Frontiers in Computing and Intelligent Systems
ISSN: 2832-6024 | Vol. 8, No. 3, 2024

Study on the Development of Property Insurance Industry Based on ARIMA and CGDAM-WRIR Models

Yibo Huang [1, *], Guoli Sun [1], Xiaoyin Huang [2], Cen Lin [1], Hang Li [1]

[1] The Department of Business Administration, Tianjin University of Commerce Boustead College, Tianjin, China
[2] Department of Economics and Trade, Tianjin University of Commerce Boustead College, Tianjin, China

*** Corresponding author:** Yibo Huang (Email: 3065687330@qq.com)

Abstract: This study aims to address the climate change challenges faced by the property-casualty insurance industry by using the ARIMA model to assess the potential impacts of extreme weather events on the insurance industry and the CGDAM-WRIR model to analyze the regional impacts of tornadoes and the impacts of climate change on communities. With these models, we provide accurate risk management and underwriting strategy adjustment recommendations. Future research will further explore complex models, overcome limitations, integrate multi-source data, and develop real-time predictive models to help the insurance industry better manage climate change uncertainty, safeguard property owner interests, and ensure market sustainability and resilience.

Keywords: ARIMA Model; CGDAM-WRIR Model; Risk Assessment.

1. Introduction

The property and casualty insurance industry are facing unprecedented challenges as the frequency and intensity of extreme weather events triggered by climate change increases [1]. This study aims to provide insight into the potential impacts of extreme weather events on the insurance industry and the impacts of climate change on communities and regions by employing the ARIMA model and the CGDAM-WRIR model. The ARIMA model will be used to assess the future losses of the insurance industry due to extreme weather events, to predict the mortality rate, and to analyze the impacts of tornadoes of varying intensities on regions. And the CGDAM-WRIR model will be used to study the impacts of climate change on communities and the regional impacts of tornadoes. Using these models, we aim to provide the property insurance industry with more accurate risk management recommendations to help the industry meet the challenges posed by climate change and ensure the sustainability and resilience of the market. The combined use of these methods will provide the insurance industry with more comprehensive insights and decision support to address the increasingly complex challenges of climate change.

2. ARIMA Model

2.1. Assessing the Potential Losses of the Insurance Industry Caused by Future Extreme Weather Events

As global climate change continues to intensify, the P&C insurance industry is facing unprecedented challenges. The increased frequency and intensity of extreme weather events such as floods, hurricanes, and droughts have led to significant property damage that has a direct impact on insurers' ability to pay and risk management strategies [2]. In order to ensure the sustainability of the P&C insurance industry, there is an urgent need for insurers to develop models that predict potential future losses. This model will help insurers price insurance products more accurately, flexibly adjust underwriting policies, and develop effective capital reserves and risk diversification measures to respond to changing natural disaster risks. The total casualty statistics for different disaster events is shown in Figure 1.

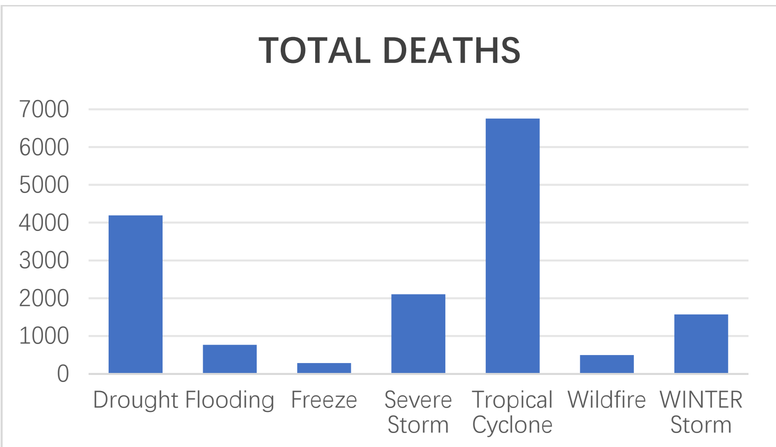

Figure 1. Total casualty statistics for different disaster events (USA)

Considering the uniqueness of the time series data, we chose the ARIMA model for analysis. This model is a widely used method for time series forecasting, especially for non-seasonal data that exhibit clear trends or patterns [3]. The model predicts future values by combining historical and error values of time series data. The purpose of this analysis is to assess the potential losses to the insurance industry caused by possible future extreme weather events, and the ARIMA model can provide us with more accurate predictions, which can help to formulate more effective risk management strategies and capital reserve decisions.

