Development trend analysis and prediction of photovoltaic building integration plate index

Linpeng Zheng[1, *, #], Weipeng Cai[2, #], Qili Fu[2, #]

[1]Jinan University, Jinan University-University of Birmingham Joint Institute, Guangzhou, Guangdong, 511443

[2]Jinan University, College of Information Science and Technology of Jinan University, Guangzhou, Guangdong, 511443

*Corresponding author:man007@stu2019.jnu.edu.cn

[#]These authors contributed equally.

Abstract. Photovoltaic building integration plate as one of the eye-catching emerging plate in the capital market this year, so the stock of related products has become the focus of current practitioners.In this paper, a moving average model of photovoltaic building integration plate index was established with N days as the calculation cycle, and the optimized model was used to draw the daily moving average, weekly moving average and monthly moving average of the next 20 days, 3 weeks and 2 months respectively.In order to measure the correlation between Shanghai Composite Index and photovoltaic building integrated plate index in recent two years, variance analysis method and Spearman correlation coefficient method are used to analyze the correlation, and it is found that the two have a positive correlation and strong correlation except for a few periods.

Keywords: BIPV Moving average, correlation analysis.

1. Introduction

All manuscripts must be in English, also the table and figure texts, otherwise we cannot publish your paper. Please keep a second copy of your manuscript in your office. When receiving the paper, we assume that the corresponding authors grant us the copyright to use the paper for the book or journal in question. Should authors use tables or figures from other Publications, they must ask the corresponding publishers to grant them the right to publish this material in their paper. Use *italic* for emphasizing a word or phrase. Do not use boldface typing or capital letters except for section headings (cf. remarks on section headings, below).

Do not number your paper: All manuscripts must be in English, also the table and figure texts, otherwise we cannot publish your paper. Please keep a second copy of your manuscript in your office. When receiving the paper, we assume that the corresponding authors grant us the copyright to use the paper for the book or journal in question.

2. Introduction

BIPV (building integrated photovoltaic) is a kind of solar (pv) products are integrated into the construction technology, application of solar power is a kind of new concept[1], simply is to install solar energy photovoltaic phalanx in the palisade structure of the building exterior surface to provide power, to our country to realize "carbon peak" and "carbon neutral" play an important role.Therefore, photovoltaic building integration plate as one of the eye-catching emerging plate in the capital market this year, has a quite good performance in the past few years, so the stock of related products has become the focus of current practitioners[2].

3. Moving average model

3.1 Model establishment

In this paper, the weight of each stock on the sector is defined by the total market value of each stock[3]. The performance of enterprises with larger scale in the capital market has a greater impact on the sector[4]. Therefore, the market value is the most intuitive indicator to measure the size of a company and naturally becomes the core element to confirm the weight of each company on the sector.The total market value of a company on a specific trading day and the total market value of the photovoltaic building integration plate are shown in Formula 1 and Formula 2:

$$MC_{ij} = GC_{ij} \times CP_{ij} \tag{1}$$

$$PMC_j = \sum_{i=1}^{37} MC_{ij} = \sum_{i=1}^{37} GC_{ij} \times CP_{ij} \tag{2}$$

MC_{ij}, and respectively represent the total market value, total share capital, closing price of the trading day and the total market value of the plate on the day j of Company I.GC_{ij}、$CP_{ij}PMC_j$This paper summarized the data needed to estimate the market value by querying the changes in the equity capital of these companies in the past two years on eastmoney.com[5], and started to build the moving average model gradually after the data was complete: The index calculation formula of photovoltaic building integrated block is as follows:

$$PI_j = \frac{PMC_j}{PMC_1} \times 1000 \tag{3}$$

PI_j(Plate Index) represents the Index of photovoltaic building integration Plate on day J, and represents the total market value of photovoltaic building integration Plate on day J and day 1.PMC_jPMC_1

Based on the obtained plate index of different days, we can confirm the moving average with different periods as a time unit, and further build the model[6]:

$$MA_N(t) = \frac{1}{N} \times \sum_{j=t-N+1}^{t} PI_j \tag{4}$$

$MA_N(t)$Represents the moving average with N days as a cycle (the moving average period is N days), and (Plate Index) represents the Index of photovoltaic building integrated Plate on the J day.PI_jTaking the 5-day, 10-day and 20-day moving average models as examples, the examples of moving average models are as follows:

$$MA_{D5}(t) = \frac{1}{5} \times \sum_{j=t-4}^{t} PI_j \tag{5}$$

$$MA_{D10}(t) = \frac{1}{10} \times \sum_{j=t-9}^{t} PI_j \tag{6}$$

$$MA_{D20}(t) = \frac{1}{20} \times \sum_{j=t-19}^{t} PI_j \tag{7}$$

After plugging in the data, the corresponding moving average curves are drawn, as shown in Figure 1 and Figure 2.

