Advances in Social Science, Education and Humanities Research (ASSEHR), volume 252
3rd International Conference on Judicial, Administrative and Humanitarian Problems of State Structures and Economic Subjects (JAHP 2018)

Study on Tourist Volume Forecasting Based on ABA-SVR Model Within Network Environment*

Binru Zhang
School of Finance and Economics
Yangtze Normal University
Chongqing, China 408100

Yulian Pu
School of Management
Yangtze Normal University
Chongqing, China 408100

***Abstract*—The search of consumers under internet environment reflects the potential tourism demands of tourists. This paper takes Sanya as an example and attempts to use the traffic flow data related to Sanya tourism from August 2009 to March 2016 and the network search data to construct the input sets of SVR model. And this paper also forecasts the domestic tourists received in Sanya by applying the ABA-SVR model, thereinto the Adaptive Bat Algorithm, ABA is used to optimize the free parameters of Support Vector Regression model, SVR. The 12-month forecasting results and the significance testing show that this method can effectively improve the forecast accuracy of the model. The forecast results can provide necessary reference for the macro-administration of policy-making departments related to tourism.**

Keywords—consumer search; adaptive bat algorithm; Support Vector Regression; tourist volume; forecast accuracy

I. INTRODUCTION

The correct forecast of tourism demand is of great importance to the scientific policy-making of tourism-related departments [1][2]. However, the time series of tourist volume and other tourist demands possesses complex nonlinear characteristics which causes the tourism-related departments to face the dilemmatic situation upon policy-making. On one hand, the coming of peak tourist season enables the tourist volume to rapidly increase within a short period which will cause the short of passenger transport, over-crowdedness in scenic areas and therefore the capacity of scenic areas will face serious challenge. On the other hand, in tourist off-season, the tourist volume will sharply decrease and the large quantity of facilities of tourism-related industries will be in idle which will cause unnecessary resource wastes [3] and the occurrence of certain emergency will cause unpredictable influence on tourists. Therefore, the tourism demand curve always represents certain nondeterminacy fluctuation and complex nonlinear characteristics which causes it hard for the traditional forecast techniques to achieve ideal forecast results (Wang, et al, 2010)[5]. The Support Vector Regression (SVR) possesses strong nonlinear forecast capacity. However, there exists the problem of critical parameters selection of model and the limitation on input set structure of model. This paper makes further discussion on the problems on the two aspects.

In recent years, the scholars have lunched the studies on machine learning forecast method including Artificial Neural Network, ANN, etc. ANN possesses favorable nonlinear forecast capacity and has been successfully applied on the forecast on tourism demands [6][7]. Cho (2003) used the exponential smoothing, ARIMA and ANN models to conduct forecast to the tourist volume from different countries to Hong Kong. And it was shown in the results that ANN is the optimal forecast model [8]. Even so, this method still has its limitations. For example, Suykens (2001) pointed out that it is hard to use this method to obtain overall optimal solution through empirical study. In addition, during the model training process, this method is relatively more time-consuming and has the problems such as overfitting, etc [9].

Vapnik (1995) introduced the Support Vector Machine, SVM to conduct classification [10]. SVM is a kind of new machine learning algorithm based on statistical learning theory. Later, with the introduction of ε - insensitive loss function, the SVM of regression version which is the SVR has been used for the nonlinear regression forecast problems [11]. Except for the forecast in the fields including finance, electricity, traffic flow, etc, SVR has been successfully applied to the forecast on tourism demands [12][13]. It is shown in those studies that the forecast capacity of SVR model is better than that of traditional nonlinear forecast methods. But, two technical problems will be faced when using this method to forecast: the first one is the problem of free parameter selection of model and the improper parameter setting will cause significant impacts on forecast results [4]. Second, the existing literature mainly uses the hysteretic observation of time series of tourism demands as the input set of model. Such data issued by the official usually has hysteretic nature which limits the application of model to a certain extent and will impact the timeliness of forecast.

