

Study On River Regulation and Emergency Management Based on SARIMA And Dam Control Models

Yihuai Zhang^{1, #}, Kenuo Li^{2, *, #}

¹ Faculty of Information Science and Engineering, Ocean University of China, Qingdao, China, 266100

² College of Food Science and Engineering, Ocean University of China, Qingdao, China, 266100

* Corresponding Author Email: 19838623691@163.com

#These authors contributed equally.

Abstract. Across the globe, with population growth and climate change, improving the accuracy and responsiveness of water resource management is crucial for ensuring water security. This study takes the establishment of an optimal water level control system for the Great Lakes in North America as its starting point. It comprehensively considers the changes in flow rates of six major rivers that affect the water level fluctuations of the Great Lakes, using SARIMA models to forecast the seasonal variations in the flow rates of these rivers. Considering the complexity of lake systems, this article introduces lag factors, sensitivity factors, and emergency control factors to establish a dam control model for the Great Lakes. After effective evaluation of the model, under this control system, the simulated water levels of the Great Lakes are close to the optimal levels and exhibit good emergency response capabilities. Applying data from the Great Lakes in 2017 to the aforementioned control system, the results show that the average difference between simulated values and optimal values is 0.125, indicating a well-regulated control system and effective enhancement of water level control efficiency for the Great Lakes.

Keywords: SARIMA; Dam Control Model; Emergency Control; Effective Evaluation.

1. Introduction

On a global scale, sustainable management and efficient utilization of water resources have become increasingly pressing challenges. With the acceleration of population growth and industrialization, the demand for water resources continues to rise. Simultaneously, extreme weather events caused by climate change, such as floods and droughts, pose higher demands on water resource management systems[1]. Therefore, enhancing the accuracy and responsiveness of water resource management is crucial for ensuring water security and disaster risk reduction. The Great Lakes are the largest group of freshwater lakes in North America which include Lake Superior, Lake Michigan-Huron, Lake St. Clair, Lake Erie, and Lake Ontario. These lakes are interconnected through a series of connecting rivers, straits, and flows between lakes. Among them, the regulation of water levels in the Great Lakes is crucially important through two major control dams (the Compensating Works of the Soo Locks at Sault Ste. and Marie and the Moses-Saunders Dam at Cornwall) which are managed by humans.

Traditional flow prediction models often struggle to accurately capture such seasonal fluctuations, whereas the Seasonal Autoregressive Integrated Moving Average (SARIMA) model is widely applied to such problems due to its robust seasonal analysis capabilities. SARIMA has been applied in various fields including atmospheric environment[2], economics[3], road traffic[4], electric power[5,6], tourism[7], aviation[8] and medicine[9]. To meet the prediction accuracy requirements of this model, this study adopts the SARIMA model to forecast water level changes in the Great Lakes of North America, aiming to achieve optimal water level control.

The data in this article is sourced from www.contest.comap.com. Firstly, based on the seasonal variations of flow rates of the six major rivers connected to the Great Lakes, this article constructs SARIMA models and utilizes them to predict the future one-year flow rate changes of each river. Then, this article establishes relationships for the water level changes of each lake, considering the

influences of normal river flows and extreme weather conditions such as heavy rainfall, and develop an intelligent dam control model to adjust dam flow rates for the next month, thereby maintaining optimal water levels for each lake. Finally, this article conducts simulation control using this model, comparing simulated, optimal, and actual water levels to evaluate the feasibility of the model.

2. Forecasting and Controlling Water Levels in the Great Lakes

Under the influence of factors such as temperature, humidity, wind intensity, tides, and evaporation, the water levels of the Great Lakes fluctuate continuously in the absence of human intervention[10]. However, due to the complexity and uncertainty of the Great Lakes, fluctuations in lake water levels within the normal range of two to three feet can potentially cause catastrophic damage. Therefore, it is necessary to make real-time predictions of the water level changes in the Great Lakes. By utilizing control systems to adjust the discharge and inflow volumes, this article maintains the optimal water levels of the Great Lakes.

