sciendo

POLISH MARITIME RESEARCH Special Issue 2018 S3 (99) 2018 Vol. 25; pp. 29-35
10.2478/pomr-2018-0109

APPLICATION OF EVALUATION ALGORITHM FOR PORT LOGISTICS PARK BASED ON PCA-SVM MODEL

Bianjiang Hu
School Of Business,Central South University,Changsha,Hunan,China

ABSTRACT

To predict the logistics needs of the port, an evaluation algorithm for the port logistics park based on the PCA-SVM model was proposed. First, a quantitative indicator set for port logistics demand analysis was established. Then, based on the grey correlation analysis method, the specific indicator set of port logistics demand analysis was selected. The advantages of both principal component analysis and support vector machine algorithms were combined. The PCA-SVM model was constructed as a predictive model of the port logistics demand scale. The empirical analysis was conducted. Finally, from the perspective of the structure, demand, flow pattern and scale of port logistics demand, the future logistics demand of Shenzhen port was analysed. Through sensitivity analysis, the main influencing factors were found out, and the future development proposals of Shenzhen port were put forward. The results showed that the port throughput of Shenzhen City in 2016 was 21,328,200 tons. Compared with the previous year, it decreased by about 1.74 %. In summary, the PCA-SVM model accurately predicts the logistics needs of the port.

Keywords: port logistics demand, support vector machine, principal component analysis, economic hinterland

INTRODUCTION

With the integration of the world economy, logistics has been widely valued by various countries and has achieved rapid development. At present, the logistics industry has become an important pillar industry and a new economic growth point in China. It is a strong support for the tertiary industry. As an important branch of logistics, urban logistics guarantees the continuous operation of various functions of the city. This is a new impetus to promote the healthy and rapid development of the regional economy. With the continuous deepening of economic transformation and the optimization and upgrading of industrial structure, the specialization of logistics and the process of socialization have obviously accelerated. The level of urban logistics services has been further improved. With the supply-side structural reform entering the tough stage, logistics companies must upgrade their own strength to adapt to the development of the times. As of the end of 2015, the total cost of social logistics was approximately 11.1 trillion. It grew by 2.9 % compared to the previous year. Among them, the transportation cost is 6 trillion. Compared with the previous year, it increased by 3.3 %. The cost of storage is 3.7 trillion. Compared with the previous year, it increased by 1.3 %. The administrative fee is 1.4 trillion. Compared with the previous year, it increased by 5.6 %. This shows that the steady growth of the logistics industry is the fundamental guarantee for the sound development of the tertiary industry and social economy. In August 2015, policies such as “Multiple measures to promote the accelerated development of modern logistics” and “Notice on Accelerating the Implementation of Major Projects in Modern Logistics” were issued. The logistics industry has been promoted to an unprecedented height. A sound modern logistics service system can promote economic transformation and upgrading of industrial structure.

Based on the research status of port logistics demand at home and abroad, the internal relationship between port logistics demand and port economic hinterland is analysed. According to the corresponding construction principles, a set of quantitative analysis indicators for port logistics demand was established. Domestic and international common port logistics demand analysis methods are compared. Combined with the characteristics of port logistics demand, the advantages of support vector machine and principal component analysis are combined and the PCA-SVM model is established. It provides a new method and idea for port logistics demand analysis.

