International Scientific Technical and Economic Research (ISSN: 2959-1309)
Vol. 4, No. 1, (March 2026), pp. 1-22
Journal homepage: https://www.istaer.online

RESEARCH ARTICLE DOI: 10.71451/ISTAER2601

Forecasting Short-Term Export Volumes with Hybrid Models Integrating SARIMA with Attention-Based LSTM

Wenhao Wang*, Siyi Shen, Ye Yuan

Department of Trade and Logistics, Daegu Catholic University, Gyeongsan, Daegu, Republic of Korea

Abstract: The short-term export volume forecast is of great significance for international trade decision-making and macroeconomic regulation, but the time series of export volume usually contains significant seasonal, trend and nonlinear fluctuation characteristics at the same time, so it is difficult to obtain ideal results with a single forecast model. In order to improve the prediction accuracy and stability, this paper proposes a hybrid prediction method combining seasonal autoregressive moving average model (SARIMA) and attention based memory network (attention based LSTM). Firstly, SARIMA model is used to describe the linear structure and seasonal components in the export volume series, and its prediction residual is modeled; Then, the attention LSTM is used to learn the nonlinear dynamic characteristics in the residual sequence, and finally the prediction results are obtained by additive fusion. The experimental results based on the monthly export volume data of Poland show that compared with the traditional SARIMA, LSTM and other comparative models, the MAE, RMSE and MAPE of the SARIMA-Attention-LSTM (Proposed) on the test set are reduced, on average, by about 20%–35%, respectively. The prediction residual fluctuation converges significantly, and shows better stability and generalization ability in repeated experiments. The results show that the effective integration of statistical model and deep learning model can significantly improve the short-term export forecasting performance, and provide a feasible and efficient solution for the prediction of complex economic time series.

Keywords: Short term export volume forecast; Time series analysis; SARIMA; Attention mechanism; Mixed forecasting model

How to Cite: Wang, W., Shen, S., & Yuan, Y. (2026). Forecasting short-term export volumes with hybrid models integrating SARIMA with attention-based LSTM. *International Scientific Technical and Economic Research, 4*(1), 1–22. **https://doi.org/10.71451/ISTAER2601**

Article history: Received: 20 Nov 2025; **Revised:** Dec 12 2025; **Accepted:** 28 Dec 2025; **Published:** 02 Jan 2026

1 INTRODUCTION

Under the background of highly interconnected global economy, export trade, as an important indicator to measure a country's economic activity and international competitiveness, has a direct impact on macroeconomic operation, industrial layout adjustment and policy-making direction. Especially on the short-term scale, the fluctuation of export volume often has

***Corresponding author:** Wenhao Wang, Department of Trade and Logistics, Daegu Catholic University, Gyeongsan, Daegu, Republic of Korea. Email: ww216886@cu.ac.kr

a significant impact on exchange rate regulation, inventory management, port logistics scheduling and enterprise production planning. Therefore, it is of great practical significance to build a high-precision short-term export forecasting model for improving the foresight of macro decision-making and the stability of enterprise operation.

From the perspective of methodology, time series forecasting is always the core technology path of export forecasting research [1]. Traditional statistical models, such as ARIMA and its seasonally extended form SARIMA, have strong theoretical basis and interpretability in depicting linear trends and periodic structures, and have been widely used in the field of economy and trade [2]. However, with the increasing complexity of the global trade environment, the export volume series gradually shows the characteristics of nonlinear enhancement, frequent sudden fluctuations and obvious structural changes. The single statistical model shows limited prediction accuracy and insufficient adaptability when dealing with complex dynamic models [3],[4]. At the same time, deep learning models, especially LSTM based on recurrent neural networks, show good performance in capturing nonlinear relationships and long-term dependence, but their explicit ability to depict seasonal structures is limited, and the interpretability of the model is relatively weak.

In this context, how to effectively integrate the advantages of traditional statistical models and deep learning models has become a key research issue to improve the performance of short-term export forecasting. Statistical models can provide stable structural constraints for the prediction process, while deep learning models are good at automatically learning complex nonlinear features from data [5],[6]. Through a reasonable model fusion strategy, it is expected to maintain the ability of trend and seasonal modeling, and further mine the dynamic change rules hidden in the data, so as to achieve more accurate and robust prediction results [7],[8].

Based on the above research motivation, this paper proposes a hybrid prediction method combining SARIMA model and LSTM network with attention mechanism. Firstly, the SARIMA model is used to model the linear and seasonal components in the export volume series, and then the remaining nonlinear information is learned through the attention based LSTM, and the results of the two parts are effectively fused in the prediction stage. Compared with the existing research, the innovation of this paper is mainly reflected in: the structural complementarity between models is achieved by residual modeling; Attention mechanism is introduced to enhance the model's ability to focus on key historical information; The hybrid model is systematically applied to the short-term export volume prediction scenario to verify its advantages in prediction accuracy and stability.

The rest of this paper is structured as follows: the Section 2 reviews the related research work, and analyzes the research progress and shortcomings of the existing methods; The Section 3 introduces the hybrid forecasting model and its methodological basis in detail; The Section 4 describes the data source and preprocessing process; The Section 5 describes the experimental design and evaluation index; In Section 6, the experimental results are analyzed and discussed; Finally, in Section 7, the conclusion of the full text is summarized, and the future research direction is prospected.

2. LITERATURE REVIEW

With the continuous expansion of the scale of international trade, export volume prediction has gradually become a cross research hotspot in many fields, such as economics, management science and information technology [9],[10]. Early studies were mostly based on classical econometric theory, using time series model to model export scale, trade volume and related macroeconomic indicators [11],[12]. This kind of research emphasizes the stationarity assumption of data, reveals the internal evolution law of economic variables through trend and cycle analysis, and has achieved some results in the medium and short-term prediction. However, with the complexity and uncertainty of the trade environment increasing, the export volume series gradually shows the characteristics of expanding volatility and frequent structural changes, prompting researchers to continue to explore more adaptive prediction methods.

Among many traditional models, SARIMA model is widely used in the field of short-term economic forecasting because it can explicitly describe the seasonal structure [13],[14],[15]. Relevant studies show that SARIMA has a good fitting ability in processing export volume data with stable periodic characteristics, especially for monthly or quarterly statistical data [16]. However, the model is still linear in nature, and its ability to describe nonlinear relationships and sudden fluctuations is limited [17]. When there are obvious structural changes or external shocks in the data, the prediction error of the model tends to increase significantly. In addition, SARIMA model is sensitive to parameter setting and differential order, and the model selection process is subjective in complex scenes, which limits the further improvement of its prediction performance to a certain extent.

In recent years, deep learning methods have made significant progress in the field of time series prediction. Among them, the long-term and short-term memory network (LSTM) is widely used in financial market forecasting, energy load forecasting and macroeconomic index analysis because it can effectively alleviate the gradient disappearance problem in the traditional circular neural network [18],[19]. Relevant studies show that LSTM has obvious advantages in capturing the characteristics of nonlinear relationship and long-term dependence [20]. However, when processing long time series, the standard LSTM may still not allocate enough key information, resulting in limited prediction accuracy. Therefore, the attention mechanism is introduced into the time series prediction model, which can dynamically adjust the importance weights of different time steps to enhance the attention ability of the model to the key historical information, so as to further improve the prediction effect. However, deep learning models usually rely on a large number of data for training, and the internal mechanism of the model is not easy to explain, so there are still some obstacles in the application of economic forecasting scenarios.