The current main flow regulation facility for the Great Lakes is the Compensating Works of the Soo Locks at Sault Ste Marie and the Moses-Saunders Dam at Cornwall. Dams are the primary means of water level regulation, and to simplify the model, this article considers the flow control of dams.

2.1. Seasonal-ARIMA Model

The SARIMA model can effectively capture the trend and seasonal variations of time series. It includes the following three aspects: Auto Regressive Model, Integrated and Moving Average Model. Therefore, this article applies the ARIMA model to forecast the water level changes in the Great Lakes for some time in the future.

2.1.1 Seasonality test

The Great Lakes primarily include six major rivers: St. Mary's River, St. Clair River, Detroit River, Niagara River, Ottawa River and St. Lawrence River. According to the monthly flow data of the six major rivers from 2017 to 2021, this article constructed the time series plot as shown in Figure 1:

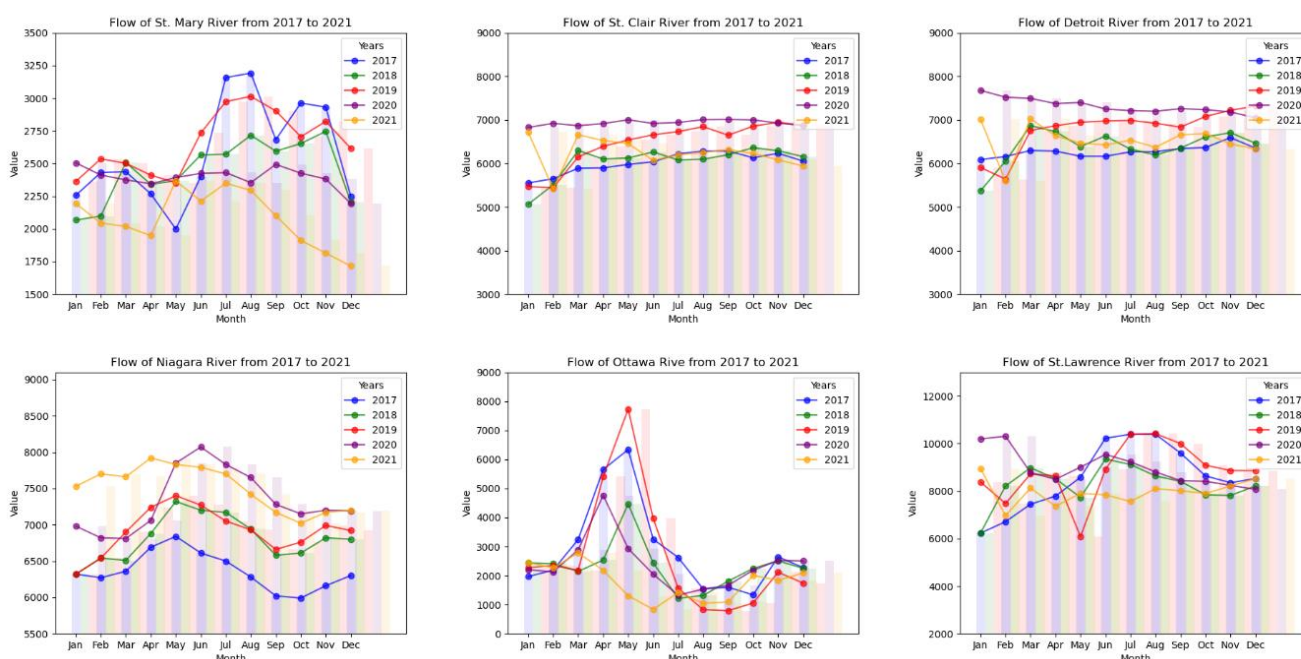


Fig.1 Periodic discharge changes of six major rivers

As depicted in Figure 1, for most rivers, there is a significant increase in flow from spring to summer, reaching its peak during the summer season. Subsequently, the flow gradually diminishes

during the autumn and winter seasons. Therefore, the flow variations of the six major rivers exhibit seasonality.

