J Inf Process Syst, Vol.18, No.6, pp.719~728, December 2022
https://doi.org/10.3745/JIPS.02.0185

ISSN 1976-913X (Print)
ISSN 2092-805X (Electronic)

JOURNAL OF INFORMATION PROCESSING SYSTEMS JIPS

Application of an Optimized Support Vector Regression Algorithm in Short-Term Traffic Flow Prediction

Ruibo Ai[1], Cheng Li[2,*], and Na Li[3]

Abstract

The prediction of short-term traffic flow is the theoretical basis of intelligent transportation as well as the key technology in traffic flow induction systems. The research on short-term traffic flow prediction has showed the considerable social value. At present, the support vector regression (SVR) intelligent prediction model that is suitable for small samples has been applied in this domain. Aiming at parameter selection difficulty and prediction accuracy improvement, the artificial bee colony (ABC) is adopted in optimizing SVR parameters, which is referred to as the ABC-SVR algorithm in the paper. The simulation experiments are carried out by comparing the ABC-SVR algorithm with SVR algorithm, and the feasibility of the proposed ABC-SVR algorithm is verified by result analysis. Continuously, the simulation experiments are carried out by comparing the ABC-SVR algorithm with particle swarm optimization SVR (PSO-SVR) algorithm and genetic optimization SVR (GA-SVR) algorithm, and a better optimization effect has been attained by simulation experiments and verified by statistical test. Simultaneously, the simulation experiments are carried out by comparing the ABC-SVR algorithm and wavelet neural network time series (WNN-TS) algorithm, and the prediction accuracy of the proposed ABC-SVR algorithm is improved and satisfactory prediction effects have been obtained.

Keywords

Artificial Bee Colony Algorithm, Optimization, Prediction Algorithm, Short-time Traffic Flow, Support Vector Regression

1. Introduction

The traffic congestion in large and medium-sized cities proceeds to be a severe problem due to the increasing number of car buyers during the process of the continuous acceleration of urbanization and the rapid development of economy. The traffic congestion has not only polluted the environment, but also surged the probability of traffic accidents. Hence, intelligent transport system (ITS) has emerged in response. Traffic information can be collected and analyzed in real time by ITS so as to induce the traffic control. The key technology of ITS is short-term traffic flow prediction, which has a sampling interval of no more than 15 minutes [1]. In practical application, how to effectively and precisely carry out the traffic flow prediction is still a great challenge due to the characteristics of traffic information and the various interference factors in collection process [2-4].

Manuscript received December 8, 2021; first revision February 15, 2022; second revision April 29, 2022; third revision June 16, 2022; accepted July 8, 2022.

* **Corresponding Author:** Cheng Li (licheng@qqhru.edu.cn)
[1] College of Sciences, Qiqihar University, Qiqihar, China (lcxsnh1@163.com)
[2] College of Computer and Control Engineering, Qiqihar University, Qiqihar, China (licheng@qqhru.edu.cn)
[3] Public Foreign Language Teaching and Research Department, Qiqihar University, Qiqihar, China (hhln2000@163.com)

The research on short-term traffic flow prediction algorithm has partly become one of research focuses at home and abroad. As early as the 1970s, the mature prediction algorithms from the various disciplines ranging from economics to physics were employed widely by some scholars. Among them, the linear theory algorithm and the statistical theory algorithm were mainly adopted, which included ARIMA model [5] and SARIMA model [6], so as to enhance the accuracy and performance of prediction. In addition, it also included historical average algorithm [7], time series algorithm [8], Kalman filter algorithm [9], etc. Although the application of these conventional algorithms is relatively mature, the calculation error is pretty larger. Currently, various intelligent algorithms have been employed to carry out modeling for prediction by scholars, for making up for the drawbacks existing in these algorithms above. A novel KNN algorithm was utilized in the prediction domain, aiming at reducing the prediction errors [10]. At the same time, neural network algorithm [11] as well as support vector machine algorithm [2] has emerged. The prediction results of the intelligent algorithm are common to generate the local optimal solutions and influence the preciseness. Recently, the deep learning model has been employed in terms of enhancing the accuracy of various prediction results [12], but it requires the support of data sets with the abundant samples. Consequently, the deep learning model is not suitable for the prediction. A multiple of researches have indicated that the combination prediction algorithm is more effective for predicting the short-term traffic flow. For the algorithm, two or more prediction methods are used aiming at short-term traffic flow prediction. And advantages of various methods are exploited fully, which has improved the prediction accuracy and extensibility [1,2,13-15].

