Short-Term Traffic Flow Prediction With Wavelet and Multi-Dimensional Taylor Network Model

Shanliang Zhu, Yu Zhao, Yanjie Zhang, Qingling Li, Wenwu Wang, *Senior Member, IEEE*, and Shuguo Yang

***Abstract*—Accurate prediction of the traffic state has received sustained attention for its ability to provide the anticipatory traffic condition required for people's travel and traffic management. In this paper, we propose a novel short-term traffic flow prediction method based on wavelet transform (WT) and multi-dimensional Taylor network (MTN), which is named as W-MTN. Influenced by the short-term noise disturbance in traffic flow information, the WT is employed to improve prediction accuracy by decomposing the time series of traffic flow. The MTN model, which exploits polynomials to approximate the unknown nonlinear function, makes full use of periodicity and temporal feature without transcendental knowledge and mechanism of the system to be predicted. Our proposed W-MTN model is evaluated on the traffic flow information in a certain area of Shenzhen, China. The experimental results indicate that the proposed W-MTN model offers better prediction performance and temporal correlation, as compared with the corresponding models in the known literature. In addition, the proposed model shows good robustness and generalization ability, when considering data from the different days and locations.**

***Index Terms*—Traffic flow prediction, multi-dimensional Taylor network, wavelet transform.**

I. Introduction

ACCURATE prediction of traffic flow information is of great importance in modern transportation systems. It is a booster for numerous intelligent applications that require reliable future traffic information. For example, traffic prediction is a significant reference for dynamic path planning, which allows travelers to choose better routes and avoid traffic congestion [1]. Predicting where and when congestion will happen plays a vital role in traffic management and transportation [2]. With its huge potential on numerous applications, traffic prediction has become an important research topic [3]. Traffic flow data is characterized by strong nonlinearity, time-variation, and randomness [4], and is thus difficult to predict. However, short-term rather than long-term prediction of traffic flow is possible, and is thus the focus of this paper. The key of this study is to improve the accuracy of short-term prediction.

The problem of short-term traffic flow prediction has been studied for decades, and many prediction algorithms have been developed, such as the Kalman filter [5], the time series model [6], the K-nearest neighbor algorithm [7], the support vector machine (SVM) [8], [9], the Bayesian theory [10], the dynamic tensor completion [11], and various neural network (NN) [4], [12]–[14] approaches. Among those prediction techniques, the SVM and the NNs are popular methods used by many researchers due to their unique characteristics such as nonlinear approximation, giving desirable prediction results. The SVM has good self-learning and nonlinear prediction ability, and it can provide better prediction accuracy in the case of limited training samples. The NNs not only have powerful ability in function approximation and pattern classification, but also has good self-organization, self-adaptability, and incorrectness tolerance. However, they also have several limitations such as the difficulty in determining appropriate kernel functions for the SVM [15], and choosing an appropriate threshold for the NNs [16] to prevent them from getting stuck in local minimum. Improper parameter selection will adversely affect the performance of those methods.

To overcome the limitations of those models, this paper proposes a simple but effective prediction method based on multi-dimensional Taylor network (MTN), which is often used to solve the problem of nonlinear system identification and prediction [17]. The MTN [18], proposed by Yan, is a machine learning method for short-term prediction based on nonlinear time series relationship of observed data, which has been applied in several areas. For example, the MTN has been used in [19] to predict the changes in financial and stock markets and in [20], [21] for studying a range of nonlinear system control problems. In these works, the MTN has been shown to give excellent results.

Given a certain location, the MTN can be used to provide an accurate prediction of traffic status without transcendental knowledge and mechanism of the system. With the original traffic information given to the input layer, the MTN obtains the description of the system state equation by polynomial array approximation. This offers advantages over the SVM and the NN models, which are unable to express explicitly the impact of historical observation data on forecasting data, because these models can only obtain implicit nonlinear state relations. To this end, we concentrate on the explicit expression of the nonlinear time-varying system by the polynomial approximants of the MTN, and employ a conjugate gradient method [18] to learn the parameters of the MTN model. Therefore, this paper proposes a method for system dynamics prediction based on the WT and the MTN, which is named as the W-MTN model. With this model, we are able to obtain the explicit state equations for the nonlinear systems using only sample data of traffic information without requiring prior knowledge of system parameters, and to also capture the existence of short-term noise disturbance in traffic flow information. We employ the WT to decompose the sample data