The ARIMA model (Autoregressive Integral Moving Average Model) is a powerful time series forecasting tool that captures and leverages autocorrelations in time series data to predict future trends by integrating information from historical data [4]. By building the ARIMA model, we can make full use of the intrinsic regularity of time series data to make accurate predictions about future data. It combines the advantages of three main components: autoregressive (AR), differential (I), and moving average (MA) to describe the dynamic changes in time series data.

Autoregressive (AR) Section: This section predicts the current value by using past values of time series data. The order p represents the number of past values used for prediction. By adjusting the order p, the ARIMA model can capture different autoregressive relationships in time series data.

Differential (I) part: For non-stationary time series data, we can make it stationary by differential operation. Differential operations eliminate trend or seasonal effects by subtracting the time series from itself. Differential operations are usually expressed as $(1-L)^d$, where L is the lag operator and order d is the number of differential times to ensure smooth data.

Moving Average (MA) Section: This part involves a linear combination of prediction error terms. Order q represents the number of past prediction errors used to predict the current value. By adjusting the order q, the ARIMA model can capture the short-term correlation of time series data.

The expression for the basic equation of the ARIMA model is:

$$\left(1-\sum_{i=1}^{p} \varphi_i L^i\right)(1-L)^d y_t = \left(1+\sum_{i=1}^{q} \theta_i L^i\right)\epsilon_t \quad (1)$$

In the ARIMA model, L is the lag operator, which represents a delay of one step back in the time series. When L acts on y_t, the lag value of y_t can be obtained, which is the value of the previous period of y_t. This can be expressed as $L^k y_t = y_{t-k}$. Each component in the model has its specific function: φ are the autoregressive coefficients, which are multiplied by p historical values of the time series and used to predict the current value. θ_i are the moving average coefficients, which are multiplied by q prediction error terms of the model and are used to correct the prediction error. d is the number of times the model needs to be differentiated to ensure the stationarity of the time series. Through the combination of these components, the ARIMA model can effectively capture and predict the dynamic changes in time series data.

The parameters p, d, and q of the ARIMA model can usually be analyzed in terms of the autocorrelation function (ACF) and partial autocorrelation function (PACF) of the time series. These functions help us visualize the correlations between data points, as well as their independent correlation with other data points.

Autoregressive order p : This can be determined by looking at the PACF plot. In general, we can find the last significant non-zero lag in the PACF plot as the autoregressive order p of the ARIMA model.

Differential degree d : can be determined by the ADF test (Augmented DickeyFuller). If the time series becomes stationary after one or more differences, then the number of differences d is the corresponding number of differences.

Moving average order q : This can be determined by looking at the ACF plot. In general, we can find the last significant non-zero lag in the ACF plot as the moving average order q of the ARIMA model.

The ARIMA model predicts future values, such as the total annual loss for the next seven years in the insurance industry, which is valuable for insurer strategy development. To ensure the validity of the predictions, the model needs to be evaluated, including residual analysis, confidence intervals for predictions, and fit of historical data.

2.2. Predicting the Number of Future Deaths

Impact of Disaster Events on Lives: The far-reaching impact of disaster events on human life is clearly demonstrated through relevant datasets. This information is critical for contingency planning, resource allocation, and risk assessment and management in the insurance industry. Understanding and analyzing this impact can help agencies and companies develop strategies more effectively to mitigate the possible losses and impacts of future disaster events.

Predicting future risks: Through time-series analysis and prediction of future deaths, authorities can better prepare for potential extreme weather events, so that disaster relief and relief plans can be planned and implemented in advance. This forecasting approach not only enhances the ability to respond to disasters, but also ensures that resources can be mobilized quickly and orderly in the event of an emergency, thereby minimizing casualties and property damage.

Given that the dataset provided is an annual record, this may not reveal any seasonal fluctuations, but it is still possible to show long-term trends or cyclical patterns. It is based on this that the ARIMA model is selected for analysis. The ARIMA model is particularly suitable for identifying and dealing with trends and non-seasonal changes in data and can effectively convert non-stationary sequences into stationary sequences through differential methods.

Trends in seasonal extreme weather events, as shown in Figure 2.

