Hybrid CSA optimization with seasonal RVR in traffic flow forecasting

Zhangguo Shen [1,2], Wanliang Wang[1], Qing Shen[2] and Zechao Li[3]

[1]College of Computer Science and Technology，Zhejiang University of Technology，Hangzhou 310023, China
[2]School of Information Engineering, Huzhou University, Huzhou, Zhejiang 313000, China
[3]School of Computer Science and Engineering, Nanjing University of Science and Technology,
Nanjing 210094, China
[e-mail:szgxx@zjhu.edu.cn, zjutwwl@zjut.edu.cn, sq@zjhu.edu.cn, zechao.li@njust.edu.cn]
*Corresponding author: Zechao Li

Abstract

Accurate traffic flow forecasting is critical to the development and implementation of city intelligent transportation systems. Therefore, it is one of the most important components in the research of urban traffic scheduling. However, traffic flow forecasting involves a rather complex nonlinear data pattern, particularly during workday peak periods, and a lot of research has shown that traffic flow data reveals a seasonal trend. This paper proposes a new traffic flow forecasting model that combines seasonal relevance vector regression with the hybrid chaotic simulated annealing method (SRVRCSA). Additionally, a numerical example of traffic flow data from The Transportation Data Research Laboratory is used to elucidate the forecasting performance of the proposed SRVRCSA model. The forecasting results indicate that the proposed model yields more accurate forecasting results than the seasonal auto regressive integrated moving average (SARIMA), the double seasonal Holt-Winters exponential smoothing (DSHWES), and the relevance vector regression with hybrid Chaotic Simulated Annealing method (RVRCSA) models. The forecasting performance of RVRCSA with different kernel functions is also studied.

1 Introduction

Traffic flow forecasting is a hot research topic in intelligent transportation and urban big data processing [1, 2, 3, 4, 5]. In addition, it is critical to intelligent transportation systems. Real-time and accurate traffic flow forecasting can provide precise synchronous transportation status information for transportation management departments, and it is a precondition of real-time traffic signal control, transportation distribution, route guidance, automatic navigation, and accident detection in an intelligent transportation system. It also provides better-optimized route options for drivers and, thus, saves energy and improves urban transportation conditions [6]. Traffic flow real-time data can be provided by different kinds of sensor systems that are installed in roads; this data usually includes the flow, speed, and lane occupancy rate of a transportation network. Recently, researchers have been heavily focused on short-time traffic flow forecasting and travel time delay estimation, and these two aspects are primarily used in assisting transportation control centers in forecasting traffic congestion and improving traffic flexibility.

  Traffic flow forecasting refers to making predictions about traffic flow in next decision time and in later time periods when making control variable decisions. Usually, the forecasting of traffic flow can be characterized by using a time series of historical and current traffic flow data points [7,8,9]. Using historical traffic flow data, we can predict future traffic flow values. Many models and systems have been proposed for addressing this problem over the past several decades, including approaches based on time series models, Kalman filter theory, k-nearest neighbor(KNN), artificial neural network approaches(ANN), nonparametric methods, support vector machine regression models(SVR), and the Markov chain models. Many Hybrid integration techniques have also been used effectively to this area.

  Tipping [10] proposed a sparse Bayesian leaning model in 2001, which is a relevance vector machine (RVM). Compared to SVM, RVM is sparser; thus, it is more suitable to real-time detection and has a better regression accuracy rate. Therefore, the relevance vector machine regression forecasting model is now widely used in every aspect and has achieved great effect, such as electricity load forecasting, tracking of objects, 3D model recovery, software reliability forecasting, 3D pose estimation, and channel equalization prediction [11,12,13]. Practice reveals that in relevance vector regression estimation, the selection of kernel function parameters has a significant effect on forecasting effect. Currently, there are no structural methods for guiding the parameter selection of kernel function, and the most commonly used methods are simulated annealing, genetic algorithm, particle swarm algorithm, and other evolutionary methods to optimize parameters. Experiments from Juang etal. [14] show that every evolutionary method has its advantages and disadvantages, employing hybrid different algorithms will be more powerful than using a single algorithm. In this paper, the author introduced RVM into traffic flow forecasting, and focusing on the drawbacks of SA algorithm, the author proposes the CSA to determine the values of parameters in a RVR model, thus building a relevance vector regression traffic flow forecasting model using chaos-simulated annealing optimization(RVRCSA).
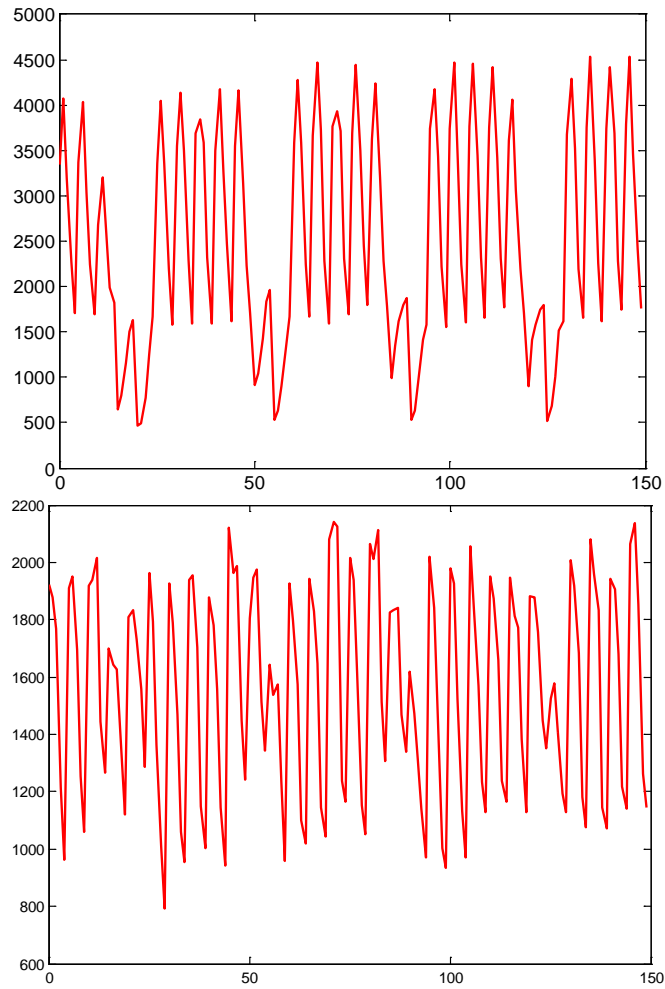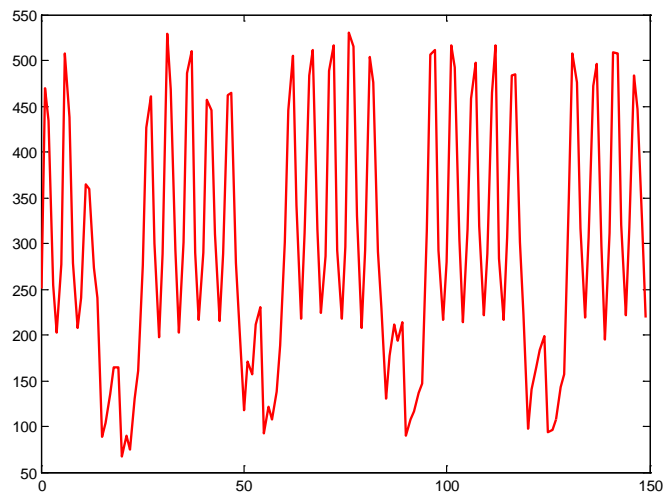
Fig. 1. Plot of the traffic volumes for week days and week end on road TH252SB, station id: N89123
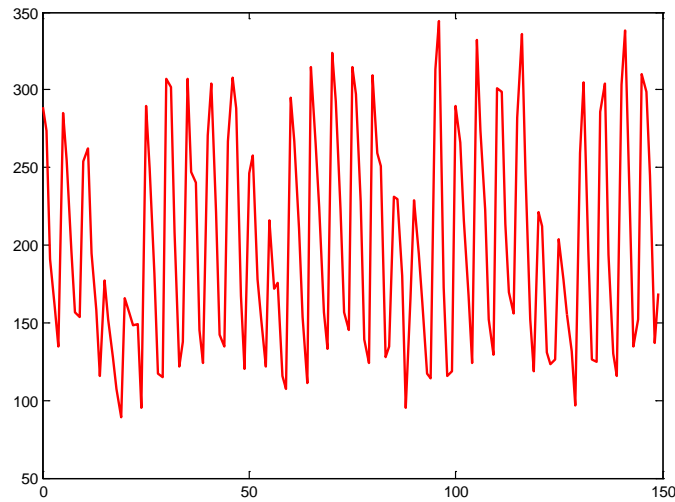
Fig. 2. Plot of the traffic volumes for week days and week end on road TH77 CDSB, station id: N88125

