WEP Warwick Evans Publishing

Transactions on Environment, Energy and Earth Sciences
ISSN: 2960-1924, eISSN: 2960-2300 | Volume 3
GRET 2024

Prediction of Extreme Weather's Frequency Based on Seasonal ARIMA and Correlation Analysis Methods

Lei Weng *

Department of China School of Banking and Finance, University of International Business and Economics, Beijing, China, 100020

* Corresponding author: ingram20040508@163.com

Abstract. The extreme-weather events have exposed high risks to many aspects of life to the current world, especially for those places with more frequency like Japan and the southeastern coast of America. High frequency of such natural disasters will do lots of damages to every aspect of human's life. However, the current research hasn't paid enough attention to the seasonal and frequency characteristic of the extreme weather. Hence, to analyze and predict the occurrence of the extreme weather in those places, this paper introduces correlation analysis and seasonal ARIMA model to figure out the seasonal characteristic and future trends of the extreme weather's frequency. It turns out that extreme weather in both sides is more frequent in 2nd and 3rd quarter of the year, and shows high similarity in the two places, which can be explained by the climate characteristic of both places. The research is expected to provide frequency information for prevention of extreme weather.

Keywords: Occurrence Predication; Correlation Analysis; Seasonal ARIMA.

1. Introduction

Extreme weather, which can be deleterious to life and property safety, has been laid stress on for many years. Japan is a country with many severe extreme weathers, typhoon and earthquake for instance [1], and the southeastern coast of the United States suffers quantities of extreme weather as well [2]. In fact, the frequent extreme weather has caused a myriad of losses to humans, and such a situation may have a negative impact on resources, construction and many other aspects [3, 4].

In order to reduce the influence of the extreme weather, several kinds of time-series model have been introduced to predict the weather, which is proved to be a workable method [5, 6]. Appropriate utilization of these time-series techniques is of great significance to learn the trend of the extreme weather, and make preparations for it ahead of time. However, the accuracy of the prediction hasn't been ensured, and there's not much research on the extreme weather's frequency of Japan. Besides, the seasonal characteristic of the extreme weather has aroused more and more attention [7].

To predict a reliable frequency of the extreme weather, this paper will be based on the seasonal ARIMA model to make a prediction of the number of extreme weather's occurrence. Meanwhile, relative seasonal characteristic of the natural disasters will be uncovered to reflect the law of the occurrence in a numerical way. Hopefully, this paper will contribute to filling the blank of frequency research in extreme weather, and provide reliable information for weather risk aversion.

2. The extreme weather prediction model

The data are taken from five representative regions in Japan: Tokyo, Nagoya, Fukuoka, Sapporo, and Kagawa for extreme weather's assessment. Through statistical analysis of the occurrence of local extreme weather, five types of natural disasters that have the most serious impact in Japan are determined: blizzard, typhoon, flood, thunderstorm and high temperature. After that, the occurrence of extreme weather between 2014 and 2023 are counted.

2.1. Statistical analysis

Through statistical analysis of the weather-driven losses in Japan during the past 10 years, five extreme weather conditions that have the most significant impact in the representative regions $y_j(j=1,2,3,4,5)$ are found. And the number of extreme weather events per quarter in the region for 10 years is counted as:

$$y_{ij}(j=1,2,3,4,5, \quad i=1,2,3,\ldots,40) \tag{1}$$

Next, the probability of each extreme weather event by quarter is calculated. The probability of the first extreme weather event in the first quarter of each year:

$$P_{11}=\frac{\sum_{i=0}^{9} y_{4i+1,1}}{900} \tag{2}$$

The probability of the first extreme weather event in the second quarter of each year:

$$P_{21}=\frac{\sum_{i=0}^{9} y_{4i+2,1}}{900} \tag{3}$$

Similarly, the probability of 5 extreme weather events in the remaining quarters can be calculated.

After that, the number of seasonal extreme weather events in the region over the past 10 years can be calculated as:

$$Y_i=\sum_{j=1}^{5} P_{kj} * y_{ij} \tag{4}$$

Where k is the remainder of i divided by 4, and if it is 0 then k=4.

