Journal of
Water & Climate Change

IWA PUBLISHING

Journal of Water and Climate Change Vol 15 No 8, 4016 doi: 10.2166/wcc.2024.252

Meta-learning applied to a multivariate single-step fusion model for greenhouse gas emission forecasting in Brazil

Liriam Michi Enamoto [a], Andre Rufino Arsenio Santos[b], Li Weigang [c,*], Rodolfo Meneguette[d] and Geraldo P. Rocha Filho[e]

[a] University of Brasilia, Brasilia, Brazil
[b] Faculty of Law, State University of Rio de Janeiro, Rio de Janeiro, Brazil
[c] Department of Computer Science, University of Brasilia, Brasilia, Brazil
[d] Computer Science Department, University of São Paulo, Sao Paulo, Brazil
[e] Department of Exact and Technological Sciences, State University of Southwest Bahia, Vitoria da Conquista, Brazil
*Corresponding author. E-mail: weigang@unb.br

LME, 0000-0003-0188-5966; LW, 0000-0003-1826-1850

ABSTRACT

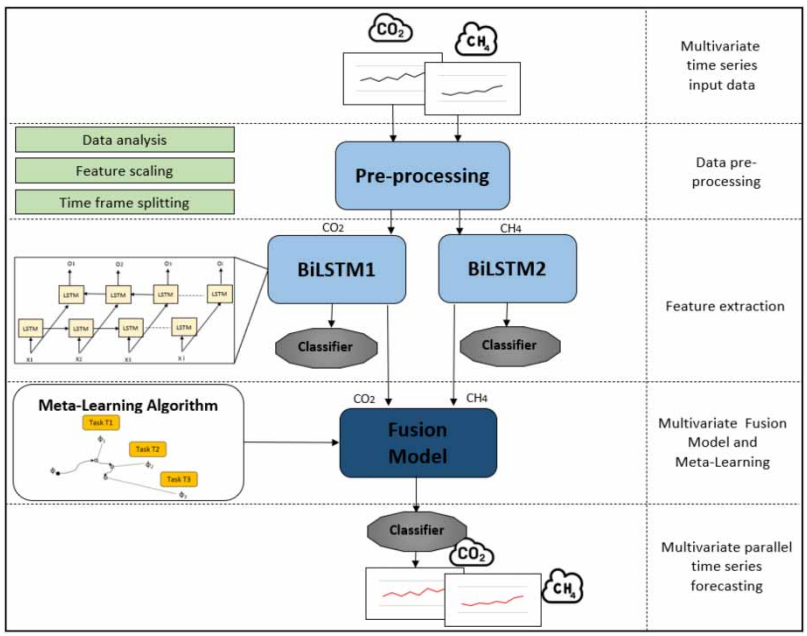
Climate change, driven by greenhouse gas (GHG) emissions, causes extreme weather events, impacting ecosystems, biodiversity, population health, and the economy. Predicting GHG emissions is crucial for mitigating these impacts and planning sustainable policies. This research proposes a novel machine learning model for GHG emissions forecasting. Our model, Meta-Learning Applied to Multivariate Single-Step Fusion Model, utilizes historical GHG emissions from Brazil over the past 60 years to predict CO_2 and CH_4 emissions. Additionally, the model employs a unique combination of two techniques in time series forecasting: (i) in the Fusion Model, each substance is individually extracted and trained based on a specific decision task, then integrated into the same feature space; (ii) Meta-Learning allows the model to learn from past prediction tasks, leading to better generalization. Our model was compared with state-of-the-art time series models using the same dataset. The results show that our approach reduces the mean absolute percentage error by 49.06% with 95% confidence compared to the Transformer-based TST model, demonstrating its superior performance and low estimated CO_2 emissions of 0.01 kg CO_2eq. Furthermore, the model's flexibility allows it to be adapted for various environmental studies and general time series forecasting.

Key words: climate change, data fusion, greenhouse gas, meta-learning, multivariate forecasting, Reptile

HIGHLIGHTS

- Multivariate fusion model and meta-learning for GHG emissions forecasting.
- BiLSTM for data extraction and Reptile for model optimization.
- Analyses Brazil's CO_2 and CH_4 emissions over the last 60 years.
- Summarizes the main Brazilian legislation for climate change.
- Simple model architecture adaptable for other substances.

GRAPHICAL ABSTRACT



1. INTRODUCTION

Globalization promotes local economic growth and investment in emerging countries to reduce poverty and hunger, stimulate education, scientific research, economy, and improve public health. Likewise, the growth of anthropogenic activities in one country contributes to increased greenhouse gas (GHG) emissions that can affect the climate in other regions or continents.

The effects of global warming are responsible for extreme heatwaves, floods, droughts, hurricanes, and biodiversity loss in several countries (Rolnick *et al.* 2022). However, the emergent countries in the southern hemisphere, including Brazil, are the most affected by climate change (Zhang *et al.* 2023) due to the lack of sanitation infrastructure, urban plans, natural disaster mitigation measures, and lack of investment in sustainable development.

A historical flash flood occurred in the southern region of Brazil in 2008 affected the most vulnerable population in the rural area (Wink Junior *et al.* 2023). The extreme precipitation in the mountain region of Rio de Janeiro in 2011 caused landslides resulting in irreparable losses for the most vulnerable population (Lopez *et al.* 2023).

Despite the economic difficulties and the climate impacts suffered in the last years, Brazil is committed to contributing to the low-carbon emission agenda with the following goals: (i) to achieve 45% of overall renewable sources in the energy mix by 2030 (Werner & Lazaro 2023) and (ii) to achieve neutral net emission by 2050 (UNFCCC 2023). According to the EDGAR – Emissions Database for Global Atmospheric Research data provided by the European Commission, the 2022 GHG emission in Brazil was 91.64% lower than China, 78.22% lower than the United States, and 66.77% lower than India, which are countries that most contribute to the GHG emissions (Crippa *et al.* 2021).

Moreover, the geographic, climatic and natural resource diversity allows Brazil to have the electricity energy matrix with 83% renewable in 2020, in which hydropower energy accounted for 60.7% (De Toledo *et al.* 2023). However, due to climate variability such as low precipitation levels and droughts, the percentage of hydropower energy decreased to 53.4% in 2021 (Werner & Lazaro 2023).

Efforts to address climate change from a legal perspective are taking place in Brazil. Law 12187 (BRAZIL 2009b) establishes the National Policy on Climate Change (PNMC) providing guidelines to mitigate the global warming impacts, preserve ecosystems, and promote sustainable development. The Brazilian Congress is analyzing the bill PL 412/2022 (BRAZIL 2022c), which regulates the trade of carbon credits in Brazil. The approval of this bill could create legal instruments to encourage transactions with carbon assets, contributing to the country's sustainable development.

In addition to creating the National Policy on Climate Change and regulating the carbon credit trade, it is essential to monitor GHG emissions. Joint monitoring and regulation action can allow Brazil to achieve the target of 45% overall renewable

energy mix by 2030 and keep contributing to the low-carbon agenda. GHG forecasting can help analyze, plan, and adopt a sustainable agenda to promote development, encourage carbon credit policy, and mitigate the impacts of climate change.

The problem of forecasting GHG emissions can be addressed through time series modeling, which has been applied in various fields. Lv *et al.* (2024) proposed a multi-temporal correlation feature fusion network for machinery fault diagnosis in the manufacturing sector. In the same year, Wang *et al.* (2024) presented a deep learning framework for acoustic modulation for autonomous underwater vehicles. Turning to food science, Natsume & Okamoto (2024) used the Echo State Network (ESN) to predict changes in food preferences. In the field of climate change, Misra *et al.* (2024) used a Convolutional Short-Term Memory Network for rainfall prediction.

In this research, we devised the **M**eta-**L**earning Applied to a Multivariate Single-Step Fusion Model **for G**reen**H**ouse **G**as Emission **(ML4GHG)** Forecasting. The model analyses Brazil's CO_2 and CH_4 emissions over the past 60 years. The model uses multivariate GHG emission data from Brazil to learn the time series forecasting. In the training process, the Fusion Model learns the multivariate series, and the Meta-Learning algorithm helps ML4GHG in the generalization process, outperforming the baseline models.

The contributions of this work are:

- Develop and evaluate a novel model based on a multivariate single-step approach leveraging the combination of the Fusion Model for data alignment and optimization-based Meta-Learning tailored for time series;
- Adapt the Reptile Meta-Learning algorithm to improve the model generalization;
- Develop a model that can be easily adapted to other substances;
- Evaluate the proposed models with other models from the literature.

The remainder of this work is organized as follows. Section 2 describes the main Brazilian legislation related to GHG emissions; Section 3 describes recent machine learning models applied to climate change. Section 4 presents the methodology used in ML4GHG Forecasting. Section 5 details the experiment results, and Section 6 details the evaluation analysis. Finally, Section 7 presents the conclusion and indicates future works.

2. BRAZILIAN LEGISLATION

This section summarizes the main Brazilian legislation related to climate change to create incentives for renewable energy, preserve ecosystems, and guide Brazil toward zero emissions by 2050. The main federal and state laws enacted in the recent years are:

- Law 12114 of December 9, 2009 (BRAZIL 2009a) establishes the National Fund on Climate Change (FNMC) to finance studies and measures to reduce and repair the effects of climate change;
- Law 12187 of December 29, 2009 (BRAZIL 2009b) establishes the National Policy on Climate Change (PNMC), providing the institutional framework and outlines to mitigate the global warming impacts, preserve and restore ecosystems, and promote sustainable development;
- Law 12305 of August 2, 2010 (BRAZIL 2010) establishes the National Policy on Disposal of Solid Waste, which is essential for the reduction of CH_4 emitted in the final disposal of organic waste;
- Law 14300 of January 6, 2022 (BRASIL 2022a) establishes tariff benefits for distributed small-scale renewable electricity production;
- Brazil's Constitution was amended to mandate that bio-fuels (BRAZIL 2022b) and low-carbon hydrogen (BRAZIL 2023) be taxed less than fossil fuels;
- Several state laws including (RJ 2015), provide tax reductions or exemptions for electric, hybrid or natural gas-powered vehicles; and
- The legislation regarding the limit of pollutant emissions by motor vehicles is increasingly strict (CNMA 2018).