In the ARIMA model analysis, the stationarity of the time series is first evaluated using the ADF test. Nonstationary sequences are differentially transformed. Subsequently, by observing the autocorrelation plot and the partial autocorrelation plot, the appropriate AR and MA parameters were determined. The AR parameter (p) is based on the number of significant lag periods after censoring of the PACF plot, and the MA parameter (q) is based on the significant lag period of the ACF plot. This ensures that the ARIMA model captures key features, improving prediction accuracy and efficiency.

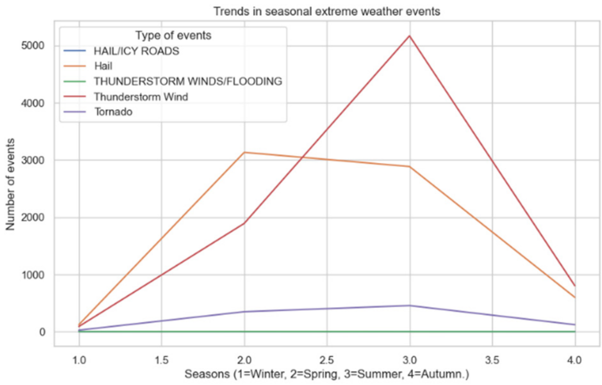

Figure 2. Trends in seasonal extreme weather events

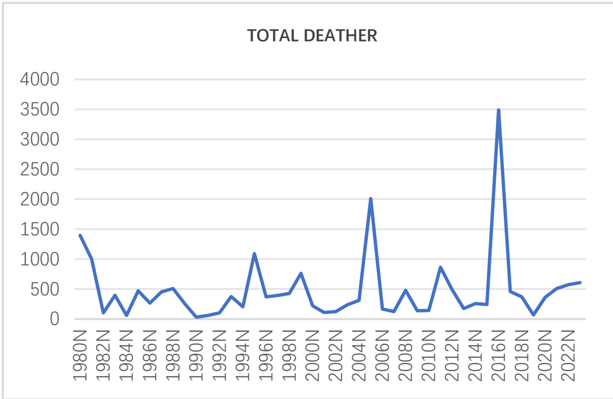

Figure 3. (US) According to the annual accident statistics

Based on these analysis steps, we were able to define the parameters of the ARIMA model, where: (p) represents the order of the autoregressive term in the model, (d) represents the number of differences that need to be performed to smooth the data, and (q) is the order of the moving average term. Therefore, the ARIMA model can be expressed by the formula ARIMA (p, d, q) as:

$$\left(1-\sum_{i=1}^{p} \varphi_i L^i\right)(1-L)^d y_t = \left(1+\sum_{i=1}^{q} \theta_i L^i\right)\epsilon_t \quad (2)$$

After successfully fitting the model, we will have the ability to predict future deaths. These predictions will reveal the level of risk we may face in the future, which will provide important decision-support information for government departments and the insurance industry. In addition, the confidence intervals attached to the forecasts show the uncertainty range of the forecasts, which is important for assessing future risks and preparing for them.

The fluctuations and peaks in Figure 3 suggest that traditional ARIMA models may not be able to fully capture the characteristics of the data. Outliers need to be dealt with or more complex models such as outlier detection or time series decomposition techniques need to be employed. Residual analysis helps to determine whether the model is extracting sufficient information, and if there is a clear pattern or autocorrelation, the model needs to be adjusted.

2.3. Analysis of ACF and PACF

In this scenario, analyzing the autocorrelation function (ACF) and partial autocorrelation function (PACF) has become the core step to understand the autocorrelation in time series data. This is because they are able to show regularities within the data, such as trends, periodicity, and other forms of correlated structure. Based on the patterns inherent in these data, the ARIMA model is designed to capture these features and apply them to the prediction process of future values.

ACF plots provide us with a view of the intrinsic correlation of data by revealing the autocorrelation of different lag periods [5]. If the ACF value goes to zero quickly after a certain point, it usually means that the time series is affected by a short-term correlation, which is especially critical when choosing the Moving Average (MA) parameter. Conversely, if the ACF exhibits a slowly declining or persistent fluctuation pattern, this may imply that the time series contains more complex correlated structures, or that the time series is non-stationary and therefore requires differential processing to reach a stationary state.

The PACF plot provides independent autocorrelation information for each lag period after excluding the effects of other lag periods. By observing the truncation phenomenon of PACF, we can determine the appropriate order of autoregressive (AR) parameters. This analysis ensures that the model accurately reflects the intrinsic dynamics of the time series, providing a solid basis for future predictions.

