IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 17 August 2023, accepted 28 August 2023, date of publication 31 August 2023, date of current version 7 September 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3310821

RESEARCH ARTICLE

Bayesian Combination Approach to Traffic Forecasting With Graph Attention Network and ARIMA Model

JINYUAN LIU[1], GE GUO[2], (Senior Member, IEEE), AND XINMING JIANG[1]
[1]Technical Development Department, FAW-Volkswagen Company Ltd., Changchun 130011, China
[2]State Key Laboratory of Synthetical Automation for Process Industries, Northeastern University, Shenyang 110819, China

Corresponding author: Jinyuan Liu (jinyuan.liu@faw-vw.com)

This work was supported by the National Natural Science Foundation of China under Grant 62173079 and Grant U1808205.

ABSTRACT To better capture the spatio-temporal characteristics and reduce unbalanced errors in short-term traffic prediction, an advanced Bayesian combination model with graph neural network (ABCM-GNN) is proposed. A new ABCM framework involving an error correction mechanism is established, based on the analysis of distance correlation between historical and current traffic volumes. Two sub-predictors built, respectively, on the graph attention gated recurrent unit (GAGRU) network, which captures the spatial correlation of road network, and autoregressive integrated moving average method (ARIMA), are incorporated into the ABCM framework to enhance the strength and capability of the framework. The effectiveness and superiority of the proposed model are demonstrated in various scenarios with experiments conducted using real-time traffic data collected on the California freeway. The overall results show that the ABCM-GNN with ARIMA method is superior to state-of-the-art methods in terms of precision and stability.

INDEX TERMS Short-term traffic prediction, Bayesian combination, error correction mechanism, graph attention gated recurrent units networks, ARIMA.

I. INTRODUCTION

Traffic flow forecasting is one of the core problems in intelligent transportation systems. Accurate real-time traffic prediction can help maximize the utilization of the road network capacity, improve traffic efficiency and safety, optimize traffic distribution, thereby alleviating congestion and reducing air pollution [1], [2], [3], [4]. There are various methods for short-term traffic forecasting, e.g., statistical models, machine learning methods and big data-driven deep learning methods, or broadly classified into three types: parametric, non-parametric and hybrid methods.

Typical statistical models for traffic prediction include autoregressive integrated moving average (ARIMA) approach [5] and its variants such as seasonal ARIMA and space-time ARIMA [6]. Kalman filters are also a powerful statistical method for traffic prediction. Machine learning methods used for traffic forecasting include hybrid wavelet analysis [7], support vector regression model [8], neural network models [9], and others. Statistical models require prior knowledges whereas traditional machine learning methods cannot deal with complicated spatio-temporal dependencies of road traffic networks.

In order to better extract traffic features from voluminous data, deep learning models, as an advanced non-parametric method, have been increasingly utilized for traffic prediction [10]. Long-short term memories (LSTM) [11], [12] and its variants gated recurrent units (GRU) [13], [14] were used to capture sequential dynamics evolution in time from the traffic data. Hybrid models combining convolutional neural networks (CNNS) and recurrent neural networks (RNNS) [15], [16], [17], [18] were utilized to explore the spatial-temporal dependencies to improve traffic forecasting performance.

The performance of deep learning models depends on a large set of high-quality traffic data, and perturbed data may produce inaccurate or even erroneous prediction [19], [20]. In recent years, fusion models were used to take advantages

The associate editor coordinating the review of this manuscript and approving it for publication was Tao Huang.

of all models involved to improve the prediction performance and stability [21], [22], [23], [24]. To name some, Zheng et al. [25] presented a Bayesian combination method (BCM) for traffic forecasting based on two neural network predictors. A BCM framework was given in [26], using gray correlation analysis to integrate the outputs of a back-propagation neural network, ARIMA, and Kalman filter, to deal with the long operation period and insensitivity to prediction error fluctuations. In particular, an improved BCM method was associated with a deep learning model GRU in [27] to address the error amplification phenomenon and improve the prediction performance.

It is worth noting that the above state-of-art techniques still leave much room for improvement. i). The deep learning networks used cannot capture the spatial characteristics of complex road networks satisfactorily, resulting in poor prediction performance. ii). The existing Bayesian combination models cannot deal with unbalanced sub-model errors with non-identical sign (i.e., negative vs positive errors). iii). The traffic data sequences are nonlinearly correlated and the correlation is ignored in determining the key parameters of the Bayesian combination method.

In this paper, we present an advanced Bayesian combination method (ABCM) in association with a graph attention gated recurrent unit network (or graph neural network, GNN) for traffic prediction. The model, denoted as ABCM-GNN, adopts co-integration and error correction to correct short-term unbalanced errors through long-term co-integration in the Bayesian combination model. While in the deep learning model, we use graph attention gated recurrent unit networks to effectively capture the temporal and spatial characteristics in the traffic network. The ABCM-GNN also integrates sub-predictors GAGRU and ARIMA, the latter of which deals with the nonlinear characteristics in short-term traffic dynamics. The combined GAGRU, GAT and GRU model can deal with massive data information and obtain the spatial characteristics of complex traffic networks. Meanwhile, the time interval correlation parameters are obtained by distance correlations.

The main contributions are as follows:

1) An advanced Bayesian combination method (ABCM) is proposed based on co-integration and error correction, which can effectively combine the sub-predictors and reduce short-term unbalanced errors.
2) By combining an ARIMA model and a GAGRU neural networks using this ABCM, the nonlinear characteristics and spatial correlations in the road network can be effectively captured.
3) Distance correlation analysis is introduced to capture the nonlinear temporal correlations in the time series data, which can also reduce the computation time cost of the algorithm.

The remaining parts of this article are organized as follows. Section II gives a further detailed description of ABCM-GNN model. Sub-predictors are introduced in Section III. Section IV presents the validation data, evaluation criterion, the processes of model implementation, and experimental results, following the conclusion and future work in Section V.

II. METHODOLOGY

This section presents the description of the ABCM-GNN model. The structure of ABCM-GNN is shown in Fig. 1. First, the correlation analysis is performed using the distance correlation coefficient. Next, the cointegration analysis and error correction model are presented. Then, we propose a new combinatorial framework named ABCM. Finally, the process of implementing the ABCM-GNN methods is presented in detail.

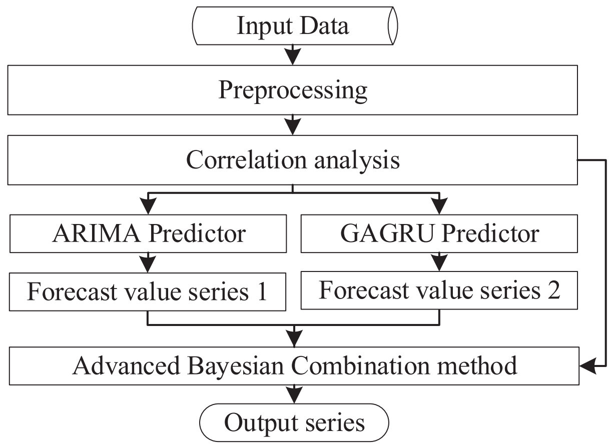

FIGURE 1. The architecture of the ABCM-GNN model.