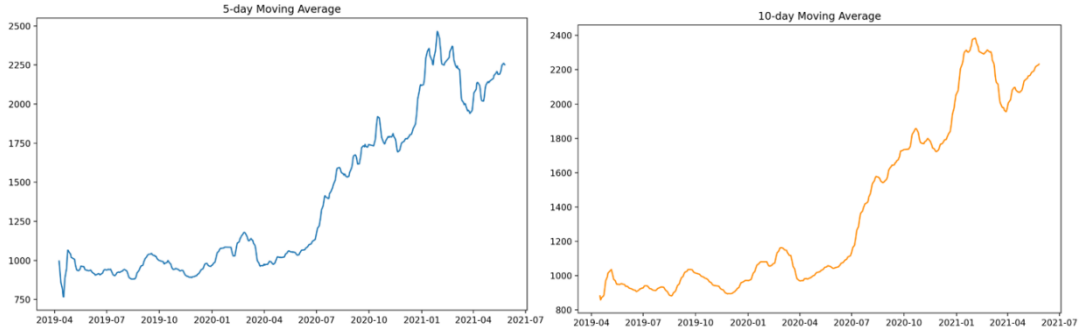

Figure 1. 5 - and 10-day moving averages

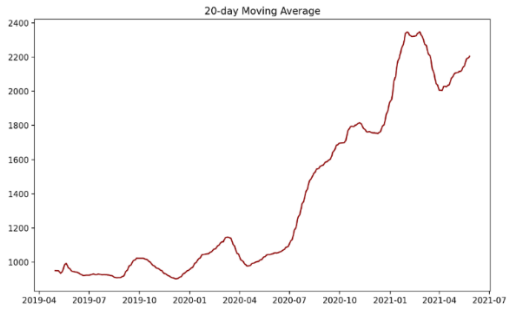

Figure 2. 20-day moving average

From the 5-day average to the 10-day average and then to the 20-day average, the curve is getting smoother and smoother. While the inflection point is not so obvious, the general shape of the three averages is basically consistent with the trend[7].This is actually because the longer the cycle[8], the more inclusive the running data is to the overall data, the smoother the curve is naturally;But it's the same data, and no matter how you average it, the shape of the curve doesn't change much.

3.2 Error analysis and correction of the model

$MA_N(t)$Represents the moving average with N days as a cycle (the moving average period is N days), and (Plate Index) represents the Index of photovoltaic building integrated Plate on the J day.PI_j

$$MA_N(t) = \frac{1}{N} \times \sum_{j=t-N+1}^{t} PI_j \tag{8}$$

Since this model is a moving average model including time variables, the essence of the model after error analysis and modification is still a moving average, so the change before and after modification can only be N days.Based on this idea, the team decided to make calculations according to the data of 17 trading days from May 6 to May 28, 2021, namely, N=1, 2, 3... 16 and 17 are respectively put into the error analysis model for calculation, and then N with the smallest error is selected as the revised model.The error analysis model and correction model will be introduced in detail below.

$$E_{N+1} = \left(\frac{\sum_{i=1}^{N} PI_i}{N} - PI_{N+1}\right)^2 \tag{9}$$

$$\sigma_j = \sqrt{\frac{\sum_{j=N+1}^{17} {E_j}^2}{17}} \tag{10}$$

Let's say we have a period of N days.Where, represents the error of the model's estimated data on the N+1 day, (Plate Index) represents the Index of the integrated photovoltaic building Plate on the I

day, and represents the overall standard deviation of the model's estimated data on the J day.$E_{N+1}PI_i\sigma_j$

After python program writing and data substitution, we get N=1, 2, 3..... Error analysis data of 16 and 17, namely, the overall standard deviation of the model's estimated data when N= N.As can be seen from the table, when N=5, the overall standard deviation predicted by the model is the smallest, which is only 60.37.Based on the above error analysis, N=5 days is selected as the modified model parameter in this paper, and the accuracy of the moving average model constructed on this basis is the highest, which is:

$$MA_5(t) = \frac{1}{5} \times \sum_{j=t-4}^{t} PI_j \tag{11}$$

$MA_5(t)$Represents the moving average of a 5-day cycle, and Plate Index represents the Index of photovoltaic building integrated Plate on the j-day.PI_j

3.3 Modified model

Since the essence of the moving average model is a time series model, according to the definition, 3-4 time units of data predicted by the model are more accurate. Generally, the error of data after the fifth cycle is relatively large, which is not desirable.In other words, the source of these errors is not the low optimization degree of the built model, but the limitation of the time series model itself.