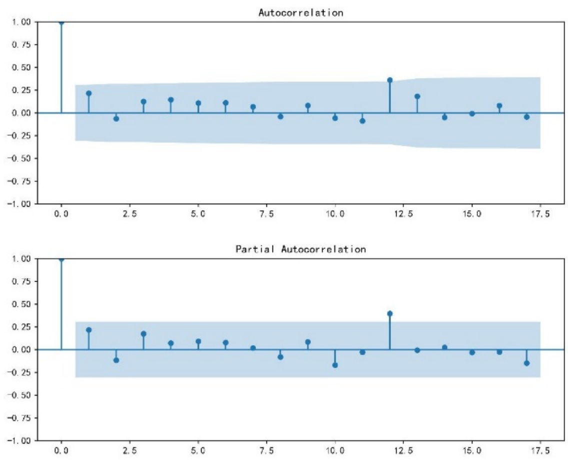

Figure 4. (US) autocorrelation and partial autocorrelation diagram

Based on the clues provided by Figure 4, we may choose an ARIMA model to predict future disaster impacts. For example, if a partial autocorrelation graph (PACF) exhibits a truncated feature after the first order, and an autocorrelation graph (ACF) exhibits decreasing fluctuations, then we might consider an ARIMA $(1, d, 0)$ or ARIMA $(1, d, 1)$ model. where the dd value is determined by the stationarity of the data.

After fitting the model, we should check the ACF and PACF of the residuals to ensure that the model has not missed any important information. If the residuals exhibit randomness, i.e., there is no significant autocorrelation or partial autocorrelation, then the model fits well.

By examining the ACF and PACF plots of the residuals, we can verify that the model is able to capture all relevant structures in the data. If the residuals are found to still exhibit some degree of autocorrelation, it may mean that the model could be further improved, such as adding higher-order AR or MA parameters or considering other models.

Once we confirm that the residuals of the model behave as random, i.e., there are no systematic patterns or trends, we can use the model to make future predictions. These forecasts can be used to guide decisions such as policy making, preparedness measures, and pricing of insurance products. By constantly monitoring and updating the model's predictions, we can adjust these decisions to respond to changes in the impact of disasters.

3. CGDAM-WRIR Model

3.1. Model Basic Framework

With the frequent occurrence of extreme weather events, especially tornadoes, which are unpredictable and destructive, they have attracted widespread attention. It is critical for communities to accurately identify and assess these risks, which involves the development of emergency response plans, the strengthening of infrastructure, and the safety precautions taken by residents. For insurers, understanding tornado risk levels in different regions can help calculate risk, adjust premiums, and develop insurance policies more precisely.

Given the frequency of extreme weather events around the world and the far-reaching impact these events can have on human societies and economies, it is important to choose an analytical tool that provides a clear picture of the level of risk in the region. This concise, descriptive statistical model provides policymakers and insurers with an intuitive map of risk by counting and visualizing the number of tornado events that occur in each state. This approach provides an important reference for predicting future risk levels and formulating corresponding countermeasures by visualizing the frequency of events in history. As shown in Figure 5.

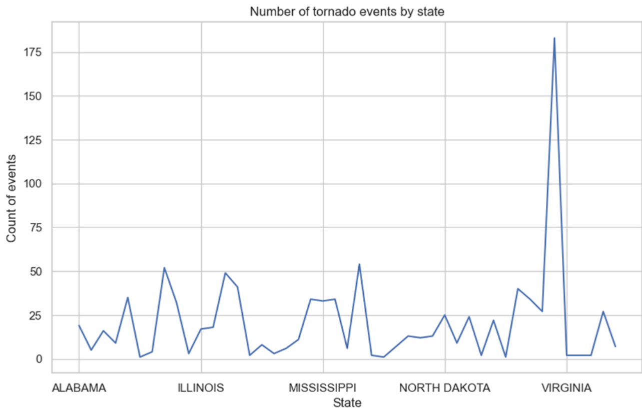

Figure 5. number of tornado events by state

3.2. Parameter Estimation and Calibration

As the impact of climate change intensifies, the frequency and intensity of extreme weather events are increasing significantly. It is critical for the insurance industry to identify areas that are more vulnerable to such events, as well as to estimate the amount of economic damage that these events may cause. This critical information can help insurers adjust insurance rates, design more accurate insurance products, and develop more effective prevention and response measures for potential disaster events.