Aiming at the selection of free parameters, the scholars mainly use the Genetic Algorithm, GA and Particle Swarm Optimization, PSO to optimize the hyper-parameter of SVR. Chen Rong, et al (2014) used the PSO to adjust parameters of SVR and conduct forecast to the tourist volume of Huangshan Scenic Resort [13]; Gu, et al (2011) applied GA to adjust parameters of SVR and conduct forecast to the house price [14]. However, such parameter optimization algorithms will easily be caught in the risk of local optimum [15] under certain

*Project program: Fund Program of Chongqing Social Science Planning of China (2017YBGL137): Funding project of Sichuan Educational Committee of China (17ZB0375)).

conditions. Yang (2010) introduced the Bat Algorithm, BA [12]. Its main origin of thought comes from the process of bat foraging in natural world. Like the Intelligent optimization algorithms for particle swarm, etc, BA is also the random search mechanism based on groups and the difference is that the bat algorithm possesses stronger randomness and therefore possesses the advantages including fast convergence rate and strong robustness. The application of SVR in tourism forecast is relatively less. This paper introduces the Adaptive Bat Algorithm, ABA [19] on the basis of literature mentioned above so as to optimize the free parameters of SVR model and inspect the effectiveness of such algorithm.

In term of the input set of SVR model, the existing studies mainly use the hysteretic observation of tourism demand series as the input set of model (in short: Classic-SVR model). However, it is required to assume that the economic time series possesses stable economic structure when using this method. Once the economic environment is damaged or there exists occurrence of emergency which may generate impact on the economic system, this will seriously impact the correctness of forecast [2]. With the development of information technology and the popularization of internet, the internet search data reflects the potential consumption demands of consumers and such data has been successfully applied in the forecast of tourism demands. For example, Yang, et al (2015) used search engine data to forecast the tourist volume in the tourist destination of Hainan and compared the forecast capacity of Google Trends and Baidu Index [16]. And other scholars also made similar studies [17-18]. Their study achievements proved that the internet search data possesses the characteristics including timeliness and low acquisition costs, etc and can also effectively improve the forecast capacity of model compared with traditional forecast methods. However, for the input set structure of model, such studies mainly focus on the historic forecast of tourism demands and there have not been any related literature that takes the internet search data as input set of SVR to conduct forecasting study. While there are innumerable links among tourism and aviation, railway and related industries [20]. The traffic flow is the premonition reflection of tourist volume in the scenic area and has significant impacts on the tourist volume in the scenic area. Therefore, the traffic flow index is also taken into the input set of model.

This paper conducts forecast to the tourist volume in Sanya, Hainan by introducing ABA-SVR model on the basis of existing literature. Thereinto, ABA is the hyper-parameter for optimizing SVR model. Firstly, construct the experimental dataset based on internet search data and the data of transportation industry related to tourism and divide the dataset into two parts including training set and test set which are separately used for training model and forecast experiment. Then conduct contrastive analysis and significance testing to the 12-month forecast results. It is shown in the empirical results that the forecast results of ABA-SVR are more precise compared with BA-SVR and Classic-SVR models.

II. STUDY METHOD

A. Principle of Support Vector Regression (SVR) Algorithm

In essential, SVR is to solve the constrained quadratic programming problem and provide overall optimal solutions. The basic principle of SVR algorithm is briefly introduced below.

For given datase: $(x_i, y_i), y_i \in R, x_i \in R^n (i=1,2,...N)$ where x_i indicates input vector; y_i indicates the output value corresponding to x_i. Defined nonlinear mapping: $\varphi : R^n \to F$ through such mapping, the dataset x_i is mapped to the high dimensional feature space F; on such feature space, theoretically there will be a linear function f (SVR function); this function describes the non-linear relationship between x_i and y_i:

$$f(x) = \omega^{\mathrm{T}} \varphi(x) + b \tag{1}$$

Where, $f(x)$ is the predicted value and coefficients ω and b are obtained from minimizing the regular risk function:

$$\min\{\frac{1}{2}\|\omega\|^2 + C\sum_{i=1}^{N}(\zeta + \zeta^*)\} \tag{2}$$

Meet following constraint conditions:

$$\begin{cases} y_i - \omega^{\mathrm{T}}\varphi(x_i) - b \le \varepsilon + \zeta_i \\ -y_i + \omega^{\mathrm{T}}\varphi(x_i) + b \le \varepsilon + \zeta_i^* \\ \zeta_i \ge 0, \zeta_i^* \ge 0, i = 1,2,3,...,N \end{cases} \tag{3}$$

The first item on the right of equal mark in Function (2) indicates Euclidean norm which controls the complexity of model; the second item indicates empirical risk which uses the errors of punishment training for ε insensitive loss function [11][12]; the constant C is used for compromising the complexity and empirical risk of model. In function (3), ζ and ζ^* indicate the positive slack variable which is sued to guarantee the existing of solution and ε is used to measure the insensitive loss function.

Generally, convert the above problems into dual problems and solve a quadratic programming problem; the linear function $f(x)$ can be manifested as:

$$f(x)=\sum_{i=1}^{N}(\beta-\beta^{*})K(x_i,x)+b \tag{4}$$

Where, lagrangian multiplier β and β^{*} are determined by solving the quadratic programming problem; $K(x_i,x)$ indicates the kernel function used by SVR. Any functions that can comply with the Mercer's conditions can be deemed as the kernel functions of SVR [12]. This paper adopts the most common Radial Basis Function (RBF) kernel, because there are three parameters (C, ε and nuclear parameter δ) to be determined in SVR. This paper applies ABA to conduct adjustment to the free parameters of model.

B. Improved Bat Algorithm

The bat algorithm (BA) was initially proposed by Yang (2010)[12]. Because BA possesses the advantages including simple concept, easy realization and simple structure, etc which enables it to be applied in multiple disciplines and engineering fields. Sheng Mengling, He Xingshi and Wang Huimin (2014) improved it and put forward the improved BA[19] which is ABA and its algorithm principles are summarized as follow:

Unlike BA, ABA introduces a kind of crossbar transition mode to solve the limitation on the complex and high-dimensional problems of BA so as to impact the location update of bat:

$$f_i=f_{\min}+(f_{\max}-f_{\min})\beta_i \tag{5}$$

Where, $f_{\min}$ and $f_{\max}$ indicate the minimum and maximum frequency determined by specific problems; $\beta_i\in[0,1]$ is the random vector following Beta distribution, while it is the random vector following uniform distribution in BA algorithm. The Beta distribution coverage rate covers the uniform distribution, normal distribution and various kinds of unsymmetrical distributions which possesses better universality.

To enable the location update of BA to possess better diversity, change its velocity updating function into following form:

$$\upsilon_i^{t}=\varepsilon\upsilon_i^{t-1}+(x_i^{t}-x^{*})f_i \tag{6}$$

Where, ε is the newly added learning factor and can follow the standard normal distribution; x_i^{t} and υ_i^{t} separately indicate the location and velocity of t generation of item i bat in the bat groups under the d -dimension space; x^{*} indicates the current overall optimal solution.

With the continuous development of search, ABA possesses more sufficient local search so as to achieve higher precision compared with BA.

C. ABA-SVR Forecast Procedure

According to the SVR and ABA algorithm principles, the forecast procedures of ABA-SVR model will be completed through following steps:

Step 1: Parameter setting. Before optimizing the SVR free parameters, the related ABA parameters must be set in advance. This paper uses Taguchi method [21] (Ghani, 2004) to set such parameters which is easy and practicable.

Step 2: Initialization of bat species. To obtain the optimal solution to SVR, the bat species randomly distributed in the search space must be initialized. There are three parameters in this paper to be determined. In essential, it is to solve a 3D optimization problem.