According to the moving average model based on N=5 days, the team predicted the daily moving average of 20 trading days after May 28, the weekly moving average of 3 weeks and the monthly moving average of 2 months, and drew three corresponding moving average line charts, as shown in Figure 3.

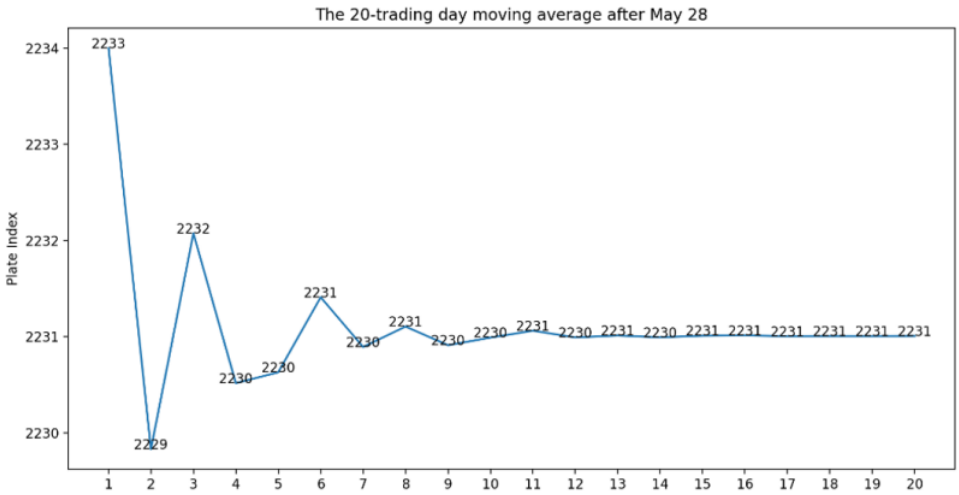

Figure 3 Daily moving averages for the next 20 trading days

The unit of time on the horizontal axis is days, and the horizontal axis is 1,2,3...It is the first and second trading days from May 28.As mentioned above, the predicted plate index of the first three trading days is relatively accurate and of great reference value, but the data from the fourth trading day is of low reliability.The weekly moving average for the last three weeks is shown in Figure 4.

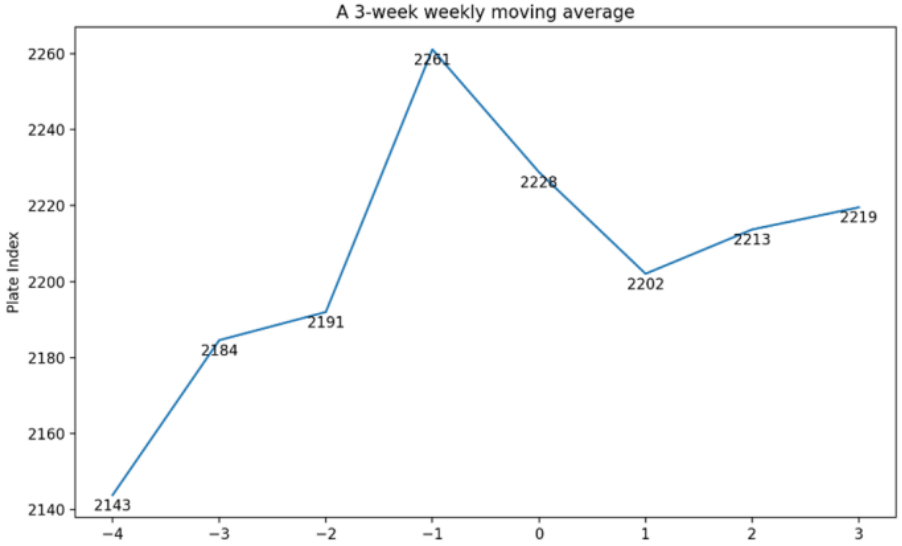

Figure 4 3 weekly moving average

The unit of variable time on the horizontal axis is the number of weeks, and the complex part of the horizontal axis represents the real plate index in the weeks leading up to May 28, 1,2,3...Represents the projected plate data for weeks 1, 2 and 3 beginning May 28.The forecast results show that the pv building integration plate index will rise continuously after a relatively rapid decline in the next three weeks, but the index has not recovered to the starting week of May 28 after the end of the third week.The monthly moving average for the next 2 months is shown in Figure 5.