Based on summarizing the current research findings, in order to effectively and precisely carry out the traffic flow prediction, an intelligent combination prediction algorithm is proposed in the paper, which is named as the artificial bee colony-support vector regression (ABC-SVR) algorithm. The major contributions of the paper are listed below.

- Through the literature analysis, it is concluded that the intelligent combination prediction algorithm is more suitable short-term traffic flow prediction algorithm for small samples.
- An SVR algorithm optimized by an ABC algorithm is applied to the field of the short-term traffic flow prediction for the first time.
- Through a large number of experiments, it is proved that the ABC-SVR algorithm is feasible, the optimization effect is good, and the prediction accuracy is high.

The remainder of the paper is organized as follows. In Section 2, based on ABC optimization algorithm, the short-term traffic flow prediction model by SVR is presented. In Section 3, the simulation experiments are carried out. In Section 4, the result discussion is demonstrated. Finally, conclusions are given with the importance and the practical value of the optimal algorithm as well as its future research directions.

2. SVR Model Optimization by Artificial Bee Colony Model

According to the optimization process of the ABC algorithm, the ABC-SVR prediction model is constructed, which is illustrated in Fig. 1.

The optimization process is presented in detail. Firstly, the historical data is preprocessed to form the training sample. Then, the swarm is initialized and the parameters are set. Next, the *SN* solution is generated randomly and the fitness value of each solution is calculated. Then, a new solution is generated and the fitness value of the new solution is calculated. If the fitness value of the new solution is higher, the old one is updated, food sources are selected according to the calculation, and a new solution is generated again. Else if the fitness value of the new solution is lower, the old solution remains unchanged, and the *limit* is equal to *limit*+1. If the *limit* reaches maximum, new solution is generated randomly and optimized continuously; if maximum iteration number is reached, the optimization process is ended. Finally, the optimal solution is assigned to SVR as a parameter for prediction.

Parameter optimization is to seek out the optimal SVR parameter set (C, σ, ε), so that the error between the prediction values and the true values can be minimized. Hence, the objective function of fitness value is calculated, which is shown in formula (1).

$$\underset{C,\delta,\varepsilon}{Min}\ f(C,\sigma,\varepsilon) = \frac{1}{N}\sum_{i=1}^{N}(y_i - y_i^*)^2$$
$$s.t.\begin{cases} C \in [C_{\min}, C_{\max}] \\ \sigma \in [\sigma_{\min}, \sigma_{\max}] \\ \varepsilon \in [\varepsilon_{\min}, \varepsilon_{\max}] \end{cases} \tag{1}$$

where, N is the sample number, and y_i and y_i^* are the actual traffic flow value and the predicted values of the i-th sample, respectively.