Step 3: Generation of new solution. On the search space, generate new solutions according to location of bat and velocity update formula. This process shall be performed by iterations until obtaining the optimal parameter set.

Step 4: Adaptability assessment. ABA must adopt corresponding measures to obtain more optimal parameter set. This paper adopts Mean Squared Error, MSE as the adaptability function and its values shall be obtained through 5-convertion cross validation technique. This method can effectively avoid the overfitting problem [22] (Hsu, Chang and Lin, 2003). the MSE on the aspect of cross validation is manifested as follow:

$$MSE_{C-V}=(1/N)\cdot\sum(y_t-\hat{y})^2 \tag{7}$$

Where, $\hat{y}$ is the predic N ted value and N is the training sample capacity.

Step 5: Stopping rules. Cycle the process of step 3 and step 4. When the performed algebra achieves the set value, the algorithm will stop. At this moment, find the optimal parameter set on search space and then substitute it into the model to conduct forecast, otherwise convert into step 3. The detailed flow diagram of above process is as shown in "Fig. 1".

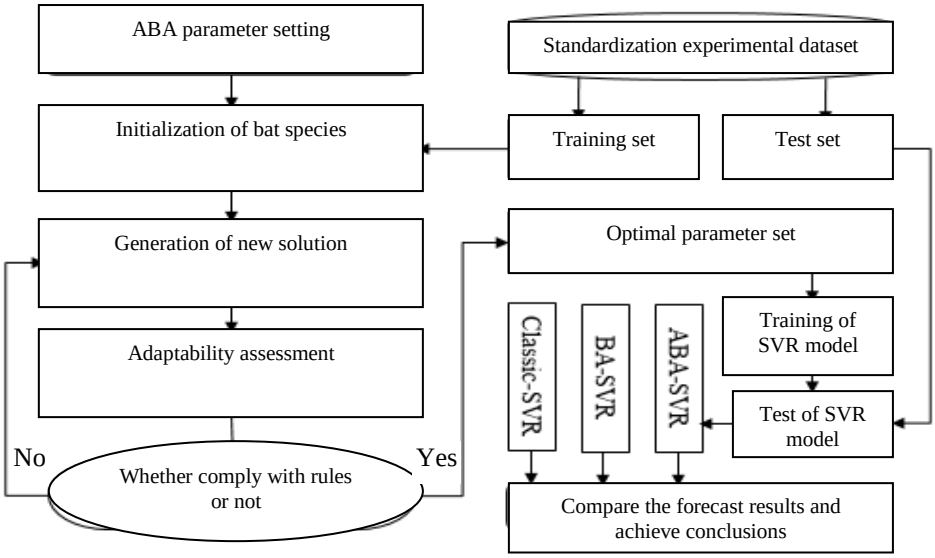

Fig. 1. ABA-SVR model forecast flow diagram.

III. Application of ABA-SVR Model

A. Construction of Experimental Dataset and Input Set

To test the effectiveness of introduced method, this paper takes the tourist destination of Sanya, Hainan as example to conduct forecast to domestic tourist volume (ten thousand person times) received. Sanya is located in southernmost point of Hainan Island and is the central city and tourism city in the south of Hainan. In 2015, Sanya received 14.957 million person times of tourists staying overnight and the total tourist income reached 30.23 billion yuan which separately increased by 10.6% and 12.1% compared with that in last year. The model input dataset includes the data of tourism-related transportation industry and the internet search data. Considering the data loss problem, the time span of data collection shall be from August 2008 to March 2016.

The domestic tourist volume data of Sanya and the data of tourism-related traffic flow of Sanya all come from Wind, the leading financial database in Chinese Mainland. This paper preliminarily selects four variables in the data of tourism-related transportation industry data of Sanya, including aviation, hotel and railway, etc. The internet search data comes from Baidu Index which is the data sharing platform based on the data of related behaviors of mass netizen. It is shown in the existing studies that such kinds of search data can effectively promote the model forecast accuracy[17,18]. Use the four-stage key word extraction method adopted by Yang (2015) et al to select 11 key words. During this process, we conduct mean weighted summation treatment to such key words so as to obtain the monthly search data because the internet search data is the weekly data.