Manuscript received April 10, 2019; revised September 30, 2019; accepted February 24, 2020. This work was supported in part by the Fundamental Research Funds for the Central Universities under Grant 2242019k30017, in part by the Natural Science Foundation of Shandong Province under Grant zr2019mee030, and in part by the Project of Shandong Province Instructing Ability Improvement for Graduate Supervisor under Grant SDYY17042. The Associate Editor for this article was N. Bekiaris-Liberis. *(Corresponding author: Shuguo Yang.)*

Shanliang Zhu is with the School of Mathematics and Physics, Qingdao University of Science and Technology, Qingdao 266061, China, and also with the College of Electromechanical Engineering, Qingdao University of Science and Technology, Qingdao 266061, China (e-mail: zhushanliang@qust.edu.cn).

Yu Zhao is with the College of Economics and Management, Qingdao University of Science and Technology, Qingdao 266061, China (e-mail: zhaoyu_yep@163.com).

Yanjie Zhang is with the School of Information Science and Technology, Qingdao University of Science and Technology, Qingdao 266061, China (e-mail: dogac1234@163.com).

Qingling Li is with the College of Electromechanical Engineering, Qingdao University of Science and Technology, Qingdao 266061, China (e-mail: qustmaths@126.com).

Wenwu Wang is with the Department of Electrical and Electronic Engineering, University of Surrey, Guildford GU2 7XH, U.K. (e-mail: w.wang@surrey.ac.uk).

Shuguo Yang is with the School of Mathematics and Physics, Qingdao University of Science and Technology, Qingdao 266061, China (e-mail: ysg_2005@163.com).

Digital Object Identifier 10.1109/TITS.2020.2977610

into several subsequences, in order to effectively reduce the noise disturbance of the traffic flow. On the basis of that we establish the MTN model for short-term traffic flow prediction, which expresses explicitly the nonlinear relationship between historical traffic data and future data. The experimental results show that predictions based on the proposed model are more accurate than several baseline methods, and the model also has good generalization capabilities in short-term traffic flow prediction.

The remainder of this paper is organized as follows. Section II introduces the proposed W-MTN model. The traffic information in a certain area of Shenzhen is used in the experiments in Section III to illustrate the performance of our method. Finally, some conclusions are drawn in Section IV.

II. Methodology

In this section we briefly describe the correlation theory of the W-MTN and the method for determining the parameters. See the literatures [17], [18], [22]–[25] for details.

A. Multi-Dimensional Taylor Network

The MTN with a three-layer topological structure is shown in Fig. 1. The input of the MTN is $\boldsymbol{x}(t) = [x_1(t), x_2(t), \ldots, x_n(t)]^T$, which is reconstructed by the observation data in the n dimensional state space. The middle layer is the polynomial combination of the inputs and corresponding connection weights. The output of the MTN is $\boldsymbol{x}(t+1)$, which is the future status. An unknown smooth nonlinear function $\boldsymbol{f}(t)$ can be approximated by the MTN, as follows,

$$x_n(t+1) = f_n(x_1(t), x_2(t), \ldots, x_n(t)) = \boldsymbol{\omega}_n^T \boldsymbol{P}_m(t), \tag{1}$$

where $\boldsymbol{\omega}_n = (\omega_{n,1}, \omega_{n,2}, \ldots, \omega_{n,l})^T$ is the weight vector, and $\omega_{n,\kappa}$, $\kappa = 1, 2, \ldots, l$ is the weight of the variable-product items. m is the highest power of the middle layer, and the value of l can be determined by m and n. The elements of the vector $\boldsymbol{P}_m(t)$ are $\prod_{i,j=1}^{n} x_i^{\lambda_i} x_j^{\lambda_j}$, where λ_i, λ_j are nonnegative integers and satisfy $1 \leq \lambda_i + \lambda_j \leq m$.

B. Parameter Learning Algorithm of the MTN

The parameter learning algorithm is aimed to determine the parameters of the product items, which establish the dynamic relationship between near-term traffic flow information and future traffic flow information.

To find the optimal parameters $\boldsymbol{\omega}_n^T = (\omega_{n,1}, \omega_{n,2}, \ldots, \omega_{n,l})^T$, we use the minimum error square sum given by

$$E = \frac{1}{2} \sum_{t=1}^{s} (x(t+1) - \boldsymbol{\omega}^T \boldsymbol{P}_t)^2, \tag{2}$$

where $\{(\boldsymbol{x}(t), x(t+1))\}_{t=1}^{s}$ is the sample set, and $\hat{x}(t+1) = \boldsymbol{\omega}^T \boldsymbol{P}_t$ is the prediction value of the model.