2.2. Seasonal ARIMA model

With the purpose of predicting the frequency of the extreme weather, this paper introduces the seasonal ARIMA model to fit in with the frequency data of extreme weather in Japan. Seasonal ARIMA is a time-series model that has an advantage in prediction works of data with a seasonal characteristic [8]. It adds four seasonal parameters to the traditional ARIMA model, namely p, d, q and the seasonal cycle parameter, s. The basic formula of the Seasonal ARIMA model is as follows:

$$\phi_p(B)\Phi_P(B_s)(1-B_s)^d(1-B_s)^D x_t=\theta_q(B)\Theta_Q(B_s)\epsilon_t \tag{5}$$

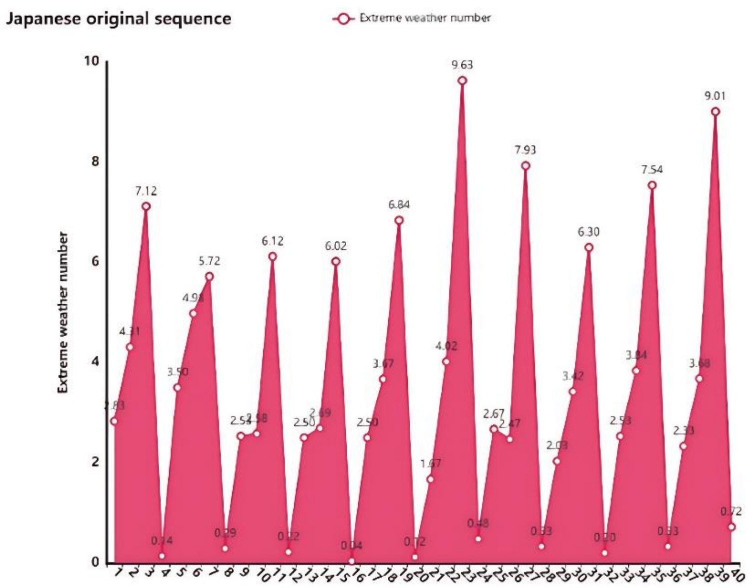
In formula (5), B is the delay operator of the time-series model, and typically B_s is the seasonal delay operator. $\phi(B),\theta(B)$ are the delay polynomial operators while $\Phi(B)$ and $\Theta(B)$ are the seasonal delay polynomial operators [9]. The participation of seasonal parameters in this model uplifts its performance for data with a potential seasonal trend.

The seasonal ARIMA model has been proved to have a better performance on data with seasonal fluctuation [10]. Based on the seasonal ARIMA model above, a frequency prediction model is introduced: First, the time series is divided into trend data, seasonal data and random data to preliminarily judge the seasonal effect of the data. Then the stationarity test of the time series is carried out, and the final difference series diagram is viewed for autocorrelation analysis. Finally, the fitting effect of our model is evaluated. For convenience, the predicted Y is denoted as F.

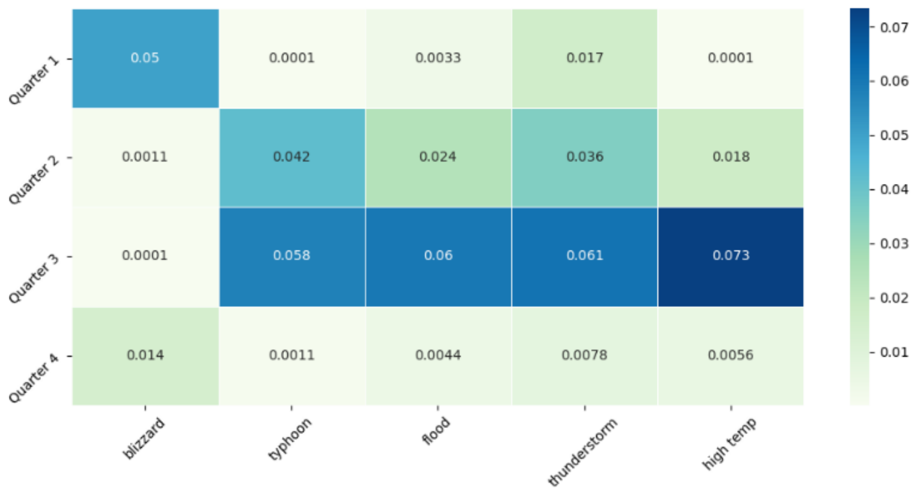
3. Results and discussion

3.1. Results in Japan

Firstly, representative regions in Japan: Tokyo, Nagoya, Fukuoka, Sapporo, and Kagawa are selected for extreme weather's assessment. Through statistical analysis of the occurrence of local extreme weather and corresponding insurance payouts, five types of extreme weather that have the most serious impact in Japan are determined: blizzard, typhoon, flood, thunderstorm and high temperature. By combining the probability of extreme weather occurrence in each quarter with its statistical number and substituting it into the corresponding formula of the model, the average number of extreme occurrences in each quarter for each region in Japan in recent 10 years is calculated:



