oh and S. Park, "Investigating the effects of daily travel time patterns on short-term prediction," *KSCE Journal of Civil Engineering*, vol. 15, pp. 1263–1272, 2012.

[18] K. Jha, J. P. Wikander, W. L. Eisele, M. W. Burries, and D. L. Schrank, "Estimating freeway route travel time reliability from data on component links and associated cost implications," *International Journal of Urban Sciences*, vol. 22, pp. 414–430, 2018.

[19] N. G. Polson and V. O. Sokolov, "Deep learning for short-term traffic flow prediction," *Transportation Research Part C: Emerging Technologies*, vol. 79, pp. 1–17, 2017.

[20] H.-C. Park, D.-K. Kim, and S.-Y. Kho, "Bayesian network for freeway traffic state prediction," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2672, no. 45, pp. 124–135, 2018.

[21] I. Guyon and A. Elisseeff, "An introduction to feature extraction," *Feature Extraction: Foundations and Applications, Springer, Berlin*, vol. 207, pp. 1–25, 2006.

[22] Y. Wei and M. C. Chen, "Forecasting the short-term metro passenger flow with empirical mode decomposition and neural networks," *Transportation Research Part C: Emerging Technologies*, vol. 21, pp. 148–162, 2012.

[23] C. Dong, C. Shao, S. H. Richards, and L. D. Han, "Flow rate and time mean speed predictions for the urban freeway network

using state space models," *Transportation Research Part C: Emerging Technologies*, vol. 43, pp. 20–32, 2014.

[24] D. Wang, H. Luo, O. Grunder, and Y. Lin, "Multi-step ahead wind speed forecasting using an improved wavelet neural network combining variational mode decomposition and phase space reconstruction," *Renewable Energy*, vol. 113, pp. 1345–1358, 2017.

[25] Y. Zhang, Y. Zhang, and A. Haghani, "A hybrid short-term traffic flow forecasting method based on spectral analysis and statistical volatility model," *Transportation Research Part C: Emerging Technologies*, vol. 43, pp. 65–78, 2014.

[26] Y. Sun, B. Leng, and W. Guan, "A novel wavelet-SVM short-time passenger flow prediction in Beijing subway system," *Neurocomputing*, vol. 166, pp. 109–121, 2015.

[27] H. Wang, L. Liu, Z. Qian, H. Wei, and S. Dong, "Empirical mode decomposition–autoregressive integrated moving average: hybrid short-term traffic speed prediction model," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2460, no. 1, pp. 66–76, 2014.

[28] X. Jiang, L. Zhang, and M. X. Chen, "Short-term forecasting of high-speed rail demand: a hybrid approach combining ensemble empirical mode decomposition and gray support vector machine with real-world applications in China," *Transportation Research Part C: Emerging Technologies*, vol. 44, pp. 110–127, 2014.

[29] K. Dragomiretskiy and D. Zosso, "Variational mode decomposition," *IEEE Transactions on Signal Processing*, vol. 62, pp. 531–544, 2014.

[30] S. Kumar, *Neural Networks: A Classroom Approach*, McGraw-Hill Education, New Delhi, 2005.

[31] M. G. Karlaftis and E. I. Vlahogianni, "Statistical methods versus neural networks in transportation research: differences, similarities and some insights," *Transportation Research Part C: Emerging Technologies*, vol. 19, no. 3, pp. 387–399, 2011.

[32] M. Lorenz and L. Elefteriadou, "Defining freeway capacity as function of breakdown probability," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1776, pp. 43–51, 2001.

[33] H. J. Kim, H. Mahmassani, and J. Dong, "Likelihood and duration of flow breakdown: modelling the effect of weather," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2188, pp. 19–28, 2010.

[34] E. Jianwei, Y. Bao, and J. Ye, "Crude oil price analysis and forecasting based on variational mode decomposition and independent component analysis," *Physica A: Statistical Mechanics and its Applications*, vol. 484, pp. 412–427, 2017.

[35] E.-J. Kim, H.-C. Park, S.-W. Ham, S.-Y. Kho, and D.-K. Kim, "Extracting Vehicle Trajectories Using Unmanned Aerial Vehicles in Congested Traffic Conditions," *Journal of Advanced Transportation*, vol. 2019, pp. 1–16, 2019.