Further conduct cross correlation analysis to the 15 forecast variables mentioned above. Respectively calculate the Pearson correlation coefficient between 0-12 phase lagged variables and forecasted variables of each forecast variables. Take 0.85 as threshold value and select the corresponding time series with maximum correlation coefficient. Finally, obtain the 6 forecast variables with largest forecast capacity based on stepwise regression thought where 4 are the key word variables; the traffic flow data is “Sanya airplane sorties” and the 12-phase lags of forecasted variables are also included in the forecast variables.

See “Table I” for its correlation analysis, where the time interval is the corresponding time quantum after time series has been aligned based on lag order. It is shown in “Table I” that there exists significant correlativity between selected 6 forecast variables and the forecasted variables on the 1% level. Manifest the experiment dataset as: $\{y_{-12}, x_1, x_2, x_3, x_4, x_5; y\}$, where $\{y_{-12}, x_1, x_2, x_3, x_5\}$ is the constructed SVR input set, y indicates output variable. There are total 80 data points in the dataset after aligning based on lag order. To promote the model forecast accuracy and remove the impact of data absolute amount, this paper applies the formula: $(x_t - x_{\min})/(x_{\max} - x_{\min})$ to standardize the data on the section of [0, 1], where x_t indicates the value when experiment data concentrates each time series on time t, and $x_{\max}, x_{\min}$ respectively indicate the maximum value and minimum value of each time series on their respective time period. Finally, divide the experiment dataset into training set and test set; use the first 68 data on training model and use the last 12 data points on forecast test.

TABLE I. CORRELATION ANALYSIS ON SELECTED FORECAST VARIABLES AND FORECASTED VARIABLES

Forecasted Variable y : Domestic Tourist Number of Sanya (August 2009- March 2016)						
Forecast variable	***Code***	***Lag order***	***Correlation coefficient***	***Adjusted time interval***	***t statistic***	***P value***
Lagged observation of y	y-12	12	0.906***	2008.08-2015.03	30.07	<2.2E -16
Airplane sorties of Sanya	X1	12	0.904***	2008.08-2015.03	28.99	<2.2E -16
Sanya travel strategy	X2	11	0.876***	2008.09-2015.04	15.46	<2.2E -16
Sanya weather	X3	12	0.873***	2008.08-2015.03	15.34	<2.2E -16
Hainan map	X4	11	0.857***	2008.09-2015.04	14.54	<2.2E -16
China Comfort Travel	X5	6	0.854***	2009.02-2015.09	14.46	<2.2E -16

[a.] Note: Those with ***, ** and * respectively indicate the significance on the levels including 1%, 5% and 10%, same below.

B. Forecast Experiment and Discussion

According to ABA-SVR forecast procedures, firstly obtain the optimal hyper-parameter of SVR on the training set to construct the forecast model; then conduct forecast experiment on the test set. To test the effectiveness of introduced model, this paper compares the forecast results of ABA-SVR and reference model BA-SVR and the Classic-SVR. Here, ABA is still applied to conduct optimization to the parameters of Classic-SVR model. Finally, use the Mann-Whitney U method to conduct significance testing to the forecast accuracy ($error_i = |(f_i - y_i)/y_i| \times 100\%$) of three models, where f_i and y_i respectively indicate the predicted value and actual value on test set. Thereinto, for ABA-SVR, see "Fig. 2" for the evolution curve of adaptability function value on training set. This figure shows the evolution process of optimal adaptability function value and mean adaptability function value on the aspect of cross validation. When the evolution code is 300, the adaptability function value has clearly been restrained which means that ABA possesses excellent parameter optimization performance. The optimal parameters of ABA-SVR are: C=22.471, ε=1.662, δ=0. 108. Similarly, BA-SVR (C=6.641, ε=0.944, δ=0. 019), Classic-SVR (C=47.441, ε=14.734, δ=1. 309).