This is an unconstrained optimization problem for quadratic function, which can be solved by a conjugate gradient method. Let $\boldsymbol{P} = (\boldsymbol{P}_1, \boldsymbol{P}_2, \ldots, \boldsymbol{P}_s)$, we seek the partial derivation of (2) with respect to $\boldsymbol{P}$ to obtain $\nabla E(\boldsymbol{\omega}) = \boldsymbol{P}\boldsymbol{P}^T \boldsymbol{\omega} - \boldsymbol{P}x$.

Assuming the initial conjugate vector $\boldsymbol{u}_0$ is taken as the negative gradient $-\nabla E(\boldsymbol{\omega}_0)$ at the random initial point $\boldsymbol{\omega}_0$. For each iteration, the value of $\boldsymbol{\omega}$ is corrected by the negative gradient direction with the sample data. This leads to the following iteration

$$\boldsymbol{\omega}_{k+1} = \boldsymbol{\omega}_k + \vartheta_k \boldsymbol{u}_k, \tag{3}$$

where ϑ_k is a step size, and $\boldsymbol{u}_k$ is the conjugate direction.

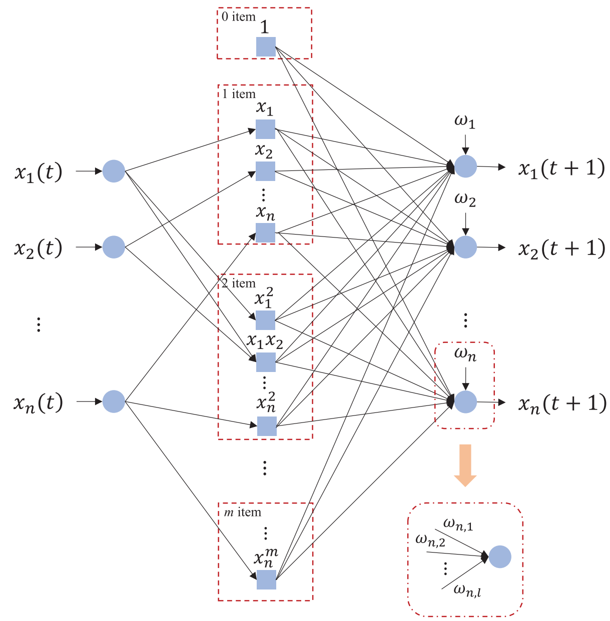


Fig. 1. The structure of the MTN.

According to the steepest descent method, we have

$$\vartheta_k = \frac{\nabla E(\boldsymbol{\omega}_k)^T \nabla E(\boldsymbol{\omega}_k)}{\boldsymbol{u}_k^T (\boldsymbol{P}\boldsymbol{P}^T) \boldsymbol{u}_k}. \tag{4}$$

The conjugate direction can be obtained as

$$\boldsymbol{u}_k = \begin{cases} -\nabla E(\boldsymbol{\omega}_k), & k = 0, \\ -\nabla E(\boldsymbol{\omega}_k) + \iota_{k-1} \boldsymbol{u}_{k-1}, & k \geq 1, \end{cases} \tag{5}$$

where

$$\iota_{k-1} = \frac{\nabla E(\boldsymbol{\omega}_k)^T (\boldsymbol{P}\boldsymbol{P}^T) \boldsymbol{u}_{k-1}}{\boldsymbol{u}_{k-1}^T (\boldsymbol{P}\boldsymbol{P}^T) \boldsymbol{u}_{k-1}}. \tag{6}$$

The main procedures of this parameter learning algorithm are summarized as follows:

step1: Set the parameters of the MTN, i.e., the dimension n, the highest degree m, the initial parameter $\boldsymbol{\omega}_0$, the error threshold ε, and the iterative number max_k. Input the training sample set and set the iterative number $k = 0$.

step2: Calculate the objective function $\min_{\vartheta \geq 0} E(\boldsymbol{\omega}_k + \vartheta \boldsymbol{u}_k)$.

step3: Check whether the stop criterion is met. If $\left\| \nabla E(\boldsymbol{\omega}_{k+1}) \right\| < \varepsilon$, stop the iterations and $\boldsymbol{\omega}_{k+1}$ is the optimal parameters; otherwise, turn to *step4*.

step4: Calculate ϑ_k and $\boldsymbol{u}_{k+1}$.

step5: Let $k = k + 1$.

step6: Determine the maximum number of iterations max_k by the error threshold ε. If $k > max\_k$, stop the process and output the optimal parameters; otherwise, turn to *step2*.