A. CORRELATION ANALYSIS

Short-term traffic volume forecasting is a typical time sequences prediction question. The time interval is the unit of time in which the sensor collects traffic flow data. Considering the set of time intervals in the NBCM model where the traffic volume in the prediction traffic sequence intervals t is determined strongly related to the historical traffic flow [26]. Let Z denote the time period collection and $Z = \{t-1, t-2, \ldots, t-z\}$. We present a distance correlation analysis approach to obtian the correlation of present and historical traffic volumes [28].

Assume that y_t is the traffic flow in a time interval, which is influenced by the J previous traffic flow for time period $(y_{t-1}, y_{t-2}, \ldots, y_{t-J})$, where J is set sufficiently large to include most related time intervals. Let $Y_t = \{y_t(l)|l \in E\}$ denotes the traffic volumes target sequence to be predicted, and $Y_{t-z} = \{y_{t-z}(l)|l \in E\}$ is a traffic volume alternative series whose data are the time interval z before the relevant data in Y_t. The collections of all period of the traffic volume series are indicated as L, where $E = \{1, 2, \ldots, L\}$ and L are the lengths of the traffic volume series.Sequence $Y_t, Y_{t-1}, \ldots, Y_{t-J}$ is derived from the identical traffic data series the relations between the alternative series Y_{t-z} and the target series Y_t can be calculated by the distance correlation coefficient, denoted as (1), and $r(t-z)$ denotes the correlation between Y_t and Y_{t-z}.

$$r(t-z)=\frac{d\,\mathrm{cov}(Y_t,Y_{t-z})}{\sqrt{d\,\mathrm{cov}(Y_t,Y_t)d\,\mathrm{cov}(Y_{t-z},Y_{t-z})}} \tag{1}$$

where $d\mathrm{cov}^2(Y_t,Y_{t-z})=A_1+A_2-2A_3$.

$$A_1=\frac{1}{J^2}\sum_{i=1}^{J}\sum_{n=1}^{J}\|y_{t-i}(j)-y_{t-n}(j)\|_{y_t(j)} \|y_{t-z-i}(j)-y_{t-z-n}(j)\|_{y_{t-z}(j)} \tag{2}$$

$$A_2=\frac{1}{J^2}\sum_{i=1}^{J}\sum_{n=1}^{J}\|y_{t-i}(j)-y_{t-n}(j)\|_{y_t(j)} \frac{1}{J^2}\sum_{i=1}^{J}\sum_{n=1}^{J}\|y_{t-z-i}(j)-y_{t-z-i}(j)\|_{y_{t-i}(j)} \tag{3}$$

$$A_3=\frac{1}{J^3}\sum_{i=1}^{J}\sum_{n=1}^{J}\sum_{l=1}^{J}\|y_{t-i}(j)-y_{t-n}(j)\|_{y_t(j)} \|y_{t-z-i}(j)-y_{t-k-i}(j)\|_{y_{t-z}(j)} \tag{4}$$

The size of the set Z is formulated as:

$$R(Z)=\arg\min\{r(t-Z)|k\geq\delta,\forall k\in\{1,2,\ldots J\}\} \tag{5}$$

where δ is a measurement that determines the dimensionality of the set Z and the set $\delta\in[0,1]$.

B. ERROR CORRECTION MECHANISM

Traffic flow series are typical time series. Regression analysis of a non-stationary time series as a stable one can lead to pseudo-regressions, where there is no correlation between the variables, but erroneous conclusions are drawn that the regression results are given indeed correlated. According to the cointegration theory, a number of stable variable sequences, if their single integration order is the same, some of their linear combinations are stable, it demonstrates the existence of long-term equilibrium relations between these variable series cointegration relationship.

Let the sequence of two variables x_i and y_i be a first-order integral process, where $y_t\sim I(1)$ and $x_t\sim I(1)$., and if the following formula is true.

$$y_t=\beta x_t+u_t\ \ u_t\sim I(0) \tag{6}$$

The linear combination of the two nonstationary time series is called cointegration. According to Granger and Siklos [29] theorem, if there is a cointegration relationship between two non-stationary time series, then these variables have error correction expressions. When the stable relationship between these variable series will have some imbalance in the short term, i.e., the variables deviate from the cointegration relationship, the error correction model introduces the cointegrating variables reflecting the long-term equilibrium relations into the dynamic equation and uses the long-term equilibrium error as the correction term for short-term fluctuations, which compensates for the shortcomings of traditional statistical analysis models. The error correction model incorporates the long-term equilibrium nexus with the short-term disequilibrium state to improve the stability of the forecasting model.

Due to the time variable x_t and y_t there is a long-term equilibrium relationship as:

$$y_t=k_0+k_1x_t+\varepsilon_t \tag{7}$$

The short-term disequilibrium relationship is:

$$\Delta y_t=\beta_1\Delta x_t-\lambda(y_{t-1}-\alpha_0-\alpha_1x_{t-1}) \tag{8}$$

The error correction expression is:

$$\Delta y_t=\beta_1\Delta x_t-\lambda(y_{t-1}-\alpha_0-\alpha_1x_{t-1})+\mu_t \tag{9}$$

where $y_{t-1}-k_0-k_1x_{t-1}$ refers to nonequilibrium error term of period time $t-1$; $\lambda(y_{t-1}-\alpha_0-\alpha_1x_{t-1})$ represents the error correction term; α_0,α_1 are the long-term reaction parameters; β_1, λ are the short-term reaction parameters; ε_t is the residual value.

C. ADVANCED BAYESIAN COMBINATION MODEL

For a certain time interval t, usually only the n^{th} sub-model with the highest prediction accuracy is selected as the best model. We obtain posterior probability as:

$$\begin{aligned}p_t^n&=P(U=n|y_t,y_{t-1},\ldots,y_1)\\&=\frac{P(y_t,U=n|y_t,y_{t-1},\ldots,y_1)}{\sum_m^N P(y_t,U=m|y_{t-1},\ldots,y_1)}\end{aligned} \tag{10}$$

where p_t^n is considered as the weight of the n^{th} predictor at time period t; And N denotes the count of component predictors.

Based on Bayes' rule, the following recursion can be obtained:

$$\begin{aligned}&P(y_t,U=n|y_{t-1},\ldots,y_1)\\&\quad=\frac{P(y_t,U=n,y_{t-1},\ldots,y_1)}{P(y_{t-1},\ldots,y_1)}\\&\quad=\frac{P(y_t,U=n,y_{t-1},\ldots,y_1)}{P(U=n,y_{t-1},\ldots,y_1)}\\&\qquad\cdot\frac{P(U=n,y_{t-1},\ldots,y_1)}{\sum_{m=1}^N P(y_{t-1},U=m|y_{t-1},\ldots,y_1)}\\&\quad=P(y_t|U=n,y_{t-1},\ldots,y_1)\cdot p_{t-1}^n\end{aligned} \tag{11}$$

Assume that the sub-predictor error obeys a Gaussian white noise $e_t^n=(y_t-y_t^n)\sim N(0,\sigma_n)$, then we have:

$$\begin{aligned}&P(y_t|U=n,y_{t-1},\ldots,y_1)\\&\quad=P(e_t^n=y_t-y_t^n|U=n,y_{t-1},\ldots,y_1)\\&\quad=\frac{1}{\sqrt{2\pi\sigma_n}}\exp(-[e_t^n/\sqrt{2}\sigma_n]^2)\end{aligned} \tag{12}$$

where e_t^n denotes the forecasting error of the predictor over the time interval; y_t^n is the traffic forecasting of the n^{th} predictor at the time interval t.