In order to make up for the deficiency of single model, hybrid prediction model has gradually become an important development direction of time series prediction research [21]. This kind of model usually realizes the collaborative modeling of linear structure and nonlinear characteristics through the combination of statistical model and machine learning model. Previous studies have verified the advantages of hybrid model in prediction accuracy and robustness in the fields of financial time series and energy demand forecasting. However, the existing hybrid methods still have room for improvement in the model fusion strategy. For example, some studies only simply superimpose the prediction results, and do not fully mine the structural complementarity between models; At the same time, when the deep learning model is introduced, the dynamic modeling of time-dependent weights is still insufficient.

Based on the existing research, it can be found that although statistical models, deep learning models and their hybrid forms have achieved some results in time series forecasting, under the specific scenario of short-term export volume forecasting, how to effectively improve the modeling ability of nonlinear fluctuations while maintaining the ability to depict seasonality and trend remains to be further studied. Based on this, from the perspective of residual modeling, this paper organically integrates SARIMA model and LSTM network with attention mechanism, striving to improve the prediction accuracy and enhance the stability of the model, providing a more practical solution for short-term export volume prediction.

3. METHODOLOGY AND MODEL BUILDING

This study proposes a hybrid forecasting method combining statistical time series model and deep learning model for the high-precision prediction of short-term export volume. The overall method uses SARIMA model to depict the linear structure and seasonal law in the export volume series, introduces the attention mechanism of LSTM network to learn the nonlinear dynamic characteristics, and realizes the effective combination of complementary advantages through residual modeling and result fusion.

3.1 Overall research framework and technical route

The time series of export volume usually contain the characteristics of trend, periodicity, seasonality and irregular fluctuation at the same time, and a single model is often difficult to fully describe its complex structure. Based on this, this paper adopts the overall technical route of "decomposition modeling fusion". Firstly, the SARIMA model is used to model the linear and seasonal of the original series, and the basic prediction results are obtained; Then, the prediction residuals of SARIMA model are calculated, and the residual sequence is input into the attention based LSTM model to learn the potential nonlinear mode; Finally, the two parts of the prediction results are superimposed to obtain the final prediction value.

The overall framework is shown in Figure 1, which reflects the sequential relationship and information flow between models.

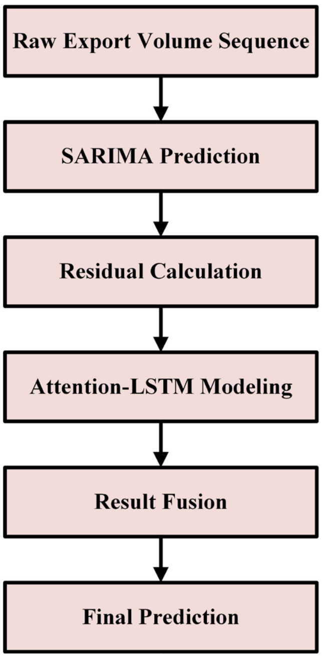

Figure 1. Overall framework of hybrid forecasting model

Let the original export volume time series be:

$$\{y_t\}, t = 1,2, \dots, T \tag{1}$$

SARIMA model predicts it and obtains the linear predictive value $\hat{y}_t^{(S)}$. The corresponding residual sequence is defined as:

$$e_t = y_t - \hat{y}_t^{(S)} \tag{2}$$

The residual sequence e_t is regarded as the time series containing the main nonlinear information and is used as the input of the deep learning model.

3.2 Theoretical basis and modeling process of SARIMA model

SARIMA (seasonal autoregressive integrated moving average) model is an extension of ARIMA model on seasonal time series, which is suitable for export volume data with significant periodic fluctuations [22]. SARIMA model is usually expressed as SARIMA$(p, d, q)(P, D, Q)_s$, where s is the length of seasonal cycle. Its general form can be written as:

$$\Phi_P(B^s)\phi_p(B)(1 - B)^d(1 - B^s)^D y_t = \Theta_Q(B^s)\theta_q(B)\varepsilon_t \tag{3}$$

B is the lag operator, $\phi_p(B)$ and $\theta_q(B)$ represent the non seasonal autoregressive and moving average polynomials respectively, $\Phi_P(B^s)$ and $\Theta_Q(B^s)$ are the seasonal polynomials, ε_t is the white noise sequence.

In the actual modeling process, the difference orders d and D are determined by unit root test and seasonal test, and then the model order is determined by combining the autocorrelation function (ACF) and partial autocorrelation function (PACF) graph. The model parameters are usually solved by maximum likelihood estimation. The prediction results of SARIMA model mainly reflect the linear trend and periodic structure in the export volume series.

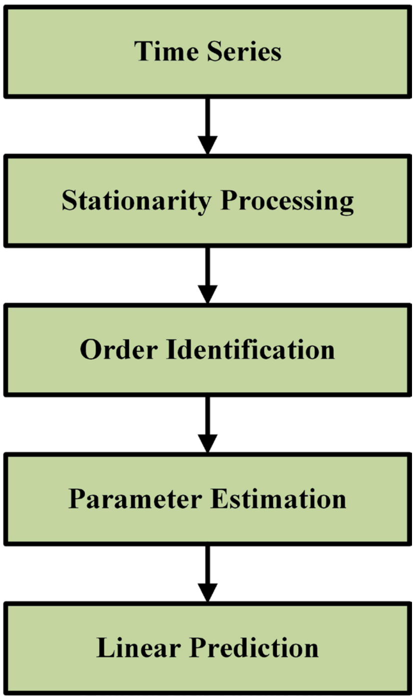

Figure 2. SARIMA modeling and prediction flow chart

3.3 Attention based LSTM model structure and working mechanism

In order to effectively capture the nonlinearity and long dependence in the export volume residual series, this paper introduces the LSTM network with attention mechanism. LSTM alleviates the gradient disappearance problem in traditional RNN through gating structure, and its core calculation process can be expressed as follows:

$$\begin{aligned} f_t &= \sigma\left(W_f[h_{t-1}, x_t] + b_f\right) \\ i_t &= \sigma(W_i[h_{t-1}, x_t] + b_i) \\ \tilde{c}_t &= \tanh(W_c[h_{t-1}, x_t] + b_c) \\ c_t &= f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \\ o_t &= \sigma(W_o[h_{t-1}, x_t] + b_o) \\ h_t &= o_t \odot \tanh(c_t) \end{aligned} \tag{4}$$

Where, f_t, i_t and o_t represent the forgetting gate, input gate and output gate respectively, and h_t is hidden.