2.1.2 Stationarity Test

To utilize the ARIMA model, it is necessary to check the stationarity of the time series. This article employed the augmented Dickey-Fuller unit root test to examine the stationarity of the data. If the obtained p-value is less than 0.05, the time series is considered stationary. However, if it is found to be non-stationary, this article need to perform differencing. Based on observations, it is evident that the original series lacks stationarity and thus needs to be differenced. Considering the actual circumstances, we set the parameters as shown in Table 1.

Table 1: Determination method for p and q values

Difference order d	Cycle m	Seasonal difference order D D
2	12	0

After differencing the original sequences (as shown in the top row of Figure 1), the second differenced sequences of the six major rivers demonstrate stationarity.

2.1.3 Determining the parameters p, q

Determining the values of p and q based on the ACF and PACF is done as shown in Table 2. The second and third rows of Figure 2 illustrate the truncation patterns of the ACF and PACF after second-order differencing.

Table 2: Method of determining p and q values

Model	ACF	PACF
AR(p)	Attenuation tends to zero	truncation at p lags
MA(q)	truncation at q lags	Attenuation tends to zero
ARMA(p,q)	q	truncation at p lags

Note: Truncation indicates within the confidence interval.

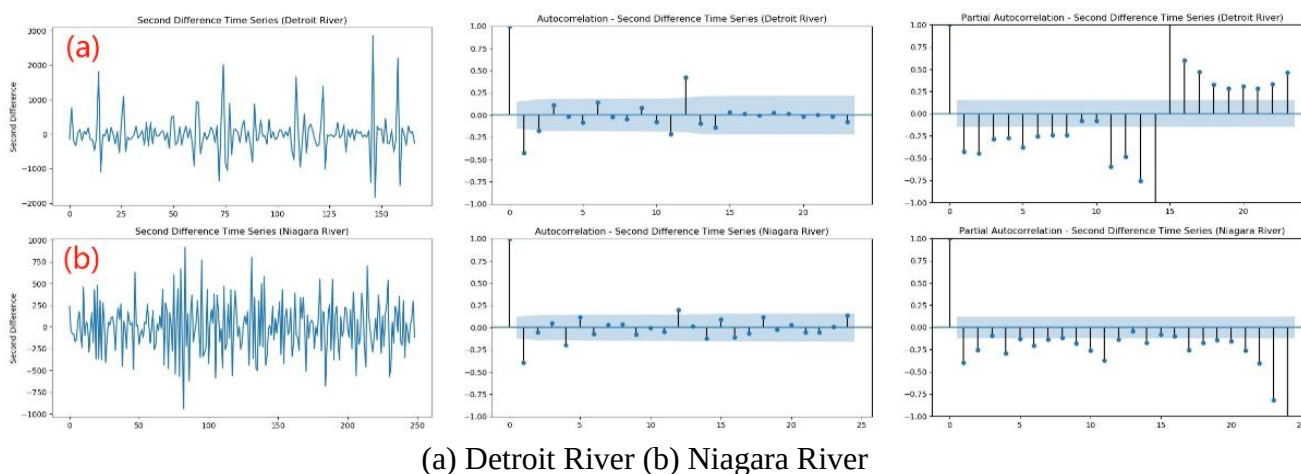


Fig. 2 Differential result series diagram, ACF, PACF of two major rivers.

According to the implementation of automatic SARIMA parameter selection, the optimal values of p, d, P, and Q for the six major rivers are summarized in Table 3:

Table 3: Determining the parameters p, q, P, and Q

	p	q	P	Q
St. Lawrence River	1	1	2	0
Niagara River	2	2	1	1
Detroit River	0	2	0	1
St. Clair River	1	2	1	0
St. Mary's River	0	0	2	0
Ottawa River	1	0	2	0

Note: p value is Autoregressive Order; q value is Moving Average Order; P value is Seasonal Autoregressive Order; Q value is Seasonal Moving Average Order.