C. Wavelet Transformation

In recent years, the WT has been applied in time scales analysis of the variables in traffic prediction systems [22]–[24], [25] for denoising the traffic data. This paper uses the Mallat algorithm [26] to decompose the original signal into high-frequency part D_1 and low-frequency part A_1, in which A_1 can be further decomposed into D_2 and A_2, and so on. The original signal X is finally decomposed into:

$$X = D_1 + D_2 + \cdot + D_N + A_N, \tag{7}$$

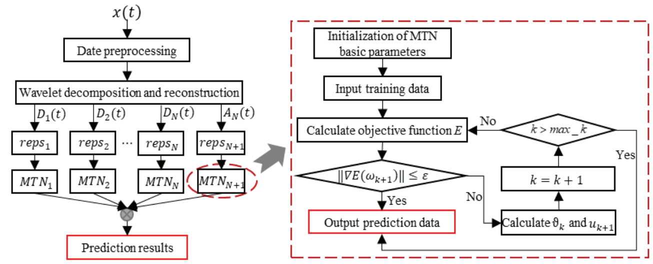


Fig. 2. The process of the W-MTN model.

where N is the decomposition level. The determination of wavelet structure is carried out by trying various mother wavelet functions and the scale parameters.

To summarize, the process for optimizing the proposed W-MTN prediction model is shown in Fig. 2.

III. Experimental Results

In this section, the traffic flow information from Traffic Police Bureau of Shenzhen Public Security Bureau is used to evaluate the proposed W-MTN model by comparing it to several baseline prediction methods. All these tests are implemented by using a Lenovo R720 (CPU: i5-7300HQ, GPU: NVIDIA GeForce GTX 1050 2 GB).

A. Data Description

Traffic flow data collected from Jingtian Road East in Futian district of Shenzhen, China are used in this experiment. Jingtian Road East is a commercial street in the centre of Shenzhen from Jingtian North Third Street to Lianhua Road. This road has two lanes of traffic in both directions and the length is about 412 meters. As pointed out in [27], an update window of 10 minutes should be used to ensure a satisfactory performance for updating the traffic flow information. Consequently, the average traffic speed excluding waiting time of traffic lights in 25-31 March 2018 was obtained for each 10-min time interval in the whole road. We had 144 records for each day and 1008 data records overall. In addition, the model was also evaluated by the data of the same time period in different roads.

B. Metrics of Performance

The mean absolute error (MAE), the mean absolute percentage error (MAPE), and the root mean square error (RMSE) were used to evaluate the performance. We also employ the average correlation (AC) to evaluate the model for temporal distribution prediction. The prediction accuracy of the model can be measured more comprehensively through those performance metrics. The metrics are expressed as follows:

$$MAE = \frac{\sum_{i=1}^{144} |P_i - T_i|}{144}, \tag{8}$$

$$MAPE = \sum_{i=1}^{144} |(P_i - T_i)/T_i| \times \frac{100}{144}, \tag{9}$$

$$RMSE = \sqrt{\frac{\sum\limits_{i=1}^{144} |P_i - T_i|^2}{144}}, \tag{10}$$

$$AC = \frac{1}{I_t} \sum_{t=1}^{I} Corr(P_t', T_t'), \tag{11}$$

where P_i and T_i are the prediction value and the true value, respectively. P_t' and T_t' represent respectively the predicted traffic flow vector and the actual traffic flow vector at time point t, I is the time window size and I_t is the number of predicted vectors.

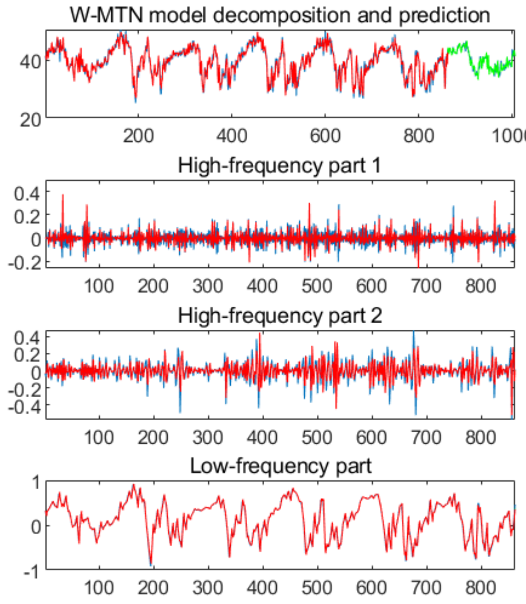


Fig. 3. The W-MTN model decomposition and prediction.