(a) Average number of extreme weather's occurrence



(b) Heat plot of the extreme weather and the year quarters

Figure 1. The average number and heat plot of extreme weather occurrences in each quarter in Japan

As shown in Figure 1, the count of the extreme weather and correlation in each quarter in Japan indicates the high occurrence of extreme weather in summer and autumn. As for the correlations between year-quarters and the five-type extreme weather, the third quarter has a strong relationship with most of the extreme weathers. The information shows a strong seasonal characteristic in the extreme weather's occurrence in Japan, which indicates the necessity of more investment in risk aversion during the 2nd and 3rd quarter in Japan.

On top of that, the seasonal ARIMA (SARIMA) time series analysis is then performed on the data in order to make a prediction of the number of extreme weather events in the future.

Table 1. Summary of the SARIMA model

variable	sequence	t	P	AIC	Threshold		
					1%	5%	10%
Extreme weather number	Original sequence	-2.477	0.121	83.479	-3.661	-2.961	-2.619
	1st order difference	-2.895	0.046**	82.388	-3.689	-2.972	-2.625
	1st order difference - 1st order seasonal difference	-5.809	0.00***	74.546	-3.661	-2.961	-2.619
	2nd order differential	-5.405	0.00***	86.853	-3.661	-2.961	-2.619
	2nd order difference - 1st order seasonal difference	-6.238	0.00***	79.283	-3.689	-2.972	-2.625

Note: ***, **, and * represent the significance levels of 1%, 5%, and 10%, respectively

It is clear seen in Table 1 that through the ADF test, the first-order difference series based on the number of the extreme weather has a P-value of 0.046**, showing much significance at the horizontal level. The null hypothesis is rejected, and the series is a stationary time series. Based on the Extreme weather number variable, it can be concluded from the analysis of the residual Q statistic results that Q6 does not show significance in level, and the hypothesis that the residual of the model is a white noise sequence cannot be rejected, which means the model basically meets the requirements. The goodness of fit R^2 of the model is 0.629, indicating the model performs well.

Finally, this paper evaluates the degree of fitting and prediction of the model, and give the forecast data (shown in Figure 2).