2.1.4 Forecast

Using the fitted SARIMA models for each river obtained earlier, the future one-year flows for each river are predicted as shown in Figure 3:

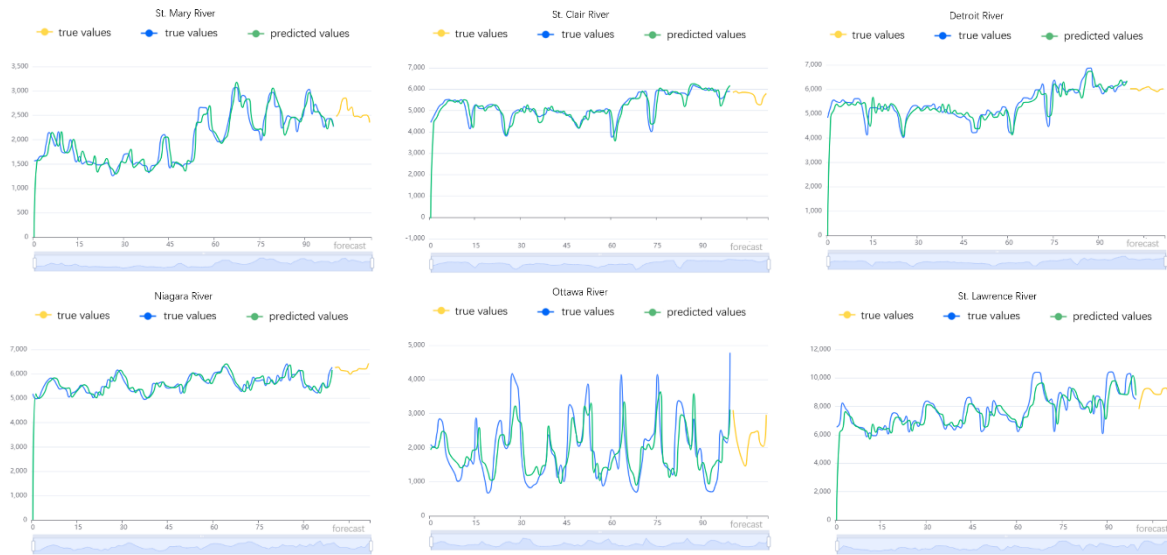


Fig.3 Flow predictions for the six major rivers

2.2. The Dam Control Model

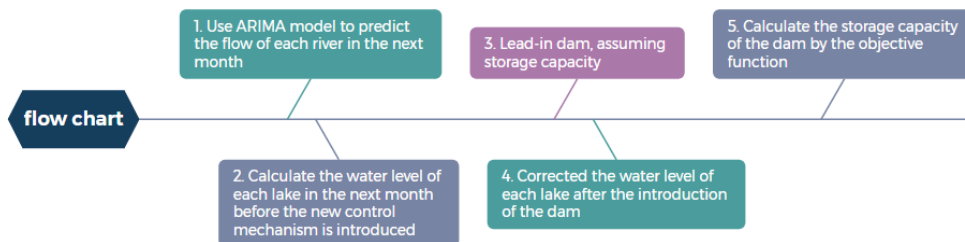


Fig.4 Flowchart of 2.2 The Dam Control model

The flowchart of the Dam Control model is shown in Figure 4. In this model, this article use N1 to represent Lake Superior, N2 to denote Lake Michi-gan-Huron, N3 to represent Lake St. Clair, N4 to denote Lake Erie, and N5 to represent Lake Ontario. The Compensating Works of the Soo Locks (*dam1*) is located between N1 and N2, and The Moses-Saunders Dam (*dam2*) is located downstream of N5. Assuming that the change in water storage volume for *dam1* and *dam2* until next month is represented by $\Delta V1$ and $\Delta V2$ respectively. It can be inferred that:

$$p_{i-no_dam} = (Inflow_{(t)} - Outflow_{(t)} + NBS_{(t)} + \varepsilon_{else}) \cdot \beta_{transfer} \quad (1)$$

Where, p_{i-no_dam} represents the predicted water level of Lake Ni without the implementation of any dam control mechanisms for the next month; $Inflow_{(t)}$ is an inflow of the lake; $Outflow_{(t)}$ is an outflow of the lake; $NBS_{(t)}$ is Net Basin Supply of the Great Lakes; ϵ_{else} is the elasticity coefficient; $\beta_{transfer}$ is the conversion coefficient.