On this basis, the attention mechanism is introduced to strengthen the attention ability of the model to the key time steps. The calculation process of attention weight is as follows:

$$\alpha_t = \frac{\exp(u_t)}{\sum_{k=1}^{T} \exp(u_k)}, u_t = v^\top \tanh(W_a h_t + b_a) \tag{5}$$

Where h_t represents the hidden state at time step t, W_a is the weight matrix, b_a is the

Wang et al., ISTAER. 2601., 02 Jan 2026 https://doi.org/10.71451/ISTAER2601

bias term, and $v^{\top}$ is the output mapping vector. Then, u_t is normalized using the softmax function to obtain the attention weights α_t for each time step. Finally, the context vector c is obtained by weighted summation of all hidden states h_t based on the attention weights:

$$c = \sum_{t=1}^{T} \alpha_t h_t \tag{6}$$

This mechanism enables the model to weight the importance of the prediction results according to different time steps, so as to improve the modeling ability of complex dynamic changes.

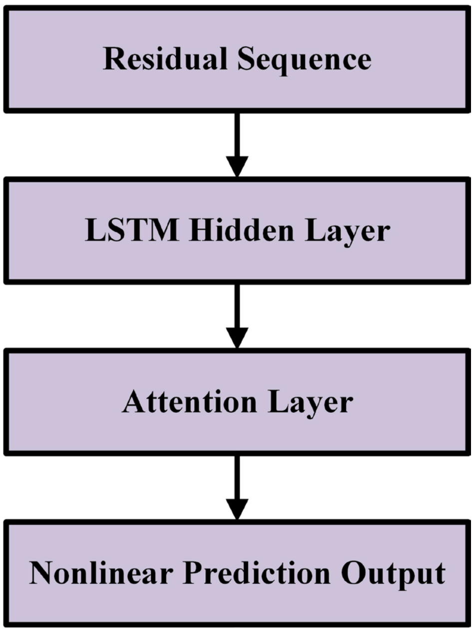

Figure 3. Attention based LSTM network structure

3.4 Hybrid modeling idea of SARIMA and attention LSTM

SARIMA model has good interpretability in depicting seasonal and linear trends, while attention LSTM has significant advantages in dealing with nonlinear patterns and long-term dependence. The combination of the two models can achieve structural complementarity and reduce the prediction bias of single model.

The core idea of the hybrid model is to decompose the original time series into linear and nonlinear parts

$$y_t = L_t + N_t \tag{7}$$

Where L_t represents the linear component captured by SARIMA, and N_t represents the nonlinear component of attention LSTM learning. The accuracy and stability of short-term export volume prediction can be effectively improved by modeling separately and combining in the prediction stage.

3.5 Model fusion strategy and prediction process design

The final predicted value is obtained through additive fusion:

$$\hat{y}_t = \hat{y}_t^{(S)} + \hat{e}_t^{(L)} \tag{8}$$

Where $\hat{y}_t^{(S)}$ is the output of SARIMA model, and $\hat{e}_t^{(L)}$ is the prediction result of attention LSTM for residual series.

Wang et al., ISTAER. 2601., 02 Jan 2026 https://doi.org/10.71451/ISTAER2601

The overall prediction process is shown in Figure 4, which clearly shows the complete steps from data input to final prediction output.

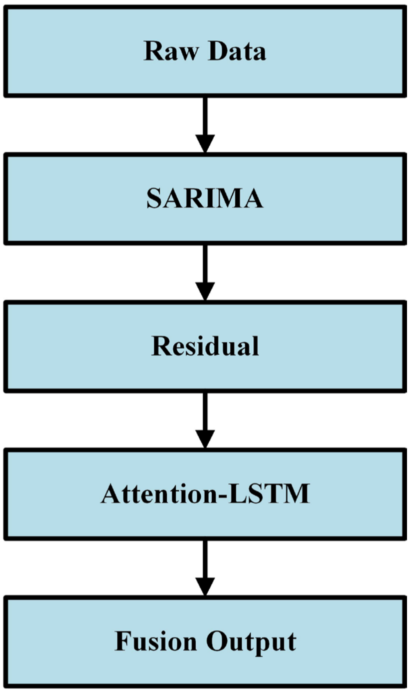

Figure 4. SARIMA – attention LSTM mixed prediction flow chart

The fusion strategy not only maintains the simplicity of the model structure, but also effectively enhances the ability to describe the characteristics of complex export volume fluctuations, providing a solid methodological basis for subsequent experimental verification.

4. DATA DESCRIPTION AND PREPROCESSING

This study takes export volume time series data as the modeling object, and the data quality and preprocessing process directly affect the performance of the mixed forecasting model. This section describes the data sources, statistical characteristics, cleaning processing and data division strategies, laying the foundation for subsequent model training and experimental analysis.

4.1 Export volume data source and time range description

The data used in the study is the monthly total commodity export data of Poland, covering a total of 120 observation samples from January 2015 to December 2024. The export volume is expressed in the standard unit of measurement (US $100 million). The data has the characteristics of strong continuity and stable cycle, which is suitable for short-term forecasting research.

Table 1 shows the statistics of some original export volume data to illustrate the time structure and magnitude distribution of the data.

Table 1. Time series of partial original export volume

Time	Export volume
2015-01	182.4
2015-02	175.9

Wang et al., ISTAER. 2601., 02 Jan 2026 https://doi.org/10.71451/ISTAER2601

2015-03	189.7
2015-04	193.2
2015-05	201.5
2015-06	198.6
2015-07	205.1
2015-08	202.8
2015-09	196.3
2015-10	199.4

From the overall time range, the data cover multiple complete annual cycles, providing sufficient samples for seasonal modeling and deep learning training.

4.2 Data feature analysis

Before modeling, it is necessary to analyze the statistical characteristics of export volume time series. Through descriptive statistics of the whole sample, we can preliminarily identify its trend, seasonality and fluctuation characteristics.

Let the time series of export volume be $\{y_t\}$, and its mean and variance are respectively defined as:

$$\mu = \frac{1}{T}\sum_{t=1}^{T} y_t \,, \sigma^2 = \frac{1}{T}\sum_{t=1}^{T} (y_t - \mu)^2 \tag{9}$$

Table 2 summarizes the annual statistical characteristics of export volume in different years.

Table 2. Analysis of annual statistical characteristics of export volume

Year	Mean	Maximum	Minimum	Standard deviation
2015	195.6	205.1	175.9	9.4
2016	198.3	210.4	182.7	8.9
2017	203.7	218.9	190.2	9.7
2018	209.5	225.3	196.8	10.1
2019	214.2	232.1	201.5	10.8
2020	208.6	228.7	189.4	12.6
2021	221.9	245.8	205.7	11.9
2022	229.4	258.3	210.2	13.2
2023	235.8	265.1	215.4	14.0

2024	242.3	272.6	221.9	14.5

The statistical results show that the overall export volume shows a long-term trend of slow rise, and the annual internal fluctuation is obvious. Further, through autocorrelation analysis, it is found that there is a significant peak near the lag period of 12, indicating that the series has a stable annual seasonal structure, which provides a basis for the subsequent use of SARIMA model.

4.3 Data cleaning and outlier handling methods

As macroeconomic data may be affected by policy adjustments, emergencies and other factors in the process of statistics or collection, there may be missing values or abnormal fluctuations in the original series. In order to reduce noise interference, the data were systematically cleaned.