STATE OF THE ART

Lin et al. analysed the freight negotiations in France and assessed the effectiveness of the urban freight policy. The combination of qualitative and quantitative methods was used to analyse the negotiation of the institutional framework of the French region at three levels. Specific consultations on freight cost issues need to be implemented, and important consultation results are translated into effective policies and behaviours [1]. Taking Poland as an example, Fresno et al. analysed the role of government in the development of urban logistics. A collaborative process model for urban logistics policy makers was constructed. Studies have shown that government departments not only lack a comprehensive and effective plan to develop urban logistics, but also lack cooperation with urban logistics practitioners [2]. Hope et al. analysed the relationship between the construction of transportation infrastructure and employment growth in the region. The regression model was used to analyse the employment data from 1992 to 2008. The study showed that the transportation infrastructure construction increased the employment rate [3]. Taking the export of the Spanish region as an example, Wu et al. studied the logistics infrastructure and trade import and export volume. Data from 19 Spanish regions to 64 destinations for the period 2003–2007 were analysed using a bilateral export model. The survey results show that the level and quality of logistics infrastructure have a positive impact on trade export flows [4]. Qi et al. studied the relationship between logistics development and economic development for a long time and established an econometric model. WGDP represents economic growth. Transportation turnover and inventory represent logistics development. Tests prove that economic growth is positively related to logistics development [5]. Li et al. studied the relationship between traffic flow and economic growth in Indonesia from 1988 to 2010 to analyse the relationship between logistics and economic development. Shipping, air and rail freight volumes are used as logistics indices, and GDP is used as an economic index to build a series of models. Experiments show that logistics plays a supporting role in maintaining economic growth. Logistics infrastructure construction is one of the means to ensure sustainable economic development [6]. To reduce the inventory cost brought by the bullwhip effect, Feng et al. used the DWT-ANN model to predict regional logistics demand. In contrast to the ARIM method, data from three different manufacturing companies were collected for analysis. Experiments show that the DWT-ANN model has higher precision [7]. Foraker et al. pointed out that the evaluation method of the effectiveness of logistics measures is crucial. The O-D matrix is established from three aspects of transportation service type, delivery time and vehicle type. The collected data was used for simulation and satisfactory results were obtained [8]. A scholar used the Delphi method to predict the gross domestic product of the road freight traffic from Finland and concluded that different economic developments would produce different transportation needs [9].

In general, developed countries such as Europe and the United States attach great importance to the forecasting of logistics demand, and the research time is much earlier than China. Foreign studies have achieved theoretical success. Relevant laws and regulations guarantee the healthy and sustainable development of the logistics industry [10].

PORT LOGISTICS DEMAND FORECASTING MODEL BASED ON PCA-SVM

The logic of the support vector machine learning process is expressed in mathematical formulas as follows:

If there is an unknown joint probability $F(x, y)$ between the input variable x and the output variable y, according to several independent and identically distributed samples (x_1, y_1), (x_2, y_2), (x_3, y_3),…(x_n, y_n), an optimal $f(x', y')$ is found in a set of functions to minimize the expected risk [11]:

$$R(X) = \int L(y, f(x', y'))dF(x, y) f(x', y') F(x, y) \tag{1}$$

In the formula (1), $L(y, f(x', y'))$ represents the loss function caused by the function performing regression prediction. In general, when the function is fitted, the calculation error of the least square difference is used to represent the loss function. Minimizing risk is the main purpose of machine learning. However, in many practical problems, the information available is very limited. It is difficult to calculate the expected risk. The principle of empirical risk minimization is as follows [12]:

$$R_{emp}(W) = \frac{1}{n}\sum_{i=1}^{n} L(y, f(x', y')) \tag{2}$$

The principle of empirical risk minimization has been used to replace the principle of expected risk minimization, which has achieved good application in many fields [13]. It provides a new way of thinking and perspective for problem solving and has gradually become a common guideline.

The operation steps of the PCA-SVM prediction model are shown in Figure 1:

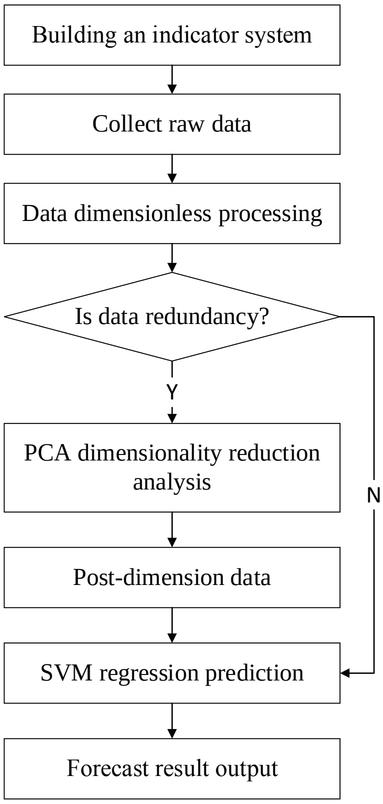

Fig. 1. Operation steps of the PCA-SVM model

In the PCA-SVM model, there is a problem of parameter optimization. The appropriateness of the parameter selection directly determines the prediction accuracy of the model. There are two parameters in the model that need to be optimized, namely the penalty parameter C and the kernel function selection parameter g. At present, the academic community generally uses the grid search method and the particle swarm algorithm to optimize the parameters.