In our Comprehensive Geospatial and Weather-Related Insurance Risk Damage Assessment Model (CGDAM-WRIR), we place a high value on the accuracy and upto-date nature of property loss data. The data pre-processing phase is a critical step in ensuring that the loss amount is correctly identified and converted into a numerical format, laying the foundation for subsequent in-depth analysis. By combining the loss data with specific geographic location (latitude and longitude information), we can pinpoint the areas that are most affected and further analyze the specific types of risks that these areas face.

This approach allows us not only to identify the areas most affected by disasters, but also to gain a deeper understanding of the risk characteristics of different regions. This provides a scientific basis for formulating more effective risk management strategies, adjusting premium policies, and designing more targeted insurance products.

In this visualization, each point on the graph represents a recorded weather event, and the color and size of the point adjusts to the corresponding property damage value. In this way, the larger the loss value of the event, the more vivid the color and size of the point it represents. This design provides a visual indication of which areas have suffered significant economic losses, which may be due to frequent weather events in the region or the extremely high impact of a single event.

3.3. The Impact Degree of Different Intensity Tornadoes on Regions

In the field of meteorological disaster management, an in-depth understanding of tornadoes of different intensities and their impact on each region is critical. This information is important for risk assessment, effective allocation of resources, development of emergency response plans, and planning for long-term infrastructure. At the same time, this data is also critical for insurers when designing insurance products and assessing the associated risks.

In today's increasingly frequent extreme weather events, using pie charts to show the distribution of tornado levels can help governments, rescue agencies, insurance companies and other relevant agencies better understand the characteristics and patterns of tornado events, so as to formulate more effective prevention, response and recovery measures [6].

In this model, we focus on the Fujita scale of tornadoes - a measure of tornado intensity, and by calculating the number of events under each criterion and converting it into percentages, we can get a clear picture of the relative frequency of tornado events under each criterion. to help us identify which types of tornado damage are most prevalent and provide recommendations and plans for different types of tornadoes to help residents and institutions prepare for and respond.

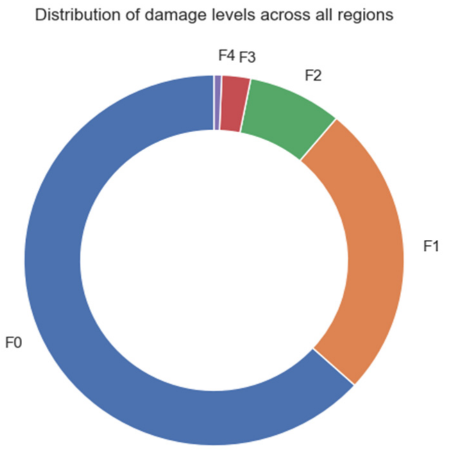

Figure 6. Distribution of damage levels across all regions

The pie chart in Figure 6 shows the distribution of different levels of tornadoes in each region, which can help us understand the level of tornado risk that different regions face. According to Fujita Scale, F0 is the weakest tornado, while F4 and F5 are the strongest. The most recorded events belong to lower-grade tornadoes (e.g., FO), while highergrade tornadoes such as F4 and F5 are rarer. This kind of information is important for optimizing resource allocation. For example, in areas that are frequently hit by highlevel tornadoes, increased construction intensity or improved early warning systems may be needed to reduce damage and safeguard public safety.

3.4. The Impact of Extreme Weather Events Caused by Climate Change on Communities

As climate change causes extreme weather events to occur more frequently and have a significant impact on communities, local governments need to optimize emergency preparedness and resource allocation based on historical data. Alabama, located in the southeastern part of the United States, can face several meteorological disasters such as tornadoes, floods, and storms. By understanding the historical data of these natural disasters, we can better understand their frequency and impact, so that we can develop sound plans, including budgeting, infrastructure planning, and public education activities, to reduce damage and improve public safety. This work is crucial because it helps ensure that communities are better able to cope and recover in the face of natural disasters.