C. Determination of the Parameters

In order to select the most appropriate parameters for the W-MTN model, the dataset was divided to training and testing subsets. The length of the training data is 864 and the length of the prediction data is 144, which means that 6 days historical data are used to implement the traffic flow prediction for next one day.

According to the preset error threshold for stopping the iteration $\varepsilon = 10^{-4}$, the maximum number of iterations $max\_k = 100$ was determined empirically during performance testing of the model. Different combinations of the maximum number of expansion m and the spatial dimension n were tested in order to find out the appropriate parameters. Finally, $m = n = 3$ were chosen among different combinations based on their error metrics values and the running time. Determining the parameters for the WT consists of selecting the best mother wavelet function and the decomposition level. Various orthogonal functions including Symlet, Coifmann, and Daubiches, were compared under same conditions, and the Db 3 mother wavelet function was chosen in terms of their error metrics values. Different decomposed parts were tested to obtain the best level of decomposition. The wavelet decomposition level N is set as 2, according to their error metrics values and the running time. Fig. 3 shows the results of decomposing the historical data into two high-frequency parts and a low-frequency part by the WT. The prediction results of the traffic flow at 144 time points of a day are shown in Fig. 4.

D. Model Accuracy Comparison

The proposed W-MTN model is compared with several prediction methods including the ARIMA [6], the SVM [9], the ANN [12], the SAE model [14], and the WT-ANN [23]. Those models consist of prediction techniques with different structures and depths. For fair comparisons, we make a slight modification so as to adapt to the test dataset. To avoid overfitting of the ANN and the SAE in the training process, we assign 15% training data to a validation set. A three-layer ANN model with three neurons in the first layer, three neurons in the second layer and one neuron in the third layer was chosen. The number of iterations in neural network training is set as 1000. For the SVM, a simple linear kernel function is used, in which

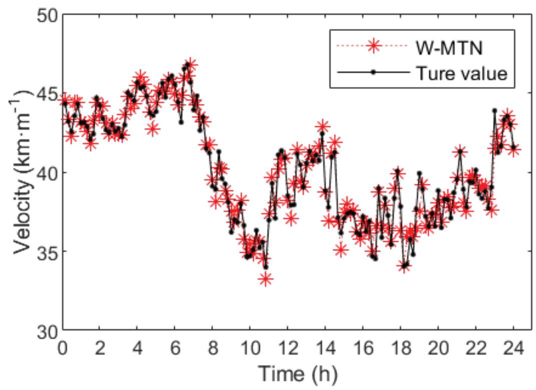


Fig. 4. The prediction results of the model at 144 time points of a day.

TABLE I
PERFORMANCE METRICS OF VARIOUS MODELS

Model	MAE	MAPE (%)	RMSE	AC
W-MTN	**0.5218**	**1.3234**	**0.6901**	**0.9784**
WT-ANN	0.6350	1.6025	0.8105	0.9704
ANN	0.6892	1.7362	0.8790	0.9646
SAE	1.3885	3.5361	1.7623	0.8818
SVM	0.6752	1.7002	1.0966	0.9489
ARIMA	1.2521	3.1603	1.6232	0.8822

the dimension of the feature space and the input space is the same, that is, 130, with fewer parameters and faster speed. The ARIMA model is implemented on a linear regression framework, which is the first-order difference including a first-order autoregressive component and a first-order moving average component. TABLE I shows the evaluation results of these models.

It can be observed that the W-MTN model outperforms the baseline prediction models on all the evaluation metrics. This shows that the proposed model gives promising performance in extracting temporal information of the traffic flow. Among the baseline prediction models, the WT-ANN performs better than the ANN, the SAE, the SVM, and the ARIMA. Nevertheless, the W-MTN model outperforms the WT-ANN with improvements of 0.1132, 0.2791%, 0.1204, and 0.008 in terms of the MAE, the MAPE, the RMSE, and the AC, respectively. Comparing the error metrics of the WT-ANN with the ANN shows that denoising the traffic data has significant effect on improving the prediction accuracy.