Outliers are identified using a statistical threshold-based method. First, standardized residuals are calculated to measure the degree to which each observation deviates from the overall distribution. The standardized residual is defined as the difference between the current observation y_t and the series mean μ, divided by the series standard deviation σ.

$$z_t = \frac{y_t - \mu}{\sigma} \tag{10}$$

This standardization process ensures that the residuals follow a distribution with a mean of 0 and a standard deviation of 1, facilitating the setting of a uniform threshold. When identifying outliers, a threshold of $\mid z_t \mid > 3$ is set. If the absolute value of the standardized residual at a certain moment is greater than 3, it indicates that the observation point deviates significantly from the normal range and is identified as an outlier. For identified outliers, a local linear interpolation method is used for correction. The correction method involves linear estimation using the observation values y_{t-1} and y_{t+1} of the two adjacent points of the outlier, with the following formula:

$$y_t^* = \frac{y_{t-1} + y_{t+1}}{2} \tag{11}$$

Here, y_t^* represents the corrected value. This method assumes that the time series changes approximately linearly locally, and thus uses the arithmetic mean of two adjacent points to smoothly replace outliers.

Table 3 shows the comparison results of some abnormal values before and after processing.

Table 3. Abnormal value identification and correction

Time	Original value	Z value	Abnormal	Corrected value
2019-02	176.4	-3.21	yes	198.1
2020-04	184.7	-2.98	no	184.7
2020-05	171.2	-3.45	yes	190.6
2021-01	205.7	-0.82	no	205.7
2021-02	212.3	-0.34	no	212.3
2022-03	268.4	3.18	yes	238.7

2022-04	210.2	-1.05	no	210.2
2023-01	271.6	3.32	yes	245.9

Through the above processing, the time series not only maintains the overall trend and seasonal characteristics, but also effectively weakens the adverse impact of extreme values on model training.

4.4 Data standardization and division principle of training/test set

Before inputting the data into the attention based LSTM model, the sequence needs to be standardized to avoid the instability of gradient update caused by dimensional differences [23],[24]. This paper uses the Min–Max normalization method, which is defined as follows:

$$y_t^{(norm)} = \frac{y_t - y_{\min}}{y_{\max} - y_{\min}} \tag{12}$$

Where y_t represents the observation value of the original sequence at time step t, and $y_{\min}$ and $y_{\max}$ are the minimum and maximum values of the sequence, respectively. Through this transformation, the normalized data $y_t^{(norm)}$ is mapped to the interval [0, 1], which helps to improve the convergence efficiency and training stability of deep networks.

In the aspect of sample division, considering the time-series dependence of time series prediction, the division strategy based on time sequence is adopted. The first 80% of the samples are used for model training, and the remaining 20% are used as test sets to evaluate the prediction ability of the model on the unseen data. This division method not only ensures the sufficiency of training data, but also avoids the problem of information leakage, and provides guarantee for the reliability of subsequent experimental results.

5. EXPERIMENTAL DESIGN AND EVALUATION METHOD

In order to verify the effectiveness and superiority of the proposed SARIMA – attention LSTM hybrid model in short-term export volume prediction, this paper systematically designs the experimental environment configuration, comparison model selection, evaluation index setting and the overall experimental process to ensure the scientificity and reproducibility of the experimental results.

5.1 Experimental environment and parameter setting

All experiments were completed in a unified software and hardware environment to eliminate the influence of external factors on the performance of the model. The experimental platform is constructed based on Python deep learning ecosystem, and the statistical model and neural network model are respectively implemented by calling mature time series analysis and deep learning library. The main experimental environment configuration is shown in Table 4.

Table 4. Experimental environment and hardware and software configuration

Project	Configuration description
Operating system	Windows 11 / Ubuntu 20.04
Processor	Intel Core i7-12700
Memory	32 GB

Python version	Python 3.9
Deep learning framework	TensorFlow 2.10
Time series Library	Stats models 0.13
GPU	NVIDIA RTX 3080
CUDA version	CUDA 11.6

In terms of parameter setting, the order of SARIMA model is determined by grid search combined with AIC criterion; The attention based LSTM network adopts a single hidden layer structure, the number of hidden units is set to 64, the time step length is 12, and the learning rate is 0.001. Adam is selected as the optimizer. All the deep learning models trained 200 epochs and adopted the early stop strategy to prevent over fitting.

5.2 Comparison model selection and experimental setup

In order to comprehensively evaluate the prediction performance of the hybrid model, this paper selects a variety of representative baseline models as the comparison object, covering the traditional statistical model and the deep learning model. These models are widely used in export volume forecasting and related time series research, and have strong representativeness.

Table 5 lists the main comparison models used in this paper and their modeling characteristics.

Table 5. Comparison Model and its characteristics

Model name	Model type	Main features
ARIMA	Statistical model	No seasonal term, linear modeling
SARIMA	Statistical model	Explicit characterization of seasonality
LSTM	Deep learning	Capture nonlinearity and long dependence
GRU	Deep learning	Simplified structure and high training efficiency
Attention-LSTM	Deep learning	Introduction of attention mechanism
SARIMA-LSTM	hybrid model	Linear and nonlinear separation modeling
Proposed Model	hybrid model	SARIMA + Attention-LSTM
Naive	Baseline model	Historical mean prediction

All models run under the same conditions of dividing the training set and the test set to ensure the fairness of the comparison results. For the deep learning model, the input characteristics are all historical export volume series, and no additional exogenous variables are introduced to highlight the difference in the prediction ability of the model structure itself.

5.3 Prediction performance evaluation index

In order to quantitatively evaluate the prediction accuracy of different models, this paper selects three commonly used and complementary error indicators: mean absolute error (MAE), root mean square error (RMSE) and mean absolute percentage error (MAPE) [25],[26]. It is defined as follows:

Mean absolute error:

$$\text{MAE} = \frac{1}{N}\sum_{t=1}^{N} | y_t - \hat{y}_t | \tag{13}$$

Where y_t represents the actual value, $\hat{y}_t$ represents the model's predicted value, and N represents the total number of samples. This metric reflects the absolute average level of the prediction error.

Root mean square error:

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(y_t - \hat{y}_t)^2} \tag{14}$$

Average absolute percentage error:

$$\text{MAPE} = \frac{100\%}{N}\sum_{t=1}^{N} \left| \frac{y_t - \hat{y}_t}{y_t} \right| \tag{15}$$

Table 6 summarizes the prediction error results of different models on the test set.

Table 6. comparison results of prediction performance of different models

Model	MAE	RMSE	MAPE (%)
Naive	12.84	15.92	6.31
ARIMA	10.27	13.45	5.02
SARIMA	8.94	11.76	4.38
LSTM	8.31	10.92	4.11
GRU	8.56	11.14	4.25
Attention-LSTM	7.62	9.98	3.74
SARIMA-LSTM	7.18	9.43	3.51
SARIMA-Attention-LSTM (Proposed)	6.41	8.67	3.12

The results show that the hybrid model is better than the single statistical model and the single deep learning model as a whole, and the SARIMA – attention LSTM model proposed in this paper achieves the best performance in three indicators, which shows that this method has higher accuracy and stability in the short-term export volume prediction.