We combine (5) and (10)-(12) to obtain:

$$p_t^n = \frac{\frac{1}{\sqrt{2\pi}\sigma_n} p_{t-1}^n \exp(-[e_t^n/\sqrt{2}\sigma_n]^2)}{\sum_{m=1}^N \frac{1}{\sqrt{2\pi}\sigma_m} p_{t-1}^m \exp(-[e_t^m/\sqrt{2}\sigma_m]^2)} \tag{13}$$

Altering $p_{t-1}^n, p_{t-2}^n, p_{t-3}^n, \ldots, p_1^n$ in (13), then:

$$p_t^n = \frac{(\frac{1}{\sqrt{2\pi}\sigma_n})^{t-1} \exp(\sum_{i=0}^{t-1} -[e_t^n/\sqrt{2}\sigma_n]^2)}{\sum_{m=1}^N (\frac{1}{\sqrt{2\pi}\sigma_m})^{t-1} \exp(\sum_{i=0}^{t-1} -[e_t^m/\sqrt{2}\sigma_m]^2)} \tag{14}$$

Eq. (14) denotes an unrealistic assumption that the traffic volumes in the forecast interval are relevant with past traffic volumes. As referred to in a few previous research studies [15], [16], traffic volumes are susceptible to disturbances from the external environment, especially during peak hours. Usually only the traffic volumes of the last few periods are strongly correlated with the traffic volumes of a given forecast period. The smaller the interval, the greater the influence of the traffic on the current flow.

Therefore, we have the weight p_t^n of the n^{th} predictor as:

$$p_t^n = \frac{(\frac{1}{\sqrt{2\pi}\sigma_n})^{R(Z)} \exp(\sum_{t\in K} -[e_{t-i}^n/\sqrt{2}\sigma_n]^2)}{\sum_{m=1}^N (\frac{1}{\sqrt{2\pi}\sigma_m})^{R(Z)} \exp(\sum_{t\in K} -[e_{t-i}^m/\sqrt{2}\sigma_m]^2)} \tag{15}$$

The prediction results can be obtained as a linear combination of the output of each predictor with the weights of each sub-predictor over the time interval. [26]], denoted as:

$$\hat{y}_{t+1} = \sum_{m=1}^N p_t^m y_{t+1}^m \tag{16}$$

For the condition of non-stationary time series combined prediction modeling, to establish an integrated prediction model of non-stationary time series, the effectiveness of combined prediction modeling must be judged first. Since the traffic flow series are non-stationary, the cointegration theory is adopted into the fused traffic volumes forecasting and the cointegration is verified for the single forecasting series and the forecasting series. Test whether the predicted sequence $\hat{y}_{t+1}$ corresponds to m single item.

The traffic flow sequence $y_{t+1}^m, y_{t+1}^{m-1}, \ldots, y_{t+1}^1$ is predicted to have a cointegration relationship, and the Engle-Granger two-step test [29] is used here.

Step one, the least square method is used to estimate the regression formula as follow:

$$\hat{y}_{t+1} = \alpha + \beta_i y_{t+1}^i + \varepsilon_{t+1}^i (i = 1, 2, \ldots, m) \tag{17}$$

$\hat{\alpha}$ and $\hat{\beta}_i$ are used to denote the estimated value of the regression coefficients, so that the estimated value of the model residuals as:

$$\varepsilon_{t+1}^i = \hat{y}_{t+1} - \hat{\alpha} - \hat{\beta}_i y_{t+1}^i (i = 1, 2, \ldots, m) \tag{18}$$

Step two, check the stationarity of ε_{t+1}^i.

If $\varepsilon_{t+1}^i \sim I(0)$, there is cointegration relationship between the predicted sequences $\hat{y}_{t+1}$ and the m output sequences in the sub-prediction period. If it is found that there is a cointegration relationship between each single traffic flow prediction series and the predicted series, and there is a long-term equilibrium relation between them, it can be proved that these single traffic flow prediction series can be used to establish an effective combined traffic flow prediction model.

$$\hat{e}_{t+1} = y_{t+1} - \hat{y}_{t+1} = y_{t+1} - \sum_{m=1}^N p_t^m y_{t+1}^m \sim I(0) \tag{19}$$

According to Granger representation theorem, error correction model can be obtained in the form of:

$$\begin{aligned}\Delta y_{t+1} &= \sum_{i=1}^m \beta_i \Delta y_{t+1}^i \\ &\quad - \lambda(y_{t-1} - \alpha_0 - \sum_{i=1}^m \alpha_i \Delta y_t^i) + \mu_{t+1} \\ &= \sum_{i=1}^m \beta_i \Delta y_{t+1}^i - \lambda\varepsilon_t + \mu_{t+1}\end{aligned} \tag{20}$$

The first order hysteresis y_t is used as error correction model. At last, combined equation (16) takes the form:

$$\hat{y}_{t+1} = \sum_{m=1}^N p_t^m y_{t+1}^m + p_{m+1} y_t + \hat{e}_{t+1} \tag{21}$$

Through the cointegration of the sub-model and the prediction model, combined with the long-term equilibrium correlation of the prediction time series, combined with the idea of error correction, the error correction of the final prediction model is carried out to improve the prediction accuracy.

D. STEPS OF ABCM-GNN MODEL IMPLEMENTATION

This part summarizes the steps of ABCM-GNN implementation:

Step1: The raw traffic volume data are normalized using the Min-Max method so that the traffic data take values between zero and one.

Step2: Introduce the distance correlation analysis coefficients to derive the historical time series that are highly correlated with the traffic volume within the target time interval.

Step3: Each sub-predictor is calibrated using the past traffic volume data to select the best settings of the predictors. Then the trained sub-predictors are used to forecast the test traffic flow.

Step4: The traffic volume is predicted by the ABCM model using the weights and error correction model.

III. TWO COMPONENT PREDICTORS

A. GRAPH ATTENTION GATED RECURRENT NETWORK

1) GRAPH ATTENTION MECHANISM

Attention mechanism in time series task can extract more valuable information accurately and efficiently [17]. The attention mechanism is employed to capture the propagation characteristics and spatial correlation of traffic volume data seen in complex traffic road networks.

The input data to this part is A group of node characteristics, $H = \{h_1, h_2, \ldots, h_n\}$, $h_i \in \mathbb{R}^F$ where N is the number of nodes characteristics in the traffic road networks graph. Masked attention is usually used to deal with real-world graphs. In adjacency matrix, A_{ij}=1means the road nodes and in the range are connected, others mean disconnected.

$$e_{ij} = \begin{cases} a(T(h_i), T(h_j))\ A_{ij} = 1 \\ 0 \qquad otherwise \end{cases} \tag{22}$$

In equation (22), adjacency matrix is obtained through threshold Gaussian core.

