Contents lists available at Science-Gate

International Journal of Advanced and Applied Sciences

Journal homepage: http://www.science-gate.com/IJAAS.html

Enhancing global methane emissions forecasting using hybrid time series models

Maryam Habadi *, Mona Alshehri, Ibtesam Alsaggaf

Department of Statistics, King Abdulaziz University, Jeddah, Saudi Arabia

ARTICLE INFO

Article history:
Received 4 September 2024
Received in revised form
6 January 2025
Accepted 8 April 2025

Keywords:
Global warming
Methane emissions
Forecasting models
Hybrid models
Climate change

ABSTRACT

Global warming is a major environmental issue that raises the average air temperature on Earth's surface. Human activities have played a key role in increasing greenhouse gas emissions, which contribute to higher temperatures and climate change. Methane is the second most significant greenhouse gas driving global warming. This study focuses on predicting global methane emissions using the SARIMA (Seasonal Autoregressive Integrated Moving Average) statistical model and three machine learning models: MLP (Multilayer Perceptron), LSTM (Long Short-Term Memory), and GRU (Gated Recurrent Unit). Two hybrid models, SARIMA-MLP and SARIMA-GRU, were also applied. The models' accuracy was assessed using statistical metrics, including Mean Square Error (MSE), Root Mean Square Error (RMSE), and Mean Absolute Percentage Error (MAPE). The findings show that the SARIMA model outperformed the standalone machine learning models. However, the hybrid models demonstrated better forecasting performance, with SARIMA-GRU emerging as the most effective model for predicting global methane emissions. The forecast results indicate a continuous rise in methane emissions over time.

1. Introduction

Global warming is a critical environmental problem that increases the average atmospheric temperature on the Earth's surface. This trend began with the Industrial Revolution in the late 18th century. While temperature changes occur naturally over certain periods, they have increased due to human activities from the Industrial Revolution, including the emissions of greenhouse gases (Nunes, 2023). The concentration of greenhouse gases has significantly increased due to these activities, leading to climate imbalance and global warming. Major sources of greenhouse gases from human activity include carbon dioxide (CO_2) and methane (CH_4). Methane is the second most impactful greenhouse gas contributing to global warming. It has contributed to approximately 30% of global temperature increases since the Industrial Revolution (IEA, 2022). Methane emissions originate from various sources, primarily burning fossil fuels for power generation and transportation. Reducing methane emissions is essential to mitigate climate change. Many initiatives have been implemented in the past few years to reduce methane emissions. The Global Methane Pledge, which was launched in November 2021 during COP26 and signed by countries representing 45% of global methane emissions, intends to decrease global methane emissions by 30% by 2030. Saudi Arabia has committed to this pledge as part of its ambition to create a clean, green future. In 2021, Saudi Arabia introduced the Middle East Green Initiative (MGI) to address climate change in the Middle East and North Africa region (MENA).

In October 2023, MENA nations agreed to collaborate to reduce methane emissions. Furthermore, corporations in the MENA region have pledged to support Aiming for Zero, an Oil and Gas Climate Initiative program to reduce the industry's methane emissions (OGCI, 2023). Egypt and Iraq intend to regulate and monitor methane emissions from the oil and gas industries. Meanwhile, the Gulf Cooperation Council (GCC) countries have pledged to achieve zero emissions by 2050 or 2060. The UAE and Saudi Arabia have already taken steps to limit emissions by banning the burning of methane and other gases (Rousset and Araissi, 2024).

Given the increasing levels of greenhouse gases in the atmosphere, it is critical to develop models that

* Corresponding Author.
Email Address: mhabadi@kau.edu.sa (M. Habadi)
https://doi.org/10.21833/ijaas.2025.04.005
Corresponding author's ORCID profile:
https://orcid.org/0000-0002-6392-1697

can reliably forecast methane emissions over time. These models provide insights that enable governments and communities to take action to reduce emissions. This study investigates global methane emissions using different time series forecasting methods to improve our understanding of methane emissions and predict their trends. However, forecasting time series data is complex due to linear and nonlinear patterns in the data. While the Autoregressive Integrated Moving Average (ARIMA) model is widely employed for time series forecasting, it has difficulty handling nonlinear patterns. In contrast, machine learning algorithms excel at detecting patterns in datasets. To address the complexity of real-world data, a hybrid strategy combining linear and nonlinear models can improve forecasting accuracy (Nawi et al., 2021). As a result, this study aims to develop and analyze two hybrid ARIMA models based on ANN (artificial neural network) and RNN (recurrent neural network) to improve predictions of methane emissions.

Previous studies have successfully used ARIMA and machine learning algorithms to forecast greenhouse gas emissions. A study by Habadi and Tsokos (2017) utilized multiplicative SARIMA models to forecast atmospheric CO_2 levels in the Middle East and the temperature in Saudi Arabia. The performance of the two SARIMA models was assessed using different statistical measures, and the results indicate that both models have good forecasting performance.

Chowdhury et al. (2021) conducted a study utilizing ANN models to forecast CH_4 and CO_2 emissions in Bangladesh based on agricultural data collected between 1972 and 2019. They experimented with various combinations of neurons and layers to identify the effective ANN model for emission forecasting. The models' performance was evaluated using the sum of squared errors (SSE) and RMSE. The results revealed that an ANN model with a layer consisting of three neurons and employing a sigmoid activation function achieved an impressive prediction accuracy rate of 95%.

A study by Kumari and Singh (2023) aimed at predicting CO_2 emissions in India using statistical, machine learning, and deep learning methods. It was concluded that LSTM, SARIMAX, and Holt-Winters were identified as the best performing models out of six based on nine evaluation criteria. The results indicated that LSTM effectively predicted CO_2 emissions with the lowest error value.

Another study by Diaz et al. (2023) used the univariate ARIMA, the multivariate vector autoregressive (VAR), and ARIMA with exogenous inputs (ARIMAX) to forecast the methane concentrations of several coal mines in the USA. The effectiveness of these models was assessed and compared using various statistical metrics like RMSE, MAE, and MAPE. The results indicated that multivariate models have higher forecasting accuracy by incorporating multiple factors. The VAR model outperformed both ARIMA and ARIMAX models in forecasting methane concentrations, and it successfully captured the dynamic interactions of environmental variables with methane emissions over time.