As shown in TABLE I, although these models all obtain satisfactory forecasting results in short-term traffic prediction, the SAE model offers lower prediction accuracy compared with the other models. One of the reasons may be that the amount of test data is not enough for deep learning approaches such as the SAE which normally requires a large amount of data to prevent overfitting.

E. Testing the W-MTN in Different Days and Locations

The short-term prediction performance of the proposed model may vary with different conditions, such as accidents, weather conditions, different days and locations. In this section, we only analyze the prediction performance of the model in different days and locations without considering other factors.

In order to evaluate the performance of the model in different days, Friday and Saturday are chosen, which represent respectively workdays and weekdays. Their MAPE values are 2.6% and 1.3%, which clearly exhibit satisfactory prediction ability. The MAPE per hour is shown in Fig. 5.

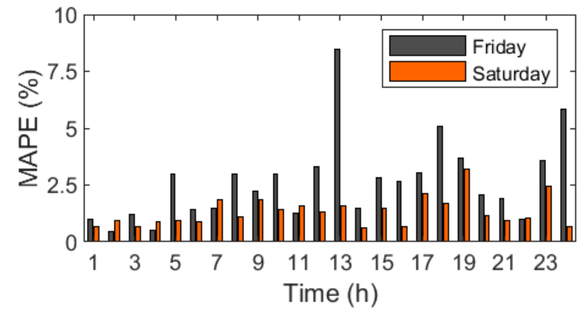


Fig. 5. Performance comparison of the W-MTN model in different days.

As shown in Fig. 5, the MAPE of about 90 percent of the time is less than 3%.The results indicate the model performs well not only on the workdays, but also on the weekdays, which show that the model has a good generalization ability.

In addition, Xiangmei Road, which is another road with different features in Futian district of Shenzhen, China, was used for testing the performance of the proposed model. Traffic flow data in the same time period was collected from this road located from Municipal Services Building to a residential area named Jingfa Garden. This road has four lanes of traffic in both directions and the length is about 520 meters. The results of the road are 0.5590, 1.6184%, and 0.7131 in the terms of the MAE, the MAPE and the RMSE, which are higher than the results of Jingtian Road East. The prediction accuracy of the proposed model decrease slightly, possibly due to the influence of the parameters of the proposed model and random events, but the model still yields highly reliable results, which show that the proposed model has good accuracy and robustness.

IV. CONCLUSION AND FUTURE WORK

We have presented a new traffic flow prediction model W-MTN for the short-term traffic flow prediction by combining the time series prediction of a wavelet transform and multi-dimensional Taylor network. The WT helps improve prediction accuracy by reducing the effect of the short-term uncertainty factors and stochastic noise. The MTN model identifies nonlinear relations among the traffic flow information and temporal feature.

(I) The model exploits polynomials for nonlinear approximation, which provides advantages over the SVM and the NNs in several aspects, including structural simplicity, high training efficiency, strong learning ability, and fast convergence with only a few control parameters.

(II) The MTN is a topological structure for learning a nonlinear function relationship via an array of polynomials, which is able to use historical data to predict the traffic flow without transcendental knowledge and mechanisms. The model has good generalization ability and approximation accuracy.

(III) The use of the WT to decompose the time series of the traffic flow helps reduce the adverse impact of random noise on the prediction accuracy.

The proposed model not only gives excellent prediction performance, but also offers advantages over traditional neural networks in the sense that it has a closed-form expression, which is different from the "black-box" model used in traditional neural networks. In contrast to most works in the field, which focus on improving prediction accuracy, our proposed models also offer improved efficiency and robustness, based on the traffic data captured in a certain area of Shenzhen, China. In the future, it would be meaningful to study further the performance of the MTN model for traffic flow, traffic density and other transportation related fields such as intelligent traffic control and handling of random events.

REFERENCES

[1] C. Guo, D. Li, G. Zhang, and M. Zhai, "Real-time path planning in urban area via VANET-assisted traffic information sharing," *IEEE Trans. Veh. Technol.*, vol. 67, no. 7, pp. 5635–5649, Jul. 2018.

[2] Q.-J. Kong, Y. Xu, S. Lin, D. Wen, F. Zhu, and Y. Liu, "UTN-model-based traffic flow prediction for parallel-transportation management systems," *IEEE Trans. Intell. Transp. Syst.*, vol. 14, no. 3, pp. 1541–1547, Sep. 2013.