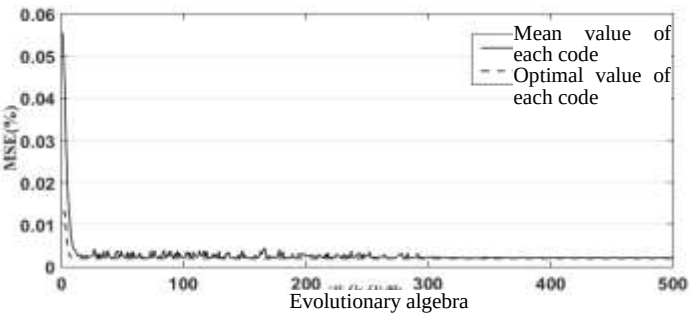

Fig. 2. Evolution curve of adaptability function.

See "Table II" for the forecast results of three models. It is shown in this table that in the 12-month forecast and the forecast value with minimum deviation with the actual value (in black body): ABA-SVR accounting for 7; BA-SVR accounting for 3 and Classic-SVR only accounting for 1. This shows that ABA-SVR possesses better forecast accuracy compared with the other two reference models. However, in October 2015, the forecast accuracy of three models is relatively low which may be because the fitting capacity on peak tourist season of the model is relatively weak.

This paper uses the Mean Absolute Percentage Error, MAPE and the two common statistical indexes of correlation coefficient R between predicted value and actual value to conduct further test to the above conclusions "Table II". Thereinto, MAPE can measure the deviation between predicated value and actual value. The less this value is, the more precise the forecast will be; R can measure the correlation between predicted value and actual value. If this value is closer to 1, it will mean that the correlation is stronger and the model fitting capacity is better. It is shown in the results that the MAPE value of ABA-SVR is less than that of the other models but R value will be the maximum which is close to 1; at the same time, it is shown in P value in the Mann-Whitney U significance testing that the forecast accuracy (MAPE) of ABA-SVR is significantly higher than that of the other models, which further proves the excellent forecast capacity of ABA-SVR and that the introduction of internet search data can effectively promote the forecast accuracy of models and this can comply with results of Yang, et al (2015)[16]. Fig.3 more intuitively shows the forecast results of the model and the three forecast curves are all very close to actual curve, but ABA-SVR can be closer to the expected curve and can better be similar to the experiment data, while deviation between forecast curve and actual curve of Classic-SVR is the largest. The empirical analysis shows that ABA can effectively avoid the situations including local optimum and precocious convergence.

TABLE II. FORECAST RESULTS AND FORECAST MEASUREMENT OF ABA AND REFERENCE MODEL

Month	Actual value	ABA-SVR	BA-SVR	Classic-SVR
2015.04	105.50	105.27	105.78	106.10
2015.05	98.61	94.89	95.93	96.12
2015.06	85.77	87.09	88.12	89.27
2015.07	99.18	99.85	101.61	105.52
2015.08	116.25	111.87	114.45	113.50
2015.09	91.90	93.06	94.93	96.08
2015.10	113.18	120.95	118.39	123.14
2015.11	151.07	152.05	156.99	147.99
2015.12	187.52	186.37	183.87	181.22
2016.01	144.19	145.82	149.62	147.87
2016.02	159.99	163.36	160.54	154.57
2016.03	140.37	146.23	147.67	143.34
MAPE(%)		2.2311	2.7331**	3.5613***
R value		0.9942	0.9936	0.9903
p value			0.0266	0.0094