The updated predicted water levels for each lake after implementing dam control mechanisms are as follows:

For Lake N1, when there is a change in the storage volume of dam1 by ΔV_1 :

$$p_{1-dam} = p_{1-no_dam} + \frac{\Delta V_1}{S_1} \quad (2)$$

Where p_{i-dam} represents the predicted water level of Lake Ni in the next month after the introduction of the dam control mechanism. S_i represents the cross-sectional area of Lake Ni.

For Lake N2, when there is a change in the storage volume ΔV_1 of dam1, the inflow rate of Lake N2 changes. Therefore, after the introduction of the control mechanism, the predicted water level of Lake N2 in the next month is:

$$p_{2-dam} = p_{2-no_dam} + \frac{\Delta V_1}{S_2} \quad (3)$$

For Lake N3, the regulation of dam1 will affect the inflow rate of N3 after a certain period. This process has a certain lag, so this article introduce a lag factor to represent the lagging water level. After the introduction of the control mechanism, the predicted water level of Lake N3 in the next month is as follows:

$$p_{3-dam} = p_{3-no_dam} + \sigma_3 \quad (4)$$

Where σ_i represents the lagging water level of the lake caused by the lag phenomenon of dam regulation.

For Lake N4, the regulation of dam1 will not affect the inflow rate of N4 in the short term. Moreover, due to the significant difference in elevation between N4 and N5 lakes, the regulation of dam2 can be considered to have no impact on the water level of N4. Therefore, the updated water level of Lake N4 is:

$$p_{4-dam} = p_{4-no_dam} + \sigma_4 \quad (5)$$

For Lake N5, the storage volume ΔV_2 of dam2 directly affects the water level of N5. The Montreal area downstream of Lake Ontario (N5) is simultaneously affected by the outflow from Lake Ontario and the Ottawa River. Improper regulation of dam2 could potentially harm the Montreal area.^[8] Additionally, from Figure 3, it can be observed that the outflow from the Ottawa River increases significantly during the summer. Therefore, to ensure that the water level in the Montreal area remains within a safe range, this article need to introduce a sensitivity factor, denoted as ($\beta_{sensitivity}$), to dam2 that accounts for the variation in water volume in the Ottawa River. To respond in real-time to the changes in the flow rate of the Ottawa River, the updated water level of Lake N5 would be:

$$p_{5-dam} = p_{5-no_dam} + \frac{\Delta V_2}{S_5} + \beta_{sensitivity} \quad (6)$$

It can be observed that the water level changes in the Great Lakes are unique for each lake. The conditions that need to be satisfied when regulating dam1 and dam2 are:

$$\min \sum_{i=1}^5 |p_{i-dam} - b_i| \quad (7)$$

Where, p_{i-dam} represents the predicted water level of Lake Ni in the next month after introducing the control mechanism, and represents the optimal water level of Lake Ni in the next month.

Based on the above formula, this article can calculate the changes in the storage volumes of dam1 ($\Delta V1$) and dam2 ($\Delta V2$). When $\Delta V > 0$, it means that the dam needs to increase its storage volume in the next month, with an increase of V in storage volume. When $\Delta V < 0$, it means that the dam needs to increase its drainage capacity in the next month, with an increase of V in drainage capacity. Moreover, since the water level changes in lakes are directly regulated by river flow, and the flow of rivers takes some time, there is a lag factor involved. This lag factor has been incorporated into the aforementioned system dynamics model.