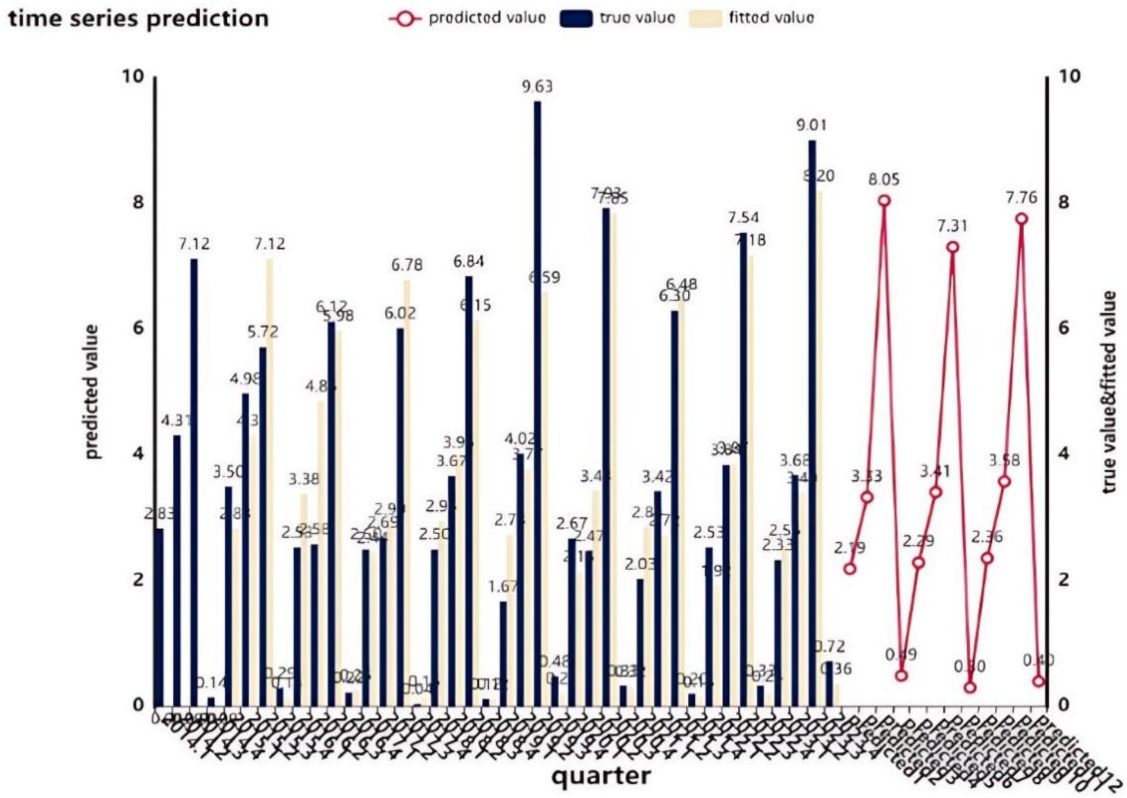
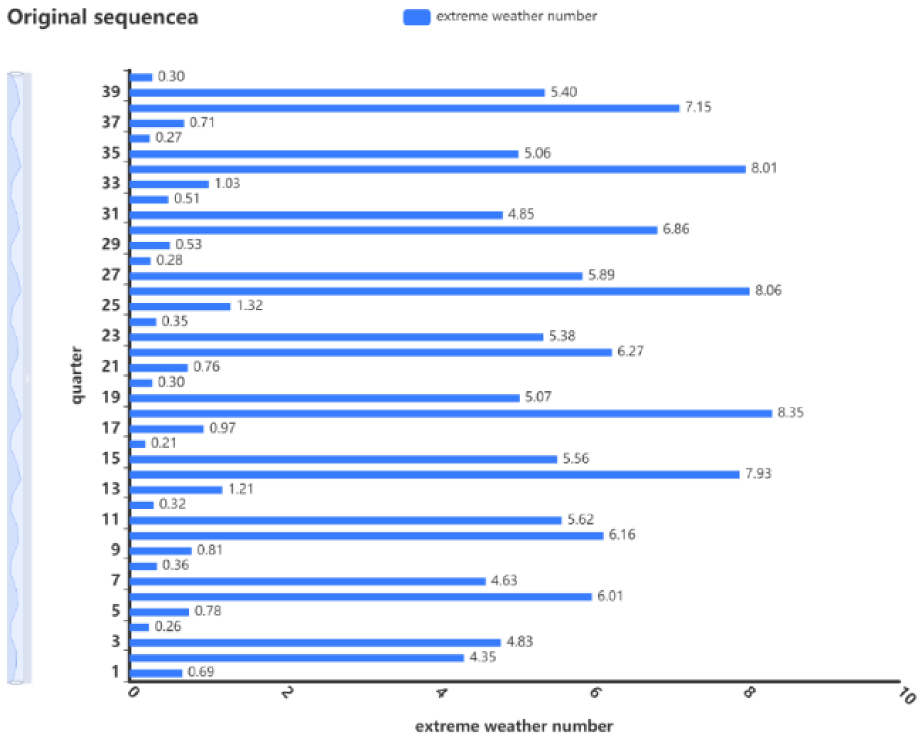