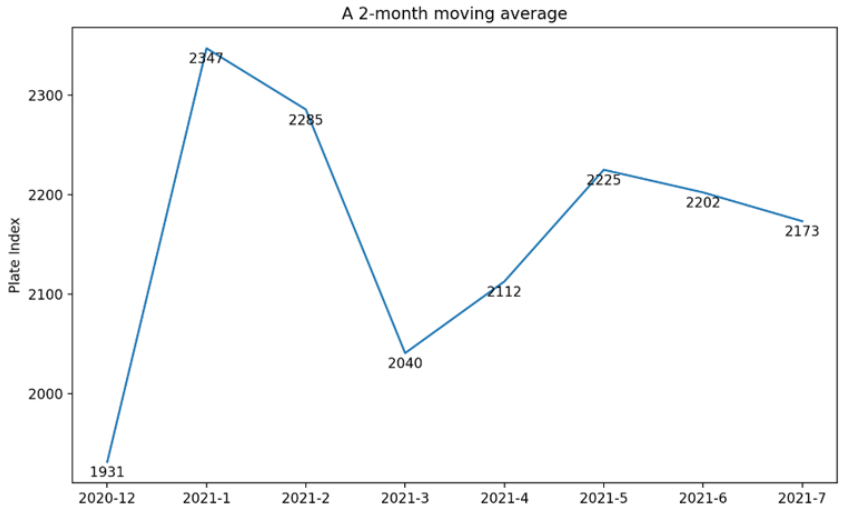

Figure 5. 2-month moving average

The horizontal variable is time. The average line before the horizontal coordinate of 2021-5 represents the real value, while the line chart for the following two months from 2021-5 is estimated.The forecast results show that in the next two months, the monthly average of the pv building integration plate index is on a downward trend, and the index decline rate is about 16 points per month.

4. Correlation analysis of the two indices

In this paper, the period from April 1, 2019 to May 28, 2021 is divided into a two-month cycle, which can be roughly divided into 13 "February". The data of each "February" includes two groups of data of Shanghai Composite Index and PHOTOVOLTAIC Building Integrated sector Index, based on which correlation analysis is conducted.In order to fully understand the correlation degree of the

two indicators, this paper adopts the variance analysis method and spearman correlation coefficient to analyze the correlation.

4.1 Static scatter diagram and broken line diagram

The scatter chart of Shanghai Composite Index and pv building integration plate Index is shown in Figure 6.

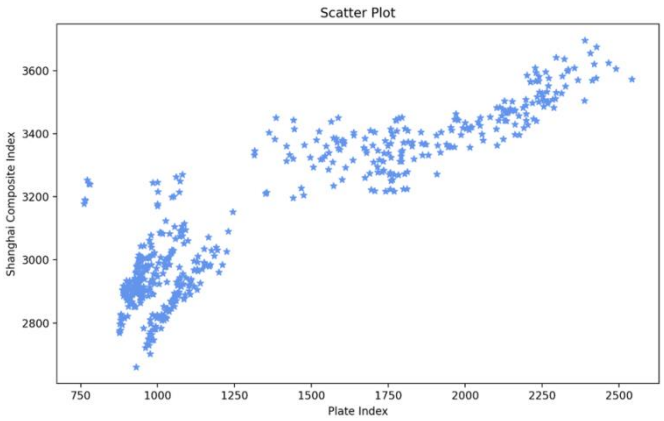

Figure 6 Scatter chart of Shanghai Composite Index and pv building integration plate Index

The abscissa and ordinate respectively represent the photovoltaic building integration plate index and the Shanghai Composite Index[9]. The results in the figure show that: When China's A-share market is relatively depressed, that is, when the Shanghai Composite Index is low, the index of the photovoltaic building integration new plate changes slowly with the changes of the Shanghai Composite Index;On the contrary, when the market of our country capital market begins to rise, the rate of change of the index of this plate gradually accelerates.No matter how high the index is, the data of the two groups are positively correlated.The line chart of Shanghai Composite Index and photovoltaic building integrated plate Index is shown in Figure 7.