The Brazilian Congress is analyzing the bill PL 412/2022 (BRAZIL 2022c), which regulates the trade of carbon credits in Brazil. The approval of this bill could create legal instruments to encourage transactions with carbon assets, contributing to the country's sustainable development. Other bills that provide relevant changes are PL 639/2015 (BRAZIL 2015) and PLS 302/18 (BRAZIL 2018), which create incentives for the generation of energy in landfills resulting in large capture and utilization of CH_4 for electricity generation.

3. RELATED WORKS

This section describes traditional methodologies and state-of-the-art recent machine learning models for time series forecasting applied to climate change and GHG emissions.

We can find several approaches in the literature for time series forecasting using statistical methods. In the last decade, deep learning models have gained attention in Natural Language Processing, computer vision, and sequential signal processing and have successfully been applied to time series forecasting. More recently, transformer-based methods have presented as a promising option when the computational resource and financial cost are flexible. Alternatively, a meta-learning-based model has been used to help time series model generalization with simple architecture and low computational resource consumption.

Statistical methods designed for time series, such as Auto-Regressive models (AR) and their extensions Auto-Regressive Moving Average (ARMA) and Auto-Regressive Integrated Moving Average (ARIMA) (Hillmer & Tiao 1982) have been successfully used. Time series forecasting relies on stationary data to efficiently predict the future behavior of a certain feature. However, real-world data may have trends or seasonality, which needs to be prepossessed and transformed into stationary data. The prepossessing becomes very challenging in case of disruption of a trend or seasonality in the data, such as an unprecedented global threat: COVID-19. The CO_2 emissions in 2020 were at their lowest level (Meng & Noman 2022) compared to prior decades, causing a never seen disruption in the data trend. Meng & Noman (2022) used a statistical approach with Seasonal Auto-regressive Integrated Moving Average (SARIMA) to forecast the global CO_2 emission in China for the post-COVID-19 period. COVID-19 had a strong impact on the air quality around the world. Gupta *et al.* (2023) used SARIMAX (SARIMA modeling with exogenous factor) to predict air quality improvement in India during the nationwide lockdown imposed by the COVID-19 pandemic. Teggi *et al.* (2020) proposed InFORM to forecast the daily weather (temperature, humidity, and visibility) in Bangalore (India) using the ARIMA statistical method.

In recent years, deep learning models, such as Long Short-Term Memory (LSTM) (Hochreiter & Schmidhuber 1997), Recurrent Neural Networks (RNN), and Convolutional Neural Networks (CNN) have been used for time series forecasting in the climate change domain. Kumari & Singh (2023) compared LSTM with statistical models (ARIMA and SARIMAX), classical machine learning models (Linear Regression and Random Forest) for CO_2 forecasting in India. LSTM model outperformed statistical models in CO_2 forecasting (Kumari & Singh 2023) and outperformed Prophet – an additive regression model – in air temperature forecasting in Indonesia (Haris *et al.* 2022).

Climate change has caused an impact on the agricultural sector, which is very sensitive to the weather and temperature oscillation. Alex & Sobin (2021) applied LSTM and ARIMA for temperature forecasts to help harvest planning and reduce losses in the agricultural sector. In response to global warming, efforts to convert the energy matrix to renewable energy, including wind and solar are growing. Almalaq *et al.* (2021) used RNN, LSTM, and GRU (Cho *et al.* 2014) for solar and wind energy forecasts, helping supply-demand energy planning in Saudi Arabia.

Deep learning-based models often need a large amount of training data to learn the complex relationship between past and future time series data. To address this problem, we can apply meta-learning. Mo *et al.* (2023) used Model-Agnostic Meta-Learning (MAML) (Finn *et al.* 2017) with parameter initialization and Euclidean distance for similarity matching to forecast the remaining useful life of mechanical equipment. Thi Kieu Tran *et al.* (2020) used a meta-learning-based genetic algorithm for hyperparameter optimization of deep learning models to perform temperature forecasting. Reptile (Nichol *et al.* 2018), the first-order meta-learning algorithm, can be used to improve time series forecasting. Tian *et al.* (2021) leverage Reptile with transfer learning, Leelakittisin & Sun (2021) and Gupta & Raghav (2020) applied Reptile to a CNN-based model.

Driven by the country's low-carbon agenda, we also found several works using machine learning models in the context of climate change in Brazil. Hydroelectricity represents around 60% of the electricity generated in Brazil, and extreme weather events impact the country's main source of electrical energy. De Toledo *et al.* (2023) applied a classical machine learning model (Random Forest, Support Vector Regression, Kernel Ridge Regression) and statistical model (SARIMAX) for stream-flow prediction based on climate indices. Galvão Filho *et al.* (2020) proposed an LSTM-based model for water flow forecasting used in a hydroelectric power plant in Rondonia state of Brazil.

The difference between our model and the aforementioned works are: (i) Meng & Noman (2022), Gupta *et al.* (2023), Teggi *et al.* (2020), and De Toledo *et al.* (2023) used statistical methods while our model uses meta-learning approach; (ii) Kumari & Singh (2023), Alex & Sobin (2021), Haris *et al.* (2022), Galvão Filho *et al.* (2020), Almalaq *et al.* (2021) proposed deep learning-based models while we used deep learning model combined with meta-learning; (iii) Mo *et al.* (2023) used MAML, Thi

Kieu Tran *et al.* (2020) used genetic algorithm while we used Reptile for meta-learning-based model optimization; Tian *et al.* (2021) leverage Reptile with transfer learning whereas we trained the model from scratch; Leelakittisin & Sun (2021) and Gupta & Raghav (2020) applied Reptile to a CNN-based model while we used a Bidirectional Long Short-Term Memory (BiLSTM)-based model and Fusion Model for multivariate data processing.

4. METHODOLOGY

This section presents the methodology used for GHG emission forecasting in Brazil, called ML4GHG – **M**eta-**L**earning Applied to a Multivariate Single-Step Fusion Model **for G**reen**h**ouse **G**as Emission Forecasting. ML4GHG aims to analyze Brazil's CO_2 and CH_4 emissions over the past 60 years. Therefore, ML4GHG is based on a multivariate single-step approach leveraging the combination of the Fusion Model for data alignment and optimization-based Meta-Learning tailored for time series. Moreover, ML4GHG adapted the Reptile Meta-Learning algorithm to improve the model. In the next sections, we will describe ML4GHG in more depth.

4.1. ML4GHG overview

The overview of ML4GHG architecture is illustrated in Figure 1. ML4GHG is a multivariate single-step parallel time series forecasting model in which two variables are observed at each time step. Concretely, the model takes two features (CO_2 and CH_4) as input data and predicts two features (CO_2 and CH_4) in parallel as output data. As a single-step model, it forecasts the next time step based on previous values.

First, we used EDGARv8.0 as a multivariate time series input data for GHG emission prediction. Next, we performed three tasks in the data pre-processing: (i) data analysis to understand the trend and seasonality; (ii) future scaling for data normalization; and (iii) time frame splitting.

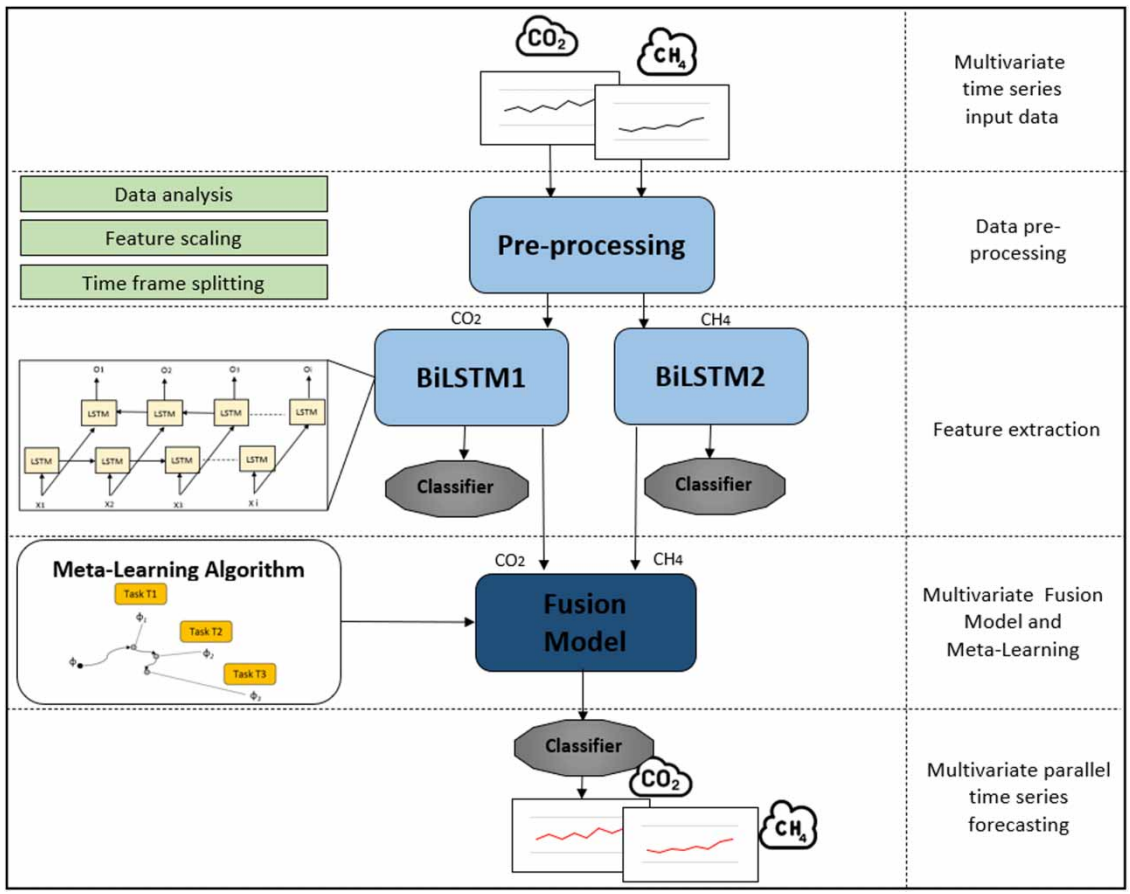


Figure 1 | Meta-learning applied to multivariate single-step fusion model for greenhouse gas emission forecasting.