On the other hand, urban traffic flow data is usually seasonal due to rush hours, holidays, and rest days as well as changes in the weather and other factors. For example, **Fig. 1** and **2** illustrate the flow data collected by the University of Minnesota Duluth in Minnesota Twin Cities and Rochester Manchester City on roads labeled with link numbers TH252SB, station id: N89123 and TH77 CDSB, station id: N88125, from April 1-30, 2015, from 6 a.m. to 11:00 a.m. and from 3:00 p.m. to 8:00 p.m. [15] From this figure, it can be seen that the traffic flows in these four figures all exhibit obvious seasonal change. In recent traffic flow forecasting papers, many researchers added seasonal factors to the models. For example, Williams et al. [16] proposed a multivariate ARIMA model that included upstream traffic flow data and considered seasonal change as well as distinguished seasonal flow on peak/non-peak flow; thus, it greatly improved the forecasting capacity of the ARIMA traffic flow model. Hong W et al. [17] considered the seasonal condition in the SVR traffic flow forecasting model, and Lippi M et al. [18] used the RBF seasonal kernel and linear seasonal kernel in the support vector regression (SVR) traffic flow forecasting model by combining the seasonal kernel with a RBF and linear kernel. However, applying the RVR model to generate a seasonal prediction model is currently rare. This paper refers to some papers and proposes a hybrid chaotic-simulated annealing optimization with a seasonal RVR model in traffic flow forecasting (SRVRCSA). Furthermore, practice example analysis is applied in order to demonstrate that the proposed method is effective.

The contribution of this paper is as follows. First, for RVR, we propose a new CSA optimization model for estimating hyper-parameters. To the best of our knowledge, this approach has not been attempted. Second, the SRVRCSA model is, for the first time, applied to the traffic forecasting problem and successfully outperform some state-of the art approaches. At last, the traffic flow forecasting performance of RVRCSA with different kernel functions is studied, and some suitable kernels are selected.

2   Related works

In existing research, the main point of interest in traffic flow forecasting is the urban expressway. Researchers have proposed numerous research models centered on urban

expressway traffic flow forecasting problems, of which, time series analysis modeling is one of the most common methods [4, 19]. The autoregressive integrated moving average model (ARIMA) [4] is one of the mostly used time series method. In the early research, researchers were more willing to build a multivariate time series traffic flow forecasting model, which could simultaneously consider time and space. For example, Stathopoulos and Karlaftis [20] used traffic flow data collected by upstream detectors and combined it with ARIMA to build a real-time traffic flow-forecasting model. In addition, they compared the univariate ARIMA model with the multivariate ARIMA model and found that the multivariate ARIMA model is superior to the univariate ARIMA model. Kamarianakis and Prastacos [21] used the time and space ARIMA model to forecast the traffic flow in an urban area and to describe the traffic flow conditions of a road network by introducing the intersection turns ratio matrix and using that to describe the variation law of traffic flow. They also proposed an ARIMA model based on jump in order to improve the forecasting accuracy of short-term traffic flow. However, the biggest disadvantage of the traditional ARIMA model is that the forecasting value tends to the average value of the past time sequence and, thus, cannot capture the fast multilateral movement of urban traffic flow.

Other methods that are often used for traffic flow forecasting including the K-nearest neighbor (KNN) models, non-parameter regression, local linear regression, and Kalman filters [5,22,23,24]. Cai P et al. [25] improved the KNN algorithm based on time and space relationship and achieved short-time traffic flow multi-step forecasting as well as improved forecasting accuracy. Xia D et al. [26] proposed time and space weighted KNN model and improved the real-time characteristic of traffic flow forecasting using the parallel processing of MapReduce in a Hadoop-distributed computation platform. Zheng Z et al. [5] improved the KNN algorithm by introducing a time window and local minimum value in order to reduce overlapping candidate neighbors in the neighbor selection stage. Chen C et al. [27] designed a distributed data processing traffic flow forecasting system framework based on MapReduce and used time and space characteristics and discrete characteristics of traffic flow to improve forecasting accuracy. Daraghmi Y etal. [28] used negative binomial additive models (NBAMs). Tchrakian T et al. [29] built a corresponding mode function by building a covariance matrix from historical data and decided the number of mode functions used to forecast through a real-time traffic flow feature, and he built a real-time traffic flow forecasting algorithm based on a spectrum. Sun S et al. [30] proposed the selective random subspace predictor (SRSP) method, which can maintain a robust feature even if part of the traffic flow data is lost. Furthermore, Oh S et al. [31] combined the Gauss model with artificial neural network clustering to propose a multi-factor pattern recognition urban traffic flow forecasting system and to forecast urban traffic flow by considering the road geometry factor, the environment factor, and the traffic flow feature. Sun S et al. [32] proposed a new variational approximation unlimited Gaussian process and applied it to predict traffic flow problems. Experiments demonstrate that this method is clearly more effective than other methods.

As discussed above, accurately describing and forecasting the changing traffic flow of urban roads is a very complex process. Artificial neural networks (ANNs) are considered to be able to approximate any arbitrary function to an arbitrary degree of accuracy, and become another commonly used traffic flow forecasting modeling method. Currently, researchers have proposed many ANN traffic flow forecasting models with static or dynamic structures, including multi-layer feed forward neural networks(FFNN), Radial basis function (RBF) neural network model, time delay neural network model, and regression neural network model. Furthermore, ANN and simulated annealing (SA), combined with other optimized

algorithms, can perfectly replace the traditional traffic flow forecasting model [33]. The latest research trend involves using these new evolutionary algorithms or wavelet function and artificial neural network to explore global optimization. For example, Vlahogianni etal. [34] repeatedly reflect the traffic flow data of time and space features through selection part and using the Genetic Algorithm (GA) combined with multi-layer structure optimization to determine the appropriate neural network structure. Yin H et al. [35] proposed a fuzzy neural network model (FNN) to forecast the traffic flow on urban roads, and experiments have shown that the FNM model can obtain more accurate forecasting results than the back-propagation neural network (BPNN). Li C et al. [36] proposed the combination of SA with the GA algorithm to solve local optimization and built the new simulated annealing genetic back-propagation neural network (BPNN-SA) based on the BPNN traffic flow forecasting model. Chan K et al. [37] applied multi-scale analysis to divide a traffic flow parameter sequence into a low-high frequency parameter and combined ANN with GA to optimize these multiscale parameters. Yang W et al. [38] used wavelet transform decomposition and reconstructed traffic flow data to propose the wavelet neural network model that combines wavelet transform with RBF. Chan K et al. [39] proposed the intelligent swarm-based neural networks (APSO-ANN) algorithm. Although these methods are more accurate than the pure neural network parameters, and thus they require more computation capacity and memory [17,40,41]. The neural network is limited by the fact that the determination of the appropriate network structure cannot be interpretative as well as by its in poor ability to promote.