We employ active the function normalizes the attention coefficients so that they are comparable between first-order neighborhoods. The normalization function is shown below.

$$\hat{e}_{ij} = soft\ \max(e_{ij}) = \frac{\exp(e_{ij})}{\sum_{k \in N_i} \exp(e_{ik})} \tag{23}$$

where *soft* max(.) refers to a nonlinearity function. The attention coefficient can be acquired by

$$h_i = \sigma(\sum_{j \in i} \hat{e}_{ij} W h_j) \tag{24}$$

where W is the weight matric of node characteristics.

In order to further extract spatial features of traffic networks efficiently, a multi-head attention mechanism is designed. The number of independent attention heads parameter is denoted by K.The node collects information at the same time through multiple parallel channels from adjacent nodes.

$$h_i = ||_{k=1}^{K} \sigma(\sum_{j \in i} \hat{e}_{ij}^{k} W^k h_j) \tag{25}$$

where, $\|$ denotes concatenate operation, $\hat{e}_{ij}$ is a learnable weight vector.

2) GRAPH ATTENTION GATED RECURRENT NETWORK

Based on spatio-temporal information aggregation, GRU [13], [14] unit is taken as the main body. Gated recurrent Unit (GRU) is a popular recurrent neural network that can effectively capture long-term and short-term time series in recent years. We replace the original linear connection layer of GRU units with graphical note operations functions. The architecture of GAGRU memory block is illustrated in Fig. 2. The GAGRU element is established and its internal tensor calculation conforms to the principle of information processing of spatio-temporal graph nodes. The architecture of the GAGRU model is shown in Fig. 3.

$$r_t = \sigma(GA(x_t) + GA(h_{t-1}) + b_r) \tag{26}$$

$$z_t = \sigma(GA(x_t) + GA(h_{(t-1)}) + b_z) \tag{27}$$

$$n_t = \tanh(GA(x_t) + (GA(r_t * h_{(t-1)})) + b_n) \tag{28}$$

$$h_i = \sigma(\sum_{j \in i} \hat{e}_{ij} W h_j) \tag{29}$$

where x_t and y_t. denote the input and output at time step t, and h_t is hidden state of model. z_t and r_t denote the update and reset gate, σ is the activate function, and * is the Hadamard product. $GA(\cdot)$ is a graph attention mechanism function.

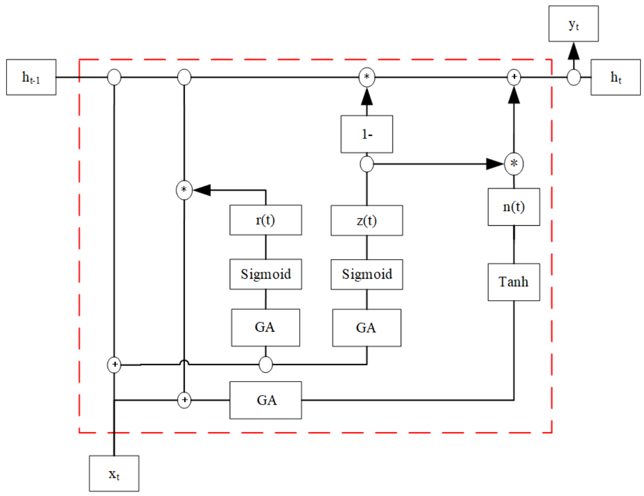

FIGURE 2. The architecture of GAGRU memory block.

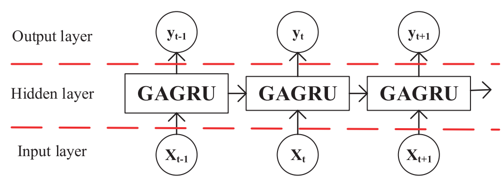

FIGURE 3. The architecture of the GARNN model.

B. ARIMA METHOD

ARIMA [5] model is one of the most widely applied time sequence prediction methods for traditional parametric methods. It can perform an arithmetic fit to past moment time data to predict the present. The ARIMA model usually refers to ARIMA(p, d, q)model, where AR is autoregressive, MA is sliding average, p is the number of autoregressive terms, d is the count of differences made to make it a smooth series, and q is the count of sliding average terms The model is expressed as follows:

$$\varphi(B)(1-B)^d \hat{y}_t = \theta(B)\varepsilon_n \tag{30}$$

$$B\hat{y}_t = \hat{y}_{t-1} \tag{31}$$

$$\varphi(B) = 1 - \varphi_1(B) - \varphi_2(B^2) - \ldots \varphi_p(B^p) \tag{32}$$

$$\theta(B) = 1 - \theta_1(B) - \theta_2(B^2) - \ldots \theta_q(B^q) \tag{33}$$

where $\varphi(B)$ denotes the autoregressive process of order p; ε_n is random error that follows a normal distribution and has a mean value of 0, variance σ^2, and $\text{cov}(\varepsilon_n, \varepsilon_{n-d}) = 0$, $\forall d \neq 0$; $\theta(B)$ denotes the moving average process of order q.

IV. EXPERIMENTS AND DISCUSSIONS

A. DATA DESCRIPTION

To verify the proposed short-term traffic volume prediction model, we validated our model on California Highway traffic data set PeMS. The traffic data set was captured by the positioned remote flow microwave sensor (RTMS). These data sets were collected in real-ime every 30 seconds by the Caltrans Performance Measurement System (PeMS) [30] is shown in Fig. 4. Traffic volume data is assembled from real-time information every five minutes. The system deploys outstrip 39,000 sensors on the freeways of California's major metropolitan areas. Geographic information about the sensor station is recorded in the dataset. The PeMS data is SAN Bernardino traffic data for June 1st to August 31st 2022 and includes 1,979 detectors on eight roads. We used the data of the first 50 days as the training set and the data of the last 12 days as the test set. We aggregated the traffic speed every 5 minutes. We trained this architecture on a server with a NVIDIA 2080Ti and an Interi9-9980XE CPU. The GAGRU and GRU model were implemented on Pytorch. And we use pyramid function for ARIMA model. The predictors were combined with NBCM or ABCM.

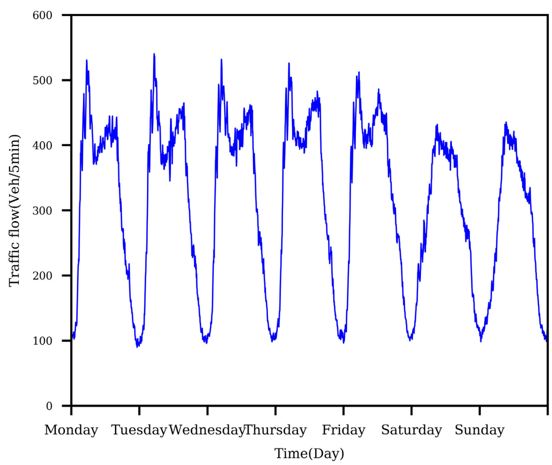

FIGURE 4. Weekly traffic volume data sampled at 5-minute intervals.