After the data preprocessing, the BiLSTM1 model performs the univariate feature extraction for CO_2 and BiLSTM2 for CH_4. Once the input features are extracted separately, the multivariate Fusion Model performs the CO_2 and CH_4 data alignment, and the meta-learning algorithm helps in the model generalization. Finally, the multivariate parallel time series forecasting is performed by the Fusion Model's classifier, which predicts the next time step values for CO_2 and CH_4.

To facilitate the understanding, Table 1 summarizes the notations used in ML4GHG.

4.2. Multivariate dataset

The multivariate time series data can be represented as $T = \{t_1,...,t_n\}$, in which n is the number of elements in the time series and t_i are the values measured at time step i. Each t_i contains the tuple $\{(x_{1i}, x_{2i}, .., x_{ji})\}_{j=1}^{m}$ where x_{1i} represents the value of the first feature at time step i, and m is the number of features.

The dataset used in our work is EDGARv8.0. It is a publicly available GHG (CO_2, CH_4, N_2O, and F-gases) emissions database for global atmospheric research reported by the European Member States and by Parties under the United Nations Framework Convention on Climate Change (UNFCCC). The dataset provides annual and monthly GHG emissions data for the time span of 1970–2022 by country (Crippa *et al.* 2021). We used the monthly data from Brazil, which contains CO_2 and CH_4 emissions for 636 consecutive months from 1970 to 2022.

4.3. Data preprocessing

Data preprocessing plays an important role in time series forecasting to help the model capture the trend and seasonality of data, avoid gradient spikes, and handle missing data. The data preprocessing comprises the following steps: data analysis, feature scaling, and time frame splitting.

We used three graphs for data analysis: auto-correlation, seasonal decomposition, and residual graphs (Kumari & Singh 2023). Figure 2 illustrates the auto-correlation graph for CO_2 and CH_4 considering the past 50 steps or months, i.e., lag = 50. In the CO_2 graph, the blue region represents values that have no significant correlation with the most recent value of CO_2. The vertical lines in the *y-axis* represent the correlation of CO_2 with the previous values, in which values near 1

Table 1 | Summary of notations

Notations	Description
T	Multivariate time series data
n	Number of elements in T
t_i	Values measured at time step i
x_{ji}	Value of the jth feature at time step i
m	Number of features
	Fusion model
Z	Concatenation operation
$x1i$	Represents CO_2
$x2i$	Represents CH_4
$f\varphi$	Function that returns CO_2 embedding vector
f_ϑ	Function that returns CH_4 embedding vector
	Meta-learning
ϕ	Fusion model weights
ε	Step-size of meta-learning
S_{tr}	Multivariate fusion training set
S_{ts}	Multivariate fusion test set
$\tilde{\phi}$	Fusion model new weights
l	Fusion model loss
t	Loss threshold

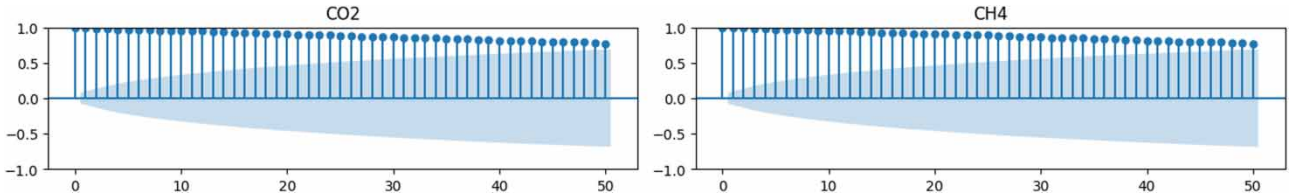


Figure 2 | Auto-correlation of CO_2 and CH_4.

represent a high correlation. The *x-axis* represents the time step in months. The graph shows that previous values of CO_2 have a high influence on the current value, but the significance of that influence decreases steadily with time.

The auto-correlation graph of CH_4 can be visualized in the second graph of Figure 2. Similarly, to the CO_2, the graph shows that only the previous few time steps values of CH_4 have a high influence on the current value.

Figure 3 illustrates the seasonal decomposition and the residual for CO_2 and CH_4 considering the entire series, i.e., 636 months. Seasonal decomposition captures the repetitive patterns and cycles within the time series. The *y-axis* represents the normalized CO_2 and CH_4 values, and the *x-axis* represents the 636 consecutive months. We can observe that both CO_2 and CH_4 have irregular cycles in the seasonal decomposition. The residual graph represents unexpected variation that does not follow a trend or seasonality. The *y-axis* represents the residual error, which ideally should be around zero. The residual graph of CO_2 and CH_4 shows periods of unexpected variation represented by the points in the *y-axis* different from zero. CO_2 and CH_4 are expressed in Kton substance/month unit with a high range of variation. We used standard normalization to minimize the effects of this variation in the neural network's gradient. In order to avoid data leaks, first the dataset was divided into training and test sets. Then we applied standard normalization to the training set, and was scaled on the test set. The EDGARv8.0 dataset does not have missing values, and no data interpolation was needed.

The last step in our multivariate single-step parallel time series forecasting data pre-processing is the time frame splitting. We used the sliding window considering the prior three time steps (lag = 3) to predict the next time step (single-step). We considered lag = 3 because the auto-correlation graph of both substances demonstrated that the current value is highly correlated only with the previous few time steps.

4.4. Feature extraction

After the data pre-processing, the CO_2 and CH_4 normalized data are extracted to produce the corresponding embedding vector. Each extractor model is an independent BiLSTM network with its own classifiers and trained as univariate time series data, as illustrated in Figure 1. BiLSTM is effective in many application areas, such as Natural Language Processing (Enamoto *et al.* 2022; Costa *et al.* 2023; Gou & Li 2023) and time series forecasting. BiLSTM comprises forward and backward LSTM (Hochreiter & Schmidhuber 1997). LSTM in turn detects an important feature from the input sequence in the early stage and transmits the information over a long distance, thus capturing potential long-term dependencies (Zrira *et al.* 2024). In time series forecasting, BiLSTM helps to capture the context of past and future time steps (Schuster & Paliwal 1997). In our model, the BiLSTM layer is followed by a Time-Distributed Layer and three Dense layers, which one is a customized Dense layer that produces the embedding of the univariate time series to be used in the Fusion Model.

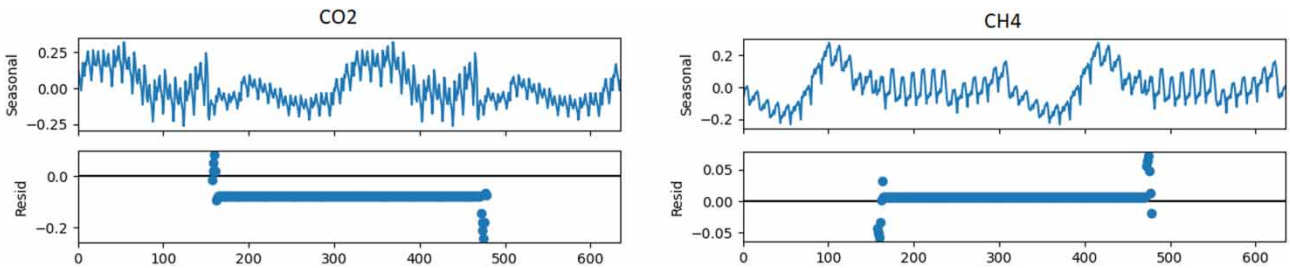


Figure 3 | Seasonal decomposition and residual plot of CO_2 and CH_4.

Both BiLSTM extractors have the same architecture composed of the following layers, as illustrated in Figure 4: (i) BiLSTM layer activated by RELU function; (ii) Time-Distributed layer that applies the same weights to the output of the previous layer for one time step at a time; (iii) Flatten layer; (iii) Dense layer activated by Relu function; (iv) Customized Dense layer that extracts the univariate embedding; (v) Dropout layer to avoid overfitting; and (vi) the last Dense layer that learns the regression classifier. The Adam optimizer was responsible for minimizing the mean squared error (MSE) loss between the true value and the predicted value.

4.5. Fusion Model

The Fusion Model performs the alignment or fusion of the heterogeneous CO_2 and CH_4 data. The goal of the Fusion Model is to create an abstraction of the unified representation of different features for each tuple t_i in $T = \{t_1,...,t_n\}$ and perform one or more tasks efficiently. In this process, the heterogeneous data need to be integrated to find the relationship between them, known as data fusion (Baltrušaitis *et al.* 2018).