Recently, some researchers applied support vector machine (SVM) and other kernel function methods to traffic flow time sequence forecasting. For example, references [2,42] discussed the application of SVM to traffic flow forecasting. This type of method generally considers transportation state factors and uses existing traffic state data. A new traffic state short-term forecasting method based on SVM is built by taking the traffic flow, occupancy rate, and average speed of previous time periods in upstream and downstream as collected by a transportation detector as the input and taking the traffic flow at the corresponding time period as the output. In contrast to the non-linearity, complexity, and uncertainty of road transportation systems, SVM and other non-parametric regressions are non-parametric, portable, and high prediction accuracy algorithms with a small error and a good error distribution. By improving the search algorithm and parameter adjust rules, it can achieve the requirements of real-time traffic flow forecasting. In addition, this method is easy to operate and can be applied to a complex environment. Additionally, it can conveniently forecast in different roads. Furthermore, SVM, SA, GA, particle swarm optimization (PSO), and other intelligent algorithms combined with optimized parameter settings can obviously improve the forecasting effect of short-term traffic flow. Li M et al. [42] optimized the parameters of robust v-support vector regression model (RSVR) by CcatCSAGA and applied the kernel principle component analysis (KPCA) algorithm to determine the final input vectors from the candidate input variables. Jeong Y et al. [2] combined online learning weighted support-vector regression (OLWSVR) to propose a new short-term traffic flow forecasting model, and experiments have shown that the OLWSVR model can achieve better forecasting results than can artificial neural network models and conventional support vector regression.

3 Forecasting methodology

3.1 Double seasonal Holt-Winters exponential smoothing (DSHWES) model

The multiplicative formulation for the double seasonal Holt-Winters method is given in the following expressions:

$$l_t = \alpha(y_t / (d_{t-s_1} w_{t-s_2})) + (1-\alpha)l_{t-1} \qquad (1)$$

$$d_t = \delta(y_t / (l_t w_{t-s_2})) + (1-\delta)d_{t-s_1} \qquad (2)$$

$$w_t = w(y_t / (l_t d_{t-s_1})) + (1-w)w_{t-s_2} \qquad (3)$$

$$\hat{y}_t(k) = l_t d_{t-s_1} + k w_{t-s_2+k} + \phi^k (y_t - (l_{t-1} d_{t-s_1} w_{t-s_2})) \qquad (4)$$

where y_t is the actual value at time t; l_t is the smoothed level; d_t and w_t are the seasonal indices for the intraday and intraweek seasonal cycles, respectively; α, δ, and ω are the smoothing parameters; and $y_t(k)$ is the k step future forecast made from the forecast origin (where $k \le s_1$). In forecast function (4), the term involving parameter ϕ is a simple adjustment for first-order autocorrelation. Eq. (1) lets the actual value y_t be smoothed in a recursive manner by weighting the current level α, and then, l_t is directly adjusted for the trend of the previous period $d_{t-s_1} w_{t-s_2}$ by adding it to the last smoothed value l_{t-1}. This helps to eliminate the lag and brings l_t to the approximate base of the current data value. Eq. (2) and (3) updates the seasonal indices d_t and w_t, which modifies the trend by smoothing with δ and ω and adding that to the previous estimate of the trend multiplied by $1-\delta$ and $1-\omega$, respectively. Eq. (4) is used for future forecasting.

An important point to note regarding the double seasonal Holt-Winters exponential smoothing approach is that, in contrast to ARIMA modeling and the majority of other approaches to short-term demand forecasting, no model specification required. As a result, the method is particularly advantageous; it is both simple and robust. The initial smoothed values for the level and the seasonal components are estimated by averaging the early observations. The parameters are estimated in a single procedure by minimizing the sum of squared one step-ahead in-sample errors. We constrained the parameters to lie between zero and one. The resultant parameters for the 10 load series are presented in Table II. For many of the series, the ϕ value is very high, whereas the α value is very low, indicating that the adjustment for first-order autocorrelation has, to a large degree, made the smoothing equation redundant for the level. It is also interesting to note that, for a given series, the values are similar for the two smoothing parameters, δ and ω, for the seasonal indices.

We also implemented a version of the method with optimized δ and ω values that were constrained to be identical and with $\alpha = 0$ so that the level was set as a constant value equal to the mean of the in-sample observations. This formulation delivered predictions that were only marginally poorer than the full method given in expressions (1) through (4). This is somewhat surprising, given that this reformulation of the method only involves two parameters.

3.2 Relevance vector machine used to regression estimation

RVM is a supervised learning method, which can be described as the following:

Given a set of example data $\{x_i\}_{i=1}^{N}$ along with corresponding target values $\{t_i\}_{i=1}^{N}$, the aim is to learn a function or to describe the reliance of a new target value t_* for an unseen input value x_*, by combining a training set together with any pertinent prior knowledge, and the correspondence relationship between which conforms to the following function [10]:

$$t = y(x; w) = \sum_{i=1}^{M} w_i k(x, x_i) + w_0,$$

where k is a nonlinear kernel function, and w_i is the weight of the model and is non-zero only when x_i belongs to the relevance vector. The probability is used as follows to describe the influence of t_i in addition to the error:

$$p(t_i) = N(t_i \mid y(x_i; w), \sigma^2)$$

A reasonable assumption can be made that $\{t_i\}_{i=1}^{N}$ are independent random variables. Therefore, given $\{w_i\}_{i=0}^{N}$ and σ^2, the probability distribution of t is as follows:

$$p(t \mid w, \sigma^2) = \prod_{i=1}^{n} N(t_i \mid y(x_i; w), \sigma^2) = (2\pi\sigma^2)^{-\frac{N}{2}} \exp\left(-\frac{\|t - \Phi w\|^2}{2\sigma^2}\right)$$

$$\text{where } w = \left[w_0, w_1, \cdots w_N\right]^T, \ \Phi = \left[\phi(x_1), \phi(x_2), \cdots \phi(x_N)\right]^T$$

$\phi(x_n) = \left[1, k(x_n, x_1), k(x_n, x_2), \cdots, k(x_n, x_N)\right]^T$; α_i is a hyper-parameter that determines the prior distribution of w_i; and $\alpha = (\alpha_1, \alpha_i, \cdots \alpha_N)$.

The posterior distribution of an unknown quantity can be calculated according to the Bayesian formula, according to the prior weight distribution and the sample set likelihood function.

$$p(w, \alpha, \sigma^2 \mid t) = \frac{p(t \mid w, \alpha, \sigma^2) p(w, \alpha, \sigma^2)}{p(t)}$$

Therefore, given a new dataset x_*, the probability distribution prediction of t_* is as follows:

$$p(t_* \mid t) = \int p(t_* \mid w, \alpha, \sigma^2) p(w, \alpha, \sigma^2 \mid t) dw d\alpha d\sigma^2 \qquad (5)$$

$$p(w, \alpha, \sigma^2 \mid t) = p(w \mid t, \alpha, \sigma^2) p(\alpha, \sigma^2 \mid t) \qquad (6)$$

Then,

$$p(w \mid t, \alpha, \sigma^2) = \frac{p(w, \alpha, \sigma^2 \mid t)}{p(\alpha, \sigma^2 \mid t)} = \frac{p(t \mid w, \sigma^2) p(w \mid \alpha)}{p(t \mid \alpha, \sigma^2)} = \frac{p(t \mid w, \sigma^2) p(w \mid \alpha)}{\int p(t \mid w, \sigma^2) p(w \mid \alpha) dw}$$

where $p(t \mid w, \sigma^2)$ and $p(w \mid \alpha)$ are both products of a Gaussian function. After integration simplification, it is as follows:

$$p(w|t,\alpha,\sigma^2) = (2\pi)^{-\frac{N+1}{2}} |\Sigma|^{-\frac{1}{2}} \exp\{-\frac{(w-\mu)^T\Sigma^{-1}(w-\mu)}{2}\}$$

$$p(t|\alpha,\sigma^2) = (2\pi)^{-\frac{N}{2}} |\Omega|^{-\frac{1}{2}} \exp\{-\frac{t^T\Omega^{-1}t}{2}\}$$

where

$$\mu = \sigma^{-2}\Sigma\Phi^T t, \Sigma = (A+\sigma^{-2}\Phi^T\Phi)^{-1},$$

$$A = diag(\alpha_0,\alpha_1,\cdots\alpha_N)$$

$$\Omega = \sigma^2 I + \Phi A^{-1}\Phi^T \quad (7)$$

An approximate solution can be obtained by inserting Formula (6) into Formula (5), as follows:

$$p(\alpha_{MP},\sigma^2_{MP}) = \arg\max_{\alpha,\sigma^2} p(\alpha,\sigma^2|t)$$

$$p(t_*|t) \approx \int p(t_*|w,\alpha_{MP},\sigma^2_{MP})p(w,\alpha_{MP},\sigma^2_{MP}|t)dw$$