In a study conducted by Luo et al. (2024), a multivariate LSTM model was used to examine the impact of temperature, wind speed, and wind direction on the predictions of methane emissions based on a multivariate climate dataset spanning from 2010 to 2021 in Alberta. The findings suggest that climate variables improve the predictive abilities of the LSTM model, with temperature showing a greater impact on enhancing predictive performance compared to wind speed and direction in terms of mean absolute error (MAE) and RMSE.

Combining various models has become a popular strategy for improving forecast accuracy since the well-known M competitions. Makridakis et al. (1982) conducted these competitions to evaluate and compare the efficiency of different time series forecasting methods. It has been observed that combining forecasts from different models frequently improves forecasting accuracy.

Zhang (2003) proposed a hybrid model that combines ARIMA and ANN models to efficiently capture relationships in time series data. This approach is useful because time series data often shows linear and nonlinear patterns. Experiments with datasets have shown that this hybrid model considerably improves forecasting accuracy.

In a study conducted by Meng et al. (2022), three deep learning models (simple RNN, LSTM, and GRU) and a novel approach that combines traditional time series analysis with deep learning models were utilized to forecast methane concentrations of coal mines in Suzhou, China. The performance of the models was evaluated through criteria such as the RMSE and R^2. The results indicated that the GRU model outperformed the RNN and LSTM models. The study demonstrated that the proposed combining method significantly reduced the prediction errors and enhanced the forecasting accuracy.

Wen et al. (2023) created a hybrid ARIMA-LSTM model using an inverse error combination method to forecast CO_2 emissions in China and its three regions (east, west, and central). This model was compared with four models: Linear regression (LR), Back Propagation Neural Network (BPNN), ARIMA, and LSTM. The results show that the hybrid ARIMA-LSTM model surpasses the others in forecasting CO_2 emissions, suggesting its effectiveness in emissions prediction applications.

Another study by Sergeev et al. (2024) proposed a hybrid model that combines wavelet transformation and the LSTM model to predict the methane concentration on the surface of Pelee Island in the polar region. The performance of the models was checked using different criteria, such as RMSE and MAE. Wavelet transform was applied to decompose data into several components, which helped identify patterns and fluctuations. The results showed that the hybrid model significantly improved forecasting accuracy compared to other models. Recent research highlights the advantages of hybrid models over single models in time series forecasting. The hybrid methods have been applied successfully in various domains, such as electricity consumption (Guo et al., 2021), gold prices (Alsuwaylimi, 2023), crude oil prices (Alrweili and Fawzy, 2022), and wheat yields.

2. Data set

This study used the global methane emissions (CH_4) dataset provided by the National Oceanic and Atmospheric Administration (NOAA). The dataset describes the globally averaged monthly mean atmospheric methane emissions in parts per billion (ppb). It consists of 474 observations over 40 years, from July 1983 to December 2022, as illustrated in Fig. 1. The data exhibit a positive trend and seasonal fluctuations, indicating non-stationarity in the series.

Table 1 presents the descriptive statistics of the global methane emissions series. The average methane emission is 1774 ppb, with a standard deviation of 65. The lowest recorded methane emission was 1626 ppm in July 1983, while the highest recorded emission was 1925 ppm in December 2022.

3. Materials and methods

In order to forecast the global methane emissions, we applied one statistical model: SARIMA, three machine learning models: MLP, LSTM, and GRU, and two hybrid models: SARIMA-MLP and SARIMA-GRU. The dataset was divided into an 80% training set and a 20% testing set to assess the forecasting performance of the models. The training set was used exclusively for developing the models, while the test set was used to evaluate the models' performance. This paper implemented all SARIMA modeling using the R language (4.1.1), while neural network and hybrid models were constructed using Python via Jupyter Notebook. The Augmented Dickey-Fuller (ADF) and Kruskal-Wallis tests were used to assess the stationarity and seasonality of the series. The significance level was set at 5%.

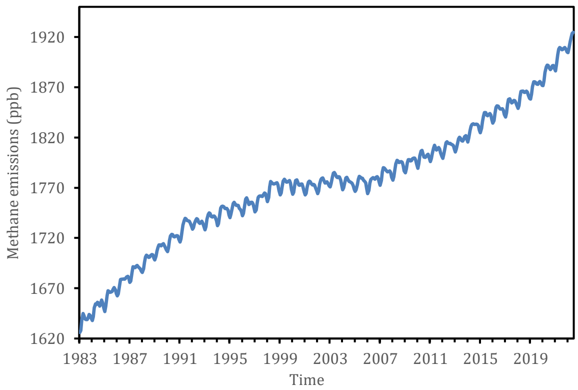


Fig. 1: Global methane emissions dataset

Table 1: Descriptive statistics for methane emissions dataset

Statistics	Length	Min	Median	Mean	Max	Standard deviation
Value	474	1626	1775	1774	1925	65

3.1. ARIMA

The ARIMA model is important and widely used in time series forecasting due to its statistical features and the well-known Box-Jenkins method, which covers the entire modeling procedure for time series (Box and Jenkins, 1970). The ARIMA model comprises two main components: The autoregressive model and the moving average model. The ARIMA model is based on the concept that the future value of a variable is a linear function of previous observations and random errors. The process of the ARIMA model includes the autoregressive, integrated, and moving average processes. The ARIMA model is represented as ARIMA(p, d, q), where p denotes the autoregressive model order, q indicates the moving average model order, and d represents the number of differences needed for stationarity. The ARIMA(p, d, q) model expressed as follow:

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \cdots + \theta_q \epsilon_{t-q} + \epsilon_t \quad (1)$$

where, y_t and ϵ_t are the actual observations and random errors at time t. $\phi_i (i = 1,2,\cdots,p)$ and $\theta_i (i = 1,2,\cdots,q)$ are the parameters of the autoregressive model and the moving average model, respectively.