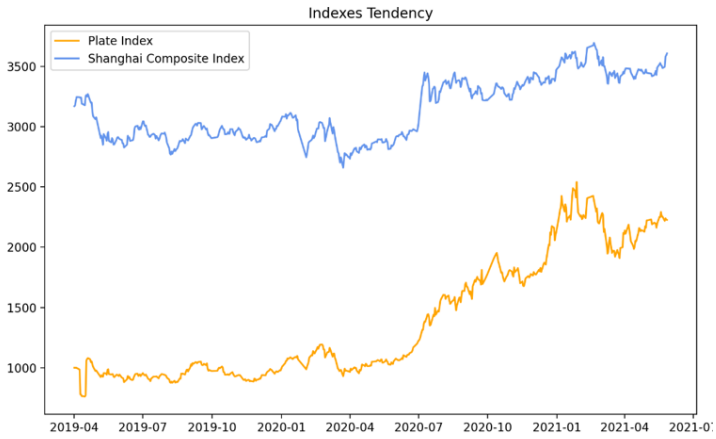

Figure 7 Line chart of Shanghai Composite Index and pv building integrated block index

Blue broken line and orange broken line respectively represent the Shanghai Composite Index and photovoltaic building integration plate index;The abscissa and ordinate represent the date and exponent of a two-month cycle, respectively.It can be seen from the figure that the overall trend of the two index line charts is relatively similar. Although the time of every inflection point of climb or dip is not the same, one index can basically keep up with the trend of the other index in a relatively short time, and the only difference is the range of change and slight time difference.Overall, the correlation between the two is strong[10].

4.2 Analysis of variance

S_{SI} And represent the standard deviation of Shanghai Composite Index and photovoltaic building integrated plate Index respectively. $S_{PI}S_{SIPI}$And respectively represent the covariance and correlation coefficient of the two groups of data.r_{SIPI}

$$S_{SI} = \sqrt{\frac{\sum_{i=1}^{13}(SI_i - \overline{SI})^2}{12}} \tag{12}$$

$$S_{PI} = \sqrt{\frac{\sum_{i=1}^{13}(PI_i - \overline{PI})^2}{12}} \tag{13}$$

$$S_{SIPI} = \frac{\sum_{i=1}^{12}(SI_i - \overline{SI}) \times (PI_i - \overline{PI})}{12} \tag{14}$$

$$r_{SIPI} = \frac{S_{SIPI}}{S_{SI} \times S_{PI}} = \frac{\sum_{i=1}^{12}(SI_i - \overline{SI}) \times (PI_i - \overline{PI})}{12 \times S_{SI} \times S_{PI}} \tag{15}$$

The final correlation coefficient can reflect the correlation of the two indices, as shown in Table 1.

Table 1 Correlation coefficients based on anOVA

time	The correlation coefficient
April 1, 2019	0.0382
June 3, 2019	0.7782
August 1, 2019	0.9601
October 8, 2019	0.7392
December 2, 2019	0.8760
February 3, 2020	0.9705
1 April 2020	0.8962
1 June 2020	0.8984
August 3, 2020	0.5244
9 October 2020	0.0030
1 December 2020	0.8110
February 1, 2021	0.9185
April 1, 2021	0.6299

It can be seen from Table 1 that the correlation coefficient between Shanghai Composite Index and photovoltaic building integrated plate Index obtained by variance analysis method is close to 0 in two periods, and basically remains above 0.6 in other periods.In other words, from April 1, 2019, to May 28, 2021, the two indices show a positive correlation in most February cycles, especially in some cycles.But in general, the average correlation is only 0.6124, which cannot be said to have a strong correlation.