The two portions of the integration formula are both products of a Gaussian function. Therefore, the result of the definite integral is as follows:

$$p(t_*|t) \approx N(t_*|y_*,\sigma^2_*)$$

$$y_* = \mu^T\phi(x_*), \sigma^2_* = \sigma^2_{MP} + \phi(x_*)^T\Sigma\phi(x_*) \quad (8)$$

$$\phi(x_*) = [1,k(x_*,x_1),k(x_*,x_2),\cdots,k(x_*,x_N)]^T$$

Finally, by solving $\alpha_{MP},\sigma^2_{MP}$ from formula (7), the following can be obtained:

$$\alpha_i^{new} = \frac{\gamma_i}{\mu_i^2} \quad (9)$$

$$(\sigma^2)^{new} = \frac{\|t-\Phi\mu\|^2}{N-\sum_{i=0}^N\gamma_i} \quad (10)$$

$$\gamma_i = 1-\alpha_i\Sigma_{ii} \quad (11)$$

where Σ_{ii} is the i-th element in the diagonal of Σ. First, the value of α,σ^2 is estimated, and then, the above formula is used to determine $\alpha_{MP},\sigma^2_{MP}$.

Let L be the number of maximum iterative cycles and th be the given threshold; then, the relevance vector regression can be summarized as follows:

Algorithm RVM Regression
Input: $\{x_i,t_i\}_{i=1}^N$, L, th
Output: y_* and σ^2_*;
1. Initialize α,σ^2;
2. **For** $l=1$ to L or convergence
3. Calculate $\{\mu,\Sigma\}$ by using (7);
4. Update α,σ^2 by using (9)-(11) respectively;

5. **If** $\alpha_i \geq th$
6. Remove the corresponding basis function;
7. **End if**
8. **End for**
9. Predicting for unseen data using (8).

4 Chaotic simulated annealing in selecting parameters and seasonal adjustment

4.1 Kernel functions used in the model

There are various kernel functions that may be used to select an arbitrary kernel function according to RVM theory. The kernel functions used in this paper are listed in **Table 1**, and the initialized values of each kernel function are also presented in **Table 1**.

Table 1. The Kernel Functions (KFs) used in the model

KFs	Abbreviation	Expression	Parameter(s)
Polynomial Function	PF	$\kappa(x, y) = (<x, y> + 1)^g$	$g > 0$, initialized as 3
Gaussian Function	GF	$\kappa(x, y) = e^{-g<x-y,x-y>^2}$	$g > 0$, initialized as 1
Symmetric Triangle Function	STF	$\kappa(x, y) = \max(1 - g \mid< x - y, x - y >\mid, 0)$	$g > 0$, initialized as 3
Cauchy Function	CF	$\kappa(x, y) = \dfrac{1}{1 + g(<x - y, x - y>)^2}$	$g > 0$, initialized as 3
Laplace Function	LF	$\kappa(x, y) = e^{-g\mid<x-y,x-y>\mid}$	$g > 0$, initialized as 3
Squared Sin Cardinal Function	SSCF	$\kappa(x, y) = \dfrac{\sin^2(<x - y, x - y>)}{(g <x - y, x - y>)^2}$	$g > 0$, initialized as 3

4.2 CSA in selecting parameters

The most important aspect of using kernel functions is determining the parameters. Small parameter values tend to cause over-learning, whereas large values can contribute to over-smoothing the data. Both situations can result in the regression having a decreased forecasting performance. The forecasting performance of a model is greatly affected by the selection of kernel function parameters in the relevance vector machine. However, there is a lack of construction methods regarding the kernel function parameter for specific data. This paper transforms the problem of kernel function parameter assignment in traffic flow forecasting to an optimizing searching problem, combines chaos optimization with a simulated annealing algorithm, and uses the randomness and uniformity characteristics of a chaos sequence. The chaos variable range is enlarged to the range of the kernel function optimization parameter, and the chaos operator is embedded into the simulated annealing solution; thus, the optimal value of the kernel function parameter is obtained.

The basic steps of relevance vector regression parameter optimization based on chaos simulated annealing (RVRCSA) include the initial setup, the computation of the objective

function, and the chaos disturbance and acceptance test. The concrete steps of the method are described in the following:

Step1:

Initial setup: Set the kernel function parameter to $g(i)$, and set the range to $g(i) \in (Max_g, Min_g)$. Then, select the initial value of the kernel function width $g(0)$ as shown in **Table 1**, and map $g(0)$ to determine the initial chaos variable $x(0) \in (0,1)$ according to Formula (12), as follows:

$$x(i) = \frac{g(i) - Min_g}{Max_g - Min_g} \qquad (12)$$

Then, use a logistic map to generate $k+1$ chaos variables $x(i+1)$, as shown in Formula (13):

$$x(i+1) = \mu \cdot x(i) \cdot (1 - x(i)), i = 0,1,2,\cdots,k \quad (13)$$

where μ is a logistic parameter. When $\mu = 4$, it is in a complete chaos state and can compute the $k+1$ kernel function width value according to Formula (14), as follows:

$$g(i+1) = Min_g + x(i+1) \cdot (Max_g - Min_g), i = 0,1,2,\cdots,k \quad (14)$$

Step 2:

Compute objective function: Use the computed $k+1$ kernel function width value to predict the relevance vector regression and obtain the optimum means average relative error (MARE) as the initial value of the objective function $E(\text{old})$, as follows:

$$MARE = \sum_{t=1}^{N} \left| \frac{q_n^d(t) - \hat{q}_n^d(t)}{q_n^d(t)} \right|$$

where N represents the ending time of the prediction, $q_n^d(t)$ is the actual flow data at time t in sub-section D_n, and $\hat{q}_n^d(t)$ is the predicted data. Record the kernel function width value with optimal MARE in the $k+1$ kernel function width values: $x(E_{\text{old}})$.

Step 3:

Add chaos disturbance: Add minor disturbance to chaos variable $x(E_{\text{old}})$ according to formula (15):

$$\tilde{x}(E_{new}) = (1 - \rho) \cdot x(E_0) \qquad (15)$$

where $0 < \rho < 1$ demonstrates the disturbance control variable: the larger the value is, the faster the disturbance. In this paper, the ρ value is 0.003. The improved kernel function parameter can be calculated according to formula (16), and the new value of the objective function, $E(new)$, can be calculated by inputting the calculated value into the relevance vector regression forecasting model.

$$\tilde{g}(E_{new}) = Min_g + \tilde{x}(E_{new}) \cdot (Max_g - Min_g) = Min_g + (1-\rho) \cdot x(E_{new}) \cdot (Max_g - Min_g)$$
$$= g(E_{new}) - \rho \cdot x(E_{new}) \cdot (Max_g - Min_g) \qquad (16)$$

Step 4:

Acceptance Test: Whether to accept the width of the kernel function can be decided according to formula (17), and the Metropolis rules are described in detail in another paper [43]. The maximum number of loops is set to 200 in this paper.