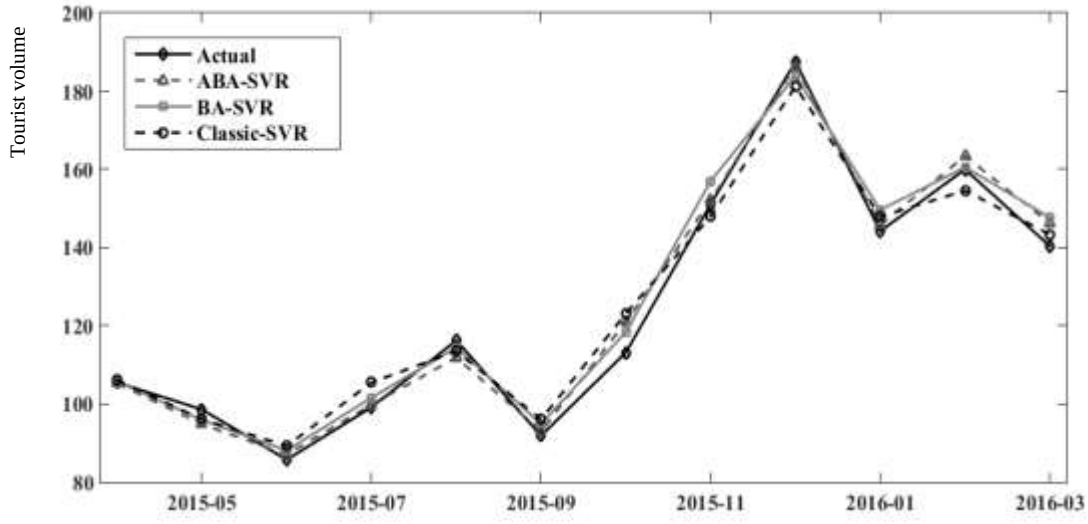

Fig. 3. Curve comparison of different SVR model tourist volume forecast.

IV. CONCLUSION

The correctness of tourism demand forecast will have significant impacts on the reasonable distribution of tourism resources. The traditional linear forecast model cannot fully describe the nonlinear characteristics of tourism demand time series. SVR model possesses good nonlinear forecast capacity but there are no universal rules to optimize its parameters. And the improper parameter setting will make it easy be caught in the limitations including local optimum and overfitting, etc. ABA can effectively avoid above problems. This paper introduces the ABA-SVR model to forecast the domestic tourist volume of Sanya, Hainan. ABA is used to optimize the model parameters. The 12-month forecast results show that ABA-SVR can effectively promote the forecast accuracy. And the excellent forecast capacity of introduced model can be ascribed to following reasons:

Firstly, the adaptivity of f_i, v_i^t enables ABA to effectively avoid the defects in traditional parameter optimization methods (such as BA) and can quickly obtain the overall optimal parameter set when facing complex high-dimensional problems. Secondly, ABA-SVR is better than Classic-SVR which means that the traditional forecast method only uses the lagged observation of tourism demand time series as input set which limits the full usage of other useful information. With the development of information techniques and popularity of the Internet, the internet search data reflects the potential demands of tourists. Such data can be easily obtained, is sensitive to the consumer behaviors and has been successfully applied in the forecast of various kinds of social and economic activities. This paper introduces the data of tourism-related traffic flow and the internet search data to further expand the information contents of input dataset which can enable the model to be much more fully close to the experiment dataset. And it is shown in the various kinds of statistical indexes and significance testing that the addition of such data significantly promotes the forecast capacity of model. Finally, before conducting forecast experiment, we have conducted the standardization treatment and other methods to the data which is favorable to promoting the accuracy to a certain extent.

The study in this paper possesses certain theoretical meanings and application values. In term of theory, this study introduces new data under internet environment and uses correlation analysis to construct the input dataset of SVR model and uses the parameters of ABA optimization model to further prove the superiority of forecast model. On the aspect of application, the construction method of input dataset of model can be applied in other related forecast fields. For example, introduce the forecast variables with forecast capacity to construct the dataset so as to conduct tourism demand forecast including hotel lodging ratio and tourism revenue, etc. The forecast results are at least 6 months before that of data released by the official which can provide necessary reference for the policy-making of related departments. In the future studies, we will encounter more complex economic situations. How to find more common parameter optimization algorithm, construct the input dataset with better forecast capacity and promote the forecast capacity at peak and trough of wave of model is the orientation of further efforts.