4.3 Spearman correlation coefficient

Spearman correlation coefficient method is a non-parametric index used to measure the degree of dependence of two variables.The principle is to use monotone equation to evaluate the correlation of two statistical variables after the original data is transformed into registration data by sorting.Its specific calculation model is as follows:

$$\rho = \frac{\sum_{i=1}^{13}(PI_i - \overline{PI}) \times (SI_i - \overline{SI})}{\sqrt{\sum_{i=1}^{13}(PI_i - \overline{PI})^2 \times \sum_{i=1}^{13}(SI_i - \overline{SI})^2}} \tag{16}$$

PI_i And respectively represent the photovoltaic building integration plate index and the Shanghai Composite Index day I plate index; $SI_i\overline{PI}$And represent the average index of 13 cycles of photovoltaic building integration plate Index and Shanghai Composite Index respectively.$\overline{SI}$According to the principle of the model, the team sorted 13 groups of data with a period of two months and then ran the program of correlation coefficient. The correlation coefficient J based on Spearman correlation coefficient method is shown in Table 2.

Table 2 Correlation coefficients based on Spearman correlation coefficient method

time	The correlation coefficient
April 1, 2019	0.3859
June 3, 2019	0.8076
August 1, 2019	0.9645
October 8, 2019	0.7601
December 2, 2019	0.8904
February 3, 2020	0.9448
1 April 2020	0.9043
1 June 2020	0.8432
August 3, 2020	0.4857
9 October 2020	0.1217
1 December 2020	0.7675
February 1, 2021	0.9223
April 1, 2021	0.6188

Similar to the data obtained by method analysis, the correlation coefficient of the two plate indices is above 0.75 in most periods, and the correlation between them is strong and positive.The correlation coefficients of the two periods from August to October and From October to December are weak and not positive.

5. Evaluation of the model

(1) During the establishment of the moving average model, different errors can be obtained by changing the number of moving periods so as to modify the model and improve the stability of the model.

(2) The variance analysis method and Spearman correlation coefficient method are adopted in the correlation analysis to make the correlation analysis more accurate and objective.

(3) When determining the portfolio, the return rate and risk of individual stocks are considered, and the optimal investment plan is formed by selecting individual stocks through the expected return rate of assuming unit risk.

References

[1] Sheikh Khozani Zohreh, Barzegari Banadkooki Fatemeh, Ehteram Mohammad, Najah Ahmed Ali,El-Shafie Ahmed. Combining autoregressive integrated moving average with Long Short-Term Memory neural network and optimisation algorithms for predicting ground water level[J]. Journal of Cleaner Production,2022,348.

[2] Tak Sehyun, Kim Sari, Yu Hwapyeong, Lee Donghoun. Analysis of Relationship between Road Geometry and Automated Driving Safety for Automated Vehicle-Based Mobility Service[J]. Sustainability,2022,14(4).

[3] Ayache Antoine, Bouly Florent. Moving average Multifractional Processes with Random Exponent: Lower bounds for local oscillations[J]. Stochastic Processes and their Applications,2022,146.

[4] Li Qi, Yang Jun, Huang Shuo,Zhao Yu. Generally weighted moving average control chart for monitoring two-parameter exponential distribution with measurement errors[J]. Computers & Industrial Engineering,2022,165.

[5] Shan Zhen, Yang Jianhua, Sanjuán Miguel A. F.,Wu Chengjin,Liu Houguang. A novel adaptive moving average method for signal denoising in strong noise background[J]. The European Physical Journal Plus,2021,137(1).

[6] ETTE HARRISON ETUK, GODWIN NWAFOR ONYEKA, LENYIE LEESIE. AN AUTOREGRESSIVE INTEGRATED MOVING AVERAGE INTERVENTION MODEL OF 2016 BRAZILIAN REAL AND NIGERIAN NAIRA EXCHANGE RATES[J]. Journal of Basic and Applied Research International,2021.

[7] Sahirul Alam, Selo Sulistyo, I Wayan Mustika, Ronald Adrian. Handover Decision for V2V Communication in VANET Based on Moving Average Slope of RSS[J]. Journal of Communications,2021,16(7).

[8] Pratap Singh Surya, Singhvi Shlok, Gupta Rahul,Gaur R.C.,Sharma M.K.. Correlation Study for the Assessment of Groundwater Quality of Sanganer Tehsil, Jaipur City, Rajasthan[J]. IOP Conference Series: Earth and Environmental Science,2021,796(1).

[9] Tian Jing. Dependence Analysis of Insurance Businesses Based on Hierarchical Archimedean Copula Function[J]. Journal of Physics: Conference Series,2021,1952(4).

[10] Sun Jiahui, Liu Chanjuan,Yu Siqin. Application of Rescaled Range Analysis Method to Ship Flow in Container Ports: Time Series Structure and Long-Range Dependence Analysis[J]. Modern Economy,2021,12(05).