5.4 Experimental process and verification scheme description

The overall experimental process follows the standard steps of "data input model training predictive output performance evaluation". Firstly, the parameters of each model are learned based on the training set; Then, rolling prediction is performed on the test set to simulate the real short-term prediction scenario; Finally, the prediction results are compared and analyzed through the multi index evaluation system.

In order to enhance the reliability of the experimental results, all the deep learning models were trained for 10 times, and the average prediction results were taken as the final output. At the same time, the prediction error is statistically analyzed to verify the stability of the model performance improvement. The experimental design ensures the objectivity and repeatability of the results, and provides a solid foundation for the subsequent analysis and discussion of the results.

6. RESULTS AND ANALYSIS

Based on the experimental design described in the previous section, this section systematically analyzes the experimental results of different prediction models, focusing on the comparison of prediction accuracy, the performance improvement effect of hybrid models, the contribution of attention mechanism, and the stability and generalization ability of models.

6.1 Comparison and analysis of prediction results of various models

In order to intuitively compare the performance of different models in short-term export volume forecasting, this paper first analyzes it from the perspective of time series forecasting trajectory. Figure 5 shows the comparison between the real export volume in the test set and the prediction results of the main models.

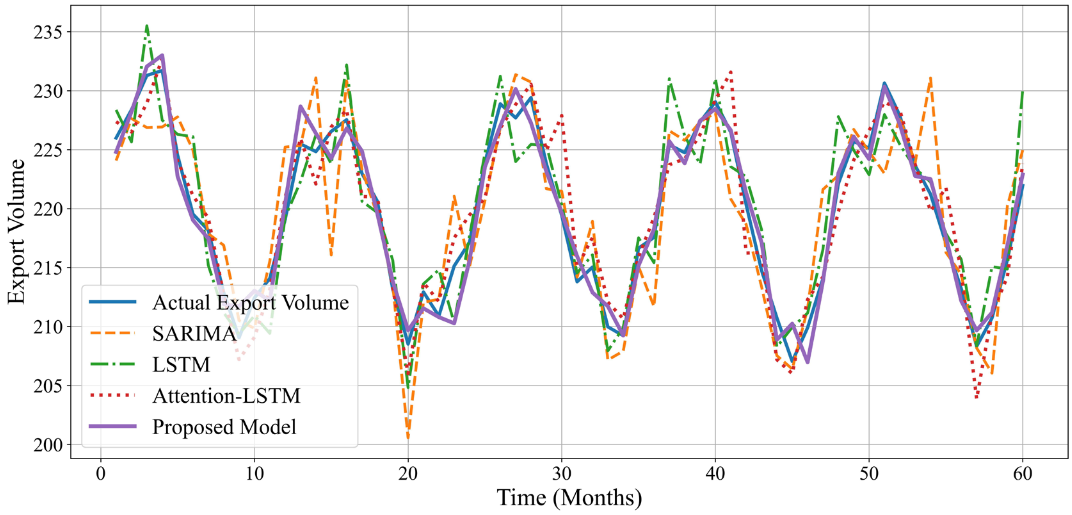

Figure 5. Comparison of prediction results of different models on the test set

It can be observed from Figure 5 that SARIMA model can better follow the overall trend and seasonal fluctuation of export volume, but there is a significant lag in the local sharp change range; LSTM models are better at capturing short-term fluctuations, but there is a certain deviation in fitting the long-term trend. In contrast, the proposed hybrid model achieves a better balance between trend tracking and local volatility characterization, and its prediction curve has the highest fit with the real value.

In order to further quantify the distribution characteristics of prediction error, the prediction residual is introduced:

$$r_t = y_t - \hat{y}_t \tag{16}$$

To further analyze the model performance from the perspective of error, Figure 6 shows the distribution of prediction residuals of different models over time. The time distribution of residual error can directly reflect the fluctuation range and concentration degree of model error.

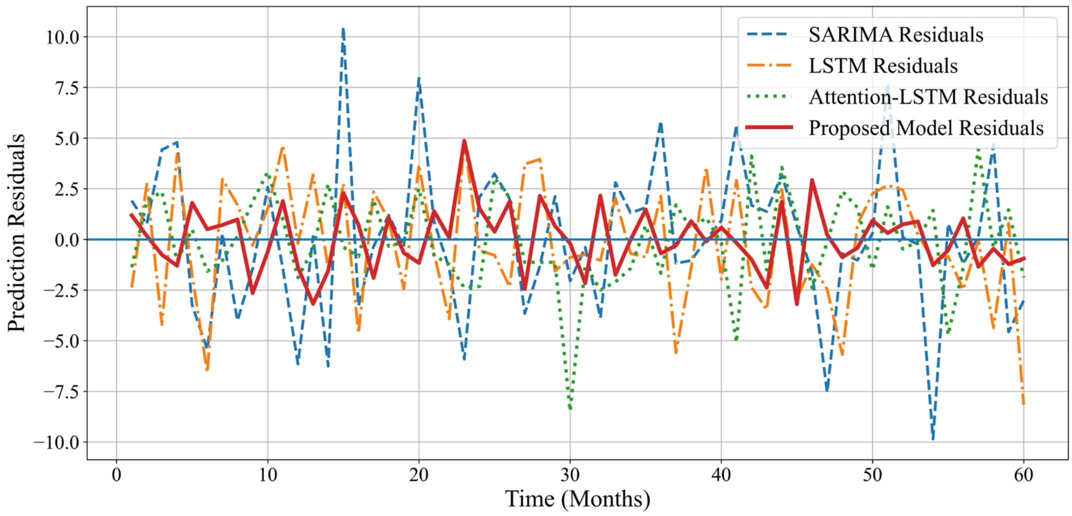

Figure. 6 Time distribution of prediction residuals of different models

The residuals of SARIMA model fluctuated significantly in several time periods, and the residuals were relatively dispersed; The residual fluctuation range of LSTM and attention LSTM models has narrowed, but there are still significant deviations at individual time points. In contrast, the residuals of the hybrid model remain within relatively small cells in most time steps, and the residuals curve is closer to the zero axis with the smallest fluctuation amplitude, indicating that the model can maintain relatively stable prediction accuracy in different time periods, and the time consistency of prediction error is stronger.

6.2 evaluation of performance improvement of hybrid model prediction

In order to quantitatively evaluate the performance improvement of the hybrid model relative to the baseline model, the relative improvement rate of prediction error is calculated. Taking MAE as an example, its promotion rate is defined as:

$$\Delta_{\mathrm{MAE}} = \frac{\mathrm{MAE}_{\mathrm{baseline}} - \mathrm{MAE}_{\mathrm{proposed}}}{\mathrm{MAE}_{\mathrm{baseline}}} \times 100\% \tag{17}$$

The $\mathrm{MAE}_{\mathrm{baseline}}$ and $\mathrm{MAE}_{\mathrm{proposed}}$ represent the mean absolute error between the baseline model and the proposed model, respectively. This metric directly reflects the extent of model improvement; a positive value indicates that the proposed model has a lower error than the baseline model, signifying improved performance.