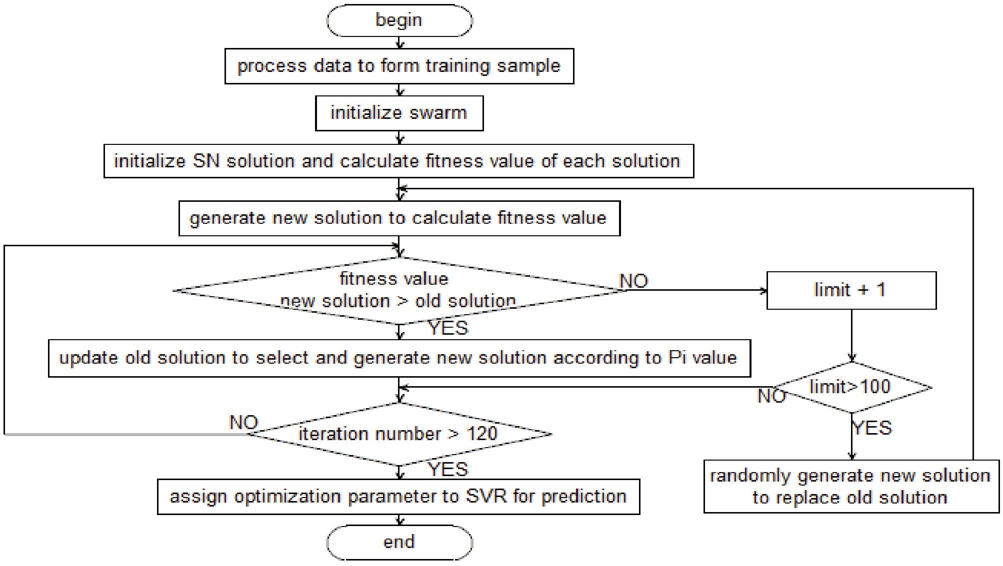


Fig. 1. ABC-SVR model construction.

3. Simulation Experiment

In the paper, three comparison simulation experiments are carried out severally.

- To verify the feasibility of the ABC-SVR model of this paper, the traditional SVR model is used to carry out comparison with the ABC-SVR model. The ABC-SVR prediction model is verified to be feasible for predicting short-term traffic flow.

- To verify the optimization effect of the ABC-SVR model, the particle swarm optimization SVR (PSO-SVR) model and the genetic optimization SVR (GA-SVR) model are adopted to carry out comparison with the ABC-SVR model. These three prediction models have the same optimization object, and the optimization object is the three parameters of SVR.
- To verify the prediction accuracy of the ABC-SVR model, the typical WNN-TS model is used to carry out comparison with the ABC-SVR model. The ABC-SVR prediction model proposed is verified to be more accurate in prediction process of short-term traffic flow.

The simulation experiments are carried out using Libsvm-3.21 toolkit, which is installed on MATLAB R2010b.

3.1 Experiment Data

The data of the experiments is downloaded at certain monitoring points from PEMS system website (California transportation performance measurement system, http://pems.dot.ca.gov/).

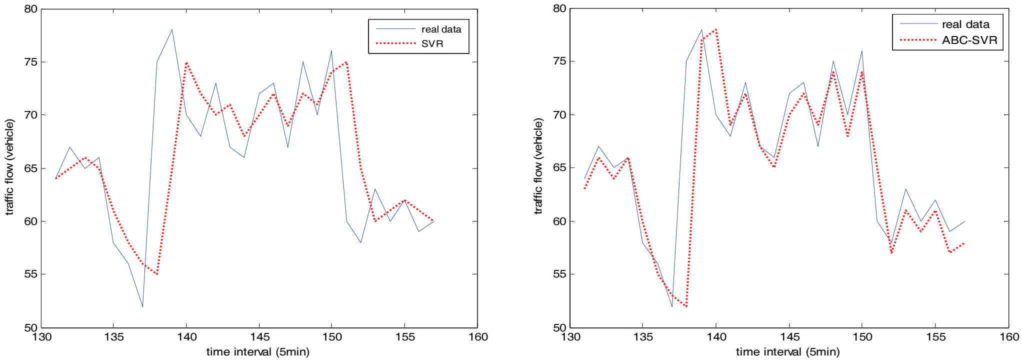
Considering the periodical changes of weekly traffic flow, and different traffic flow situations between the working days and the non-working days, the traffic flow data of the three consecutive days from September 7, 2017 (Thursday) to September 9, 2017 (Saturday) is selected as the research object. Considering the less traffic flow from 22:00 to 6:00 of the next day, and its less reference value on the actual prediction, the data sampling time of the three consecutive days is chosen from 6:00 to 22:00 every day to acquire the traffic flow data. Taking into account importance about data during traffic peak periods, the data from 7:00 to 20:00 of the 3 consecutive days is selected as the simulation experiment sample.