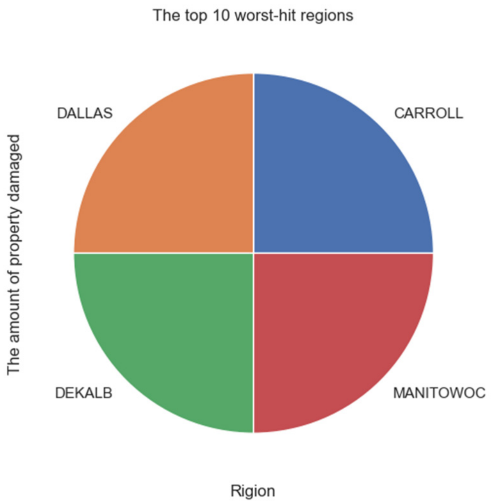

Figure 7. The top 10 worst-hit regions

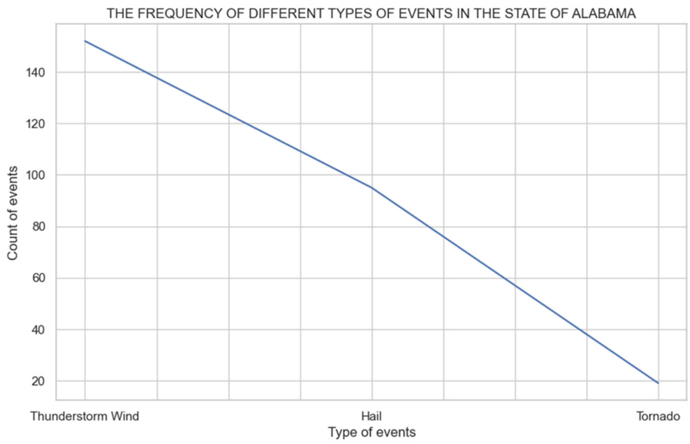

Figure 8. The frequency of different types of events in the state of ALABAMA

Firstly, filters out the data related to Alabama, then sorts it by different event types and counts the number of events of each type. By sorting by the number of events from largest to smallest, we were able to quickly identify the types of weather events that occurred most frequently within Alabama. In order to visualize the relative frequency of various events, and to facilitate comparison and interpretation, we chose a bar chart as an analysis tool.

The model leverages datasets of historical meteorological events directly, and instead of complex math, data aggregation is used to provide insights. By consolidating massive data points into a few core metrics, i.e., the frequency of various events as shown in Figure 7 and Figure 8, it provides intuitive and practical information for decision-makers.

The results of this analysis have important reference value for local government policy formulation and planning discussions. It reveals which weather events to focus on in Alabama and what specific responses may be needed. For example, if there are significantly more incidents in one type than others, then more targeted education programs or additional safeguards may be needed in urban planning. Overall, this databased approach to analysis provides a solid foundation for local governments to develop more precise and efficient policies and plans.

4. Conclusion

This study provides insights into the challenges and coping strategies of the property insurance industry in the face of increasing extreme weather events by using ARIMA and CGDAM-WRIR models. The application of ARIMA modeling helps us to assess the potential losses that may be caused by future extreme weather events, predict the mortality rate, and evaluate the impacts of tornadoes of varying intensities in different regions, which provides an important reference for risk management and underwriting strategy provides an important reference. Meanwhile, the application of the CGDAM-WRIR model reveals the significant impact of climate change on the property insurance industry, prompting insurers to re-examine their underwriting strategies, especially in high-risk areas. Future research will continue to explore more sophisticated models, address the limitations of cohort theory, and integrate data from multiple sources to develop real-time predictive models. Through continued innovation and research, we can help the property-casualty insurance industry better manage the uncertainties associated with climate change, protect the interests of property owners, and ensure the sustainability and resilience of the insurance market.

References

[1] The property and casualty insurance industry are facing unprecedented challenges as the frequency and intensity of extreme weather events triggered by climate change increases.

[2] The increased frequency and intensity of extreme weather events such as floods, hurricanes, and droughts have led to significant property damage that has a direct impact on insurers' ability to pay and risk management strategies.

[3] Kumar S V, Vanajakshi L. Short-term traffic flow prediction using seasonal ARIMA model with limited input data[J]. European Transport Research Review, 2015, 7: 1-9.

[4] Arumugam V, Natarajan V. Time Series Modeling and Forecasting Using Autoregressive Integrated Moving Average and Seasonal Autoregressive Integrated Moving Average Models[J]. Instrumentation, Mesures, Métrologies, 2023, 22(4).

[5] Boyce M S, Pitt J, Northrup J M, et al. Temporal autocorrelation functions for movement rates from global positioning system radiotelemetry data[J]. Philosophical Transactions of the Royal Society B: Biological Sciences, 2010, 365 (1550): 2213-2219.

[6] Nouri N, Devineni N, Were V, et al. Explaining the trends and variability in the United States tornado records using climate teleconnections and shifts in observational practices[J]. Scientific reports, 2021, 11(1): 1-14.