Stationarity is an important assumption in the ARIMA model. A stationary time series is one in which the mean and variance are constant over time. This study utilizes the ADF to assess stationarity in a time series. This test helps determine whether the time series is stationary or non-stationary. The null hypothesis assumes that the time series lacks stationarity, while the alternative hypothesis suggests it is stationary (Sirisha et al., 2022). When a time series is not stationary, techniques like differencing can be applied to make it stationary.

The ARIMA model is also effective for analyzing seasonal data. A seasonal ARIMA (SARIMA) model is created by incorporating additional seasonal terms into the ARIMA model. It is represented as $ARIMA(p,d,q)(P,D,Q)_s$ with s denoting the seasonal length. Seasonal components are denoted by uppercase letters, while nonseasonal components are represented by lowercase letters. P represents the seasonal autoregressive order, Q represents the seasonal moving order, and D signifies the seasonal differencing. The additional seasonal components are multiplied by non-seasonal components. To investigate seasonality in the series, the Kruskal-Wallis test was employed with a null hypothesis that assumes no seasonality in the time series.

The autocorrelation function (ACF) and partial autocorrelation function (PACF) plots are used to determine the order of the ARIMA model. The ACF plot shows how time series values are correlated at time lags, while the PACF plot reveals these correlations after removing the effects of previous lags. The ACF and PACF values vary from -1 to 1, with 1 representing a perfect positive correlation, -1 suggesting a perfect negative association, and 0 indicating no correlation between the time series values.

The ARIMA modeling process consists of several steps. Initially, the necessary difference degree to make the series stationary and the autoregressive and moving average orders are determined. Subsequently, model parameters are estimated, followed by diagnostic checks to evaluate the model's efficiency. Finally, the model is used to forecast future values. Fig. 2 provides an overview of the ARIMA modeling procedure.

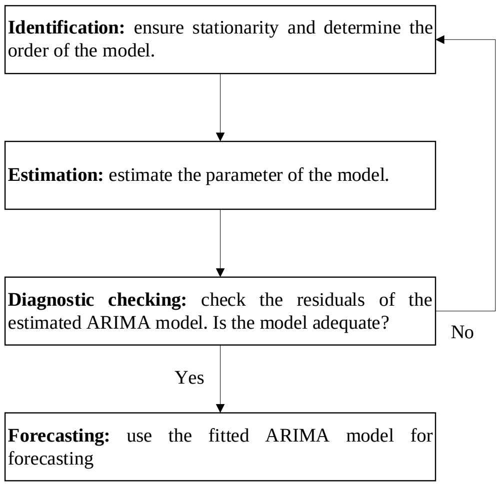


Fig. 2: Flowchart of the ARIMA model algorithm

3.2. Machine learning models

Machine learning is a subset of artificial intelligence that allows computers to autonomously learn from data and use various techniques to create mathematical models and predict outcomes based on historical information. It aims to forecast future events or circumstances unknown to the computer (Awad and Khanna, 2015). The term machine learning was first introduced by Samuel (1959), describing it as the "field of study that gives computers the ability to learn without being explicitly programmed." Deep learning, a type of machine learning, allows computer systems to improve through data and experience. It is commonly called artificial neural networks (ANNs), inspired by the human brain. ANNs have increasingly received much attention from a wide range of different science and socio-economic fields. They have been used successfully in several real-life applications. ANNs are general, flexible, and universal approximators that can estimate any nonlinear function. They can be categorized based on their topology into two main classes: Feed-forward neural networks and feedback neural networks. The feed-forward neural network is a common type that can be used for pattern recognition, classification, forecasting, and clustering. This network allows information to travel one way only. There are no feedback loops, i.e., the output of any layer does not affect its neurons. These networks tend to be straightforward in associating inputs with outputs. There are two types of feed-forward architecture: Single-layer perceptron and multilayer perceptron (MLP). A single-layer perceptron contains only input and output neurons. Unfortunately, a single-layer perceptron cannot detect complex or nonlinear relationships.

In contrast, feedback networks, such as recurrent neural networks (RNNs), have feedback paths, enabling signals to move in both directions through loops. These loops help the system retain past input data to predict outcomes. RNN was designed to analyze a sequence of values one at a time, allowing it to handle sequential data such as time series data, which is particularly useful for forecasting time series (Graves et al., 2013). RNN processes any input sequence using its internal memory. The hidden layer contains a unique backward connection that produces a recurrence in the model, allowing the network to remember information from previous stages (Ceccarelli et al., 2017). The most common recurrent models are long short-term memory (LSTM) and gated recurrent unit (GRU).

3.2.1. Multilayer perceptron (MLP)

The MLP was developed to address the limitations of the single-layer perceptron (Shanmuganathan, 2016). This model is commonly utilized for forecasting time series data (Lee and Cho, 2022). As illustrated in Fig. 3, MLP comprises three layers: One input layer, at least one hidden layer, and a single output layer. MLP trains the network using a learning technique known as backpropagation. Each neuron within the perceptron computes an output y by summing weighted input values using an activation function, as demonstrated in Eq. 2.

$$y = f(\sum_{i=1}^{n} x_i w_i + b) \quad (2)$$

where, f is a nonlinear activation function, x_i refers to the input values, w_i represents the connection weights, and b is the bias term.

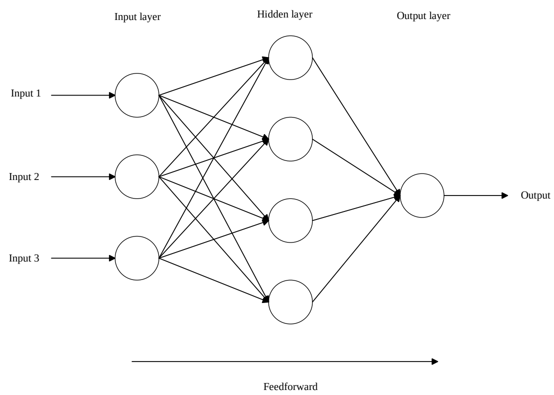


Fig. 3: Multi-layer perceptron architecture (Do et al.,2019)

3.2.2. Long short-term memory (LSTM)

Hochreiter and Schmidhuber (1997) introduced LSTM to address the issue of vanishing gradients in simple RNNs, which improved learning capabilities. LSTM is recognized for its ability to remember long-term dependencies by utilizing a cell state to store and facilitate information flow within the network. The cell state retains long-term memory, while the hidden state maintains short-term memory. The LSTM updates the information using three gates: The forget gate, the input gate, and the output gate. The cell state contains the memory content of an LSTM cell, while the gates ensure the protection and regulation of changes in this memory content. This unique architecture allows LSTM to retain information over extended periods through its chain structure of repeating cells. An overview of LSTM cells is depicted in Fig. 4, where x_t represents the input at time t, $c_{t,}$ and h_t are cell state and hidden state, respectively.