Table 7 shows the performance improvement of the model in this paper on the three evaluation indicators compared with different baseline models.

Table 7. performance improvement rate of hybrid model relative to baseline model (%)

Baseline model	Mae improvement rate	RMSE improvement rate	MAPE promotion rate
Naive	50.08	45.56	50.55
ARIMA	37.63	35.53	37.85
SARIMA	28.30	26.29	28.77
LSTM	22.86	20.60	24.09

Wang et al., ISTAER. 2601., 02 Jan 2026 https://doi.org/10.71451/ISTAER2601

GRU	25.12	22.20	26.59
Attention-LSTM	15.88	13.13	16.58
SARIMA-LSTM	10.72	8.06	11.11
Average increase	26.82	24.20	27.07

The results show that the hybrid model shows significant performance advantages in comparison with all baseline models, especially compared with the traditional statistical model, the prediction accuracy is improved more significantly. This shows that combining SARIMA's linear modeling ability with attention LSTM's nonlinear learning ability can effectively reduce the structural error of single model.

6.3 Impact analysis of attention mechanism on prediction accuracy

In order to analyze the role of attention mechanism in the hybrid model, this paper further compares the prediction performance of SARIMA-LSTM model without attention mechanism with that after the introduction of attention mechanism. Table 8 summarizes the error results of the two models under different prediction steps.

Table 8. Comparison of prediction performance of models with or without attention mechanism

Prediction step	Model	MAE	RMSE	MAPE (%)
1-step	SARIMA-LSTM	6.98	9.12	3.38
1-step	SARIMA-Attention-LSTM (Proposed)	6.41	8.67	3.12
3-step	SARIMA-LSTM	7.52	9.88	3.71
3-step	SARIMA-Attention-LSTM (Proposed)	6.94	9.21	3.45
6-step	SARIMA-LSTM	8.46	11.03	4.12
6-step	SARIMA-Attention-LSTM (Proposed)	7.89	10.26	3.86
12-step	SARIMA-LSTM	9.83	12.71	4.88
12-step	SARIMA-Attention-LSTM (Proposed)	9.12	11.84	4.41

The results show that the attention mechanism can bring stable performance improvement under different prediction steps, and the advantage is more obvious in the medium and long-term prediction scenarios. This is because the attention mechanism assigns weights dynamically:

$$\alpha_t = \frac{\exp(u_t)}{\sum_k \exp(u_k)} \tag{18}$$

Where u_t represents the attention score at time step t, and the weights α_t are obtained by normalization using the softmax function. This mechanism enables the model to adaptively focus on historical information that is more important to the current prediction, thereby improving its ability to model complex temporal dependencies, especially in maintaining prediction accuracy over a longer prediction range.

At the internal mechanism level of the model, the impact of attention mechanism on the prediction results can be analyzed intuitively through the weight distribution. Figure 7 shows the change of attention weight allocated by attention based LSTM to historical time steps.

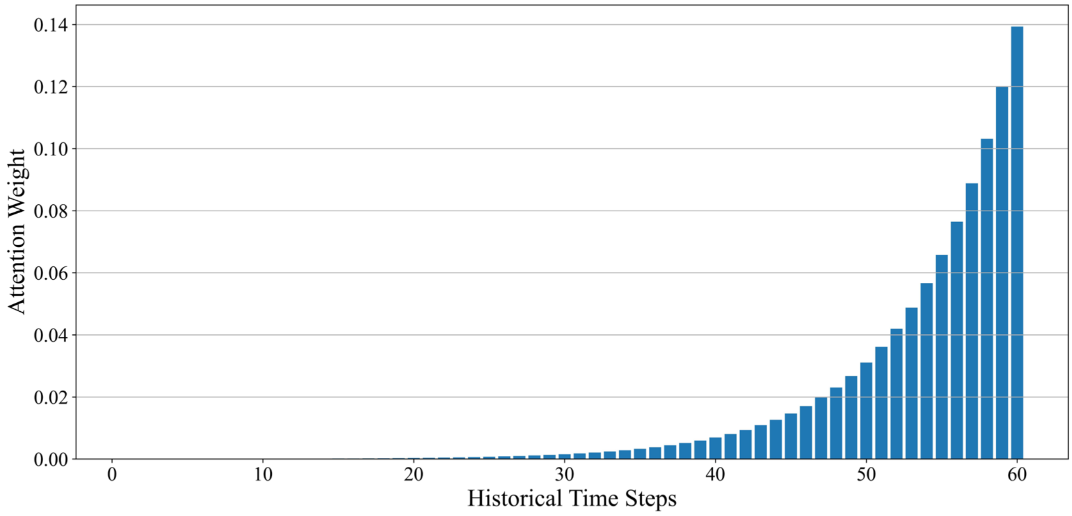

Figure 7. Variation of attention weight over time

It can be clearly seen from Figure 7 that the attention weight presents obvious non-uniform distribution characteristics. The model gives significantly higher weights to some time steps, while the weights of time steps far away from the predicted time point or with small information contribution are significantly reduced. The weight distribution usually forms a concentration area in the recent observation value and the time position with periodic significance, indicating that the model can automatically identify the historical information with the most reference value for the current prediction. This weight allocation mechanism helps to reduce the interference of redundant information, and explains the reason for the improvement of prediction accuracy after the introduction of attention mechanism from the structural level.

6.4 Discussion on model stability and generalization ability

In addition to the prediction accuracy, the stability of the model on multiple training and different data subsets is also an important indicator to measure its practical application value. This paper analyzes the statistical distribution characteristics of model error through repeated experiments. The standard deviation of MAE is defined as:

$$\sigma_{\text{MAE}} = \sqrt{\frac{1}{M}\sum_{i=1}^{M}(\text{MAE}_i - \overline{MAE})^2} \tag{19}$$

Where M represents the number of experiments repeated, MAE_i is the mean absolute error obtained in the i-th experiment, and $\overline{MAE}$ is the mean MAE of the M experiments. This index reflects the degree of fluctuation in model performance; a smaller σMAE indicates that the model has better stability and robustness.

Table 9 shows the error mean and standard deviation of the main models in 10 repeated experiments.

Table 9. Comparison of prediction error stability of different models

Model	MAE mean	MAE standard deviation
SARIMA	8.94	0.63

Wang et al., ISTAER. 2601., 02 Jan 2026 https://doi.org/10.71451/ISTAER2601

LSTM	8.31	0.57
Attention-LSTM	7.62	0.44
SARIMA-LSTM	7.18	0.36
SARIMA-Attention-LSTM (Proposed)	6.41	0.28
GRU	8.56	0.52
ARIMA	10.27	0.71
Naive	12.84	0.85

The results show that the hybrid model proposed in this paper is not only optimal in the average error, but also has the minimum fluctuation range, which shows that it has stronger stability and generalization ability in different initial conditions and training process. This feature is of great significance for the application of real short-term export forecasting.