For the 3 consecutive days, from 7:00 to 20:00, with the collection time interval of 5 minutes, 471 data is collected in all. With 157 data per day, the first 130 traffic flow data is the training data of SVR. And 27 traffic flow data left is adopted as the test set for SVR.

3.2 Feasibility Experiment

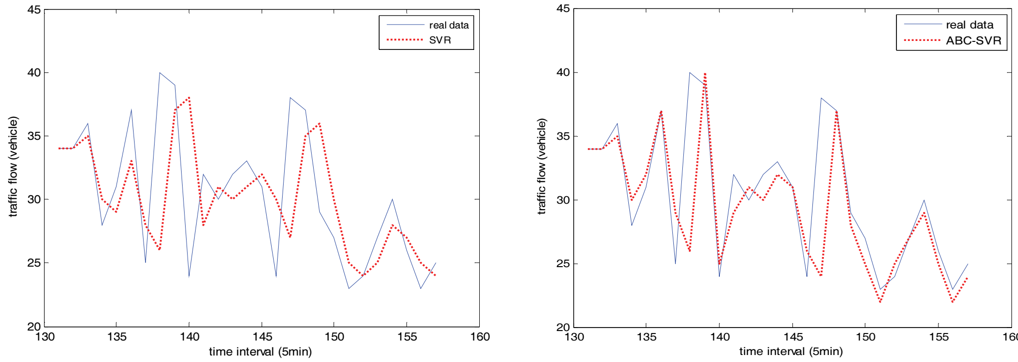
The initial value of each parameter of the optimized model are set as follows. The food source number is 20. The employed bee number is 10. The onlooker bee number is 10. *limit* is 100. And the maximum iteration number is 120.

The comparison curves between the true values and the prediction values of SVR model and the ABC-SVR model in 3 days are shown in Figs. 2–4.



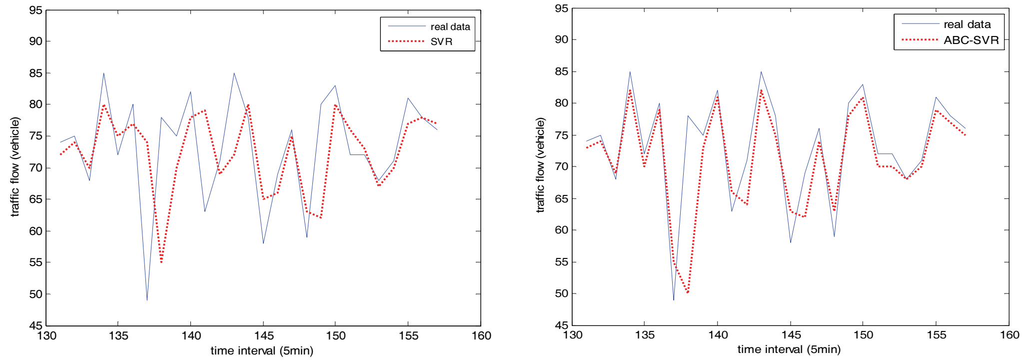
(a) (b)

Fig. 2. Prediction results of (a) SVR and (b) ABC-SVR on September 7, 2017.



(a) (b)

Fig. 3. Prediction results of (a) SVR and (b) ABC-SVR on September 8, 2017.



(a) (b)

Fig. 4. Prediction results of (a) SVR and (b) ABC-SVR on September 9, 2017.

3.3 Optimization Experiment

The specific parameter settings of each model in the simulation experiments are presented in Table 1. The three parameters of SVR are optimized, which is interpreted in Table 2.