[3] Y. Wu, H. Tan, L. Qin, B. Ran, and Z. Jiang, "A hybrid deep learning based traffic flow prediction method and its understanding," *Transp. Res. C, Emerg. Technol.*, vol. 90, pp. 166–180, May 2018.

[4] K. Y. Chan, T. S. Dillon, J. Singh, and E. Chang, "Neural-network-based models for short-term traffic flow forecasting using a hybrid exponential smoothing and levenberg-marquardt algorithm," *IEEE Trans. Intell. Transp. Syst.*, vol. 13, no. 2, pp. 644–654, Jun. 2012.

[5] J. Guo, W. Huang, and B. M. Williams, "Adaptive Kalman filter approach for stochastic short-term traffic flow rate prediction and uncertainty quantification," *Transp. Res. C, Emerg. Technol.*, vol. 43, pp. 50–64, Jun. 2014.

[6] S. V. Kumar and L. Vanajakshi, "Short-term traffic flow prediction using seasonal ARIMA model with limited input data," *Eur. Transp. Res. Rev.*, vol. 7, no. 3, pp. 1–9, Jun. 2015.

[7] S. Oh, Y.-J. Byon, and H. Yeo, "Improvement of search strategy with K-Nearest neighbors approach for traffic state prediction," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 4, pp. 1146–1156, Apr. 2016.

[8] X. Feng, X. Ling, H. Zheng, Z. Chen, and Y. Xu, "Adaptive multi-kernel SVM with spatial temporal correlation for short-term traffic flow prediction," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 6, pp. 2001–2013, Jun. 2019.

[9] M. Lippi, M. Bertini, and P. Frasconi, "Short-term traffic flow forecasting: An experimental comparison of time-series analysis and supervised learning," *IEEE Trans. Intell. Transp. Syst.*, vol. 14, no. 2, pp. 871–882, Jun. 2013.

[10] Y. Li *et al.*, "Multiple measures-based chaotic time series for traffic flow prediction based on Bayesian theory," *Nonlinear Dyn.*, vol. 85, no. 1, pp. 179–194, Feb. 2016.

[11] H. Tan, Y. Wu, B. Shen, P. J. Jin, and B. Ran, "Short-term traffic prediction based on dynamic tensor completion," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 8, pp. 2123–2133, Aug. 2016.

[12] N. Garg, S. K. Mangal, P. K. Saini, P. Dhiman, and S. Maji, "Comparison of ANN and analytical models in traffic noise modeling and predictions," *Acoust. Aust.*, vol. 43, no. 2, pp. 179–189, Jul. 2015.

[13] T. Pamula, "Impact of data loss for prediction of traffic flow on an urban road using neural networks," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 3, pp. 1000–1009, Mar. 2019.

[14] Y. Lv, Y. Duan, W. Kang, Z. Li, and F. Wang, "Traffic flow prediction with big data: A deep learning approach," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 2, pp. 865–873, Apr. 2015.

[15] B. Guo, S. R. Gunn, R. I. Damper, and J. D. B. Nelson, "Customizing Kernel functions for SVM-based hyperspectral image classification," *IEEE Trans. Image Process.*, vol. 17, no. 4, pp. 622–629, Apr. 2008.

[16] D. Chen, "Research on traffic flow prediction in the big data environment based on the improved RBF neural network," *IEEE Trans. Ind. Informat.*, vol. 13, no. 4, pp. 2000–2008, Aug. 2017.

[17] B. Zhou and H. S. Yan, "Nonlinear system identification and prediction based on dynamics cluster multi-dimensional Taylor network model," *Control Decis.*, vol. 29, no. 01, pp. 33–38, 2014.

[18] Y. Lin, H. S. Yan, and B. Zhou, "Nonlinear time series prediction method based on multi-dimensional Taylor network and its applications," *Control Decis.*, vol. 29, no. 05, pp. 795–801, 2014.

[19] B. Zhou and H. S. Yan, "Financial time series forecasting based on wavelet and multi-dimensional Taylor network dynamics model," *Syst. Eng. Theory Pract.*, vol. 33, no. 10, pp. 2654–2662, 2013.

[20] Y.-Q. Han and H.-S. Yan, "Adaptive multi-dimensional Taylor network tracking control for SISO uncertain stochastic non-linear systems," *IET Control Theory Appl.*, vol. 12, no. 8, pp. 1107–1115, May 2018.