EMPIRICAL ANALYSIS OF PORT LOGISTICS DEMAND

STRUCTURAL ANALYSIS OF PORT LOGISTICS DEMAND

Shenzhen port is the second largest port in China and the third largest container port in the world. It is a key location for China's foreign trade. With the transformation of economic growth mode and the continuous development of Shenzhen's economy, port logistics has become increasingly important in the logistics system of Shenzhen. In 2015, the container throughput of Shenzhen port increased by 0.7 % compared with the previous year, with a total of 24,240,600 TEU [14]. Among them, the export container throughput decreased by 0.9 %. The total number of standard boxes is 12.60 million. Compared to last year, it fell by 0.9 %. The cargo throughput of the port is 21,706,380 tons. Compared to last year, it fell by 2.8 %.

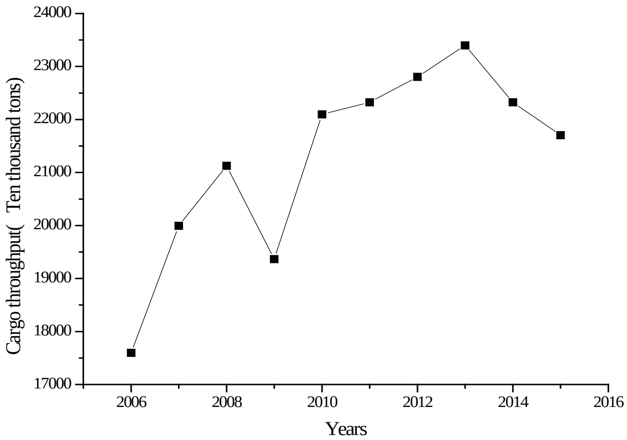

Fig. 2. Curve of port cargo throughput in Shenzhen from 2006 to 2015

In recent years, the total volume of imports and exports in Shenzhen has shown a downward trend. Since 2000, the total annual import volume of Shenzhen has been greater than the total export value. However, the growth rate of the import and export gap has continued to slow down. This shows that Shenzhen is in the transitional stage of "bringing in" and "going out". At present, Shenzhen's foreign trade is still dominated by imports. Port logistics demand will also be based on import logistics demand.

Hong Kong, the United States, and Taiwan account for the largest share of Shenzhen's imports, followed by South Korea and Japan. The main trading partners are concentrated in the Pacific Rim. With the continuous development of the global economy, the cooperation between Shenzhen and the major trading partners in the Pacific Rim continues to strengthen, which will inevitably strengthen the logistics needs of Shenzhen's foreign trade. Moreover, Shenzhen, as the largest port in South China, is the logistics point of import and export trade. With the continuous improvement of the transportation system, the connection between Shenzhen and the economic hinterland has been continuously strengthened. The development of the hinterland will inevitably provide a huge source of transit for Shenzhen. Such a large potential market will provide opportunities for the development of modern logistics in Shenzhen. It has obvious rules of logistics flow.

FORECAST OF LOGISTICS DEMAND OF BASIC PCA-GA-SVM

First, the gray correlation method is adopted. The following eight indicators were selected as input indicators for the forecast of Shenzhen port logistics demand. They are the gross domestic product of the hinterland, the gross agricultural production in the hinterland, the gross industrial production in the hinterland, the freight turnover in the hinterland, the transportation in Shenzhen, the fixed investment in warehousing and postal services, the gross domestic product of Shenzhen, total import and export volume and social freight volume in Shenzhen. Through the relevant literature research and the actual situation, the port throughput is selected as the output index to carry out the target prediction. As can be seen from Table 1, the correlation of the sample data is generally above 0.8, which means that the sample data is highly correlated. It can be seen from Table 2 that the significance is 0, which is less than 0.05, indicating that it is less than the significance level. Therefore, the null hypothesis was rejected. It is significantly different from the identity matrix. The KMO index is 0.761. Therefore, the data can be analysed by principal component analysis.

Principal component analysis was performed using SPSS 22.0 software. The cumulative contribution rate is selected to be greater than 99 % as the main component. Since the cumulative contribution rate of the first, second and third principal components reached 99.369 %, the information of the original eight indicators was well retained. Three integrated variables were used to replace the eight variables in the original data. The coefficient matrix of each principal component can be calculated according to the factor score coefficient.

The 2-RBF function is selected. At this time, the penalty parameter C and the kernel function parameter g are determined. The choice of parameters C and g determines the performance of the SVR. Appropriate parameters allow the model to achieve more accurate predictions. After multiple experiments, grid search was used for parameter optimization. K-fold cross-validation was used to evaluate model performance to avoid overfitting.