7. DISCUSSION

The comprehensive experimental results show that the proposed SARIMA – attention LSTM hybrid model shows obvious advantages in the task of short-term export volume prediction. On the one hand, SARIMA model can effectively capture the trend and seasonal structure in the export volume series, and provide a stable basic framework for forecasting; On the other hand, the LSTM network with attention mechanism strengthens the learning ability of key historical information in the residual modeling process, and makes up for the deficiency of traditional statistical models in depicting nonlinear fluctuations. The complementarity of the two models in structure and function makes the hybrid model superior to the single model in prediction accuracy, error stability and response ability to local fluctuations. This shows that hierarchical modeling and feature complementation is an effective technical path in the prediction of economic time series with both regularity and uncertainty, such as export volume.

From the perspective of practical application, the hybrid model has strong applicability and practical value. The short-term export volume forecast is widely used in many scenarios, such as macroeconomic monitoring, trade policy evaluation, and enterprise production and logistics decision-making. Without relying on additional exogenous variables, this method can achieve high-precision prediction only based on historical export volume data, reducing the cost of data acquisition and model deployment. At the same time, the trend and seasonal information provided by SARIMA model can be explained to some extent, which is helpful for decision makers to understand the basic law of export changes; The deep learning module improves the adaptability of the model to complex market changes and makes the prediction results more reliable in practical applications. Therefore, this method has potential promotion value in government departments, trade enterprises and related research institutions.

Although this method has achieved ideal prediction effect in the experiment, it still has some limitations. First of all, the model only uses the single variable export volume time series for modeling, and does not explicitly consider the impact of exogenous variables such as exchange rate, international demand or policy factors, which limits the model's ability to respond to sudden structural changes to a certain extent. Secondly, the overall complexity of the hybrid model is higher than that of the traditional statistical model. When the sample size is small or the data quality is low, the performance of the model may be affected. In addition, although the attention mechanism improves the prediction accuracy, the economic meaning of its weight distribution remains to be further explained.

In view of the above shortcomings, future research can be improved and expanded from multiple directions. On the one hand, multi-source economic indicators can be introduced into the hybrid model framework to build a multivariate prediction model to enhance the ability to describe external shocks; On the other hand, more flexible model fusion strategy or lightweight network structure can be explored to ensure the prediction performance and reduce the computational cost. In addition, combined with the interpretability analysis method, an in-depth study on the relationship between attention weight and economic meaning will also help to enhance the application value of the model in the actual decision-making scenarios.

8. CONCLUSION

This paper focuses on the research of short-term export forecasting, which is of great practical significance, and proposes a hybrid forecasting method based on SARIMA model and LSTM network with attention mechanism. Through the systematic model construction and experimental verification, the research results show that this method can effectively describe the trend, seasonal and nonlinear dynamic characteristics in the export volume time series, and is superior to the traditional statistical model and single deep learning model in terms of prediction accuracy and stability. The experimental analysis further verifies the positive role of hybrid modeling and attention mechanism in reducing the prediction error and improving the robustness of the model, indicating that the method is suitable for the short-term prediction task of complex economic time series.

From the academic point of view, the main contribution of this paper is to provide a hybrid modeling idea with structural complementarity for export volume forecasting. By combining the interpretability of statistical model with the nonlinear modeling ability of deep learning model, this paper expands the application research of mixed time series prediction model in the economic field. At the same time, attention mechanism is introduced to model the residual sequence, which provides a new perspective for understanding the relative importance of different historical time steps in the prediction process. From the perspective of engineering practice, this method achieves high prediction performance without relying on complex exogenous variables, and has good feasibility and promotion potential, which can provide technical support for trade monitoring, macroeconomic analysis and enterprise operation decision-making.

Although this study has achieved some results, there is still room for further expansion. Future work can introduce multivariate information on the basis of the existing model framework, and incorporate exogenous factors such as exchange rate, international demand index or policy variables into the prediction system, so as to enhance the adaptability of the model to sudden changes. At the same time, we can explore a more efficient or interpretable deep learning structure to further enhance the practical value of the model. In addition, the application of the hybrid forecasting method to other macroeconomic indicators or demand forecasting in different industries will also be a direction worthy of further study.

Abbreviations

ARIMA, Autoregressive Integrated Moving Average;
SARIMA, Seasonal Autoregressive Integrated Moving Average;
LSTM, Long Short-Term Memory;
GRU, Gated Recurrent Unit;
RNN, Recurrent Neural Network;
MAE, Mean Absolute Error;
RMSE, Root Mean Square Error;
MAPE, Mean Absolute Percentage Error;
AIC, Akaike Information Criterion;

ACF, Autocorrelation Function;
PACF, Partial Autocorrelation Function;
GPU, Graphics Processing Unit;
CUDA, Compute Unified Device Architecture.

Supplementary Material

Not applicable.

Appendix

Not applicable.

Ethics approval and consent to participate.

This study did not involve human participants, animal subjects, or any data requiring ethical approval. Therefore, ethics approval and consent to participate are not applicable.

Acknowledgements

The authors would like to thank the editors of this journal and all the anonymous reviewers who provided valuable comments on this work.

Competing interests

The authors declare that they have no financial or personal relationships that may have inappropriately influenced them in writing this article.

Author contributions

All authors have read and agreed to the published version of the manuscript. The individual contributions are specified as follows: W.W.: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data Curation, Writing – Original Draft, Visualization. S.S.: Methodology, Software, Validation, Investigation, Data Curation, Writing-Original Draft, Visualization. Y.Y.: Conceptualization, Resources, Writing – Review & Editing, Supervision, Project administration, Funding acquisition.

Funding information

The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Data availability

The data that support the findings of this study are available upon request from the corresponding authors, W.W.

Disclaimer

The views and opinions expressed in this article are those of the authors and are the product of professional research. It does not necessarily reflect the official policy or position of any affiliated institution, funder, agency, or that of the publisher. The authors are responsible for this article's results, findings, and content.

Declaration of AI and AI-assisted Technologies in the Writing Process

During the preparation of this work the authors used ChatGpt-5 in order to check spell and grammar. After using this tool, the authors reviewed and edited the content as needed and takes full responsibility for the content of the publication.