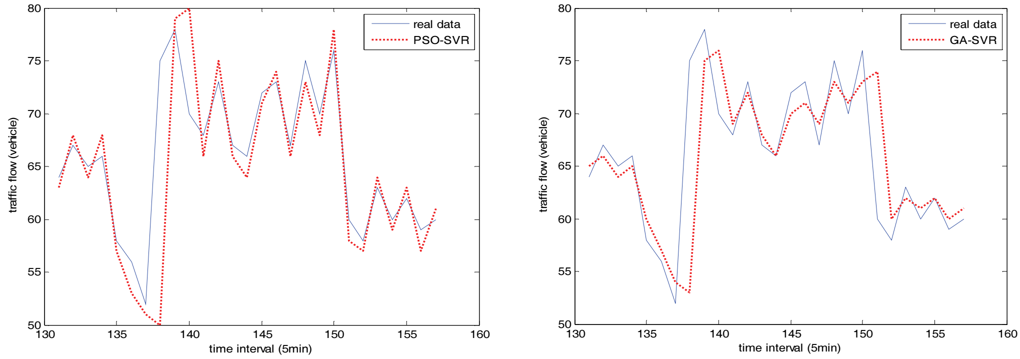
The comparison curves between the true values and the prediction values of PSO-SVR model and GA-SVR model in 3 days are illustrated in Figs. 5–7. The comparison curves between the true values and the prediction values of the ABC-SVR model in three days are illustrated in Figs. 2(b), 3(b), and 4(b).

Table 1. Initial parameter settings of each model

Optimization model	Parameter	Initial value
PSO	Weight ω	1.0
	$c1$, $c2$	1.7, 2.0
	Elasticity coefficient	1.0
GA	Pc	0.85
	Pm	0.02
ABC	Quantity of food source	20
	Employed bee, onlooker bee	10, 10
	limit	100
	Maximum number of iterations	120

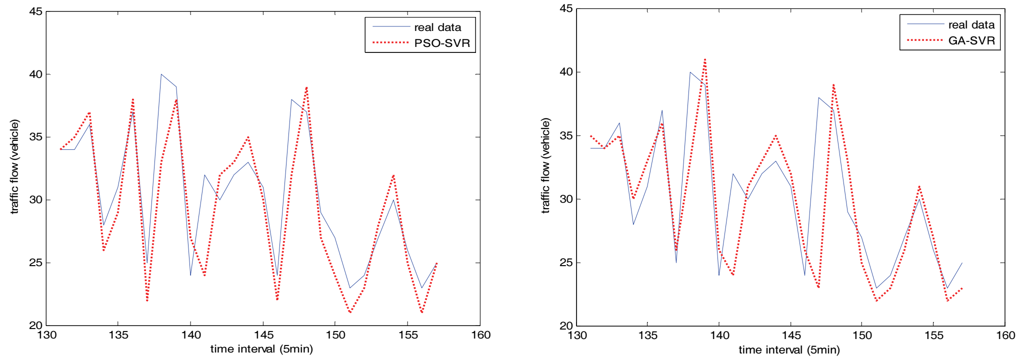
Table 2. Parameter optimization results of each optimized algorithm

Parameter	PSO-SVR	GA-SVR	ABC-SVR
C	14.31	20.19	48.22
ε	0.10	0.16	0.23
σ	0.64	0.68	0.56



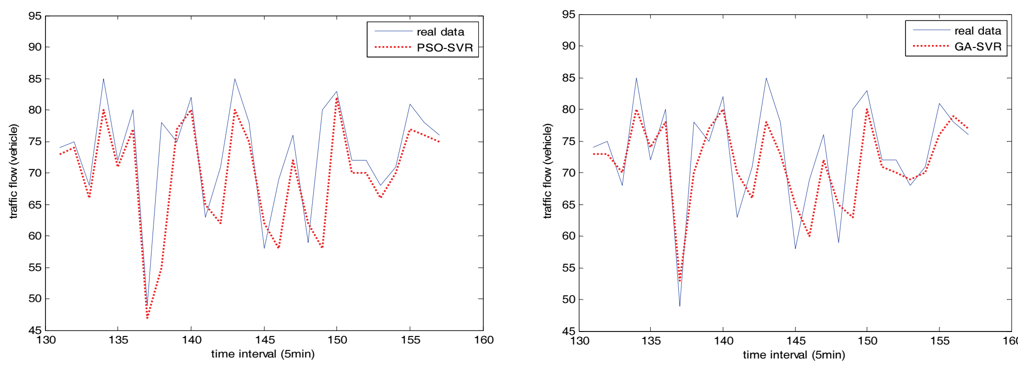
(a) (b)