There are two steps to improving the grid optimization method. First, the parameters are roughly selected, and the approximate range of the optimal parameters is selected using 0.5 steps to narrow the search range. Second, within the range of the optimal parameters determined by the rough selection of the parameters, the search is performed using a smaller step size. In this paper, 0.05 steps are used to find the final parameters. If multiple sets of parameters appear in the search results as the optimal result, the group with the smallest penalty parameter C is selected, because this can achieve a stronger generalization ability of the model. If the penalty parameter C is the same, the parameter that was first searched is taken as the optimal parameter.

The logistics demand structure of Shenzhen port and the main body of demand were analysed. With the continuous upgrading of the industrial structure, the proportion of high value-added and high-tech goods will increase in the logistics demand of Shenzhen ports. The logistics demand structure of Shenzhen port will be changed. In the case of slower port growth, the development of Shenzhen port has shifted from port throughput growth and traditional port industry layout to industrial chain and value chain. The port logistics service system is the main efficient industrial development model and establishes a new port and urban development relationship.

ANALYSIS OF FORECAST RESULTS AND FUTURE PREDICTIONS

The PCA-GA-SVM model is compared with four models to evaluate the prediction accuracy of the model. The PCA-PSO-SVM model is compared to the PCA-SVM model. The characterization of the training set is shown in Figures 3 and 4:

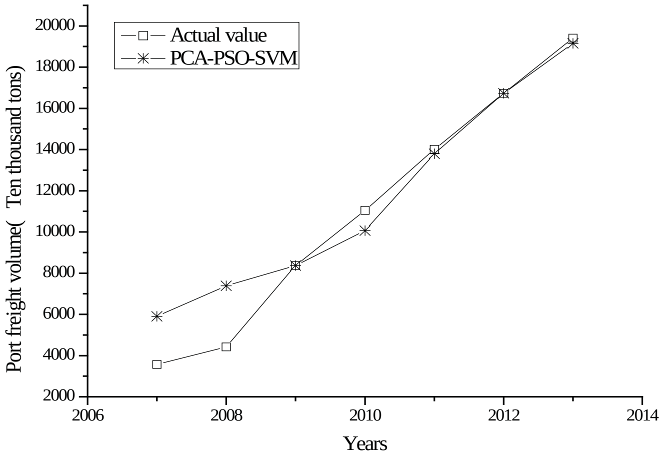

Fig. 3. The fitting of PCA-PSO-SVM model to the training set

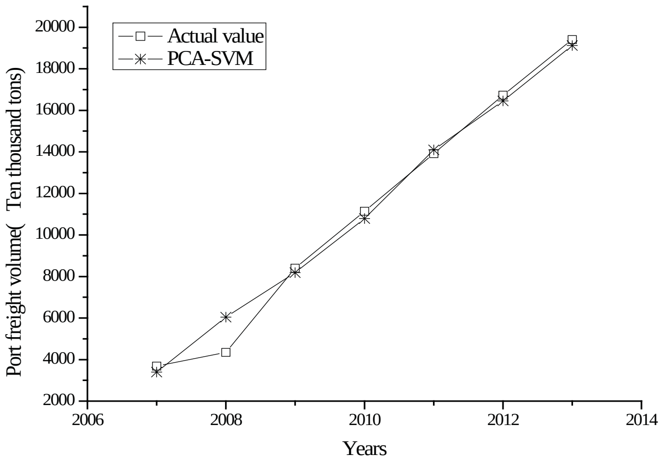

Fig. 4. The fitting of PCA-SVM model to the training set

The linear regression and BP neural network are used as the comparison model. The fitting of the training set is shown in Figures 5 and 6. The prediction results of the prediction set are shown in Table 2.