In the literature (Baltrušaitis *et al.* 2018), we can find three types of data fusion methods: (i) late fusion or decision-level fusion in which each feature is individually extracted and trained based on a specific decision task and then integrated into the same feature space; (ii) early fusion or feature-level fusion which exploits the low-level features just after the extraction, creating a strong interaction between modalities (Wang *et al.* 2024); and (iii) hybrid fusion or intermediate-level fusion which learns a joint representation of different features by combining the decision-level and feature-level fusion.

We applied the decision-level fusion in our model. As illustrated in Figure 1, each BiLSTM extractor and the Fusion Model have its own classifier (Baltrušaitis *et al.* 2018). The advantage of this fusion type is that it enables separate calibration according to the data quality.

The Fusion Model is illustrated in Figure 5 comprising the following layers: (i) input layer that concatenates the embedding features extracted by BiLSTM1 and BiLSTM2 models; (ii) batch normalization layer responsible for normalizing the input of each layer for every mini-batch; (iii) followed by two Dense layers activated by Relu function; (iv) Customized Dense layer that extracts the multivariate embedding; (v) and the last Dense layer activated by Swish function to learn a third regression classifier.

Equation (1) details the concatenation operation, in which i is the i_{th} element of the dataset, x_{1i} represents the CO_2 and x_{2i} is the CH_4. The function f_φ returns the features learned by the BiLSTM1 model representing the CO_2 embedding, and the

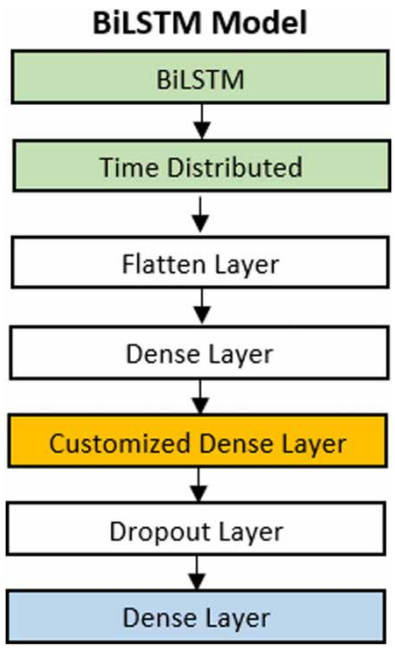


Figure 4 | BiLSTM model for time series data extraction.

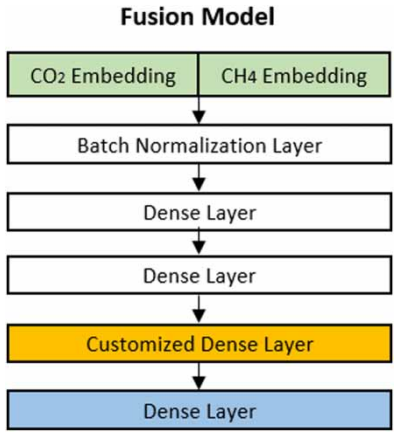


Figure 5 | Fusion Model.

function f_ϑ returns the features learned by BiLSTM2 representing the CH_4 embedding.

$$Z = concat(f_\varphi(x_{1i}), f_\vartheta(x_{2i})) \tag{1}$$

4.6. Multivariate Meta-Learning

Time series forecasting can be challenging in case of disruption of a trend or seasonality in the data. One possible option to address this issue is to use Meta-Learning to optimize the model's learning capabilities (Enamoto *et al.* 2023).

In our work, we adopted the flexibility of Reptile (Nichol *et al.* 2018): a gradient-based Meta-Learning approach. Meta-Learning helps the underlying model learn from past experience, adapt, and generalize according to the new task (Enamoto *et al.* 2023). The details of the Reptile are described in Algorithm 1.

Algorithm 1 Reptile-based multivariate time series Meta-Learning

1:	Initialize Fusion Model's weights ϕ
2:	Initialize meta step-size ε
3:	Construct multivariate fusion training set S_{tr}
4:	Construct multivariate fusion test set S_{ts}
5:	**for** each meta-iteration **do**
6:	Calculate $\tilde{\phi}$ by SGD to ϕ on S_{tr}
7:	Predict the next time step with S_{ts}
8	**if** loss $l <$ threshold t **then**
9	**exit for**
10	**else**
11	Update $\phi \leftarrow \phi + \varepsilon(\tilde{\phi} - \phi)$
12	Adjust ε
13	**end if**
14	**end for**

First, the weights ϕ of the Fusion Model are randomly initialized, and the meta step-size ε is initialized with a fixed value (lines 1 and 2). The multivariate fusion training set S_{tr} and the test set S_{ts} are generated (lines 3 and 4). In the meta-iteration loop, the new weights $\tilde{\phi}$ are computed by Stochastic Gradient Descent (SGD) using the training set S_{tr} (line 6). The test set S_{ts} is used for prediction (line 7) and if the loss l is greater than the threshold t, the weights ϕ are updated moving ϕ closer to the

optimal value (lines 11 and 12). In practice, we add a little perturbation in the weights ϕ, and after a few meta-iterations, the weight update helps in the model generalization.

4.7. Model training

Applying the Reptile algorithm, the Fusion Model keeps its knowledge by updating the weight ϕ under the guidance of ε and acquires new knowledge in turn. This mechanism helps the Fusion Model adapt for a time series forecasting where an unexpected disruption in the data trend or seasonality may lead to an undesired outcome.

Before training the model, we used grid search to obtain the best hyperparameter combination, as detailed in Table 2. The first column represents the hyperparameters and the second column the corresponding values used in the grid search. In the pre-processing phase, the best results were obtained with standard scaling, which was done separately, first on the training set and then scaled on the test set to avoid data leakage. We compared LSTM and BiLSTM for univariate data extractor and BiLSTM resulted in the best performance. After the univariate feature extraction, two data alignments were tested and concatenation was the method that best preserved the univariate sequence values.

After obtaining the best values for the hyperparameter, we trained the model using the configuration detailed in Table 3. The first column represents the model, 'Loss' represents the objective function and 'Learn. Rate' represents the learning rate of each model. 'Parameters' is the number of trainable parameters, 'Batch Size' is the number of batches used for training, and 'Meta-Iteration' is the number of meta-learning repetitions using the same training/test data split. 'Loss Thres.' is the loss value used as the threshold before stopping the iteration, and 'Meta Step-size' is the initial value of epsilon used to update the Fusion Model's weights for better convergence.

5. MODEL EVALUATION

This section describes the evaluation results of ML4GHG using the EDGARv8.0 dataset. The cross- validation strategy is described in Subsection 5.1, and the results of Meta-Learning Applied to Multivariate Single-Step Fusion Model for GHG emissions forecasting are described in Subsection 5.2.

Table 2 | Hyperparameters values used in the grid search

Parameter	Values
Data normalization	{Standard, MinMax}
Feature extraction	
Model	{BiLSTM, LSTM}
Batch size	{32, 64, 128}
Epoch	{100, 200, 300, 400, 500}
Learning rate	{0.0001, 0.0005, 0.001, 0.003}
Dropout	{0.3, 0.5, 0.7}
Optimizer	{Adam, RMSprop}
Fusion Model	
Data alignment	{concatenation, average}
Batch size	{32, 64, 128}
Epoch	{100, 200, 300}
Learning rate	{0.0001, 0.0005, 0.001, 0.003}
Dropout	{none, 0.3, 0.5, 0.7
Optimizer	{adam, rmsprop, swish}
Meta-Learning	
Meta step-size	{0.15, 0.25, 0.35, 0.45}
Meta-iteration	{5, 10}
Loss threshold	{0.02, 0.03, 0.03, 0.04}

Table 3 | Main hyperparameters settings of ML4GHG. MSE (Mean Square Error), K (thousand)

Model	Loss	Optimizer	Learn. rate	Drop out	Parameters	Batch size	Meta-iteration	Loss thres.	Meta step-size
BiLSTM1	MSE	Adam	0.0001	0.3	138 K	32	–	–	–
BiLSTM2	MSE	Adam	0.0001		138 K	32	–	–	–
Fusion model	MSE	Swish	0.0005		37 K	32	–	–	–
Meta-Learning	–	–	–			10	0.36	0.45	

5.1. Cross-validation strategy

In time series forecasting, the temporal sequence of data needs to be preserved so the model learns the relationship between data from the current and previous time steps. We use the time series splitting method for the cross-validation strategy, in which we use continuous time blocks of different durations for training. We divided the EDGARv8.0 dataset into ten blocks of different time steps to make up the training set and a fixed length of 24 time steps (months) for the test set.

The data split of CH_4 can be visualized in Figure 6, which shows the entire monthly data of CH_4 emissions in Brazil. The *y-axis* represents CH_4 emissions expressed in Kton substance and the *x-axis* represents the monthly time span from January 1970 to December 2022. In this first split for model evaluation, training data goes from January 1970 to December 2002, and values from January 2003 to December 2004 are used for test data. Figure 7 illustrates the data division for CO_2 following the same time-division for CH_4. In the second data split, the training set window moves forward 24 months, from January 1970 to December 2004, and values from January 2005 to December 2006 are used for testing.

The details of the data split are described in Table 4. For example, in the split 1, 396 continuous time steps or months are used to train the model and the following 24 time steps for testing. In split 2, 420 continuous time steps are used to train and the following 24-time steps for testing. This way, after executing the ten splits, the model is evaluated using different continuous non-overlapping blocks of months. In the split 10, 612 continuous months were used for training and the last 24 months for testing, totaling all the 636 months that compose the EDGARv8.0 dataset.

5.2. Results

In this Subsection, the cross-validation results of ML4GHG with the EDGARv8.0 dataset are reported. Next, we compare the results of ML4GHG with two deep learning-based models and five recent time series forecasting models.