REFERENCES

[1] Kashpruk, N., Piskor-Ignatowicz, C., & Baranowski, J. (2023). Time series prediction in industry 4.0: A comprehensive review and prospects for future advancements. *Applied sciences*, *13*(22), 12374. DOI: **https://doi.org/10.3390/app132212374**

[2] Quang, P. D., & Thang, T. Q. (2025). Analysis and forecasting of daily global gold price: an SARIMA-LSTM approach with Random Forest technique. *Cogent Economics & Finance*, *13*(1), 2568969. DOI: **https://doi.org/10.6084/m9.figshare.30329180**

[3] Ghadami, A., & Epureanu, B. I. (2022). Data-driven prediction in dynamical systems: recent developments. *Philosophical Transactions of the Royal Society A*, *380*(2229), 20210213. DOI: **https://doi.org/10.1098/rsta.2021.0213**

[4] Slater, L., Arnal, L., Boucher, M. A., Chang, A. Y. Y., Moulds, S., Murphy, C., ... & Zappa, M. (2022). Hybrid forecasting: using statistics and machine learning to integrate predictions from dynamical models. *Hydrology and Earth System Sciences Discussions*, *2022*, 1-35. DOI: **https://doi.org/10.5194/hess-27-1865-2023**

[5] Ahmed, S. F., Alam, M. S. B., Hassan, M., Rozbu, M. R., Ishtiak, T., Rafa, N., ... & Gandomi, A. H. (2023). Deep learning modelling techniques: current progress, applications, advantages, and challenges. *Artificial Intelligence Review*, *56*(11), 13521-13617. DOI: **https://doi.org/10.1007/s10462-023-10466-8**

[6] Wu, Z., Rincon, D., Gu, Q., & Christofides, P. D. (2021). Statistical machine learning in model predictive control of nonlinear processes. *Mathematics*, *9*(16), 1912. DOI: **https://doi.org/10.3390/math9161912**

[7] Gettelman, A., Geer, A. J., Forbes, R. M., Carmichael, G. R., Feingold, G., Posselt, D. J., ... & Zuidema, P. (2022). The future of Earth system prediction: Advances in model-data fusion. *Science Advances*, *8*(14), eabn3488. DOI: **https://doi.org/10.1126/sciadv.abn3488**

[8] Sham, F. A. F., El-Shafie, A., Jaafar, W. Z. W., Sherif, M., & Ahmed, A. N. (2025). Advances in AI-based rainfall forecasting: A comprehensive review of past, present, and future directions with intelligent data fusion and climate change models. *Results in Engineering*, *27*, 105774. DOI: **https://doi.org/10.1016/j.rineng.2025.105774**

[9] Sharma, P., Rohatgi, S., & Jasuja, D. (2023). Scientific mapping of gravity model of international trade literature: A bibliometric analysis. *Journal of Scientometric Research*, *11*(3), 447-457. DOI: **https://doi.org/10.5530/jscires.11.3.48**

[10] Xue, R., Li, K., Wang, F., & Baron, C. (2024). Research progress and hot-spot analysis of the economic complexity research based on Citespace. *Advances in Management and Applied Economics*, *14*(1), 1-10. DOI: **https://doi.org/10.47260/amae%2F14110**

[11] Istaiteyeh, R., Najem, F., & Saqfalhait, N. (2023). Exports-and imports-led growth: Evidence from a time series analysis, case of Jordan. *Economies*, *11*(5), 135. DOI: **https://doi.org/10.3390/economies11050135**

[12] Jetwiriyanon, J., Susnjak, T., & Ranathunga, S. (2025). Generalisation Bounds of Zero-Shot Economic Forecasting using Time Series Foundation Models. *Machine Learning and Knowledge Extraction*, *7*(4), 135. DOI: **https://doi.org/10.3390/make7040135**

[13] Dubey, A. K., Kumar, A., García-Díaz, V., Sharma, A. K., & Kanhaiya, K. (2021). Study and analysis of SARIMA and LSTM in forecasting time series data. *Sustainable Energy Technologies and Assessments*, *47*, 101474. DOI: **https://doi.org/10.1016/J.SETA.2021.101474**

[14] Sirisha, U. M., Belavagi, M. C., & Attigeri, G. (2022). Profit prediction using ARIMA, SARIMA and LSTM models in time series forecasting: A comparison. *Ieee Access*, *10*, 124715-124727. DOI: **https://doi.org/10.1109/ACCESS.2022.3224938**

[15] He, K., Ji, L., Wu, C. W. D., & Tso, K. F. G. (2021). Using SARIMA–CNN–LSTM approach to forecast daily tourism demand. *Journal of Hospitality and Tourism Management*, *49*, 25-33. DOI: **https://doi.org/10.1016/j.jhtm.2021.08.022**

[16] Insaad, T., & Sribunrueang, P. (2025). A SARIMA-Based Time Series Analysis of Sugar Export Quantity and Price Trends in Thailand. *Journal of Business, Innovation and Sustainability (JBIS)*, *20*(2), 111-128. DOI: **https://doi.org/10.71185/jbis.2025.277638**

[17] Gonzales Martinez, R. (2024). Apocalypse unleashed: a critical perspective on complexity science, catastrophes, and black swan events in international business. *Critical perspectives on international business*, *20*(1), 94-120. DOI: **https://doi.org/10.1108/cpoib-03-2022-0014**

[18] Alizadegan, H., Rashidi Malki, B., Radmehr, A., Karimi, H., & Ilani, M. A. (2025). Comparative study of long short-term memory (LSTM), bidirectional LSTM, and traditional machine learning approaches for energy consumption prediction. *Energy Exploration & Exploitation*, *43*(1), 281-301. DOI: **https://doi.org/10.1177/01445987241269496**

[19] Alzara, M., Gihad, N., Abdou, H., Soltan, A., Hilal, A., & Ehab, A. (2025). Macroeconomic-aware forecasting of construction costs in developing countries: Using gated recurrent unit and long short-term memory deep learning framework. *PLoS One*, *20*(10), e0333189. DOI: **https://doi.org/10.1371/journal.pone.0333189**

[20] Al-Selwi, S. M., Hassan, M. F., Abdulkadir, S. J., & Muneer, A. (2023). LSTM inefficiency in long-term dependencies regression problems. *Journal of Advanced Research in Applied Sciences and Engineering Technology*, *30*(3), 16-31. DOI: **https://doi.org/10.37934/araset.30.3.1631**

[21] Hajirahimi, Z., & Khashei, M. (2023). Hybridization of hybrid structures for time series forecasting: A review. *Artificial Intelligence Review*, *56*(2), 1201-1261. DOI: **https://doi.org/10.1007/s10462-022-10199-0**

[22] Alharbi, F. R., & Csala, D. (2022). A seasonal autoregressive integrated moving average with exogenous factors (SARIMAX) forecasting model-based time series approach. *Inventions*, *7*(4), 94. DOI: **https://doi.org/10.3390/inventions7040094**

[23] Xu, B., Wang, Y., Yuan, L., & Xu, C. (2023). A novel second-order learning algorithm based attention-LSTM model for dynamic chemical process modeling. *Applied Intelligence*, *53*(2), 1619-1639. DOI: **https://doi.org/10.1007/s10489-022-03515-2**

[24] Wen, X., & Li, W. (2023). Time series prediction based on LSTM-attention-LSTM model. *IEEE access*, *11*, 48322-48331. DOI: **https://doi.org/10.1109/ACCESS.2023.3276628**

[25] Zhao, T., Chen, G., Pang, C., & Busababodhin, P. (2025). Application and Performance Optimization of SLHS-TCN-XGBoost Model in Power Demand Forecasting. *CMES-Computer Modeling in Engineering and Sciences*, *143*(3), 2883-2917. DOI: **https://doi.org/10.32604/cmes.2025.066442**

[26] Zhao, T., Chen, G., Suraphee, S., Phoophiwfa, T., & Busababodhin, P. (2025). A hybrid TCN-XGBoost model for agricultural product market price forecasting. *PLoS One*, *20*(5), e0322496. DOI: **https://doi.org/10.1371/journal.pone.0322496**