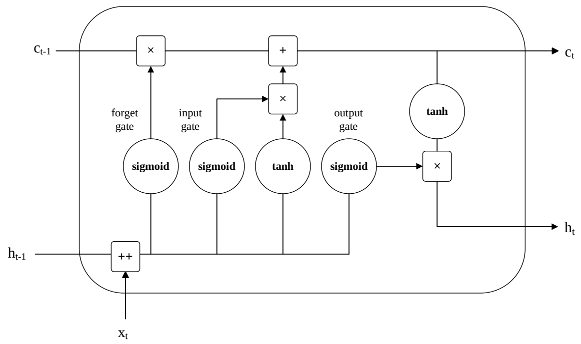


Fig. 4: LSTM cell (Krantz et al., 2019)

3.2.3. Gated recurrent unit (GRU)

Chung et al. (2014) created GRU to improve LSTM by minimizing the number of gates and using only two gates: A reset gate and an update gate. GRU is similar to LSTM, except the forget and input gates are integrated into a single updating gate. Consequently, it has fewer trainable parameters. This makes it easier to compute and apply. Fig. 5 shows the GRU unit construction, which includes an update gate z_t, a reset gate r_t, and a current memory content $\tilde{h}_t$. At each time step t, the update gate takes the input x_t and the output from the previous unit h_{t-1}, feeding it through a sigmoid function. The update gate solves the vanishing gradient problem by learning the model how much information to pass forward. The reset gate dictates how much of the previous information should be forgotten. It is also computed using the current input and the previous hidden states (Dutta et al., 2020).

3.3. Hybrid model

Hybrid methods can improve time series forecasting by combining different statistics and machine learning models. Statistical and machine learning models have significant advantages for time series forecasting. However, neither of these approaches is adequate to model a real-world time series. Table 2 shows the strengths and limitations of each method. The basic idea behind hybrid methods is that they combine the strengths of the two approaches to compensate for the limits of one. Therefore, the hybrid methods are motivated by the following considerations: First, it is difficult to determine whether a time series is linear or nonlinear in real-world situations. Second, time series data is rarely purely linear or nonlinear in the real world. They typically include both patterns. Third, no method is ideal in every circumstance because real-world data is frequently complicated, so no single model can accurately represent all patterns in the data.

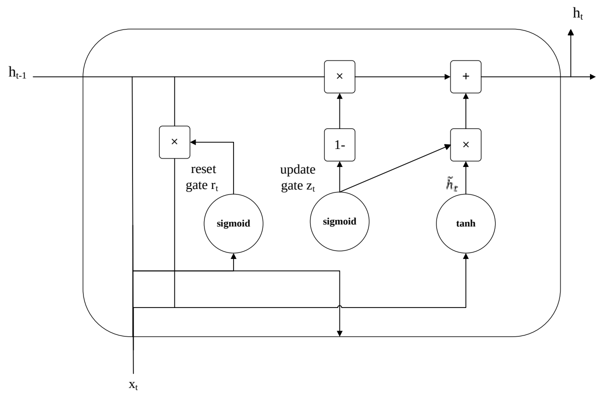


Fig. 5: GRU cell (Li et al., 2021)

According to Zhang (2003), the hybrid models can be expressed as

$$y_t = L_t - N_t \tag{3}$$

where, L_t and N_t represent the linear and nonlinear components, respectively. The residuals which contain

only nonlinear relationships can be calculated by subtracting the estimated values of the linear modeling procedure from the actual observation as follows:

$$e_t = y_t - \hat{L}_t \tag{4}$$

where, statistical models can conduct the linear modeling procedure, and machine learning models can estimate the residual. Therefore, the estimated value $\hat{y}_t$ can be represented as the sum of the estimated values of the statistical and machine learning models, as follows:

$$\hat{y}_t = \hat{L}_t - \hat{N}_t \tag{5}$$

3.4. Model evaluation

Different statistical metrics were utilized to assess the performance of the time series forecasting models. Various error measurements, such as mean square error (MSE), root mean square error (RMSE), and mean absolute percentage error (MAPE), were employed. Lower error values indicate a stronger correlation between observed and forecast values.

The mathematical formulas for these error measurements are given in the following equations, where y_i is the actual observation, $\hat{y}_i$ is the fitted value, and n refers to the number of observations.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \quad (6)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad (7)$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i-\hat{y}_i}{y_i}\right| \times 100 \quad (8)$$

Furthermore, the Akaike Information Criterion (AIC) is essential in problems related to evaluating statistical models. AIC calculates how much information the model loses; less information loss indicates greater model quality. AIC aims to select the model that best explains the variance with the fewest parameters. The model with the lowest AIC value is the most effective forecasting model. It can be expressed as

$$AIC = -2\log(L) + 2k \quad (9)$$

where, L represents the likelihood of the model, and k is the total number of parameters.

Table 2: Strengths and limitations of statistics and machine learning methods

	Strengths	Limitations
Statistical methods	Simple and flexible Work on a small data set	Assume linearity Limited information capability
Machine learning models	No assumption of linearity Universal approximator	Require large data sets Computational effort

4. Results

This section presents the modeling findings of the global methane emissions dataset along with the evaluation performance of the models. It starts with the SARIMA modeling results and then moves on to the machine learning modeling results. Finally, the results of the two hybrid models were presented and compared.

4.1. Seasonal ARIMA (SARIMA)

The results of the ADF and Kruskal-Wallis tests indicate clearly that the methane emissions data set is not stationary and has seasonal fluctuations. The series becomes stationary after first-order differencing (d = 1) and first-order seasonal differencing (D = 1). Tables 3 and 4 display the ADF and Kruskal-Wallis test results, respectively.