Figure 2. The prediction of extreme weather based on SARIMA

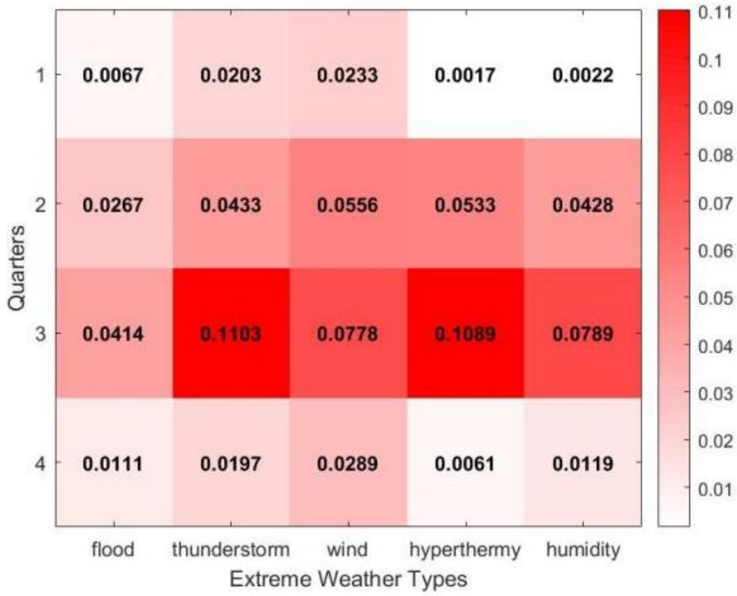
Through Figure 2, it is clear that the predicted data of extreme weather inherit the seasonal trend of the previous data, namely high in 2nd and 3rd quarter of the year. Thus, the seasonal characteristic of the extreme weather's occurrence is apparent to see.

3.2. Results in the southeastern coast of the United States

Simultaneously, the same prediction is conducted on the data in the southern coast of the United States. The same data with the consistent time period is calculated (shown in Figure 3).



(a) Quantity of the extreme weathers in southeast of America



(b) Heat plot of the extreme weather and the year quarters

Figure 3. Quantity of extreme weathers in the southeast of America and probability of occurrence

From Figure 3, it can be seen that the quarter with the highest occurrence of the extreme weather is the 3rd one (same as the result in Japan). Taking the similarity of both side's climate into consideration [11], this may indicate the same seasonal characteristic of both side's extreme weather.

Then, the seasonal ARIMA model is employed as well, whose prediction result is shown in Figure 4.

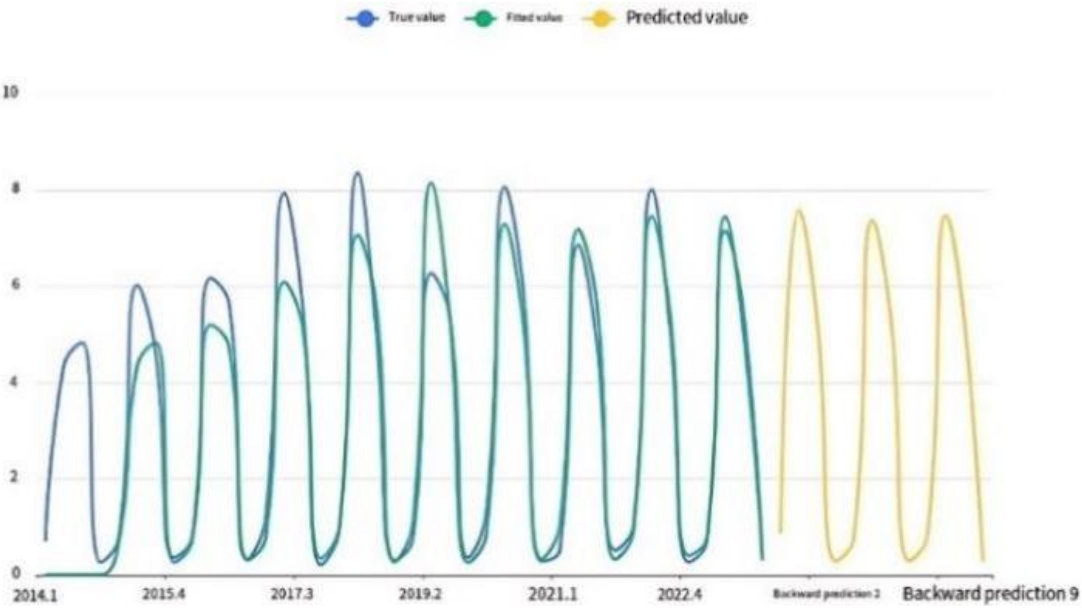


Figure 4. The result of prediction in southeast of America

As is shown in Figure 4, the prediction tends to be seasonal as well. This indicates the same frequency type of the extreme weather in the southern part of America.