B. PREDICTIVE PERFORMANCE CRITERIA

In order to be able to estimate the accuracy of the forecasting models, in this paper we adopt three error evaluation metrics to measure the capabilities of different traffic prediction methods, namely mean absolute error (MAE), mean square root error (RMSE), and mean absolute percentage error (MAPE). In addition, we also adopt three information loss criteria indicators [31], [32] for optimal model selection, namely Akaike information criterion (AIC), Bayesian information criterion (BIC), and Hannan-Quinn Information Criterion (HQIC). The formulas for each evaluation metric are expressed as below.

$$MAE(y, \hat{y}) = \frac{1}{|n|}\sum_{i=1}^{n}\left|\hat{y}_i - y_i\right| \tag{34}$$

$$RMSE(y, \hat{y}) = \sqrt{\frac{1}{|n|}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2} \tag{35}$$

$$MAPE(y, \hat{y}) = \frac{1}{|n|}\sum_{i=1}^{n}\left|\frac{\hat{y}_i - y_i}{y_i}\right| \tag{36}$$

$$AIC = \ln(\frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{|n|}) + \frac{2m}{n} \tag{37}$$

$$BIC = \ln(\frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{|n|}) + \frac{ln(n)m}{n} \tag{38}$$

$$HQIC = \ln(\frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{|n|}) + \frac{ln(ln(n))m}{n} \tag{39}$$

where n denotes the number of observed samples, m denotes the number of model parameters, $\hat{y}_i$ denotes the predicted value, y_i is the true value.

C. MODEL SETTINGS

1) GAGRU AND GRU

GAGRU and GRU [13] utilizes the data set of previous two months for training process. In the GAGRU model, there are few hyperparameters that require confirmation, such as the number of attention heads in the graph attention mechanism. Table 1 shows the effect of the number of attention heads on the performance, and it can be seen from the table that when the number of attention heads is 3, the metric of each parameter metric is the smallest on the test dataset. The initial learning rate is le-3 and the decay rate is 0.6 / 10 epochs. We set the number of model training epochs to 100, the batch size to 32, and set the number of hidden units to 64. Then, the hidden sizes of GRU are set to 128, and we utilize Adam optimizer and adaptive learning rate. The mean square error of the predicted results with respect to the ground truth is minimized to obtain the weight matrix.

TABLE 1. Effect of the GAGRU with different number of attention heads.

Number	MAE	RMSE	MAPE(%)
2	4.79/5.15/5.31	6.52/6.71/6.90	12.65/12.73/13.37
3	**4.72/5.04/5.26**	**6.35/6.61/6.79**	**12.25/12.66/13.35**
4	4.75/5.14/5.29	6.40/6.68/6.89	12.55/12.70/13.35

2) ARIMA

Model identification is performed for the selected eight road segments using historical traffic volume data for the previous two months. Fig. 5 describes the process of ARIMA model prediction. First, the series smoothness test is performed using MA and AR operators, and the unstable series are differenced to obtain the smooth series. Secondly, the parameter values of ARIMA are determined using the great likelihood estimation, and the residual statistics of the model are calculated; then, the residuals of the ARIMA model are estimated and tested to obtain the best traffic volume prediction model parameters for the road target sections. In this experiment, we adjusted auto.arima () function to automatically determine the number of autoregressive terms, the number of moving average terms, and other parameters to determine the optimal model for traffic volume prediction.

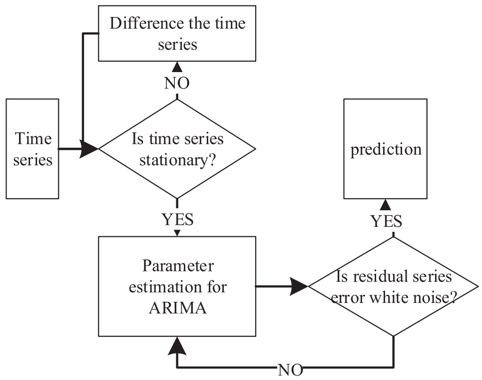

FIGURE 5. The process of setting up the ARIMA model.

3) CORRELATION ANALYSIS

In this paper, the target series Y_t is obtained by calculating the average traffic volume of eight roads in the dataset at each time period. We suppose a potential correlation between current traffic volume and traffic volume within the first 25 intervals. Therefore, the total length of the correlation interval is one hour. The distance correlation coefficient between the alternative and the object sequence is increased from 0 to 25, and the results are illustrated in Fig. 6. As depicted in Fig. 6, $R(Z)$ and the set Z and are decided by parameters. For example, set to 0.98, $R(Z) = 3$ and $Z = \{t-1, t-2, t-3\}$.

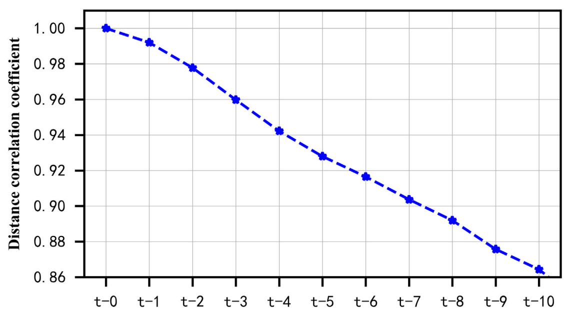

FIGURE 6. The result of the correlation analysis.

D. EXPERIMENTAL RESULTS

After the experiment, the model in Table 2 was validated. Table 2 demonstrates the detail description of testing models. And we choose ARIMA and GAGRU models as sub-predictors. The NBCM-GNN and IBCM-GNN incorporate the ARIMA and GAGRU models using the NBCM model [26] and IBCM model [27]. And ABCM-GNN fuses the same sub-predictor through our proposed traffic volume model ABCM. Meanwhile ABCM-DL fuses ARIMA and GRU model. FC-LSTM is a variant of LSTM, with input and hidden states in vector form. The temporal correlation of traffic data is captured through LSTM, but spatial correlation is not fully considered [12]. The STGCN consists of graph convolutional layers and time convolutional layers, which can capture the spatial correlation between road network nodes

TABLE 2. The detail description of the testing models.

Model	Type	Famework	Sub-predictors
ARIMA	Single	_	_
GRU	Single	_	_
GAGRU	Single	_	_
FC-LSTM	Single	_	_
STGCN	Single	_	_
NBCM_DL	Combination	NBCM	ARIMA;GRU
NBCM_GNN	Combination	NBCM	ARIMA;GAGRU
IBCM_DL	Combination	IBCM	ARIMA;GRU
IBCM_GNN	Combination	IBCM	ARIMA;GAGRU
ABCM_DL	Combination	ABCM	ARIMA;GRU
ABCM_GNN	Combination	ABCM	ARIMA;GAGRU

and the temporal characteristics of traffic data [18]. The spatial correlation between them and the temporal characteristics of traffic data.The forecasting results of the ARIMA and GAGRU models on Friday and Sunday are shown in Fig. 7 and Fig. 8.

All the results of the models with δ set as 0.98 is shown in Table 3. It shows the superior prediction performance of our proposed model under MAE, RMSE, MAPE criteria compared to other models. Comparing the GAGRU model with GRU and ARIMA, the MAE metrics improved by 2.63 and 5.53 in next 15 minutes, respectively, indicating that the GAGRU fusion graph neural network is capable of higher predictive performance compared to the traditional ARIMA and deep learning method GRU.