Table 3: ADF test result

	Training set	First-order differencing
Test statistics	-2.8306	-21.854
p-value	0.2264	0.01

Table 4: Kruskal-Wallis test result

	Training set	First-seasonal differencing
Test statistics	315.59	1.12
p-value	0	0.9999

The AIC criterion is used to determine the order of the SARIMA model accurately. As a result, SARIMA$(2,1,2)(0,1,1)_{12}$ was the appropriate model for fitting the methane emissions data set with the lowest AIC value (AIC = 807.45). The optimal SARIMA for the methane emissions dataset can be presented as

$$\hat{y}_t = -0.4214y_{t-1} - 0.4138y_{t-2} + 1.7616\epsilon_{t-1} + 0.8259\epsilon_{t-2} - 0.8531\epsilon_{t-12} + \epsilon_t \quad (10)$$

Fig. 6 illustrates the residual diagnostics of the SARIMA model. The residuals seem to have a normal distribution and fluctuate around a mean of zero. The ACF plot indicates that the residuals are not auto-correlated, which indicates that the SARIMA(2,1,2)(0,1,1)12 model is a well-fitting forecasting model of global methane emissions. The forecast values of the methane emissions for the next 24 months show an increasing trend over time, as shown in Fig. 7, which depicts the fitted and forecast values of our SARIMA model.

4.2. Machine learning models

The machine learning models were fitted to forecast methane emissions. First, the data set was transformed into supervised learning using several previous time steps as inputs and then utilizing the next time step as output to the model. For model training, the MSE was used as the loss function, and the adaptive moment estimation (Adam) with a learning rate of 0.001 was used as the optimizer for the neural network models. Tuning the parameters of the models, including batch size, number of epochs, and validation split, was essential for the training process. Different batch sizes between 32 and 128 were experimented with to balance computational efficiency and stability. All three models were trained using 500 epoch training runs, determined through experimentation and performance evaluation. Additionally, 10-20% of the dataset was allocated for validation to assess model performance on unseen data during the training process. Ultimately, the MLP model performed best with 500 epochs and a batch size of 64, while both the GRU and LSTM models excelled with 500 epochs and a batch size of 32.

Different combinations of hidden layers and nodes were applied to achieve the best network structure, and the optimal neural network model was chosen according to trial and least MSE. The appropriate MLP model for methane emissions data consists of three layers: An input layer with 16 nodes representing the input values of methane emissions, a hidden layer of 60 nodes, and an output layer with one node. The LSTM model consists of three layers: An input layer with 16 nodes, a hidden layer with 30 nodes, and a single output layer. In contrast, the GRU model comprises three layers: An input layer with 12 nodes, a hidden layer with 57 nodes, and an output layer with only one node. Fig. 8 shows the test data along with the fitted values of the MLP, LSTM, and GRU models. The results reveal that the GRU model performed better than the LSTM model, which was utilized to create the hybrid model. Machine learning models were applied to forecast the global methane emissions for the next 24 months. The forecasting values of the three neural network models are shown in Fig. 9. The forecasting shows an increase in methane emissions.

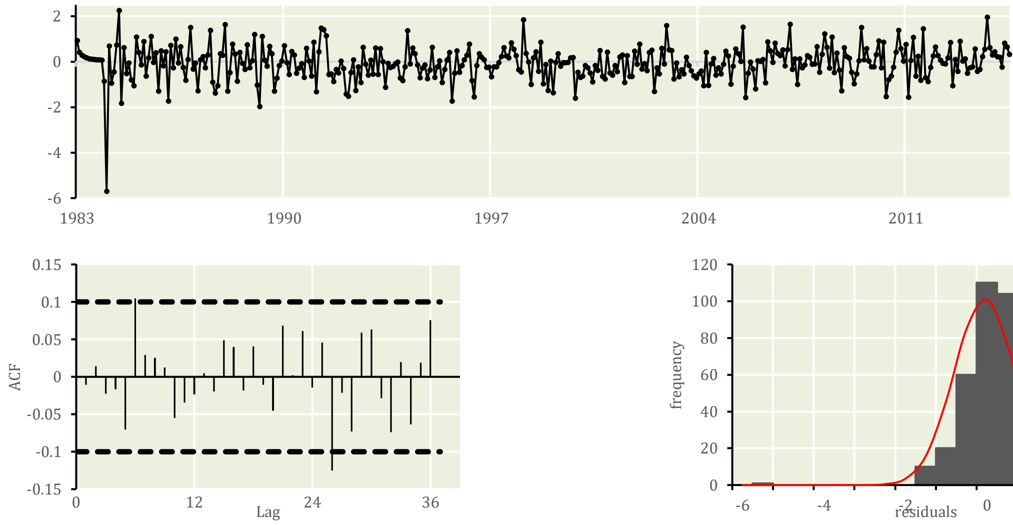
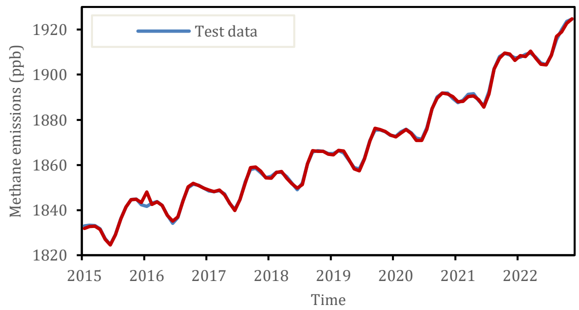