$$\begin{cases} \text{Update kernel function value,} & \text{if } E(new) > E(old) \text{ and } p<P(\text{accept } S_{new}), 0<p<1 \\ \text{Update kernel function value,} & \text{if } E(new) \leq E(old) \\ \text{The kernel function remains the same,} & Otherwise \end{cases}$$

$$\qquad (17)$$

4.3 Seasonal adjustment

As mentioned previously, traffic flow data reveals seasonal trends during daily peak periods. Therefore, for a model to produce a highly accurate forecasting performance, it is necessary that this seasonal component be estimated. There are several approaches to estimating the seasonal index of data series, including product-model type and non-product-model type. Based on the data series type consideration, this investigation employed Deo and Hurvich's approach to computing the seasonal index, which is shown in the following two steps:

Step 1: Calculate the seasonal index for each peak period. Calculate the peak difference between the actual traffic flow and the forecasting value (from the RVRCSA model) as $peak_t = \dfrac{q(t)}{\hat{q}(t)} = \dfrac{q(t)}{\sum_{i=1}^{N} w_i K(X_i, X) + w_0}$, where $t = j, l+j, 2l+j, \cdots, (m-1)l+j$ only

for the same peak time point in each period. Then, the seasonal index (SI) for each peak time point j is computed as $SI_j = \dfrac{1}{m}(peak_j + peak_{l+j} + \cdots peak_{(m-1)l+j})$.

Step 2: Compute the forecasting value of the SRVRCSA model. Eventually, the forecasting value of the SRVRCSA is obtained by $f_{N+k} = (\sum_{i=1}^{N} (w_i K(X_i, X) + w_0)) SI_k$,

where $k = j, l+j, 2l+j, \cdots, (m-1)l+j$ implies the peak time point in another period.

5   Experimental result and analysis

The forecasting performances of the various methods were evaluated using the mean average relative error (MARE). The example employed flow data used by The Transportation Data Research Laboratory (TDRL) at the University of Minnesota Duluth, which used a remote transportation microwave detector to collect data from an urban road located in Minnesota Twin Cities and Rochester City. The data included traffic flow from March 30 to April 26, 2015, from 6:00 to 11:00 and from 15:00 to 20:00. The experiments used the traffic flow from each hour to train and forecast. **Table 2** details the data used for learning and forecasting. Since the total traffic flow and distribution is very different for a weekend than a

weekday, the data was separated into weekend and weekday data, and then, each data category was used to learn and forecast separately. For example, before forecasting data for April 25 and 26 (a Saturday and Sunday), the data from April 4, 5, 11, 12, 18, and 19 (the earlier weekends in April) would be used to train the flow data in each time period and to train the model.

In the training stage, a rolling-based forecasting procedure was conducted, and in the validation and testing stage, a one hour-ahead forecasting policy was adopted. Then, several types of data-rolling were considered for forecasting the traffic flow of the next hour. Different traffic flow numbers in a time series were fed into the RVRCSA model to forecast the traffic flow in the next validation period. While training error improvement occurred, the kernel parameter of the RVRCSA model adjusted by the CSA algorithm was employed to calculate the validation error. Then, the adjusted parameters with minimum validation error were selected as the most appropriate parameters.

Table 2. Learning and prediction data from experiments

Peaks	learning			forecasting
Week days	March 30	April 6	April 13	April 20
	March 31	April 7	April 14	April 21
	April 1	April 8	April 15	April 22
	April 2	April 9	April 16	April 23
	April 3	April 10	April 17	April 24
Week end	April 4	April 11	April 18	April 25
	April 5	April 12	April 19	April 26

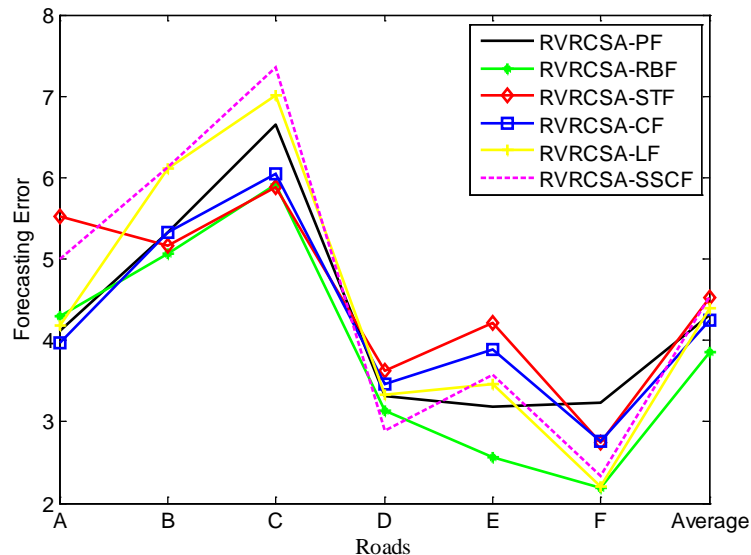## 5.1 Kernel selection of RVRCSA traffic forecasting model

Fig. 3. Plot of forecasting results for weekdays by RVRCSA with different kernel functions
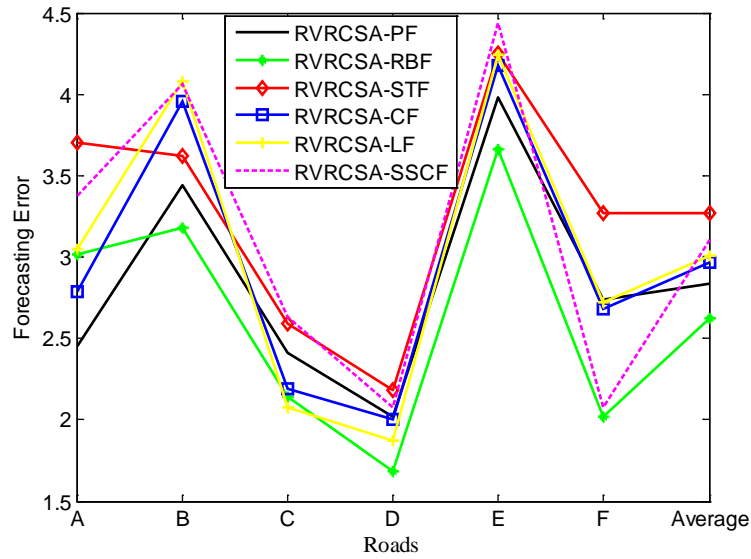
Fig. 4. Plot of forecasting results for weekend by RVRCSA with different kernel functions

First, six traffic flow datasets were used to analyze kernel function selection in RVRCSA traffic flow forecasting. The selected roads and flow detection point numbers are detailed in **Table 3**. In **Table 3**, A-F respectively represent the roads: T.H.252SB (station_id:N89123), T.H.77 CDSB (station_id:N8812), T.H.52NB (station_id:N8981), I-394 CDWB (station_id:N88425), I-35W CDSB (station_id:N87923), and I-35E CDNB (station_id:N91007).

Table 3. Comparison of RMSE value using the RVRCSA model with different kernel functions on six different road sections

road section	RVRCSA-PF		RVRCSA-GF		RVRCSA-STF		RVRCSA-CF		RVRCSA-LF		RVRCSA-SSCF	
	1	2	1	2	1	2	1	2	1	2	1	2
A	4.12	2.45	4.30	3.02	5.53	3.70	3.97	2.79	4.19	3.05	5.00	3.38
B	5.32	3.44	5.07	3.18	5.17	3.62	5.33	3.96	6.11	4.08	6.13	4.06
C	6.65	2.41	5.92	2.14	5.88	2.59	6.04	2.19	7.01	2.08	7.35	2.63
D	3.32	2.02	3.14	1.68	3.62	2.18	3.47	2.00	3.33	1.87	2.89	2.08
E	3.19	3.98	2.57	3.66	4.21	4.25	3.89	4.18	3.46	4.24	3.57	4.44
F	3.24	2.74	2.18	2.02	2.75	3.27	2.76	2.68	2.21	2.72	2.33	2.08
average	4.31	2.84	3.86	2.62	4.53	3.27	4.24	2.97	4.39	3.01	4.55	3.11

Table 3 lists the forecasting RMSE value of each kernel function used in the six roads, where 1 and 2 respectively represent work time and weekend time forecasting value. From **Table 3**, it can be seen that the forecasting capacity of the RVRCSA traffic flow forecasting varies greatly when using different kernel functions in different roads. For example, in the road section T.H. 77 CDSB (station_id: N8812) experiment, the RMSE of weekdays ranges from 5.07(GF) to 6.11(LF) when using the different kernel functions. However, the RMSE value for the weekend ranges from 3.18(GF) to 4.08(LF). From **Table 3**, **Fig. 3**, and **4**, it can be seen that when using the GF kernel function, the RMSE value of the RVRCSA traffic

flow forecasting model has a good performance for both weekday and weekend. Therefore, in the later forecasting capacity comparison experiment between SARIMA and DSHW, the kernel functions of the RVRCSA and SRVRCSA models both use GF.