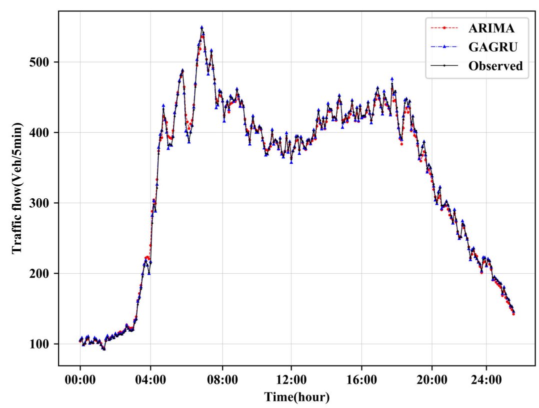

FIGURE 7. Predicted and observed traffic volume of two sub-predictors in the Next 15 Minutes on Friday.

Similarly, the MAE and MAPE criteria of ABCM-GNN and ABCM-DL improved by 0.13 and 0.44% in next 15 minutes, implying that the joint model fusion graph neural network can improve the prediction accuracy. For different combination approaches, the MAE and MAPE metrics of our proposed ABCM-GNN improve by 4.2% and 6.6% over IBCM-GNN, and the surface-based error correction ABCM

TABLE 3. Results comparison of the models in the next 15 minutes, 30 minutes and 45 minutes.

Model	MAE	RMSE	MAPE(%)	AIC	BIC	HQIC
ARIMA	9.25/10.32/11.87	12.76/13.68/14.02	19.31/19.76/20.32	5.09/5.23/5.28	5.09/5.23/5.28	5.09/5.23/5.28
FC-LSTM	7.43/7.78/8.16	9.82/10.36/10.96	17.15/17.56/17.95	6.01/6.12/6.23	11.55/11.65/11.77	6.21/6.31/6.43
GRU	7.35/7.75/8.12	9.78/10.21/10.65	16.93/17.43/17.88	5.91/5.99/6.08	11.07/11.15/11.24	6.09/6.17/6.26
STGCN	4.89/5.34/5.55	6.60/6.73/7.03	12.54/12.97/13.65	6.03/6.07/6.16	14.49/14.73/14.81	6.34/6.38/6.46
GAGRU	4.72/5.04/5.26	6.35/6.61/6.79	12.25/12.66/13.35	5.66/5.74/5.80	13.19/13.27/13.32	5.93/6.01/6.06
NBCM_DL	4.53/4.93/5.17	6.16/6.45/6.63	11.27/11.89/12.44	4.99/5.08/5.13	10.15/10.25/10.30	5.17/5.26/5.31
NBCM_GNN	4.31/4.89/5.11	5.94/6.17/6.51	11.05/11.68/12.17	5.53/5.61/5.71	13.07/13.14/13.25	5.79/5.87/5.98
IBCM_DL	4.37/4.91/5.10	5.97/6.11/6.48	11.12/11.75/12.25	4.93/4.97/5.09	10.11/10.16/10.27	5.11/5.15/5.27
IBCM_GNN	4.24/4.79/4.97	5.83/6.03/6.39	11.03/11.67/12.13	5.50/5.56/5.68	13.05/13.12/13.23	5.76/5.83/5.94
ABCM_DL	4.19/4.63/4.86	5.54/5.90/6.23	10.74/11.54/11.98	4.78/4.91/5.01	9.97/10.10/10.21	4.96/5.09/5.20
ABCM_GNN	**4.06/4.58/4.76**	**5.49/5.86/6.20**	**10.30/11.23/11.56**	**5.38/5.51/5.62**	**12.94/13.07/13.18**	**5.64/5.77/5.89**

TABLE 4. Results comparison of the models in the next 15 minutes, 30 minutes and 45 minutes on friday and sunday.

Model	Firday			Sunday		
	MAE	RMSE	MAPE(%)	MAE	RMSE	MAPE(%)
ARIMA	9.27/10.36/11.91	12.85/13.73/14.11	19.34/19.78/20.35	9.18/10.29/11.82	12.65/13.63/19.97	19.26/19.71/20.27
FC-LSTM	7.45/7.82/8.24	9.85/10.41/10.99	17.19/17.58/17.97	7.42/7.75/8.12	9.80/10.33/10.94	17.11/17.54/17.92
GRU	7.39/7.77/8.16	9.82/10.25/10.69	16.96/17.47/17.92	7.33/7.71/8.08	9.76/10.18/10.62	16.90/17.41/17.86
STGCN	4.93/5.37/5.58	6.63/6.75/7.06	12.56/12.99/13.69	4.86/5.32/5.51	6.57/6.71/7.01	12.52/12.94/13.63
GAGRU	4.72/5.09/5.31	6.38/6.64/6.81	12.27/12.69/13.37	4.69/5.02/5.23	6.32/6.58/6.77	12.22/12.63/13.34
NBCM_DL	4.57/4.9 5/5.21	6.18/6.49/6.65	11.34/11.93/12.47	4.50/4.89/5.14	6.14/6.42/6.61	11.24/11.86/12.42
NBCM_GNN	4.34/4.92/5.13	5.97/6.22/6.56	11.12/11.73/12.20	4.29/4.87/5.08	5.91/6.14/6.47	11.03/11.64/12.15
IBCM_DL	4.39/4.95/5.12	5.99/6.12/6.51	11.15/11.79/12.28	4.37/4.91/5.10	5.95/6.08/6.45	11.10/11.73/12.21
IBCM_GNN	4.26/4.83/5.01	5.88/6.05/6.42	10.78/11.57/12.08	4.21/4.73/4.94	5.81/6.01/6.37	11.01/11.64/12.10
ABCM_DL	4.21/4.66/4.87	5.57/5.94/6.25	10.78/11.57/12.08	4.15 /4.60/4.84	5.52/5.88/6.21	10.72/11.51/11.96
ABCM_GNN	**4.10/4.62/4.79**	**5.52/5.88/6.24**	**10.34/11.27/11.61**	**4.04/4.55/4.74**	**5.47/5.84/6.17**	**10.27/11.20/11.52**

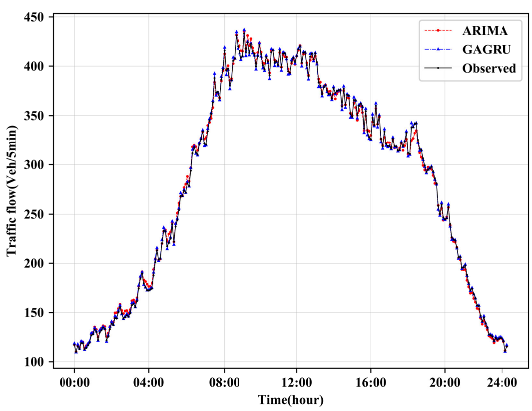

FIGURE 8. Predicted and observed traffic volume of two sub-predictors in the Next 15 Minutes on Sunday.