Fig. 6: Residuals diagnostics of the SARIMA model



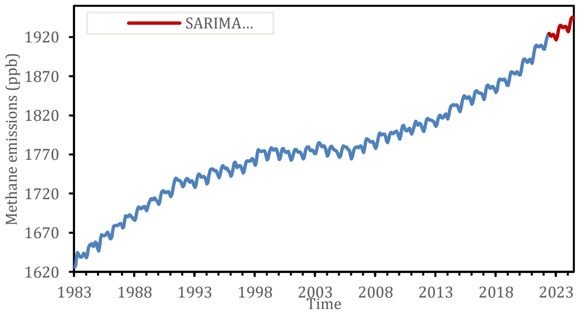
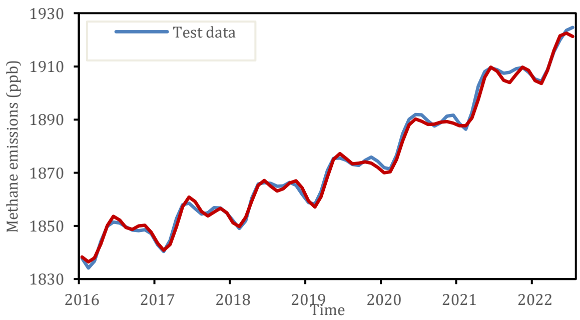
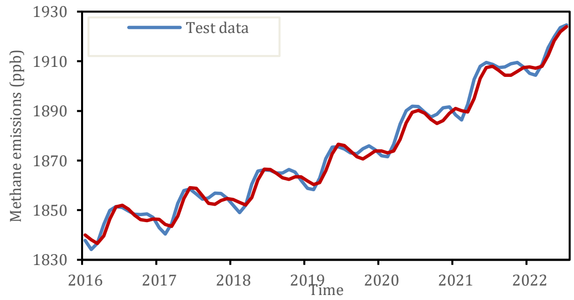


Fig. 7: Fitted and forecast values of the SARIMA model





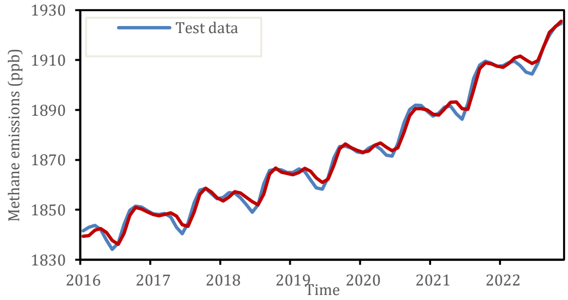


Fig. 8: Fitted values of machine learning models

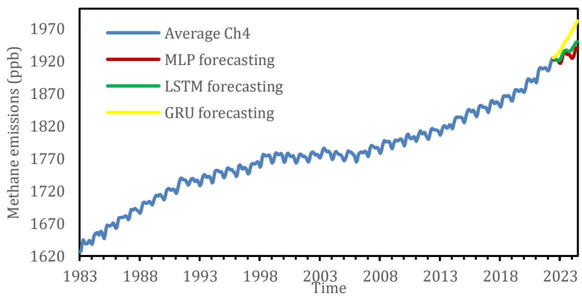


Fig. 9: Forecast values of machine learning models

4.3. Hybrid models

To construct hybrid SARIMA-MLP and SARIMA-GRU models, the residuals of the SARIMA$(2,1,2)(0,1,1)_{12}$ model, which incorporates nonlinear parts, were fitted using individual MLP and GRU models. Then, the results of the nonlinear models were summed along with fitted values of the SARIMA model. The fitted and forecast values of the two hybrid models are shown in Figs. 10 and 11, respectively. We can see that the fitted values of the hybrid models closely match the actual test data, indicating that the hybrid models have good forecasting accuracy. Thus, according to the forecast of the hybrid models, methane emissions are expected to increase over time.

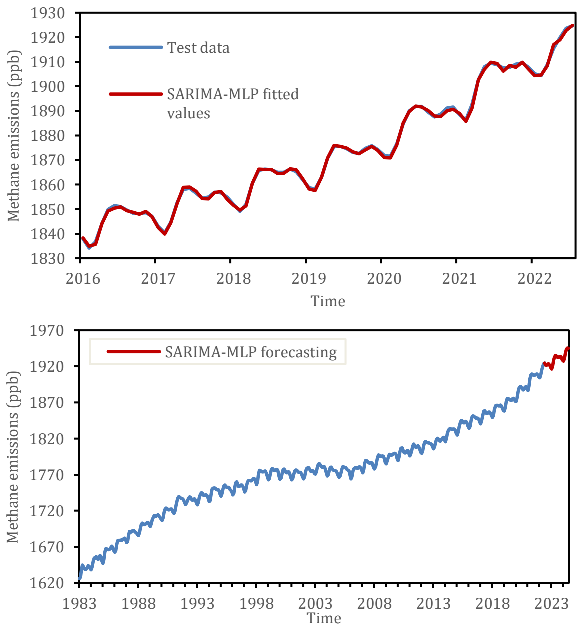


Fig. 10: Fitted and forecast values of the SARIMA-MLP model

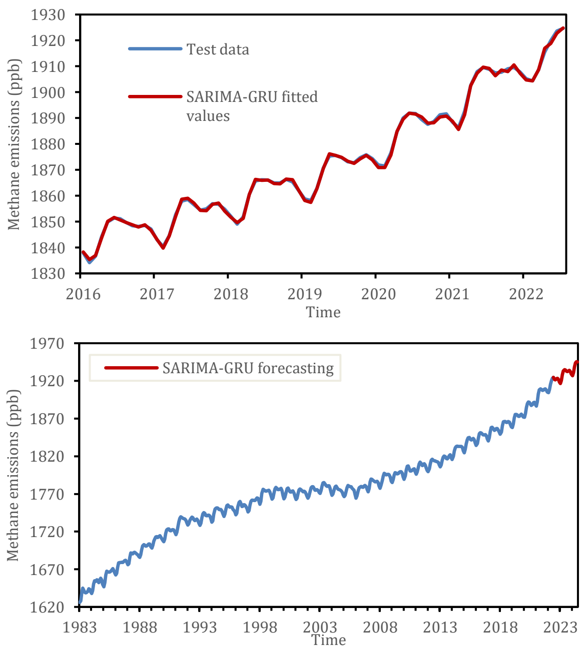


Fig. 11: Fitted and forecast values of the SARIMA-GRU model

4.4. Comparison

Table 5 presents the evaluation results of the single and hybrid models using MSE, RMSE, and MAPE. The results show that the single SARIMA model outperformed the single machine learning models. The explanation is that methane emission data has a clear linear trend, and the statistical ARIMA model is known for capturing linear patterns in the data. However, the findings of the hybrid models indicated that combining the advantages of the SARIMA model in capturing linear patterns and machine learning models in capturing nonlinear patterns can significantly improve forecasting performance. As a result, the hybrid SARIMA-GRU model produced the most accurate forecasting results for the global methane emissions dataset.