Fig. 5. Prediction results of (a) PSO-SVR and (b) GA-SVR on September 7, 2017.



(a) (b)

Fig. 6. Prediction results of (a) PSO-SVR and (b) GA-SVR on September 8, 2017.



(a) (b)

Fig. 7. Prediction results of (a) PSO-SVR and (b) GA-SVR on September 9, 2017.

3.4 Accuracy Experiment

The wavelet neural network time series (WNN-TS) model time series delay is set to 4. There are 6 hidden layer nodes, and the learning rate is *lr1*=0.09 and *lr2*=0.03. Meanwhile, the maximum iteration number is set to 120.

The comparison curves between the true values and the predicted values of the WNN-TS model in 3 days are illustrated in Figs. 8–10. The comparison curves between the true values and the predicted values of the ABC-SVR model in 3 days are illustrated in Figs. 2(b), 3(b), and 4(b).

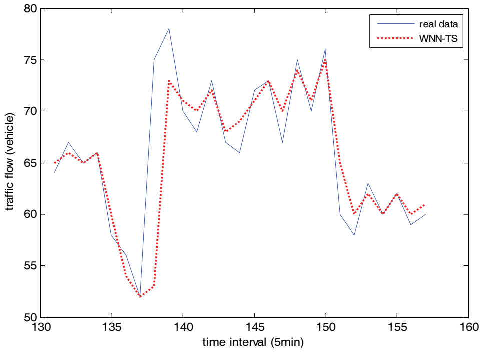


Fig. 8. Prediction results of WNN-TS on September 7, 2017.

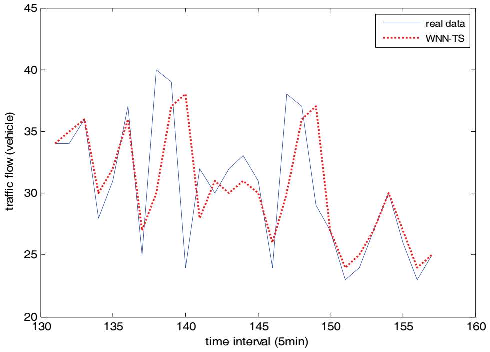


Fig. 9. Prediction results of WNN-TS on September 8, 2017.

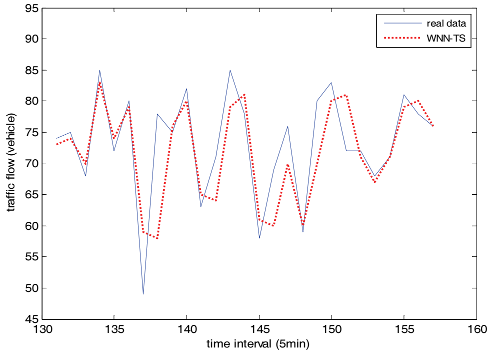


Fig. 10. Prediction results of WNN-TS on September 9, 2017.

4. Discussion

The horizontal comparison results between the true values and the predicted values of the SVR model and ABC-SVR model are illustrated in Figs. 2–4. The curve trend shows that the predicted value curve of the ABC-SVR model fits well with the true value curve, which illustrates these predicted results of the ABC-SVR are preciser than those of SVR. From the evaluation index mean square error (MSE) in Table 3, it can be proved that the MSE of the ABC-SVR is lower, indicating these predicted results of the ABC-SVR are preciser than those of the SVR.