The details of the cross-validation results are described in Table 5. The second column 'Evaluation Period' represents the period of test data, and the following columns are the regression error metrics: MSE (Mean Square Error), MAE (Mean Absolute Error), MAPE (Mean Absolute Percentage Error), and RMSE (Root Mean Squared Error). All metrics represent

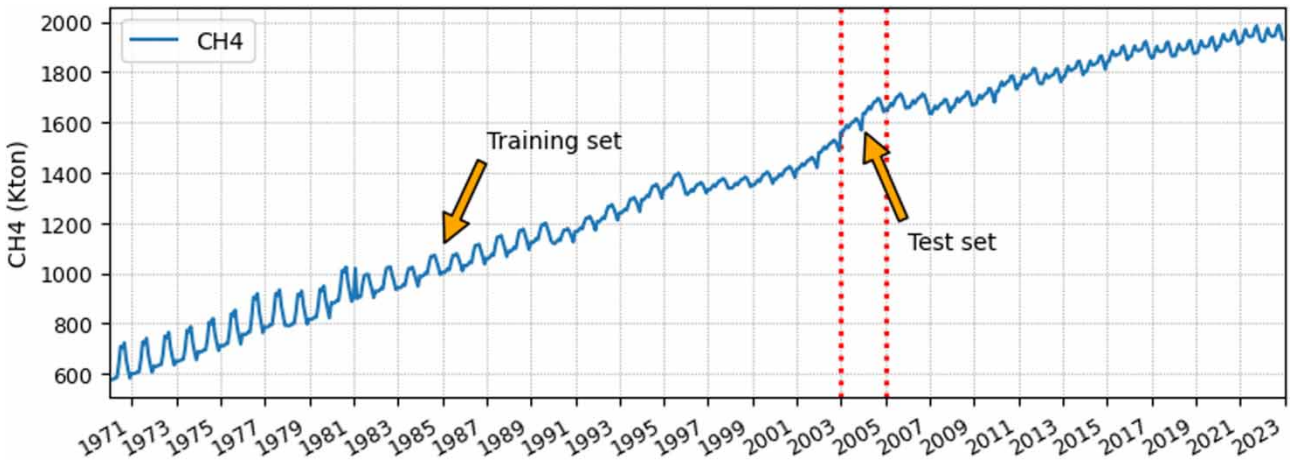


Figure 6 | CH_4 monthly emission of Brazil from 1970 to 2022 (EDGARDv.8.0 dataset).

Figure 7 | CO_2 monthly emission of Brazil from 1970 to 2022 (EDGARDv.8.0 dataset).

Table 4 | EDGARv8.0 dataset cross-validation data split

Data split	Training timestep	Test timestep
1	396	24
2	420	24
3	444	24
4	468	24
5	492	24
6	516	24
7	540	24
8	564	24
9	588	24
10	612	24

Table 5 | Results of cross-validation for each data split

Data Split	Evaluation period	MSE	MAE	MAPE (%)	RMSE
1	2003-Jan → 2004-Dec	0.022	0.115	5.691	0.141
2	2005-Jan → 2006-Dec	0.016	0.104	5.179	0.126
3	2007-Jan → 2008-Dec	0.022	0.113	5.507	0.139
4	2009-Jan → 2010-Dec	0.021	0.107	5.598	0.132
5	2011-Jan → 2012-Dec	0.022	0.112	4.928	0.134
6	2013-Jan → 2014-Dec	0.030	0.111	4.578	0.151
7	2015-Jan → 2016-Dec	0.022	0.098	5.265	0.132
8	2017-Jan → 2018-Dec	0.017	0.104	6.300	0.126
9	2019-Jan → 2020-Dec	0.020	0.121	8.578	0.143
10	2021-Jan → 2022-Dec	0.013	0.093	6.364	0.10

Note: MSE, mean square error; MAE, mean absolute error; MAPE, mean absolute percentage error; RMSE, root mean squared error.

the comparison error between the actual and the estimated values of CO_2 and CH_4, which means that the smaller the error, the better the model.

We can observe in Table 5 that MAPE achieved the highest value (8.578%) during the COVID-19 pandemic represented by data split 9. This outcome suggests that the model struggled to estimate CO_2 and CH_4 emissions due to the COVID-19 pandemic's impact on the global economy.

Table 6 details the comparison of our model with two baseline models and five recent time series forecasting models. Our model achieved the lowest MAPE of 5.799% with 1.06% of standard deviation. This result is significant in the interval of (5.030, 6.555), at 95% confidence. All models were trained as a multivariate parallel single-step using the previous three time steps (lag = 3) to predict the next time step. In addition, for a fair comparison, we used the same data split of the EDGARv8.0 dataset described in Table 4 to assess the seven models.

We compared the ML4GHG results with two baseline models: 2 layers of LSTM and 1 layer of BiLSTM. ML4GHG outperformed both deep learning models on all error metrics, where MAPE was 89.70% lower than LSTM and 116.84% lower than BiLSTM. Next, we compared our model with five recent time series forecasting models: (i) Neural Basis Expansion Analysis – N-Beats (Oreshkin *et al.* 2020); (ii) Temporal Fusion Transformer – TFT (Lim *et al.* 2021); (iii) Neural Hierarchical Interpolation for Time Series – NHiTS (Challu *et al.* 2023); (iv) Linear model – N-Linear (Zeng *et al.* 2023); and (v) Time series Dense Encoder – TiDE (Das *et al.* 2023). ML4GHG outperformed all recent models considering MSE, MAE, and MAPE errors, except for RMSE, which was 3.87% higher than TFT. It is worth noting that our model presented the lowest standard deviation of all error metrics, suggesting stability in the result. The comparison of MSE, RMSE, and MAE can be visualized in Figure 8, and MAPE can be visualized in Figure 9.

Table 6 | Comparison of ML4GHG with different models at 95% of confidence interval

Model	Method	MSE	MAE	MAPE (%)	RMSE
BiLSTM	1L BiLSTM	0.097 ± 0.05	0.250 ± 0.06	12.575 ± 2.25	0.278 ± 0.06
LSTM	2L LSTM	0.073 ± 0.04	0.213 ± 0.06	11.001 ± 2.46	0.247 ± 0.07
N-Beats (2020)	Residual links	0.034 ± 0.04	0.129 ± 0.08	13.659 ± 8.55	0.146 ± 0.09
TFT (2021)	Transformer	0.021 ± 0.01	0.110 ± 0.04	11.384 ± 4.56	**0.129 ± 0.05**
NHiTS (2023)	Hierarchical interpolation	0.075 ± 0.03	0.187 ± 0.07	23.898 ± 7.77	0.211 ± 0.08
N-Linear (2023)	Linear model	0.074 ± 0.16	0.147 ± 0.16	15.864 ± 19.15	0.172 ± 0.20
TiDE (2023)	Encoder decoder	0.038 ± 0.06	0.114 ± 0.08	12.252 ± 9.62	0.141 ± 0.12
ML4GHG	**Meta-learning**	**0.021 ± 0.01**	**0.108 ± 0.01**	**5.799 ± 1.06**	0.134 ± 0.01

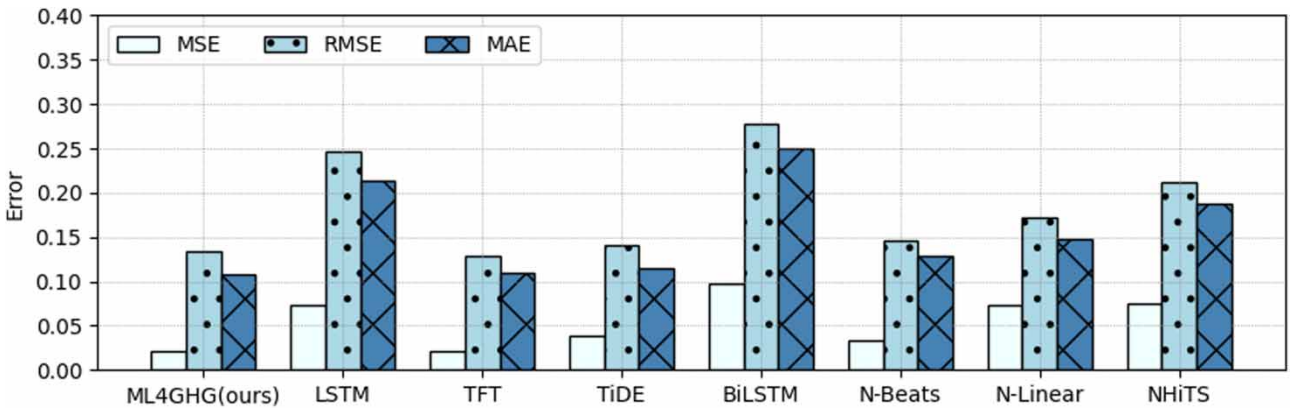


Figure 8 | MSE, RMSE, and MAE comparison between ML4GHG and other models.

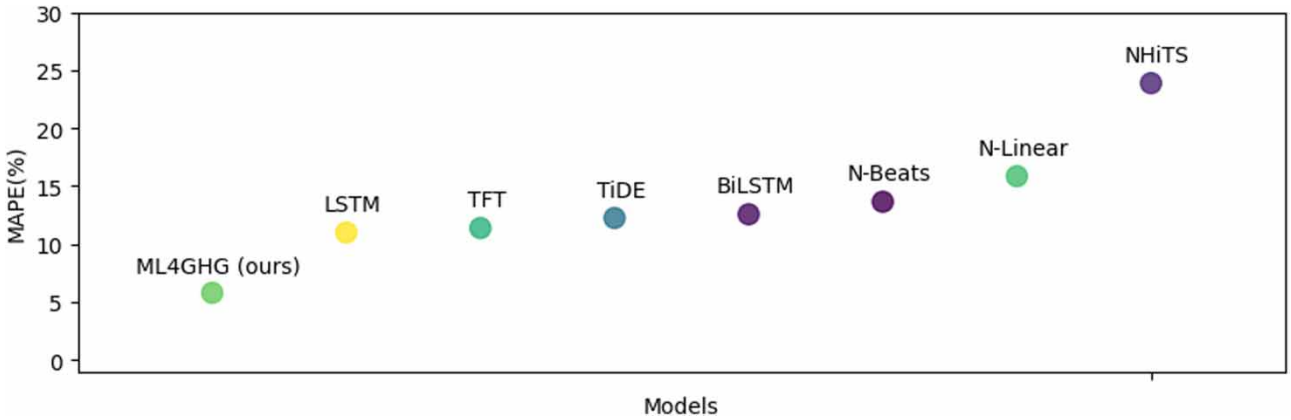


Figure 9 | MAPE comparison between ML4GHG and other models.