Table 5: Performance evaluation of models

Model	MSE	RMSE	MAPE
SARIMA	0.7535	0.8680	0.0281
MLP	2.7591	1.6610	0.0731
LSTM	8.5620	2.9260	0.1286
GRU	4.9858	2.2329	0.0952
SARIMA-MLP	0.4082	0.6389	0.0280
SARIMA-GRU	0.3516	0.5930	0.0260

5. Discussion

As the levels of methane emissions increase globally, it is essential to develop effective prediction models that can guide policymaking and implement strategies for reducing emissions. Accurate forecasting will enable decision-makers to implement timely and targeted interventions to mitigate the impact of methane on climate change. This study presented various forecasting methods to enhance the accuracy of methane emission prediction and assessed the strengths and limitations of different approaches.

In this study, we employed one statistical model (SARIMA), three machine learning models (MLP, LSTM, and GRU), and two hybrid models (SARIMA-MLP and SARIMA-GRU) to improve the prediction of global methane emissions and evaluated their efficacy. The experimental results reveal that the SARIMA model outperformed the single machine learning models in terms of forecasting accuracy. This finding suggests that the statistical properties of methane emissions were better captured by the SARIMA model, which considers seasonality and trend patterns in the data. Moreover, it is observed that the hybrid models, which combine the machine learning model with the statistical SARIMA model, can further improve the forecast accuracy. These hybrid models use the advantages of both statistical and machine learning methods, enabling the effective capture of complex relationships that may not be apparent in a single model. This highlights the potential of hybrid methods in handling forecasting challenges and emphasizes their capability to enhance predictive outcomes.

The predicted increase in global methane emissions highlighted in this study confirms the essential need for immediate action to reduce methane emissions, particularly for policymakers. This is crucial to address the issue, especially in high-emissions fields. The findings of this study can also help set emission reduction targets in line with various global initiatives, such as the Paris Agreement and the Global Methane Pledge. More importantly, they can foster international cooperation to address the global challenge of methane emissions.

To further enhance forecasting accuracy in future studies, multivariate prediction models can incorporate various influencing factors, including environmental, economic, and agricultural variables. This approach allows analysis of how different predictors, such as temperature, landfills, fossil fuel production, and livestock, impact methane emissions.

Furthermore, raising public awareness of the critical role of methane in climate change is vital for gaining community support for emission reduction efforts. Consequently, the increase in methane emissions highlights the need for improved forecasting and emphasizes the crucial role of informed policymaking in the fight against climate change.

6. Conclusions

This study aimed to create single and hybrid forecasting models to forecast global methane emissions. The study utilized the statistical SARIMA model and machine learning methods, including MLP, LSTM, and GRU. Additionally, two hybrid models, SARIMA-MLP and SARIMA-GRU, were created. Various statistical metrics such as MSE, RMSE, and MAPE were used to evaluate and compare the model performances. The results indicated that hybrid models combining the strengths of SARIMA with machine learning outperformed individual models in forecasting methane emissions. The results highlighted that neither SARIMA nor a single machine learning model could accurately capture all data patterns. Additionally, a comparison between the two hybrid models revealed that the SARIMA-GRU hybrid model was most effective at forecasting methane emissions, which are expected to increase over time. This insight is significant for policymakers and leaders as it assists in crafting strategies to manage methane emissions, achieve reduction targets, and ensure better regulation of emissions.

List of abbreviations

ADF	Augmented Dickey-Fuller
AIC	Akaike information criterion
ANN	Artificial neural network
ARIMA	Autoregressive integrated moving average
ARIMAX	Autoregressive integrated moving average with exogenous inputs
BPNN	Back propagation neural network
CO_2	Carbon dioxide
CH_4	Methane
GCC	Gulf cooperation council
GRU	Gated recurrent unit
LSTM	Long short-term memory
LR	Linear regression
MAE	Mean absolute error
MAPE	Mean absolute percentage error
MGI	Middle East Green Initiative
MENA	Middle East and North Africa
MLP	Multilayer perceptron
MSE	Mean square error
NOAA	National oceanic and atmospheric administration
OGCI	Oil and gas climate initiative
PACF	Partial autocorrelation function
ppb	Parts per billion
RNN	Recurrent neural network
RMSE	Root mean square error
SARIMA	Seasonal autoregressive integrated moving average
SARIMAX	Seasonal autoregressive integrated moving average with exogenous inputs
SSE	Sum of squared errors
VAR	Vector autoregressive

Compliance with ethical standards

Conflict of interest

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

Alrweili H and Fawzy H (2022). Forecasting crude oil prices using an ARIMA-ANN hybrid model. Journal of Statistics Applications and Probability, 11(3): 845-855. https://doi.org/10.18576/jsap/110308

Alsuwaylimi AA (2023). Comparison of ARIMA, ANN and Hybrid ARIMA-ANN models for time series forecasting. Information Sciences Letters, 12(2): 1003-1016. https://doi.org/10.18576/isl/120238

Awad M and Khanna R (2015). Efficient learning machines: Theories, concepts, and applications for engineers and system designers. Springer, Berkeley, USA. https://doi.org/10.1007/978-1-4302-5990-9

Box GE and Jenkins GM (1970). Time series analysis forecasting and control. Holden-Day, San Francisco, USA.