Table 3. MSE of feasibility experiment results

Date	MSE	
	SVR	ABC-SVR
September 7, 2017	36.48	24.59
September 8, 2017	25.67	12.67
September 9, 2017	78.67	24.11

Through the analysis of the MSE results, it has been demonstrated that the prediction based on ABC optimization SVR in our research is more feasible as well as the higher accuracy has been achieved.

By analyzing Figs. 5–7, it can be observed that the curve fitting between the true values and the predicted values of the ABC-SVR, PSO-SVR, and GA-SVR are all fairly well. And it proves the accuracy of prediction for all three algorithms is preciser than that of SVR. From the evaluation indexes of prediction results of each optimization algorithm, it has been demonstrated that ABC-SVR presents much higher prediction accuracy, and the MSE of each model prediction result is shown in Table 4.

Table 4. MSE of optimization experiment results

Date	MSE		
	PSO-SVR	GA-SVR	ABC-SVR
September 7, 2017	29.07	28.70	24.59
September 8, 2017	12.96	15.56	12.67
September 9, 2017	51.19	29.33	24.11

Through the further analysis, the experimental results confirm that the ABC-SVR of the paper has more favorable prediction accuracy as well as higher prediction property, which verifies that the ABC algorithm has the optimal optimization effect on SVR parameters. *F* test is carried out by adopting SPSS for the three groups of the predicted results, and $p<0.05$, indicating that the difference among the three groups has statistical significance. In the meantime, from Table 2, it is observed that the parameter range of the SVR obtained by the ABC optimization algorithm is larger, which demonstrates that the global search ability of ABC algorithm is stronger than those of the PSO algorithm and GA algorithm, and ABC algorithm is more suitable aiming at large randomness.

By horizontal comparison experiments among the predicted values of WNN-TS, the predicted values of ABC-SVR and the true values, it is verified the predicted results for the ABC-SVR model have higher accuracy, from MSE of the evaluation index of results of each optimization model (Table 5).

Through the analysis of MSE, it has been observed that the short-term traffic flow prediction algorithm based on ABC optimization SVR of the paper has the optimal prediction capability and prediction accuracy.

Table 5. MSE of optimization experiment results

Date	MSE	
	WNN-TS	ABC-SVR
September 7, 2017	25.89	24.59
September 8, 2017	17.56	12.67
September 9, 2017	35.00	24.11

5. Conclusion

According to the characteristics of the strong global search ability of the ABC algorithm and the suitability for small samples of the SVR algorithm, the short-term traffic flow prediction algorithm by SVR is proposed based on ABC optimization algorithm, and the satisfactory results have been attained. The intelligent combination prediction algorithm is one of research focuses at home and abroad, and the ABC-SVR proposed in the paper can also be applied to other fields. The optimal algorithm will show its more important and more practical value in the future. Meanwhile, the adaptability in different fields of prediction algorithm, the public data set construction of traffic flow, the computational complexity of prediction process as well as the own optimization of optimization algorithms need further investigations and explorations in the future research.

Acknowledgement

This research was funded by the Education Department of Heilongjiang Province of China (No. 135309463 and 145109129).

References

[1] L. Chen, L. Zheng, J. Yang, D. Xia, and W. Liu, “Short-term traffic flow prediction: from the perspective of traffic flow decomposition,” *Neurocomputing*, vol. 413, pp. 444-456, 2020.

[2] C. Li and P. Xu, “Application on traffic flow prediction of machine learning in intelligent transportation,” *Neural Computing and Applications*, vol. 33, no. 2, pp. 613-624, 2021.

[3] M. Chen, R. Chen, F. Cai, W. Li, N. Guo, and G. Li, “Short-term traffic flow prediction with recurrent mixture density network,” *Mathematical Problems in Engineering*, vol. 2021, article no. 6393951, 2021. https://doi.org/10.1155/2021/6393951

[4] Q. Cao and S. Fujita, “Cost-effective replication schemes for query load balancing in DHT-based peer-to-peer file searches,” *Journal of Information Processing Systems*, vol. 10, no. 4, pp. 628-645, 2014.