The prediction of CH_4 and CO_2 can be visualized in Figure 10, which represents the evaluation of the ML4GHG model for the first data split. The *y-axis* represents the normalized substance value, and the *x-axis* represents the test set period. The full line is the actual value, and the dotted line is the forecast value. We can observe that from January 2003 to December 2004, the predicted value of both substances follows the growth trend of actual value. However, as illustrated in Figure 11 and Figure 12, the distance between the actual and predicted values increases, reflecting the interruption and oscillations in the GHG emissions trend due to COVID-19 effects.

6. EVALUATION ANALYSIS

This section describes the analysis of our model results to verify the contribution of individual components. We performed ML4GHG ablation analyses in Subsection 6.1, check the effects of Meta-Learning in Subsection 6.2, and detailed the consumption of computational resources by estimating the CO2 emission related to the experiments in Subsection 6.3.

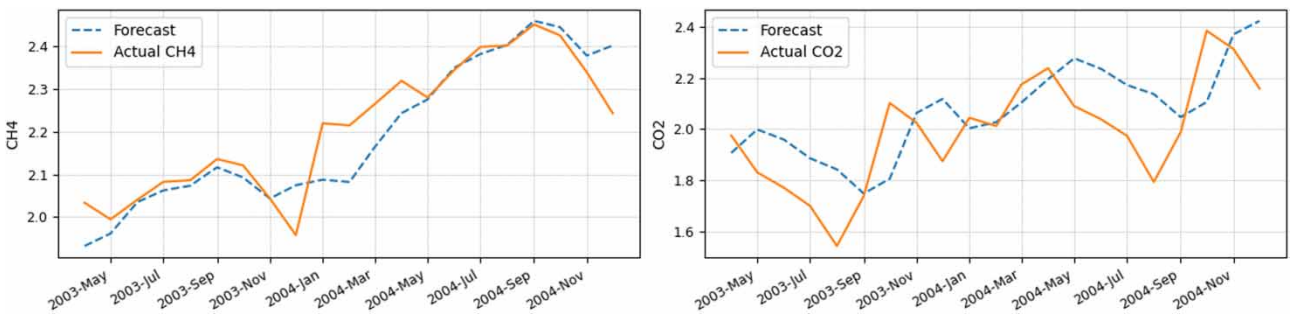


Figure 10 | CH_4 and CO_2 prediction of 2003-January to 2004-December.

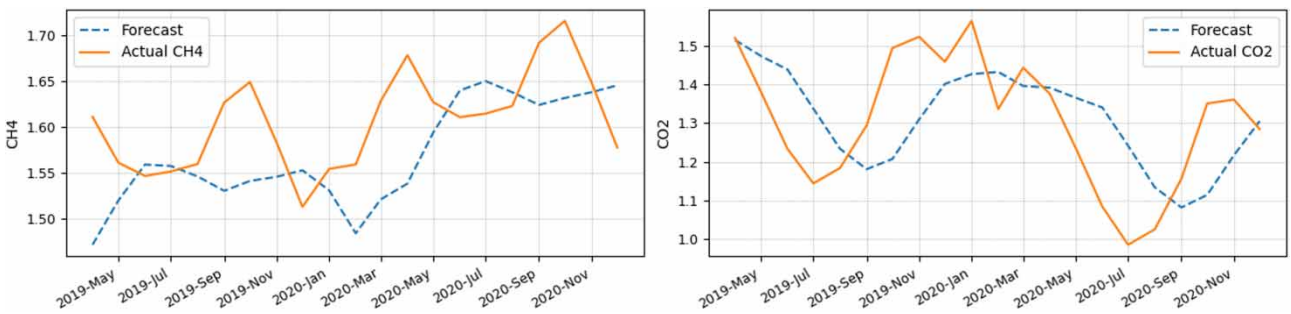


Figure 11 | CH_4 and CO_2 forecasting during COVID-19.

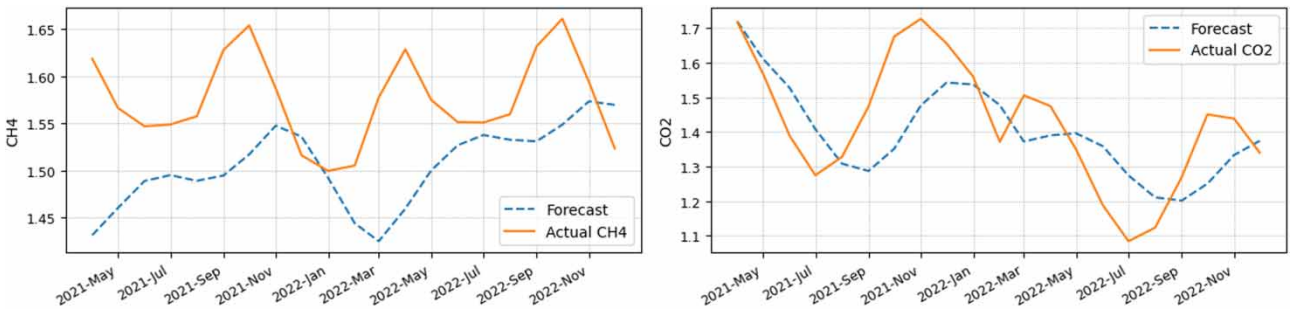


Figure 12 | CH_4 and CO_2 forecasting during and post-COVID-19 period.

6.1. Ablation analysis

In this Subsection, we performed three ablation analyses to evaluate the impact of individual components in the model: (i) the impact of the Fusion Model and Meta-Learning (Ab1); (ii) the impact of Meta-Learning (Ab2); and (iii) the impact of feature extractor (Ab3). The models used in the ablation were evaluated using the same cross-validation strategy of the EDGARv8.0 dataset described in Subsection 5.1.

The first ablation analysis (Ab1) was conducted by eliminating the Fusion Model and the Meta-learning. Then, replacing the univariate embedding extractor (BiLSTM) with multivariate BiSLTM. In this ablation, the model forecasts CO_2 and CH_4 at once using the multivariate BiLSTM. In the second ablation (Ab2), we eliminated the Meta-Learning algorithm and kept the Fusion Model. In the third ablation (Ab3), we replaced the BiLSTM feature extractor with LSTM.

The results of the ablation analysis using the EDGARv8.0 dataset are detailed in Table 7. The results are at a 95% confidence interval. The ablation type that most influenced the results was eliminating the Fusion Model and Meta-Learning (Ab1) with 12.57% MAPE, with 116.8% higher error than our model. We observed a severe degradation in the model, suggesting that the Fusion Model helps the model learn a better multivariate embedding for time series forecasting, and the Meta-Learner algorithm helps the model generalization.

The second ablation with high impact was the Fusion Model without Meta-Learning (Ab2) with 8.74% MAPE, a 50.75% higher error than our model. This result suggests that the Meta-Learner helps to avoid a lack of generalization in case of disruption of a trend in the time series data.

The impact of replacing BiLSTM with the LSTM feature extractor (Ab3) was 8.28% MAPE, representing an error 42.81% higher than our model. This result suggests that for multivariate data, BiLSTM can better capture the relationship between future and past time series data than LSTM.

The results of the ablation analysis can be visualized in Figure 13. The *y-axis* represents the MSE, and the *x-axis* represents the four models: ML4GHG is the model with all components followed by the models used in the ablation analysis 'Ab1', 'Ab2', and 'Ab3'. The points on the left-hand side of each box-plot represent the MSE for each run in the cross-validation. The ML4GHG presents the lowest median (0.021), while 'Ab1' has the highest median (0.071) with outliers, suggesting the model instability when the Fusion Model and Meta-Learning are eliminated. We can observe that ML4GHG presents fewer outliers and a less distributed box-plot compared to the other ablation versions, confirming the stability of the results when the model uses all the components.