5.2 Comparison results for each forecasting methodology

In this section, the forecasting results for SARIMA, DSHW, RVRCSA, and SRVRCSA on road T.H.252SB (station_id:N89123) is compared. First, the seasonal indexes for each peak time point are shown in **Table 4**, which are calculated based on the forecasting values of the RVRCSA as shown in **Table 2**.

Table 4. Seasonal indexes for each peak time point

| week days peak				week end peak			
Time	**Seasonal index**	**Time**	**Seasonal index**	**Time**	**Seasonal index**	**Time**	**Seasonal index**
peak 6-7 o'clock	1.0832	**peak 15-16 o'clock**	1.0012	**peak 6-7 o'clock**	0.5831	**peak 15-16 o'clock**	1.0011
peak 7-8 o'clock	1.1221	**peak 16-17 o'clock**	0.9698	**peak 7-8 o'clock**	0.8693	**peak 16-17 o'clock**	1.0101
peak 8-9 o'clock	0.9922	**peak 17-18 o'clock**	0.9486	**peak 8-9 o'clock**	0.9742	**peak 17-18 o'clock**	0.9987
peak 9-10 o'clock	0.9363	**peak 18-19 o'clock**	0.8864	**peak 9-10 o'clock**	1.0012	**peak 18-19 o'clock**	0.9184
peak 10-11 o'clock	0.8729	**peak 19-20 o'clock**	0.7382	**peak 10-11 o'clock**	1.0009	**peak 19-20 o'clock**	0.9322
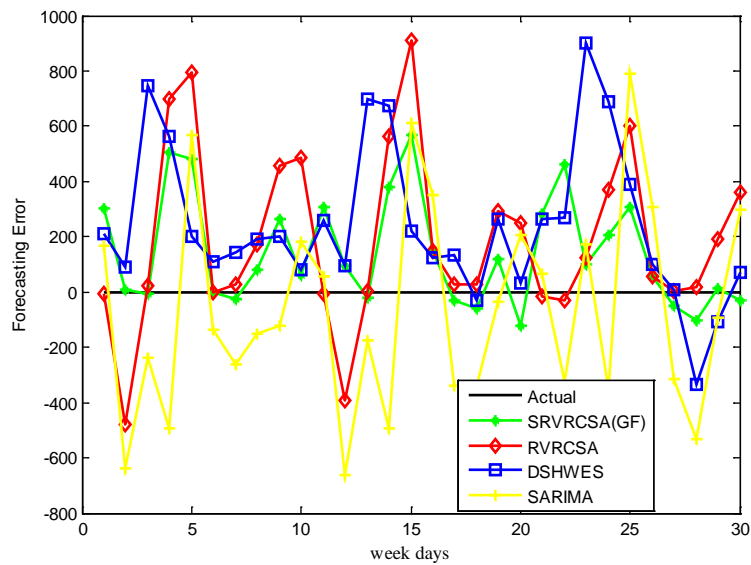
Fig. 5. Traffic flow forecasting errors for week days by different models
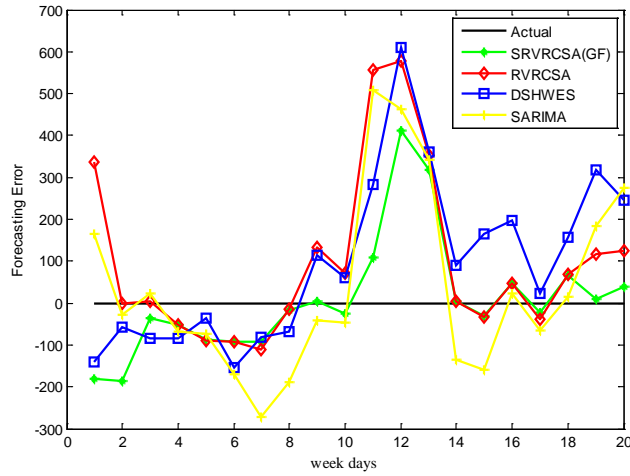
Fig. 6. Traffic flow forecasting errors for weekends by different models

Table 5. Traffic flow forecasting results for weekdays

TIME	Actual	SRVRCSA(GF)	RVRCSA(GF)	DSHWES	SARIMA
6-7 o'clock Apr.22	3770	4075	3762	3982	3939
7-8 o'clock Apr.22	4455	4461	3976	4543	3815
8-9 o'clock Apr.22	3451	3445	3472	4198	3215
9-10 o'clock Apr.22	2315	2821	3013	2876	1821
10-11 o'clock Apr.22	1661	2144	2456	1864	2230
15-16 o'clock Apr.22	2056	2057	2055	2167	1919
16-17 o'clock Apr.22	1842	1815	1872	1987	1579
17-18 o'clock Apr.22	1570	1653	1743	1762	1420
18-19 o'clock Apr.22	1233	1499	1692	1433	1110
19-20 o'clock Apr.22	1127	1189	1611	1209	1309
6-7 o'clock Apr.23	3758	4066	3754	4019	3815
7-8 o'clock Apr.23	4418	4515	4024	4512	3754
8-9 o'clock Apr.23	3416	3394	3421	4112	3242
9-10 o'clock Apr.23	2312	2693	2876	2985	1818
10-11 o'clock Apr.23	1766	2335	2675	1988	2377
15-16 o'clock Apr.23	1951	2101	2098	2074	2300
16-17 o'clock Apr.23	1875	1845	1901	2011	1537
17-18 o'clock Apr.23	1660	1602	1689	1629	1321
18-19 o'clock Apr.23	1240	1358	1532	1502	1207
19-20 o'clock Apr.23	1166	1045	1416	1198	1370
6-7 o'clock Apr.24	3613	3897	3598	3876	3680
7-8 o'clock Apr.24	4055	4515	4024	4322	3730
8-9 o'clock Apr.24	3061	3162	3187	3964	3232
9-10 o'clock Apr.24	2196	2403	2567	2885	1841
10-11 o'clock Apr.24	1712	2018	2312	2100	2502
15-16 o'clock Apr.24	1946	2004	2002	2045	2254
16-17 o'clock Apr.24	1813	1761	1816	1823	1499
17-18 o'clock Apr.24	1773	1670	1792	1437	1243
18-19 o'clock Apr.24	1383	1397	1576	1276	1289
19-20 o'clock Apr.24	1129	1098	1488	1198	1429
RMSE		2.49	4.30	3.61	4.66