Additionally, this article regulate the dam every month. However, for exceptional circumstances such as a sudden increase in lake water level due to heavy precipitation or increased runoff, this article need to immediately regulate the dam. Therefore, we need to set a safety threshold, κ . In emergencies, when the lake water level exceeds the safety threshold κ , it requires immediate regulation. This article control the outflow rate ($Outflow_{(t)}$) of the lake by regulating the discharge (g_{dam}) of the dam in emergencies. To increase the outflow rate of the lake in emergencies, this article introduces an additional emergency control factor, (E_c), to regulate outflow rate of the lake:

$$Outflow'_{(t)} = Outflow_{(t)} \cdot E_c, x > \kappa \quad (8)$$

Where $Outflow'_{(t)}$ represents the outflow rate of the lake in emergencies; $Outflow_{(t)}$ represents the normal outflow rate of the lake; E_c represents the emergency control factor.; x represents the water level of the lake; κ represents the water level safety threshold.

3. Evaluation of Regulation Effectiveness

In the previous section, this article utilized the S-ARIMA model to regulate the future one-month reservoir volume, thereby controlling the lake water level. Subsequently, simulation control was applied to the water levels of the Great Lakes in 2017. The simulated water levels (Simulation) were compared with the respective optimal water levels (Best) and actual water levels (Actual) for that year. The results are shown in Figure 5:

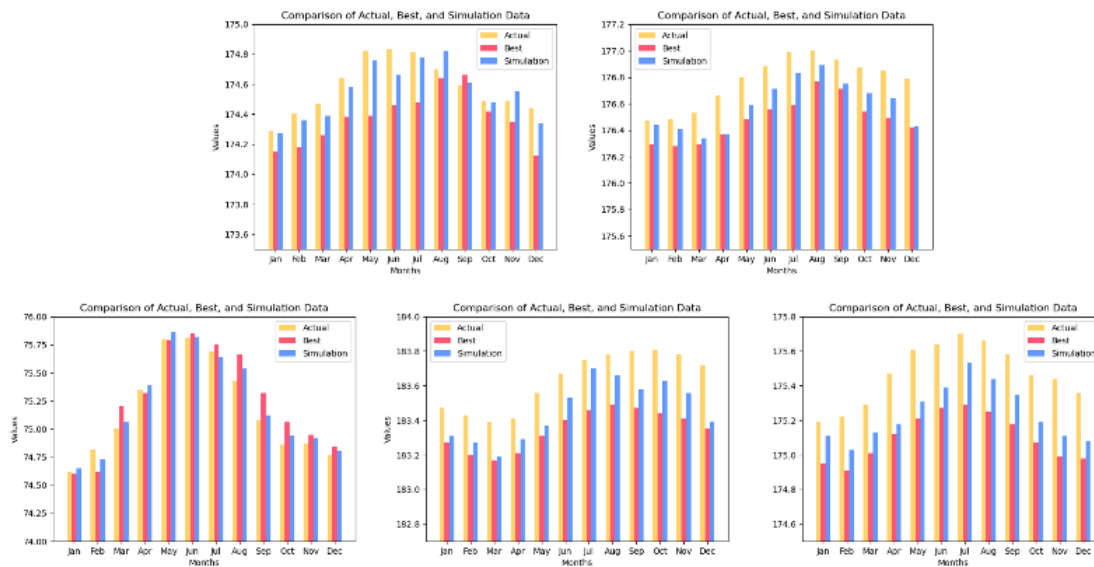


Fig.5 Comparison of actual, optimal, and simulated water levels

As shown in Figure 5, The optimal water level is generally slightly lower than the actual water level, and the simulated water level after dam regulation falls between the actual water level and the

optimal water level. Thus, this article can conclude that the water level regulation model effectively controls the lake water level, bringing it closer to the optimal value.

At the same time, to quantify the model's superiority indicators, this article use the following methods for analysis:

$$\lambda = \sum_{i=1}^5 \sum_{j=1}^{12} |a_{ij} - b_{ij}| \quad (9)$$

$$\eta = \sum_{i=1}^5 \sum_{j=1}^{12} |p_{ij} - b_{ij}| \quad (10)$$

Where, a_{ij} is the actual water level of the i th lake in the j th month of the year 2017; b_{ij} is the best water level of the i th lake in the j th month of the year 2017; p_{ij} is the predicted water level of the i th lake in the j th month of the year 2017. λ is the overall difference between the actual water level and the optimal water level without using control methods; η is the overall difference between the actual water level and the optimal water level with using control methods.