Table 7 | Ablation analysis results at 95% confidence interval

Ablation type	Predicted features	MSE	MAE	MAPE (%)	RMSE
Ab1 – without fusion model and meta-learning	CO_2, CH_4	0.097 ± 0.05	0.250 ± 0.06	12.575 ± 2.25	0.278 ± 0.06
Ab2 – without meta-learning	CO_2, CH_4	0.048 ± 0.04	0.168 ± 0.07	8.742 ± 3.17	0.193 ± 0.07
Ab3 – LSTM feature extractor	CO_2, CH_4	0.044 ± 0.03	0.156 ± 0.06	8.282 ± 2.90	0.182 ± 0.06
ML4GHG	CO_2, CH_4	**0.021 ± 0.01**	**0.108 ± 0.01**	**5.799 ± 1.06**	**0.134 ± 0.01**

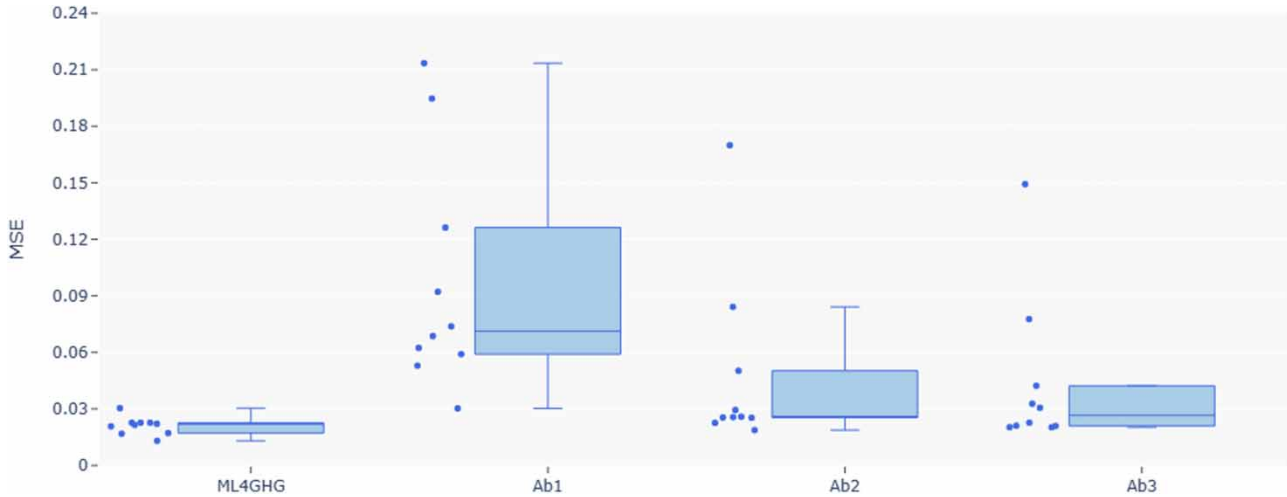


Figure 13 | Box-plot of ablation analysis.

6.2. Effects of Meta-Learning

Meta-Learning helps the underlying model learn from previous experience, resulting in task and model generalization. We used Reptile (Nichol *et al.* 2018): a gradient-based Meta-Learning approach to optimize the model learning capabilities.

Figure 14 illustrates the CH_4 and CO_2 forecast result of the first meta-iteration for January 2011 to December 2013 using all the components. The blue shadowed region in the chart represents the MSE error in the first meta-iteration. Considering the same period (January 2011–December 2013), we can visualize in Figure 15 the results of the meta-learning after performing a few meta-iterations. The shadowed region representing the forecasting error decreased notably for CH_4, confirming our finding in the ablation analysis without meta-learning (Ab2), described in Subsection 6.1.

6.3. CO_2 emissions related to the experiments

The number of machine learning models trained on the cloud providers has increased, which may collectively contribute to CO_2 emissions from data centers (Wu *et al.* 2022). Devising small machine learning models using few computational resources may contribute, even on a small scale, to reducing GHG emissions.

The estimations of carbon emissions were conducted using the Machine Learning Impact calculator (Lacoste *et al.* 2019). The CO_2-equivalents ($kgCO_2eq$) are used as a standardized measure to express how much warming a given amount of gas will have.

The ML4GHG has only 323 K parameters and was trained and evaluated in 30 minutes, considering all the cross-validation runs. The experiments were conducted using the Google Cloud Platform in the South America-east1 region, Intel(R) Xeon(R) CPU @ 2.20 GHz. The total emission is estimated to be 0.01 $kgCO_2eq$ for 0.5 hours of computation.

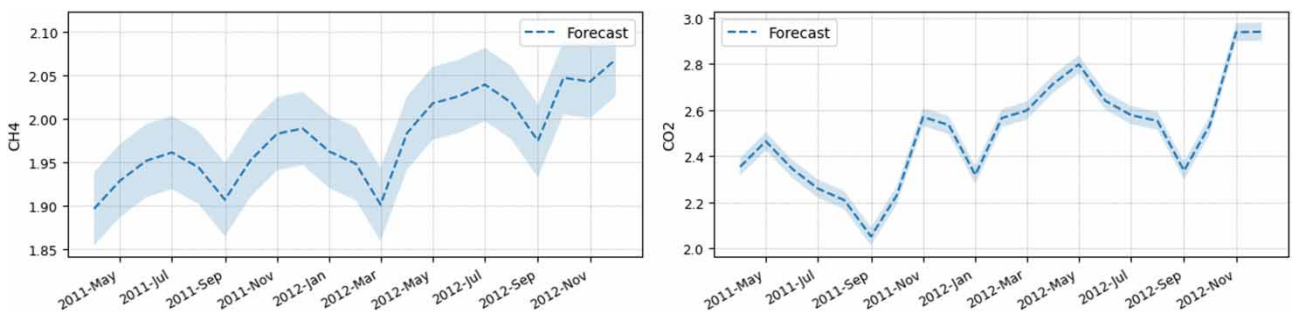


Figure 14 | CH_4 and CO_2 forecast – first meta-iteration.

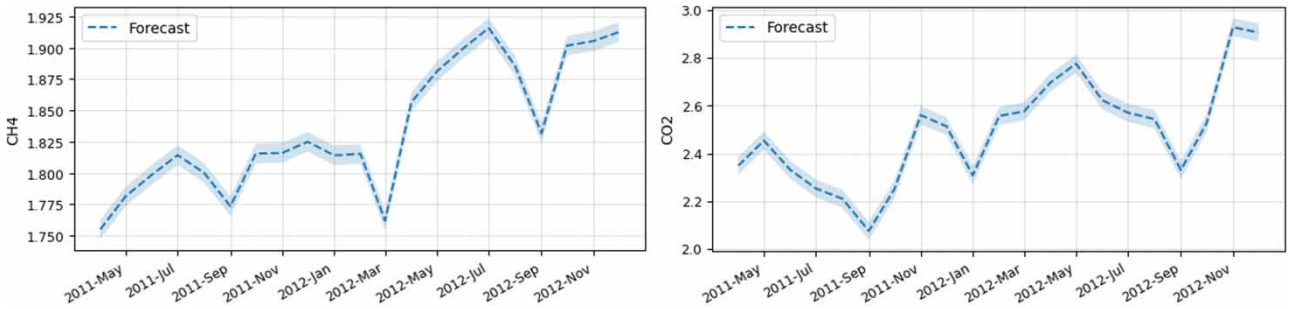


Figure 15 | CH_4 and CO_2 forecast – last meta-iteration.

7. CONCLUSIONS AND FUTURE WORKS

Despite the economic difficulties and climate impacts suffered in recent years, Brazil is committed to the low-carbon emission agenda. The geographic diversity and natural resources allow Brazil to have an electricity energy mix with 83% renewable in 2020.

From the Brazilian legislation perspective, Law 12,187 establishes the National Policy on Climate Change, and bill PL 412/2022 promotes debate to create a legal instrument for carbon credit trade regulation. In addition to legal measures, it is essential to monitor GHG emissions to keep contributing to the global low-carbon agenda. GHG forecasting can help to improve the sustainable agenda and mitigate the climate change impacts in the country.

This work proposes ML4GHG: a Meta-Learning Applied to a Multivariate Single-Step Fusion Model for GHG emission forecasting where two variables were observed: CO_2 and CH_4. These substances were extracted from the EDGARv8.0 dataset leveraging two BiLSTM models, one for each substance. Next, the multivariate Fusion Model performed the CO_2 and CH_4 data alignment, and then the Reptile algorithm provided the model generalization to the GHG forecasting.

The model was evaluated with two baseline models and five recent time series forecasting models. ML4GHG reduces MAPE by 49.06% with 95% confidence compared to the transformer-based TST model, demonstrating its superior performance and low estimated CO_2 emissions of 0.01 kg CO_2eq. After conducting an ablation analysis, it was found that removing the Fusion Model and Meta-Learning had a high impact on the model with a 116.8% increase in MAPE. Eliminating only the Fusion Model resulted in a 50.75% increase in MAPE while replacing the BiLSTM with LSTM resulted in a 42.81% increase in MAPE.

These results suggest that for multivariate data, the BiLSTM can better capture the relation between substances over time, the Fusion Model helps in the substances data alignment, and the Meta-Learner helps to avoid the lack of generalization in case of disruption in the time series data.

As a future work, we are investigating the use of the Large Language Model (LLM) and multimodal data in time series forecasting.

DATA AVAILABILITY STATEMENT

The dataset utilized in the experiment can be accessed at the following public repository: https://edgar.jrc.ec.europa.eu/dataset_ghg80.

CONFLICT OF INTEREST

The authors declare there is no conflict.

REFERENCES

Alex, N. & Sobin, C. C. 2021 A method for weather forecasting using machine learning. In *2021 5th Conference on Information and Communication Technology (CICT)*. IEEE, pp. 1–6.

Almalaq, A., Alshammarry, A., Alanzi, B., Alharbi, F. & Alshudukhi, M. 2021 Deep learning applied on renewable energy forecasting towards supply-demand matching. In: *2021 20th IEEE International Conference on Machine Learning and Applications (ICMLA)*. IEEE, pp. 1345–1349.

Baltrušaitis, T., Ahuja, C. & Morency, L. P. 2018 Multimodal machine learning: A survey and taxonomy. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **41** (2), 423–443.

BRAZIL 2009a Law no 12114, December 9, 2009. Establishes the National Policy on Climate Change Fund (FNMC). Available from: https://www.planalto.gov.br/ccivil_03/_ato2007-2010/2009/lei/l12114.htm (accessed 22 February 2024).

BRAZIL 2009b Law no 12187, December 29, 2009. Establishes the National Policy on Climate Change (PNMC). Available from: https://www.planalto.gov.br/ccivil_03/_ato2007-2010/2009/lei/l12187.htm (accessed 22 February 2024].