Table 6. Traffic flow forecasting results for weekends

TIME	Actual	SRVRCSA(GF)	RVRCSA(GF)	DSHWES	SARIMA
6-7 o'clock Apr.25	906	724	1242	765	1070
7-8 o'clock Apr.25	1414	1228	1413	1357	1387
8-9 o'clock Apr. 25	1574	1537	1578	1489	1598
9-10 o'clock Apr. 25	1742	1691	1689	1659	1673
10-11 o'clock Apr. 25	1802	1714	1713	1766	1728
15-16 o'clock Apr.25	1882	1791	1789	1729	1712
16-17 o'clock Apr.25	1877	1784	1766	1795	1604
17-18 o'clock Apr.25	1757	1741	1743	1688	1568
18-19 o'clock Apr.25	1448	1452	1582	1562	1406
19-20 o'clock Apr.25	1353	1327	1424	1413	1307
6-7 o'clock Apr. 26	516	625	1072	798	1025
7-8 o'clock Apr. 26	689	1102	1268	1298	1151
8-9 o'clock Apr. 26	1004	1321	1356	1365	1345
9-10 o'clock Apr. 26	1521	1526	1525	1611	1386
10-11 o'clock Apr. 26	1624	1593	1592	1789	1464
15-16 o'clock Apr.26	1526	1574	1573	1723	1550
16-17 o'clock Apr.26	1580	1557	1541	1604	1515
17-18 o'clock Apr.26	1410	1476	1478	1566	1425
18-19 o'clock Apr.26	1193	1203	1310	1512	1378
19-20 o'clock Apr.26	1129	1168	1253	1375	1403
RMSE		1.85	3.32	3.29	3.41

The well-trained models, SARIMA, DSHW, RVRCSA, and SRVRCSA, were applied to forecast the traffic flow during the weekday/weekend period. **Tables 5** and **6** show the actual values and the forecast values obtained using the various forecasting models in the weekday peak and the weekend peak, respectively. **Fig. 5** and **6** are provided to clearly illustrate the forecasting errors among the different models in the weekday peak and the weekend peak, respectively. The RMSE values for each peak hour were calculated in order to make a fair comparison between the proposed models with other alternative models. The proposed SRVRCSA model resulted in smaller RMSE values than the SARIMA, DSHW, and RVRCSA models for capturing traffic flow patterns on an hourly average basis. Clearly, the seasonal adjustment employed here is proficient in dealing with cyclic peak data forecasting problems.

6 Conclusions

This paper compares and analyzes the forecasting performance of the RVRCSA traffic flow forecasting model when using the Polynomial Function, Gaussian Function, Symmetric Triangle Function, Cauchy Function, Laplace Function, Squared Sin Cardinal Function, and other kernel functions, respectively. Based upon the results obtained in this study, it is recommended that the Gaussian kernel function be selected for the RVRCSA traffic flow forecasting model that is based on combining the relevance vector with the chaotic genetic optimization traffic flow forecasting model. Founded on this, RVRCSA was added to the RVRCSA model, and the seasonal change of weekday and weekend traffic flow was taken into consideration. Then, a traffic flow forecasting model by hybrid chaotic simulated annealing optimization with seasonal RVR was proposed. The experiments revealed that the employed seasonal adjustment is proficient in dealing with cyclic peak data forecasting problems.

Future research includes:
1. The general consideration of more traffic state effect factors, such as traffic flow seasonal change, average speed, and lane occupancy rate data in order to build a more precise model.
2. The use of multiple kernel functions or a seasonal kernel function structure.
3. The performance of deep research on the amount of sample selection in relevance vector machine traffic flow forecasting.

Reference

[1] Chan, Kit Yan et al., "Neural-Network-Based Models for Short-Term Traffic Flow Forecasting Using a Hybrid Exponential Smoothing and Levenberg–Marquardt Algorithm," *IEEE Transactions on Intelligent Transportation Systems,* vol. 13, no. 2, pp. 644-654, June 2012. Article (CrossRef Link)

[2] Jeong, Young-Seon et al., "Supervised weighting-online learning algorithm for short-term traffic flow prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol.14, no.4, pp.1700-1707, December 2013. Article (CrossRef Link)

[3] Li, Ming-Wei, Duan-Feng Han and Wen-long Wang, "Vessel traffic flow forecasting by RSVR with chaotic cloud simulated annealing genetic algorithm and KPCA," *Neurocomputing,* vol.157, pp.243-255, June 2015. Article (CrossRef Link)

[4] Thomas, Tom, W. Weijermars, and E. Van Berkum, "Predictions of urban volumes in single time series," *IEEE Transactions on Intelligent Transportation Systems,* vol. 11, no. 1, pp.71-80, March 2010. Article (CrossRef Link)

[5] Zheng, Zuduo and D. Su, "Short-term traffic volume forecasting: A k-nearest neighbor approach enhanced by constrained linearly sewing principle component algorithm," *Transportation Research Part C Emerging Technologies*, vol. 43, pp. 143-157, June 2014. Article (CrossRef Link)

[6] Hong, Richang et al., "Cross-Domain Multi-View Object Retrieval via Multi-Scale Topic Models," *IEEE Transactions on Image Processing*, pp. 99, September 2016. Article (CrossRef Link)

[7] http://www.google.com/patents/US20110128127.

[8] Asensio, Javier et al., "Studying Individualized Transit Indicators Using a New Low-Cost Information System," *Handbook of Research on Embedded System Design,* pp. 388-407, January 2014. Article (CrossRef Link)

[9] Fernándezares, Antonio, et al., "Comparing Wireless Traffic Tracking with Regular Traffic Control Systems for the Detection of Congestions in Streets," *ACM Symposium on Applied Computing,* pp. 552-553, June 2016. Article (CrossRef Link)

[10] Tipping, Michael E, "Bayesian inference: An introduction to principles and practice in machine learning," *Advanced lectures on machine Learning.* Springer Berlin Heidelberg, pp. 41-62, 2004. Article (CrossRef Link)

[11] Mingjun, Ji and Tang Huanwen, "Application of chaos in simulated annealing," *Chaos, Solitons & Fractals*, vol.21, no.4, pp.933-941, August 2004. Article (CrossRef Link)

[12] Khader, Abdelhaleem I. and Mac McKee, "Use of a relevance vector machine for groundwater quality monitoring network design under uncertainty," *Environmental Modelling & Software*, vol. 57, pp. 115-126, July 2014. Article (CrossRef Link)

[13] Li, Hong, Donghui Pan and CL Philip Chen, "Intelligent prognostics for battery health monitoring using the mean entropy and relevance vector machine," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 44, no. 7, pp. 851-862, July 2014. Article (CrossRef Link)

[14] Juang, Chia Feng, "A hybrid of genetic algorithm and particle swarm optimization for recurrent network design," *IEEE Transactions on Systems Man & Cybernetics Part B Cybernetics A Publication of the IEEE Systems Man & Cybernetics Society,* vol. 34, no. 2, pp. 997, March 2004. Article (CrossRef Link)

[15] http://www.dumn.edu/tdrl/traffic/

[16] Williams, Billy, "Multivariate vehicular traffic flow prediction: evaluation of ARIMAX modeling," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1776, pp. 194-200, 2001. Article (CrossRef Link)

[17] Hong, Wei-Chiang, "Traffic flow forecasting by seasonal SVR with chaotic simulated annealing algorithm," *Neurocomputing*, vol. 74, no. 12-13, pp. 2096-2107, June 2011. Article (CrossRef Link)

[18] Lippi, Marco, Matteo Bertini and Paolo Frasconi, "Short-term traffic flow forecasting: An experimental comparison of time-series analysis and supervised learning," *IEEE Transactions on Intelligent Transportation Systems*, vol. 14, no. 2, pp. 871-882, March 2013. Article (CrossRef Link)

[19] Ghosh, B, B. Basu, and M. O'Mahony, "Multivariate Short-Term Traffic Flow Forecasting Using Time-Series Analysis," *Intelligent Transportation Systems IEEE Transactions on,* vol. 10, no. 2, pp. 246-254, July 2009. Article (CrossRef Link)

[20] Stathopoulos, Anthony and M. G. Karlaftis, "A multivariate state space approach for urban traffic flow modeling and prediction," *Transportation Research Part C Emerging Technologies*, vol. 11, no. 2, pp. 121-135, April 2003. Article (CrossRef Link)