[5] J. Abdi, B. Moshiri, B. Abdulhai, and A. K. Sedigh, “Short-term traffic flow forecasting: parametric and nonparametric approaches via emotional temporal difference learning,” *Neural Computing and Applications*, vol. 23, no. 1, pp. 141-159, 2013.

[6] B. M. Williams and L. A. Hoel, “Modeling and forecasting vehicular traffic flow as a seasonal ARIMA process: theoretical basis and empirical results,” *Journal of Transportation Engineering*, vol. 129, no. 6, pp. 664-672, 2003.

[7] J. Liu and W. Guan, “A summary of traffic flow forecasting methods,” *Journal of Highway and Transportation Research and Development*, vol. 21, no. 3, pp. 82-85, 2004.

[8] B. Ghosh, B. Basu, and M. O'Mahony, "Multivariate short-term traffic flow forecasting using time-series analysis," *IEEE Transactions on Intelligent Transportation Systems*, vol. 10, no. 2, pp. 246-254, 2009.
[9] L. Cai, Z. Zhang, J. Yang, Y. Yu, T. Zhou, and J. Qin, "A noise-immune Kalman filter for short-term traffic flow forecasting," *Physica A: Statistical Mechanics and its Applications*, vol. 536, article no. 122601, 2019. https://doi.org/10.1016/j.physa.2019.122601
[10] Y. Lv, Y. Duan, W. Kang, Z. Li, and F. Y. Wang, "Traffic flow prediction with big data: a deep learning approach," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 2, pp. 865-873, 2015.
[11] M. J. Ebadi, A. Hosseini, and M. M. Hosseini, "A projection type steepest descent neural network for solving a class of nonsmooth optimization problems," *Neurocomputing*, vol. 235, pp. 164-181, 2017.
[12] Z. Yan, H. Yang, F. Li, and Y. Lin, "A deep learning approach for short-term airport traffic flow prediction," *Aerospace*, vol. 9, no. 1, article no. 11, 2021. https://doi.org/10.3390/aerospace9010011
[13] F. Zhang, T. Y. Wu, Y. Wang, R. Xiong, G. Ding, P. Mei, and L. Liu, "Application of quantum genetic optimization of LVQ neural network in smart city traffic network prediction," *IEEE Access*, vol. 8, pp. 104555-104564, 2020.
[14] Q. Chen, Y. Song, and J. Zhao, "Short-term traffic flow prediction based on improved wavelet neural network," *Neural Computing and Applications*, vol. 33, no. 14, pp. 8181-8190, 2021.
[15] Y. Shao, Y. Zhao, F. Yu, H. Zhu, and J. Fang, "The traffic flow prediction method using the incremental learning-based CNN-LTSM model: the solution of mobile application," *Mobile Information Systems*, vol. 2021, article no. 5579451, 2021. https://doi.org/10.1155/2021/5579451

Ruibo Ai **https://orcid.org/0000-0003-4259-2959**

Now, he is an associate professor at College of Sciences, Qiqihar University. His research interests include optimization algorithm, artificial intelligence, machine learning, embedded system, etc.

Cheng Li **https://orcid.org/0000-0002-9656-5385**

He received his M.S. degree in science and technology of computer from Qiqihar University in 2013. Now, he is currently pursuing the Ph.D. degree in science and technology of computer in Harbin Engineering University. He is a professor at College of Computer and Control Engineering, Qiqihar University. His research interests include artificial intelligence, optimization algorithm, network security, etc.

Na Li **https://orcid.org/0000-0002-7692-0984**

She received her M.S. degree in Qiqihar University in 2010. Now, she is a lecturer at Qiqihar University. Her research interests include optimization algorithm, artificial intelligence, machine learning, pattern recognition, network security, etc.