Tab. 1. Comparison of individual models

Years	Actual value	Predictive value					
		PCA-PSO-SVM	Error rate	PCA-SVM	Error rate	PCA-GA-SVM	Error rate
2014	24222	24261	0.1338	24231	0.000373	24311	0.00368
2015	24520	24517	0.1273	24517	0.000123	24520	0

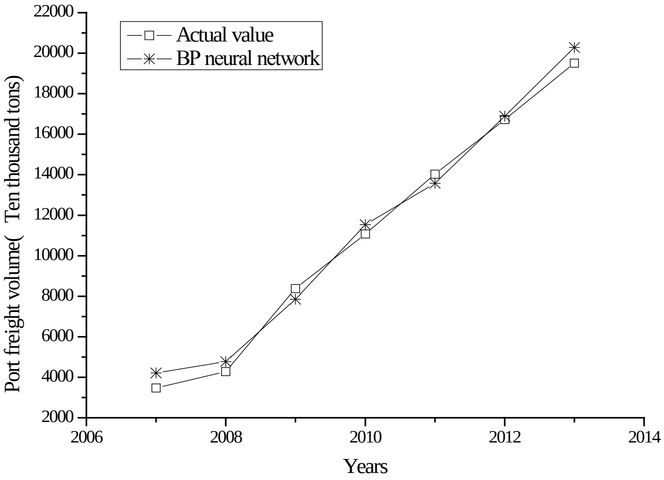

Fig. 5. The fitting of BP-Neural network model to the training set

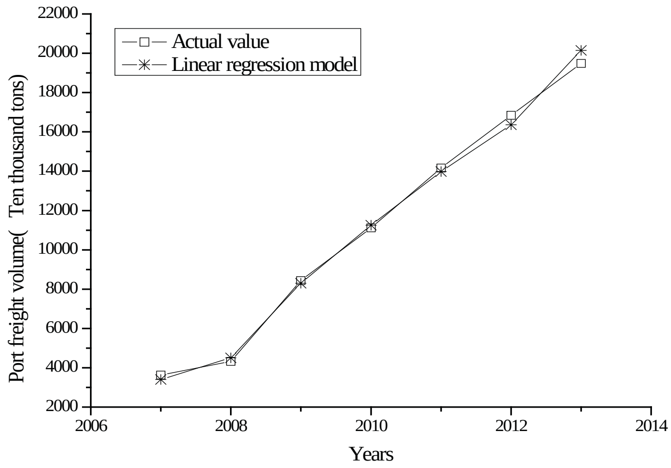

Fig. 6. The fitting of linear regression model to training set

Tab. 2. Comparison of individual models

Years	Actual value	Predictive value					
		Regression sequence	Error rate	BP model	Error rate	PCA-GA-SVM	Error rate
2014	24222	23996.94	0.00925	22791.85	0.06558	24311	0.00368
2015	24520	24520	$0.85*10^{-5}$	24272.98	0.01098	24520	0

It can be seen from Table 2 that the error rate of the PCA-GA-SVM prediction model is the smallest, which is mainly controlled within 0.5 %. Its prediction accuracy is good. According to the results of the BP neural network model, the error rate is about 5 %. The degree of prediction is worse in the prediction model. The prediction accuracy of the PCA-GA-SVM model is good, which shows that this method has certain promotion value.

The PCA-GA-SVM model is used to forecast the port cargo volume of Shenzhen in the next three years. When making predictions, the value of the independent variables of the logistics demand in the next three years should be obtained first. By using the GM model, the influencing factors of the port freight volume demand for 2019–2011 are predicted. The PCA-GA-SVM model is used to predict the freight volume of Shenzhen port. The logistics demand in Shenzhen in 2019–2011 is:

Tab. 3. Logistics demand in Shenzhen from 2019 to 2011

2019	2010	2011
34857	41648	45149

As can be seen from Table 3, the port cargo volume of Shenzhen in 2019–2011 is increasing year by year. To meet the growing demand for logistics, logistics supply capacity must be accelerated, and the transportation system improved. Logistics resources are rationally allocated, and the logistics economy is coordinated.

CONCLUSION

Based on the basic theory of port logistics demand, the relationship between port logistics demand and economic hinterland is analysed. Combined with relevant literature research and data review, the quantitative analysis index system of port logistics demand was established. Based on the comparative study of port logistics demand analysis methods, a combination of qualitative and quantitative analysis methods was established. A predictive analysis model combining support vector machine and principal component analysis is proposed. The intrinsic drive between the development of port logistics and the hinterland economy was analysed. The factors of port logistics demand were analysed. According to the principle of effectiveness, relative independence, operability and strong correlation, the index set of port logistics demand analysis is established. Through the grey correlation analysis method, a specific index set suitable for Shenzhen port logistics demand analysis is selected. Analytical methods at home and abroad are studied and compared. Based on the advantages of principal component analysis and support vector machine (SVM), the PCA-SVM model is constructed as the forecast model of the port logistics demand scale, and an empirical analysis is carried out.