[21] Y.-Q. Han and H.-S. Yan, "Adaptive multi-dimensional Taylor network tracking control for SISO uncertain stochastic non-linear systems," *IET Control Theory Appl.*, vol. 12, no. 8, pp. 1107–1115, May 2018.

[22] H. Zhang, X. Wang, J. Cao, M. Tang, and Y. Guo, "A multivariate short-term traffic flow forecasting method based on wavelet analysis and seasonal time series," *Appl. Intell.*, vol. 48, no. 3, pp. 3827–3838, 2018.

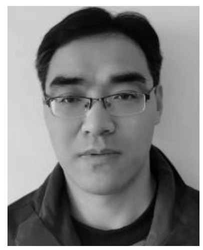
[23] S. O. Mousavizadeh Kashi and M. Akbarzadeh, "A framework for short-term traffic flow forecasting using the combination of wavelet transformation and artificial neural networks," *J. Intell. Transp. Syst.*, vol. 23, no. 1, pp. 60–71, Nov. 2018.

[24] Y. Xie, Y. Zhang, and Z. Ye, "Short-term traffic volume forecasting using Kalman filter with discrete wavelet decomposition," *Comput.-Aided Civil Infrastruct. Eng.*, vol. 22, no. 5, pp. 326–334, Jul. 2007.

[25] Z. Zheng and S. Washington, "On selecting an optimal wavelet for detecting singularities in traffic and vehicular data," *Transp. Res. C, Emerg. Technol.*, vol. 25, pp. 18–33, Dec. 2012.

[26] S. G. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, no. 7, pp. 674–693, Jul. 1989.

[27] C. Xu, Z. Li, and W. Wang, "Short-term traffic flow prediction using a methodology based on autoregressive integrated moving average and genetic programming," *Transport*, vol. 31, no. 3, pp. 343–358, Sep. 2016.



Shanliang Zhu received the M.S. degree from the School of Mathematical Sciences, Ocean University of China. He is currently pursuing the Ph.D. degree with the Qingdao University of Science and Technology. He is also an Associate Professor with the Qingdao University of Science and Technology. His research interests include data driven control, traffic system management and optimization, machine learning, and their applications.

Yu Zhao is currently a member of the Research Institute for Mathematics and Interdisciplinary Sciences. He specializes in traffic economics at the College of Economics and Management, Qingdao University of Science and Technology. His research interests include traffic management, machine learning algorithms, and their applications.

Yanjie Zhang is currently a member of the Research Institute for Mathematics and Interdisciplinary Sciences. He specializes in computer science at the School of Information Science and Technology, Qingdao University of Science and Technology. His research interests include machine learning, big data processing techniques, intelligent algorithms, and their applications.

Qingling Li received the Ph.D. degree in engineering thermodynamics from the Huazhong University of Science and Technology, Wuhan, China, in 2004. He is currently a Professor with the College of Electromechanical Engineering, Qingdao University of Science and Technology, Qingdao, China. His research interests include heat transfer and energy saving technology of process equipment, new energy development and utilization technology, and environmental protection process technology.

Wenwu Wang (Senior Member, IEEE) received the Ph.D. degree from the College of Automation, Harbin Engineering University, Harbin, China, in 2002. He was a Visiting Scholar with The Ohio State University, USA, in 2008. He has been a Guest Professor in machine perception with the Qingdao University of Science and Technology, China, since 2018. He is currently a Professor in signal processing and machine learning and the Co-Director of the Machine Audition Laboratory, Centre for Vision Speech and Signal Processing, Department of Electrical and Electronic Engineering, University of Surrey. His current research interests include blind signal processing, sparse signal processing, audiovisual signal processing, machine learning and perception, artificial intelligence, machine audition (listening), and statistical anomaly detection. He has coauthored more than 250 publications in these areas. Since 2019, he has been a Senior Area Editor of the IEEE/ACM TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING, where he has been an Associate Editor since 2020. He has been an Associate Editor of the *EURASIP Journal on Audio, Speech, and Music Processing* since 2019.

Shuguo Yang received the Ph.D. degree in navigation, guidance, and control from Harbin Engineering University, China, in 2003. From 2003 to 2006, he was an Associate Professor with Heilongjiang Bayi Agricultural University, China. From 2006 to 2019, he was a Professor with the College of Mathematics and Physics, Qingdao University of Science and Technology, China. His major research interests include machine learning, image retrieval, and moving object recognition.