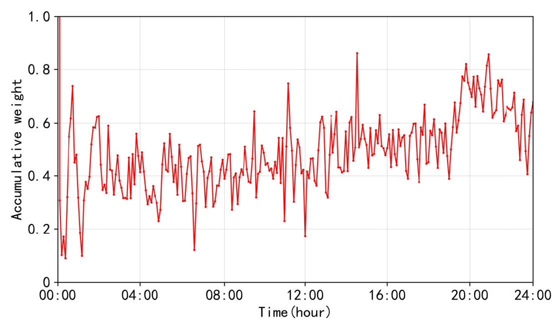

FIGURE 9. Accumulative weight of each predictor by ABCM.

model effectively corrects the prediction error compared to the IBCM model, which has better prediction results. In terms of optimal model selection, the numbers of model parameter is a disciplinary term which means that a small AIC value indicates a better model. The AIC and BIC criteria of ABCM-GNN and IBCM-DL decreased by 0.12 and 0.11 in next 15 minutes, implying the goodness of the ABCM-GNN model. Due to the large number of parameters in the deep learning model, ABCM-GNN did not demonstrate advantages in information loss criteria compared to ABCM-DL. However, in terms of error evaluation indicators, ABCM-GNN has better performance and smaller error values. The comparison of the prediction effects of ABCM-GNN and IBCM-GNN models with observations is shown in Fig. 11. Fig. 9 represents the cumulative weights calculated by ABCM for each predictor, indicating that ABCM assigns weights among sub-models according to the error magnitude. The amount of error correction in ABCM-GNN and NBCM-GNN is shown in Fig. 10, and the surface error correction mechanism is able to perform short-term equilibrium error correction in the combined model prediction process. The results of the different models on Friday and Sunday is shown in Table 4. Meanwhile, the ABCM-GNN model superior other models in the case of different traffic volume data. In addition, another essential feature of ABCM-GNN is that its prediction performed is not significantly affected by the

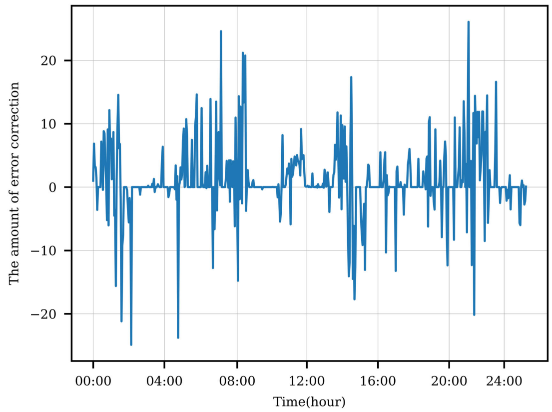

FIGURE 10. The amount of error correction of ABCM-GNN.

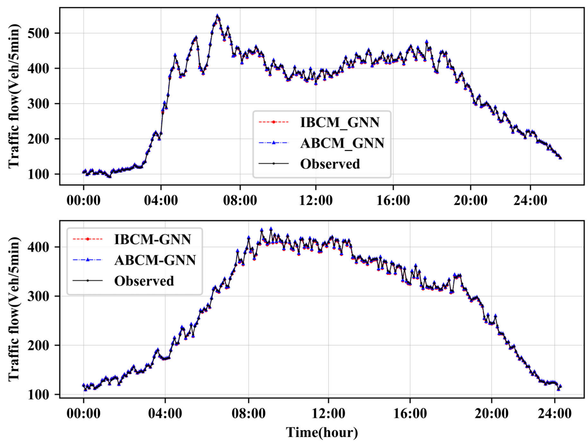

FIGURE 11. The prediction results of the ABCM-GNN model in the next 15 minutes on friday and sunday.

length of the prediction period. The results indicate that the predictive performance of our proposed ABCM-GNN model outperforms other models in the measurement of MAE, RMSE and MAPE.

V. CONCLUSION AND FUTURE WORK

This paper proposes a new short-term traffic volume forecasting model called advanced Bayesian combination model with graph neural network (ABCM-GNN) model. Firstly, an ABCM framework is established, and the cointegration analysis and error correction mechanism are introduced. Through the long-term cointegration correlation, the short-term forecasting quantity of the joint predictor is corrected to enhance the prediction accuracy. Second, the non-linear correlation between historical and current traffic flows is analyzed using distance correlation coefficients to determine the length of the previous time set sequences to be used in the ABCM framework. Then, the graph attention gating recurrent neural network model is combined with the time domain prediction model based on the advantages of graph attention mechanism in dealing with the spatial characteristics of traffic flow, and the GAGRU method is used as sub-model of the advanced Bayesian combination model. Experimental performance on a real-time dataset shows that both the proposed ABCM framework and the introduced GAGRU sub-predictor enhance the prediction accuracy of the combined model. And MAE and RMSE metrics of the ABCM-GNN model improved by 0.25 and 0.31, respectively. Meanwhile, ABCM-GNN model outperforms other models in the case of different traffic volume data.

Since the performance of ABCM-GNN model is largely dependent on sub-predictors, more high-level sub-predictors should be considered in the future. Such as SVM and KNN in ABCM framework. At the same time, more information loss standard parameters (eg: AIC and its variants) will be included in my work related to optimal model selection. Meanwhile, the next step will focus on obtaining more kinds of data information from various types of data such as climate conditions, passenger travel demand distribution, and unexpected road conditions. Further, the method can be applied to other time series predictions, such as arrival time estimation and speech recognition.

REFERENCES

[1] G. Fusco, C. Colombaroni, and N. Isaenko, "Short-term speed predictions exploiting big data on large urban road networks," *Transp. Res. C, Emerg. Technol.*, vol. 73, pp. 183–201, Dec. 2016.
[2] X. Yang, Y. Zou, and L. Chen, "Operation analysis of freeway mixed traffic flow based on catch-up coordination platoon," *Accident Anal. Prevention*, vol. 175, Sep. 2022, Art. no. 106780.
[3] Y. Zou, L. Ding, H. Zhang, T. Zhu, and L. Wu, "Vehicle acceleration prediction based on machine learning models and driving behavior analysis," *Appl. Sci.*, vol. 12, no. 10, p. 5259, May 2022.
[4] Z. Zhou, Z. Yang, Y. Zhang, Y. Huang, H. Chen, and Z. Yu, "A comprehensive study of speed prediction in transportation system: From vehicle to traffic," *iScience*, vol. 25, no. 3, Mar. 2022, Art. no. 103909.
[5] B. M. Williams, "Multivariate vehicular traffic flow prediction: Evaluation of ARIMAX modeling," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 1776, no. 1, pp. 194–200, Jan. 2001.
[6] B. M. Williams and L. A. Hoel, "Modeling and forecasting vehicular traffic flow as a seasonal ARIMA process: Theoretical basis and empirical results," *J. Transp. Eng.*, vol. 129, no. 6, pp. 664–672, Nov. 2003.
[7] Y. Xie and Y. Zhang, "A wavelet network model for short-term traffic volume forecasting," *J. Intell. Transp. Syst.*, vol. 10, no. 3, pp. 141–150, Sep. 2006.
[8] W.-C. Hong, Y. Dong, F. Zheng, and C.-Y. Lai, "Forecasting urban traffic flow by SVR with continuous ACO," *Appl. Math. Model.*, vol. 35, no. 3, pp. 1282–1291, Mar. 2011.
[9] 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.
[10] W. Huang, G. Song, H. Hong, and K. Xie, "Deep architecture for traffic flow prediction: Deep belief networks with multitask learning," *IEEE Trans. Intell. Transp. Syst.*, vol. 15, no. 5, pp. 2191–2201, Oct. 2014.
[11] X. Ma, Z. Tao, Y. Wang, H. Yu, and Y. Wang, "Long short-term memory neural network for traffic speed prediction using remote microwave sensor data," *Transp. Res. C, Emerg. Technol.*, vol. 54, pp. 187–197, May 2015.
[12] Z. Zhao, W. Chen, X. Wu, P. C. Y. Chen, and J. Liu, "LSTM network: A deep learning approach for short-term traffic forecast," *IET Intell. Transp. Syst.*, vol. 11, no. 2, pp. 68–75, Mar. 2017.
[13] N. G. Polson and V. O. Sokolov, "Deep learning for short-term traffic flow prediction," *Transp. Res. C, Emerg. Technol.*, vol. 79, pp. 1–17, Jun. 2017.
[14] D. Zhang and M. R. Kabuka, "Combining weather condition data to predict traffic flow: A GRU-based deep learning approach," *IET Intell. Transp. Syst.*, vol. 12, no. 7, pp. 578–585, Sep. 2018.
[15] L. Zhao, Y. Song, C. Zhang, Y. Liu, P. Wang, T. Lin, M. Deng, and H. Li, "T-GCN: A temporal graph convolutional network for traffic prediction," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 9, pp. 3848–3858, Sep. 2020.