The whole system makes it to conclude the time-various trends of the extreme weather in Japan, and make predictions based on the evaluated results. The high occurrence of extreme weather in the 2nd and 3rd season of the year leads to necessary actions to protect the properties and people ahead of time. Simultaneously, the findings of the prediction provide useful information about the distribution of the money to spend on preventing extreme weather.

The whole system has concluded the time-various trends of the extreme weather in Japan and the southeastern coast of the United States, and make predictions based on the evaluated results. The high occurrence of extreme weather in the 2nd and 3rd season of the year leads to necessary actions to protect the properties and people ahead of time. Simultaneous, the findings of the prediction provide useful information about the decision for distribution of the money to spend on preventing extreme weather.

4. Conclusion

To sum up, the current predictive research on the extreme weather hasn't focused on the frequency characteristic and seasonal trends of these natural disasters. Hence, this paper introduces the correlation analysis and SARIMA model to fill in the blank. Referring to the analysis and prediction, the occurrence of extreme weather in Japan and the United States keeps increasing in recent years. Via correlation analysis, it shows the 2nd and 3rd quarter of the year is inclined to have much more extreme weather, whose similarity of the result may attribute to the climate properties. Moreover, the SARIMA model is proved to perform well on the data and make a reliable prediction, which turns out to inherit the seasonal characteristic as well.

By keeping on learning the information of the climate changes, this method will perform well on the basis of the strong relationship between the extreme weather and seasonality. However, the accurate identification of the weather hasn't been introduced so far. Further studies are expected to investigate the decisional characteristic of the main kinds of the extreme weather, enabling the prediction to be more precise. This paper is expected to supply the information concerning characteristics of extreme weather events, which may be utilized to prevent the losses and make decisions in a more reasonable way.

References

[1] Nakakita, Eiichi, Yasuto Tachikawa, Tetsuya Takemi, Nobuhito Mori and Kenji Tanaka. "Future Changes of Extreme Weather and Natural Disasters due to Climate Change in Japan and Southeast Asia." [B], Bridging Science and Policy Implication for Managing Climate Extremes, 2018: 101 - 117.

[2] Jorgensen S K, Nielsen-Gammon J W. Nonstationarity in Extreme Precipitation Return Values along the US Gulf and Southeastern Coasts [J], Journal of Hydrometeorology, 2024, 25 (5): 771 - 788.

[3] Zou, X., Peng, X., Zhao, X., & Chang, C. The impact of extreme weather events on water quality: international evidence [J], Natural Hazards, 2022: 1 - 21.

[4] Choi HG, Min C, Kim SY. Air pollution increases the risk of SSNHL: A nested case-control study using meteorological data and national sample cohort data, Sci Rep, 2019 Jun 4; 9 (1): 8270.

[5] Yu, J. Weather Pollution Prediction Based on XGBoost and ARIMA: Evidence from Shanghai, HSET 2023, 76: 649 - 654.

[6] Luo, K., Li, J., & Zhao, Y. (2022). Research on Time Window Prediction and Scoring Model for Trauma-Related Sepsis [B], LISS,2021: 111 – 123.

[7] Rimi, Ruksana H. et al. Risks of seasonal extreme rainfall events in Bangladesh under 1.5 and 2.0 °C warmer worlds – how anthropogenic aerosols change the story, Hydrology and Earth System Sciences (2022): 26: 5737 - 2022.

[8] Xianqi Zhang, Xilong Wu, Guoyu Zhu, Xiaobin Lu, Kai Wang; A seasonal ARIMA model based on the gravitational search algorithm (GSA) for runoff prediction, Water Supply 1 August 2022; 22 (8): 6959 – 6977.

[9] Dimri, Tripti et al. "Time series analysis of climate variables using seasonal ARIMA approach.", Journal of Earth System Science, 2020 (129): 1 - 16.

[10] Kaur B, Kaur N, Gill K K, et al. Forecasting mean monthly maximum and minimum air temperature of Jalandhar district of Punjab, India using seasonal ARIMA model [J], Journal of Agrometeorology, 2022, 24 (1): 42 - 49.

[11] Tam K P, Chan H W, Clayton S. Climate change anxiety in China, India, Japan, and the United States [J]. Journal of Environmental Psychology, 2023, 87: 101991.