REFERENCES

[1] Athiyaman A, Robertson R W. Time series forecasting techniques: short-term planning in tourism [J]. International Journal of Contemporary Hospitality Ma.onagement, 1992(4).

[2] Song H, Li G. Tourism demand modeling and forecasting-A review of recent research [J].Tourism Management,2008,29 (2).

[3] Archer B H. Demand forecasting and estimation, in: J. R. B. Ritchie, C. R. Goeldner (Eds.), Travel. Tourism and Hospitality Research, John Wiley, New York, 1987.

[4] Andre J, Siarry P, Dognon T. An improvement of the standard genetic algorithm fighting premature convergence in continuous optimization [J]. Adv. Eng. Softw, 2001, 32 (1).

[5] Wang J Z, Zhu S L, Zhang W Y, et al. Combined modeling for electric load forecasting with adaptive particle swarm optimization [J]. Energy, 2010, 35(4).

[6] Law R. Back-propagation learning in improving the accuracy of neural network-based tourism demand forecasting [J]. Tour. Manag., 2000,21(4).

[7] Pai P F, Hong W C. An improved neural network model in forecasting arrivals [J]. Ann.Tour. Res., 2005,2(4).

[8] Cho V. A comparison of three different approaches to tourist arrival forecasting [J]. Tour. Manag.,2003,4 (3) .

[9] Suykens J A K. Nonlinear modelling and support vector machines. In: Proceedings of IEEE Instrumentation and Measurement Technology Conference, 2001.

[10] Vapnik V. The nature of statistical learning theory [M]. New York: Springer, 1995.

[11] Drucker H, Burges C J C, Kaufman L, et al. Support vector regression machines [J].Advances in Neural Information Processing Systems, 1996,28(7).

[12] Yang Xinshe. A new metaheuristic bat-inspired algorithm[M] / / Nature inspired cooperative strategies for optimization.Berlin:Springer,2010.

[13] Chen Rong, Liang Changyong, Lu Wenxing et al. Study on Tourist Volume Forecast Model Based on Seasonal SVR-PSO[J]. System Engineering Theory and Practice, 2014, 34(5).

[14] Gu J R, Zhu M C, Jiang L GY. Housing price forecasting based on genetic algorithm and support vector machine, Expert Syst. Appl, 2011, 38 (4).

[15] Duan K, Keerthi S, Poo A. Evaluation of simple performance measures for tuning SVM hyperparameters [R]. Singapore:National University of Singapore, Department of Mechanical Engineering, 2011.

[16] Yang X, Pan B, James A, et al. Forecasting chinese tourist volume with search engine data [J].Tourism Management, 2015, (46).

[17] Huang Xiankai, Zhang Lifeng and Ding Yusi. Relations between Baidu Index and Tourist Volume in Tourist Attractions and Forecasting Study[J]. Tourism Tribune, 2013, 28(11).

[18] Ren Le and Cui Dongjia. Forecast Study on Domestic Tourist Volume based on Internet Search Data[J]. Inquiry into Economic Issues, 2014, (4).

[19] Sheng Menglong, He Xingshi and Wang Huimin. The Improved Adaptive Mutation Bat Algorithm[J]. Computer Technology and Development, 2014, 24 (10): 131-135.

[20] Cho V. Tourism forecasting and its relationship with leading economic indicators [J]. Journal of Hospitality and Tourism Research, 2001, 25(4).

[21] Ghani J, Choudhury I, Hassan H. Application of Taguchi method in the optimization of end milling parameters [J]. Journal of Materials Processing Technology, 2004，145(1), 84–92.

[22] Hsu C. W, Chang C C, Lin C J. A practical guide to support vector classification. Dept. Comput. Sci., National Taiwan University, Tech. Rep. 2003.