Ceccarelli F, Sciandrone M, Perricone C, Galvan G, Morelli F, Vicente LN, Leccese I, Massaro L, Cipriano E, Spinelli FR, and Alessandri C et al. (2017). Prediction of chronic damage in systemic lupus erythematosus by using machine-learning models. PLOS ONE, 12(3): e0174200. https://doi.org/10.1371/journal.pone.0174200 **PMid:28329014 PMCid:PMC5362169**

Chowdhury S, Rubi MA, and Bijoy MHI (2021). Application of artificial neural network for predicting agricultural methane and CO_2 emissions in Bangladesh. In the 12th International Conference on Computing Communication and Networking Technologies, IEEE, Kharagpur, India: 1-5. https://doi.org/10.1109/ICCCNT51525.2021.9580106

Chung J, Gulcehre C, Cho K, and Bengio Y (2014). Empirical evaluation of gated recurrent neural networks on sequence modeling. Arxiv Preprint Arxiv:1412.3555. https://doi.org/10.48550/arXiv.1412.3555

Diaz J, Agioutantis Z, Hristopulos DT, Luxbacher K, and Schafrik S (2023). Forecasting of methane gas in underground coal mines: Univariate versus multivariate time series modeling. Stochastic Environmental Research and Risk Assessment, 37(6): 2099-2115. https://doi.org/10.1007/s00477-023-02382-8

Do LN, Taherifar N, and Vu HL (2019). Survey of neural network-based models for short-term traffic state prediction. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 9(1): e1285. https://doi.org/10.1002/widm.1285

Dutta A, Kumar S, and Basu M (2020). A gated recurrent unit approach to bitcoin price prediction. Journal of Risk and Financial Management, 13(2): 23. https://doi.org/10.3390/jrfm13020023

Graves A, Mohamed AR, and Hinton G (2013). Speech recognition with deep recurrent neural networks. In the IEEE International Conference on Acoustics Speech and Signal Processing, IEEE, Vancouver, Canada: 6645-6649. https://doi.org/10.1109/ICASSP.2013.6638947

Guo N, Chen W, Wang M, Tian Z, and Jin H (2021). Appling an improved method based on ARIMA model to predict the short-term electricity consumption transmitted by the Internet of Things (IoT). Wireless Communications and Mobile Computing, 2021: 6610273. https://doi.org/10.1155/2021/6610273

Habadi MI and Tsokos CP (2017). Statistical forecasting models of atmospheric carbon dioxide and temperature in the Middle East. Journal of Geoscience and Environment Protection, 5: 11-21. https://doi.org/10.4236/gep.2017.510002

Hochreiter S and Schmidhuber J (1997). Long short-term memory. Neural Computation, 9(8): 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735 **PMid:9377276**

IEA (2022). Global methane tracker. International Energy Agency, Paris, France.

Krantz J, Dulin M, and De Palma P (2019). Language-agnostic syllabification with neural sequence labeling. In the 18th IEEE International Conference on Machine Learning and Applications, IEEE, Boca Raton, USA: 804-810. https://doi.org/10.1109/ICMLA.2019.00141

Kumari S and Singh SK (2023). Machine learning-based time series models for effective CO_2 emission prediction in India. Environmental Science and Pollution Research, 30: 116601-116616. https://doi.org/10.1007/s11356-022-21723-8 **PMid:35780266**

Lee J and Cho Y (2022). National-scale electricity peak load forecasting: Traditional, machine learning, or hybrid model? Energy, 239: 122366. https://doi.org/10.1016/j.energy.2021.122366

Li N, Hu L, Deng ZL, Su T, and Liu JW (2021). Research on GRU neural network satellite traffic prediction based on transfer learning. Wireless Personal Communications, 118(1): 815-827. https://doi.org/10.1007/s11277-020-08045-z

Luo R, Wang J, and Gates I (2024). Forecasting methane data using multivariate long short-term memory neural networks. Environmental Modeling and Assessment, 29: 441–454. https://doi.org/10.1007/s10666-024-09957-x

Makridakis S, Andersen A, Carbone R, Fildes R, Hibon M, Lewandowski R, Newton J, Parzen E, and Winkler R (1982). The accuracy of extrapolation (time series) methods: Results of a forecasting competition. Journal of Forecasting, 1(2): 111-153. https://doi.org/10.1002/for.3980010202

Meng X, Chang H, and Wang X (2022). Methane concentration prediction method based on deep learning and classical time series analysis. Energies, 15(6): 2262. https://doi.org/10.3390/en15062262

Nawi WIAWM, Lola MS, Zakariya R, Zainuddin NH, Abd Hamid AAK, Aruchunan E, and Nazzrol NSA (2021). Improved of forecasting sea surface temperature based on hybrid ARIMA and support vector machines models. Malaysian Journal of Fundamental and Applied Sciences, 17(5): 609-620. https://doi.org/10.11113/mjfas.v17n5.2356

Nunes LJ (2023). The rising threat of atmospheric CO_2: A review on the causes, impacts, and mitigation strategies. Environments, 10(4): 66. https://doi.org/10.3390/environments10040066

OGCI (2023). Methane emissions. Oil and Gas Climate Initiative, London, UK.

Rousset F and Araissi A (2024). The Middle East is tackling the methane challenge. Arabian Gulf Business Insight, London, UK.

Samuel AL (1959). Some studies in machine learning using the game of checkers. IBM Journal of Research and Development, 3(3): 210-229. https://doi.org/10.1147/rd.33.0210

Sergeev A, Baglaeva E, and Subbotina I (2024). Hybrid model combining LSTM with discrete wavelet transformation to predict surface methane concentration in the Arctic Island Belyy. Atmospheric Environment, 317: 120210. https://doi.org/10.1016/j.atmosenv.2023.120210

Shanmuganathan S (2016). Artificial neural network modelling: An introduction. Springer International Publishing, Cham, Switzerland. https://doi.org/10.1007/978-3-319-28495-8_1

Sirisha UM, Belavagi MC, and Attigeri G (2022). Profit prediction using ARIMA, SARIMA and LSTM models in time series forecasting: A comparison. IEEE Access, 10: 124715-124727. https://doi.org/10.1109/ACCESS.2022.3224938

Wen T, Liu Y, Bai Y, and Liu H (2023). Modeling and forecasting CO_2 emissions in China and its regions using a novel ARIMA-LSTM model. Heliyon, 9(11): e21241. https://doi.org/10.1016/j.heliyon.2023.e21241 **PMid:37954263 PMCid:PMC10632432**

Zhang GP (2003). Time series forecasting using a hybrid ARIMA and neural network model. Neurocomputing, 50: 159-175. https://doi.org/10.1016/S0925-2312(01)00702-0