[21] Kamarianakis, Yiannis, and P. Prastacos, "Space-time modeling of traffic flow," *Computers & Geosciences,* vol. 31, no. 2, pp. 119-133, March 2005. Article (CrossRef Link)

[22] Clark, Stephen, "Traffic Prediction Using Multivariate Nonparametric Regression," *Journal of Transportation Engineering*, vol. 129, no. 2, pp. 161-168, March 2003. Article (CrossRef Link)

[23] Okutani, Iwao and Y. J, Stephanedes, "Dynamic prediction of traffic volume through Kalman filtering theory," *Transportation Research Part B Methodological*, vol. 18, no. 1, pp. 1-11, February 1984. Article (CrossRef Link)

[24] Sun, Hongyu, and H. X. Liu, "Short-Term Traffic Forecasting Using the Local Linear Regression Model," *Center for Traffic Simulation Studies*, 2002. Article (CrossRef Link)

[25] Cai, Pinlong et al., "A spatiotemporal correlative k-nearest neighbor model for short-term traffic multistep forecasting," *Transportation Research Part C Emerging Technologies*, vol. 62, pp. 21-34, January 2016. Article (CrossRef Link)

[26] Xia, Dawen, et al, "A distributed spatial–temporal weighted model on MapReduce for short-term traffic flow forecasting," *Neurocomputing,* vol. 179, pp. 246-263, February 2016. Article (CrossRef Link)

[27] Chen, Cheng et al., "Distributed Modeling in a MapReduce Framework for Data-Driven Traffic Flow Forecasting," *IEEE Transactions on Intelligent Transportation Systems*, vol. 14, no. 1, pp. 22-33, March 2013. Article (CrossRef Link)

[28] Daraghmi, Yousef Awwad, C. W. Yi and T. C. Chiang, "Negative Binomial Additive Models for Short-Term Traffic Flow Forecasting in Urban Areas," *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 2, pp. 784-793, April 2014. Article (CrossRef Link)

[29] Tchrakian, T. T., B. Basu and M. O'Mahony, "Real-Time Traffic Flow Forecasting Using Spectral Analysis," *IEEE Transactions on Intelligent Transportation Systems*, vol.13, no.2, pp.519-526, June 2012. Article (CrossRef Link)

[30] Sun, Shiliang and C. Zhang, "The Selective Random Subspace Predictor for Traffic Flow Forecasting," *IEEE Transactions on Intelligent Transportation Systems*, vol. 8, no. 2, pp. 367-373, June 2007. Article (CrossRef Link)

[31] Oh, Se Do, Y. J. Kim, and J. S. Hong, "Urban Traffic Flow Prediction System Using a Multifactor Pattern Recognition Model," *IEEE Transactions on Intelligent Transportation Systems,* vol. 16, no. 5, pp. 1-12, October 2015. Article (CrossRef Link)

[32] Sun, Shiliang and X. Xu, "Variational Inference for Infinite Mixtures of Gaussian Processes With Applications to Traffic Flow Prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 12, no. 2, pp. 466-475, June 2011. Article (CrossRef Link)

[33] Simroth, Axel and H. Zähle, "Travel Time Prediction Using Floating Car Data Applied to Logistics Planning," *IEEE Transactions on Intelligent Transportation Systems*, vol. 12, no. 1, pp. 243 - 253, January 2011. Article (CrossRef Link)

[34] Vlahogianni, Eleni I., M. G. Karlaftis and J. C. Golias, "Optimized and meta-optimized neural networks for short-term traffic flow prediction: A genetic approach." *Transportation Research Part C Emerging Technologies*, vol. 13, no. 3, pp. 211-234, June 2005. Article (CrossRef Link)

[35] Yin, Hongbin et al., "Urban traffic flow prediction using a fuzzy-neural approach," *Transportation Research Part C Emerging Technologies*, vol. 10, no. 2, pp. 85-98, April 2002. Article (CrossRef Link)

[36] Li, Chungui, S. Xu and X. Wen, "Traffic Flow Forecasting Algorithm Using Simulated Annealing Genetic BP Network," in *Proc. of 2010 International Conference on Measuring Technology and Mechatronics Automation*, vol. 3, pp. 1043-1046, March 2010. Article (CrossRef Link)

[37] Chan, Kit Yan, et al, "Neural-Network-Based Models for Short-Term Traffic Flow Forecasting Using a Hybrid Exponential Smoothing and Levenberg–Marquardt Algorithm," *IEEE Transactions on Intelligent Transportation Systems,* vol. 13, no. 2, pp. 644-654, June 2012. Article (CrossRef Link)

[38] Yang, Wen et al., "Traffic flow prediction based on wavelet transform and Radial Basis Function network," in *Proc. of Logistics Systems and Intelligent Management, 2010 International Conference on IEEE,* vol. 2, pp. 969-972, January 2010. Article (CrossRef Link)

[39] Chan, Kit Yan, T. S. Dillon and E. Chang, "An Intelligent Particle Swarm Optimization for Short-Term Traffic Flow Forecasting Using on-Road Sensor Systems," *IEEE Transactions on Industrial Electronics*, vol. 60, no. 10, pp. 4714-4725, August 2012. Article (CrossRef Link)

[40] Tipping, Michael E, "Sparse bayesian learning and the relevance vector machine." *Journal of Machine Learning Research,* vol. 1, no. 3, pp. 211-244, September 2001. Article (CrossRef Link)

[41] Yu, Shiwei, Ke Wang and Yi-Ming Wei, "A hybrid self-adaptive Particle Swarm Optimization–Genetic Algorithm–Radial Basis Function model for annual electricity demand prediction," *Energy Conversion and Management,* vol. 91, pp.176-185, February 2015. Article (CrossRef Link)

[42] Jeong, Young Seon, et al, "Supervised Weighting-Online Learning Algorithm for Short-Term Traffic Flow Prediction," *IEEE Transactions on Intelligent Transportation Systems,* vol.14, no. 4, pp. 1700-1707, December 2013. Article (CrossRef Link)

[43] Metropolis, Nicholas et al., "Equation of state calculations by fast computing machines," *The journal of chemical physics*, vol. 21, no. 6, pp. 1087-1092, 1953. Article (CrossRef Link)

Zhangguo Shen is a Ph.D. candidate in the College of Computer Science and Technology, Zhejiang University of Technology, Hangzhou,China. He received the B.E. degree in Computer Science from Zhejiang University City College, Hangzhou, China, in 2004 and the M.S. degree in Computer Science from Hangzhou Dianzi University, Hangzhou, China, in 2007. His research interests include traffic flow forecasting and cloud computing.
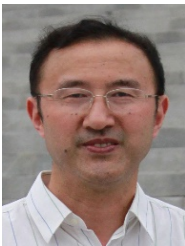
Wanliang Wang received the B.S. degree in industrial automation from Jiangsu University, Jiangsu, China, in 1982, and the Ph.D. degree in control theory and engineering from Tongji University, Shanghai, China, in 2001. He is currently a Professor with the School of Computer Science and Technology, Zhejiang University of Technology, Hangzhou, China. His current research interests include wireless network, artificial intelligence, and intelligent control.

Qing Shen received the B.E. (2004), and the M.E. (2007) degrees in Computer Science from North University of China,Taiyuan ,China. She is now a Lecturer in the School of Information Engineering of Huzhou University. Her current research interests include artificial intelligence, software testing and software reliability evaluation.

Zechao Li is a Professor in the School of Computer Science and Engineering, Nanjing University of Science and Technology, China. He received the B.E. degree from University of Science and Technology of China in 2008 and the Ph.D. degree from Institute of Automation, Chinese Academy of Sciences in 2013. His research interests include media intelligent analysis, social network analysis, etc. He was selected in the Young Talent Program of China Association for Science and Technology and received the 2015 Excellent Doctoral Dissertation of Chinese Academy of Sciences, Excellent Doctoral Theses of China Computer Federation and President Scholarship of Chinese Academy of Science.