[16] G. Guo and W. Yuan, "Short-term traffic speed forecasting based on graph attention temporal convolutional networks," *Neurocomputing*, vol. 410, pp. 387–393, Oct. 2020.
[17] T. Zhang and G. Guo, "Graph attention LSTM: A spatiotemporal approach for traffic flow forecasting," *IEEE Intell. Transp. Syst. Mag.*, vol. 14, no. 2, pp. 190–196, Mar. 2022, doi: 10.1109/MITS.2020.2990165.
[18] B. Yu, H. Yin, and Z. Zhu, "Spatio-temporal graph convolutional networks: A deep learning framework for traffic forecasting," 2017, *arXiv:1709.04875*.
[19] G. Guo, W. Yuan, J. Liu, Y. Lv, and W. Liu, "Traffic forecasting via dilated temporal convolution with peak-sensitive loss," *IEEE Intell. Transp. Syst. Mag.*, vol. 15, no. 1, pp. 48–57, Jan. 2023.
[20] Y. Zhang, T. Cheng, Y. Ren, and K. Xie, "A novel residual graph convolution deep learning model for short-term network-based traffic forecasting," *Int. J. Geograph. Inf. Sci.*, vol. 34, no. 5, pp. 969–995, May 2020.
[21] F. Guo, J. W. Polak, and R. Krishnan, "Predictor fusion for short-term traffic forecasting," *Transp. Res. C, Emerg. Technol.*, vol. 92, pp. 90–100, Jul. 2018.
[22] Y. Zhang, Y. Zhang, and A. Haghani, "A hybrid short-term traffic flow forecasting method based on spectral analysis and statistical volatility model," *Transp. Res. C, Emerg. Technol.*, vol. 43, pp. 65–78, Jun. 2014.
[23] 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.
[24] V. Petridis, A. Kehagias, L. Petrou, A. Bakirtzis, S. Kiartzis, H. Panagiotou, and N. Maslaris, "A Bayesian multiple models combination method for time series prediction," *J. Intell. Robot. Syst.*, vol. 31, nos. 1–3, pp. 69–89, 2001.
[25] W. Zheng, D.-H. Lee, and Q. Shi, "Short-term freeway traffic flow prediction: Bayesian combined neural network approach," *J. Transp. Eng.*, vol. 132, no. 2, pp. 114–121, Feb. 2006.
[26] J. Wang, W. Deng, and Y. Guo, "New Bayesian combination method for short-term traffic flow forecasting," *Transp. Res. C, Emerg. Technol.*, vol. 43, pp. 79–94, Jun. 2014.
[27] Y. Gu, W. Lu, X. Xu, L. Qin, Z. Shao, and H. Zhang, "An improved Bayesian combination model for short-term traffic prediction with deep learning," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 3, pp. 1332–1342, Mar. 2020.
[28] G. J. Székely, M. L. Rizzo, and N. K. Bakirov, "Measuring and testing dependence by correlation of distances," *Ann. Statist.*, vol. 35, no. 6, pp. 2769–2794, Dec. 2007.
[29] C. W. J. Granger and P. L. Siklos, "Systematic sampling, temporal aggregation, seasonal adjustment, and cointegration theory and evidence," *J. Econometrics*, vol. 66, nos. 1–2, pp. 357–369, Mar. 1995.
[30] C. Chen, K. Petty, A. Skabardonis, P. Varaiya, and Z. Jia, "Freeway performance measurement system: Mining loop detector data," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 1748, no. 1, pp. 96–102, Jan. 2001.
[31] M. Qi and G. P. Zhang, "An investigation of model selection criteria for neural network time series forecasting," *Eur. J. Oper. Res.*, vol. 132, no. 3, pp. 666–680, Aug. 2001.
[32] A. Chakrabarti and J. K. Ghosh, "AIC, BIC and recent advances in model selection," in *Philosophy of Statistics*, 2011, pp. 583–605.

JINYUAN LIU received the B.S. and master's degrees from Northeastern University, Shenyang, China, in 2019 and 2022, respectively. He is currently a Product Development Engineer with the Driving Assistance and Intelligent Driving Development Section, Technology Development Department, FAW-Volkswagen Company Ltd. His current research interests include traffic data analysis, mining, and forecasting.

GE GUO (Senior Member, IEEE) received the B.S. and Ph.D. degrees from Northeastern University, Shenyang, China, in 1994 and 1998, respectively. From 1999 to 2005, he was with the Lanzhou University of Technology, China, as the Director of the Institute of Intelligent Control and Robots, and has been a Professor, since July 2004. In May 2005, he joined Dalian Maritime University, China, as a Professor with the Department of Automation. He is currently a Professor with Northeastern University and the Dean of the School of Control Engineering, Northeastern University at Qinhuangdao Campus. He has published more than 100 international journal articles. His current research interests include intelligent transportation systems, cyber-physical systems, and connected automated vehicles. He was an honoree of the Ministry of Education New Century Excellent Talents, in 2004, a nominee for Gansu Top Ten Excellent Youths, and the CAA Young Scientist Award Winner. He won the First Prize of Hebei Province Natural Sciences Award. He is an Associate Editor of the IEEE Transactions on Intelligent Transportation Systems, the IEEE Transactions on Intelligent Vehicles, the *Information Sciences*, the *IEEE Intelligent Transportation Systems Magazine*, and the *ACTA Automatica Sinica*.

XINMING JIANG received the B.S. and master's degrees from Jilin University. He served as the Manager of the General Assembly Planning Section and the Electronic and Electrical Planning Section, Planning Department, FAW-Volkswagen Company Ltd; and the Manager of the Functional Electronics Section, Technical Development Department, FAW-Volkswagen Company Ltd. Currently, he is the Manager of the Technology Development Driving Assistance and the Intelligent Driving Development Section and the Director of the Intelligent Driving Development Department, FAW-Volkswagen Company Ltd. His current research interests include intelligent driving, chassis electronics and matching, and chassis traditional component development.