It can be observed that when the actual water level deviates more from the optimal water level, λ increases; and when the simulated water level gets closer to the optimal water level, η decreases. Applying the data of the Great Lakes in 2017 to the above method, this article obtain $\lambda=15.24$, $\eta=7.50$ that the average difference between the simulated value and the optimal value is 0.125. Therefore, this article can conclude that the water level control model has good regulatory effects, predicting water levels close to the optimal values, and maximizing the overall benefits for various stakeholders.

4. Conclusions

The SARIMA model, through analysis of the flow rates of the six major rivers, accurately captures seasonal variations. The Dam Control model takes into account the complexity of the lake system. By introducing lag factors, sensitivity factors, and emergency control factors, it achieves precise regulation of dam storage capacity, enabling it to have a good emergency response capability in urgent situations. By evaluating the effectiveness of the model, this article can conclude that, the simulated water level after dam regulation falls between the actual water level and the optimal water level and the water level regulation model effectively controls the lake water level, bringing it closer to the optimal value. At the same time, Applying the data of the Great Lakes in 2017 to the above method, this article obtains $\lambda=15.24$, $\eta=7.50$, that the average difference between the simulated value and the optimal value is 0.125. The water level control model has good regulatory effects, predicting water levels close to the optimal values, and maximizing the overall benefits for various stakeholders.

References

- [1] Bunting E L, Theuerkauf E J. What is a coastal hazard? Perceptions of critical coastal hazards amongst decision makers in communities across the Great Lakes[J]. Ocean and Coastal Management, Elsevier Ltd, 2024, 251.
- [2] Cui T, Shi Y, Lv B, et al. Federated learning with SARIMA-based clustering for carbon emission prediction[J]. Journal of Cleaner Production, Elsevier Ltd, 2023, 426.
- [3] Cheng J, Tiwari S, Khaled D, et al. Forecasting Bitcoin prices using artificial intelligence: Combination of ML, SARIMA, and Facebook Prophet models[J]. Technological Forecasting and Social Change, Elsevier Inc., 2024, 198.
- [4] Agyemang E F, Mensah J A, Ocran E, et al. Time series-based road traffic accidents forecasting via SARIMA and Facebook Prophet model with potential changepoints[J]. Heliyon, Elsevier Ltd, 2023, 9(12).
- [5] Ren F, Tian C, Zhang G, et al. A hybrid method for power demand prediction of electric vehicles based on SARIMA and deep learning with integration of periodic features[J]. Energy, Elsevier Ltd, 2022, 250.

- [6] Bilgili M, Pinar E. Gross electricity consumption forecasting using LSTM and SARIMA approaches: A case study of Türkiye[J]. *Energy*, Elsevier Ltd, 2023, 284.
- [7] He K, Ji L, Wu C W D, et al. Using SARIMA–CNN–LSTM approach to forecast daily tourism demand[J]. *Journal of Hospitality and Tourism Management*, Elsevier Ltd, 2021, 49: 25–33.
- [8] Carmona-Benítez R B, Nieto M R. SARIMA damp trend grey forecasting model for airline industry[J]. *Journal of Air Transport Management*, Elsevier Ltd, 2020, 82.
- [9] Lau K, Dorigatti I, Miraldo M, et al. SARIMA-modelled greater severity and mortality during the 2010/11 post-pandemic influenza season compared to the 2009 H1N1 pandemic in English hospitals[J]. *International Journal of Infectious Diseases*, Elsevier B.V., 2021, 105: 161–171.
- [10] Huang P, Sauquet E, Vidal J P, et al. Vulnerability of water resource management to climate change: Application to a Pyrenean valley[J]. *Journal of Hydrology: Regional Studies*, Elsevier B.V., 2022, 44.