Changes in port logistics demand are closely related to the development of the economic hinterland. When conducting research and analysis, in addition to relying on historical data on port logistics demand, the link between the economic hinterland and the port should be combined to establish a scientific and rational quantitative indicator set. According to the characteristics of port logistics demand, the forecasting method is selected.

REFERENCES

1. Y. K. Lin, C. T. Yeh, and C. F. Huang, *A simple algorithm to evaluate supply-chain reliability for brittle commodity logistics under production and delivery constraints*, Annals of Operations Research, Vol. 244, No. 1, pp. 67–83, 2016.

2. C. Fresno, G. A. González, G. A. Merino, A. G. Flesia, O. L. Podhajcer, and A. S. Llera, *A novel non-parametric method for uncertainty evaluation of correlation-based molecular signatures: its application on pam50 algorithm*, Bioinformatics, Vol. 33, No. 5, pp. 693, 2017.

3. J. L. Hope, A. E. Sinha, B. J. Prazen, and R. E. Synovec, *Evaluation of the dotmap algorithm for locating analytes of interest based on mass spectral similarity in data collected using comprehensive two-dimensional gas chromatography coupled with time-of-flight mass spectrometr*y, Journal of Chromatography A, Vol. 1086, No. 1, pp. 185–192, 2016.

4. J. Wu, J. Chu, Q. Zhu, P. Yin, and L. Liang, *Dea cross-efficiency evaluation based on satisfaction degree: an application to technology selection*, International Journal of Production Research, Vol. 54, No. 20, pp. 1–18, 2016.

5. X. Qi, G. Wu, K. Boriboonsomsin, and M. J. Barth, *Development and evaluation of an evolutionary algorithm-based online energy management system for plug-in hybrid electric vehicles*, IEEE Transactions on Intelligent Transportation Systems, Vol. 99, pp. 1–11, 2016.

6. M. W. Li, J. Geng, W. C. Hong, and Z. Y. Chen, *A novel approach based on the Gauss-vSVR with a new hybrid evolutionary algorithm and input vector decision method for port throughput forecasting* Neural Computing and Applications, Vol. 28, No. 1, pp. 621–640, 2017.

7. Y. C. Yang, and S. L. Chen, *Determinants of global logistics hub ports: Comparison of the port development policies of Taiwan, Korea, and Japan*, Transport Policy, Vol. 45, pp. 179–189, 2016.

8. J. Foraker, S. Lee, and E. Polak, *Validation of a strategy for harbor defense based on the use of a min-max algorithm receding horizon control law*, Naval Research Logistics, Vol. 63, No. 3, pp. 247–259, 2016.

9. J. Piao, S. Xu, Z. Wu, Y. Li, B. Qu, and X. Duan, *Su-f-t-619: dose evaluation of specific patient plans based on monte carlo algorithm for a cyberknife stereotactic radiosurgery system*, Medical Physics, Vol. 43, No. 6, pp. 3606–3606, 2016.

10. S. Y. Kim, H. Park, H. M. Koo, and D. K. Ryoo, *The Effects of the Port Logistics Industry on Port Citys Economy*, Journal of Navigation and Port Research, Vol. 39, No. 3, pp. 267–275, 2015.

11. J. Feng Jie, X. J. Liu Xiaojun, *Design of Upright Intelligent Vehicle Based on Camera*, Acta Electronica Malaysia, Vol. 1, No. 1, pp. 09–11, 2017.

12. M. S. Ibrahim, S. Kasim, R. Hassan, H. Mahdin, A. A. Ramli, M. F. Md Fudzee, and M. A. Salamat, *Information Technology Club Management System,* Acta Electronica Malaysia, Vol. 2, No. 2, pp. 01–05, 2018.

13. Z. G. He, X. N. Gu, X. Y. Sun, J. Liu, and B. S. Wang, *An efficient pseudo-potential multiphase lattice Boltzmann simulation model for three-dimensional multiphase flows*, Acta Mechanica Malaysia, Vol. 1, No. 1, pp. 08–10, 2017.

14. M. Elmnifi, M. Amhamed, N. Abdelwanis, and O. Imrayed, *Solar Supported Steam Production For Power Generation In Libya*, Acta Mechanica Malaysia, Vol. 2, No. 2, pp. 05–09, 2018.

CONTACT WITH THE AUTHOR

Bianjiang Hu

School Of Business
Central South University
Changsha,Hunan 410083
China