BRAZIL 2010 Law no 12305, August 2, 2010. Establishes the National Policy for Solid Waste. Available from: https://www.planalto.gov.br/ccivil_03/_ato2007-2010/2010/lei/l12305.htm (accessed 22 February 2024).

BRAZIL 2015 Bill n° PL 639/2015. Grants tax incentives to recycling and urban cleaning companies that have a power plant generating electricity from urban solid waste. Available from: https://www.camara.leg.br/proposicoesWeb/fichadetramitacao?idProposicao=994619 (accessed 22 February 2024).

BRAZIL 2018 Bill n° PLS 302/2018. Encourages companies that produce biogas, methane, and electricity from solid waste in landfills. Available from: https://www25.senado.leg.br/web/atividade/materias/-/materia/133684 [accessed 22 February 2024].

BRAZIL 2022a Law no 14300, January 6, 2022. Establishes the legal framework for distributed microgeneration, the Electric Energy Compensation System (SCEE) and the Social Renewable Energy Program (PERS). Available from: https://www.planalto.gov.br/ccivil_03/_ato2019-2022/2022/lei/l14300.htm (accessed 22 February 2024).

BRAZIL 2022b Constitutional amendment n° 123, July 14, 2022. Amends article 225 of the Federal Constitution to establish a competitive differential for biofuels. Available form: https://www.planalto.gov.br/ccivil_03/constituicao/emendas/emc/emc123.htm (accessed 22 February 2024).

BRAZIL 2022c Bill n° PL 412/2022. Establishes the Brazilian Commerce System of Greenhouse Gas Emissions (SBCE). Available from: https://www.camara.leg.br/proposicoesWeb/fichadetramitacao?idProposicao=2397761 (accessed 22 February 2024).

BRAZIL 2023 Constitutional amendment n° 132, December 20, 2023. Amends the National Tax System. Available from: http://www.planalto.gov.br/ccivil_03/constituicao/emendas/emc/emc132.htm (accessed 22 February 2024).

Challu, C., Olivares, K. G., Oreshkin, B. N., Ramirez, F. G., Mergenthaler-Canseco, M. & Dubrawski, A. 2023 Nhits: Neural hierarchical interpolation for time series forecasting. In: *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 37, No. 6.

Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H. & Bengio, Y. 2014 Learning phrase representations using RNN encoder-decoder for statistical machine translation. arXiv preprint arXiv:1406.1078.

CNMA 2018 Ministry of the Environment. Resolution 492, December 20, 2018. Available from: https://conama.mma.gov.br/?option=com_sisconama&task=arquivo.download&id=765 (accessed 22 February 2024).

Costa, Y. D., Oliveira, H., Nogueira Jr, V., Massa, L., Yang, X., Barbosa, A., Oliveira, K. & Vieira, T. 2023 Automating petition classification in Brazil's legal system: A two-step deep learning approach. *Artificial Intelligence and Law* 1–25.

Crippa, M., Guizzardi, D., Solazzo, E., Muntean, M., Schaaf, E., Monforti-Ferrario, F., Banja, M., Olivier, J., Grassi, G., Rossi, S. & Vignati, E. 2021 GHG emissions of all world countries. Publications Office of the European Union. Dataset available from. https://edgar.jrc.ec.europa.eu/ (accessed 22 February 2024).

Das, A., Kong, W., Leach, A., Mathur, S., Sen, R. & Yu, R. 2023 Long-term forecasting with tide: Time-series dense encoder. arXiv preprint arXiv:2304.08424.

De Toledo, J. F., Siqueira, H. V., Biuk, L. H., Sacchi, R., Azambuja, R. D. R., Junior, R. A. & Asano, P. T. L. 2023 Climate indices impact in monthly streamflow series forecasting. *IEEE Access* **11**, 21451–21464.

Enamoto, L., Santos, A. R., Maia, R., Weigang, L. & Filho, G. P. R. 2022 Multi-label legal text classification with BiLSTM and attention. *International Journal of Computer Applications in Technology* **68** (4), 369–378.

Enamoto, L. M., Weigang, L., Rocha Filho, G. P. & Costa, P. C. 2023 Generic multimodal gradient-based meta learner framework. In *2023 26th International Conference on Information Fusion (FUSION)*. IEEE, pp. 1–8.

Finn, C., Abbeel, P. & Levine, S. 2017 Model-agnostic meta-learning for fast adaptation of deep networks. In *International Conference on Machine Learning*. PMLR, pp. 1126–1135.

Galvão Filho, A. R., Silva, D. F. C., De Carvalho, R. V., Ribeiro, F. D. S. L. & Coelho, C. J. 2020 Forecasting of water flow in a hydroelectric power plant using LSTM recurrent neural network. In: *2020 International Conference on Electrical, Communication, and Computer Engineering (ICECCE)*. IEEE, pp. 1–5.

Gou, Z. & Li, Y. 2023 Integrating BERT Embeddings and BiLSTM for Emotion Analysis of Dialogue. Computational Intelligence and Neuroscience, 2023.

Gupta, A. & Raghav, Y. 2020 *Time Series Classification with Meta Learning*. AIRCC Publishing Corporation, Zurich, Switzerland.

Gupta, V., Kapadia, S. & Bhadane, C. 2023 Time Series Analysis and Forecasting of Air Quality in India. In *2023 Fifth International Conference on Electrical, Computer and Communication Technologies (ICECCT)*. IEEE, pp. 1–5.

Haris, M. D., Adytia, D. & Ramadhan, A. W. 2022 Air temperature forecasting with long short-term memory and prophet: A case study of Jakarta, Indonesia. In: *2022 International Conference on Data Science and Its Applications (ICoDSA)*. IEEE, pp. 251–256.

Hillmer, S. C. & Tiao, G. C. 1982 An ARIMA-model-based approach to seasonal adjustment. *Journal of the American Statistical Association* **77** (377), 63–70.

Hochreiter, S. & Schmidhuber, J. 1997 Long short-term memory. *Neural Computation* **9** (8), 1735–1780.

Kumari, S. & Singh, S. K. 2023 Machine learning-based time series models for effective CO_2 emission prediction in India. *Environmental Science and Pollution Research* **30** (55), 116601–116616.

Lacoste, A., Luccioni, A., Schmidt, V. & Dandres, T. 2019 Quantifying the carbon emissions of machine learning. arXiv preprint arXiv:1910.09700.

Leelakittisin, B. & Sun, F. 2021 METAHACI: Meta-learning for Human Activity Classification from IMU Data. In: *Cognitive Systems and Signal Processing: 5th International Conference, ICCSIP 2020*, December 25–27, 2020, Zhuhai, China. Revised Selected Papers 5 (pp. 105–113). Springer Singapore.

Lim, B., Arik, S. O., Leoff, N. & Pfister, T. 2021 Temporal fusion transformers for interpretable multi-horizon time series forecasting. *International Journal of Forecasting* **37** (4), 1748–1764.

Lopez, M. D. C. S., Pinaya, J. L. D., Pereira Filho, A. J., Vemado, F. L. & Reis, F. A. G. V. 2023 Analysis of extreme precipitation events in the mountainous region of Rio de Janeiro, Brazil. *Climate* **11** (3), 73.

Lv, Y., Zhang, X., Cheng, Y. & Lee, C. K. M. 2024 Intelligent fault diagnosis of machinery based on hybrid deep learning with multi temporal correlation feature fusion. Quality and Reliability Engineering International. Wiley Online Library.

Meng, Y. & Noman, H. 2022 Predicting CO2 emission footprint using AI through machine learning. *Atmosphere* **13** (11), 1871.

Misra, S., Sarkar, S., Mitra, P. & Shastri, H. 2024 Statistical downscaling of high-resolution precipitation in India using convolutional long short-term memory networks. *Journal of Water and Climate Change* **15** (3), 1120–1141.

Mo, Y., Li, L., Huang, B. & Li, X. 2023 Few-shot RUL estimation based on model-agnostic meta-learning. *Journal of Intelligent Manufacturing* **34** (5), 2359–2372.

Natsume, H. & Okamoto, S. 2024 Prediction of dynamic preference by using temporal dominance of sensations data. In *International Symposium on Affective Science and Engineering ISASE 2024*. Japan Society of Kansei Engineering.

Nichol, A., Achiam, J. & Schulman, J. 2018 On first-order meta-learning algorithms. arXiv preprint arXiv:1803.02999.

Oreshkin, B. N., Carpov, D., Chapados, N. & Bengio, J. 2020 N-BEATS: Neural basis expansion analysis for interpretable time series forecasting. In *International Conference on Learning Representations 2020*.

RJ 2015 Legislative Assembly of the State of Rio de Janeiro. Law no 7068, October 1, 2015. Available from: http://alerjln1.alerj.rj.gov.br/CONTLEI.NSF/c8aa0900025feef6032564ec0060dfff/fdadc03404aed04483257ed600606c2f?OpenDocument (accessed 22 February 2024).

Rolnick, D., Donti, P. L., Kaack, L. H., Kochanski, K., Lacoste, A., Sankaran, K., Ross, A. S., Milojevic-Dupont, N., Jaques, N., Waldman-Brown, A. & Luccioni, A. S. 2022 Tackling climate change with machine learning. *ACM Computing Surveys (CSUR)* **55** (2), 1–96.

Schuster, M. & Paliwal, K. K. 1997 Bidirectional recurrent neural networks. *IEEE Transactions on Signal Processing* **45** (11), 2673–2681.

Teggi, P. P., Natarajan, S. & Malakreddy, B. 2020 Intelligent FORecasting Model for Climate Variations (InFORM): An urban climate case study. In *2020 7th International Conference on Computing for Sustainable Global Development (INDIACom)*. IEEE, pp. 130–137.

Thi Kieu Tran, T., Lee, T., Shin, J. Y., Kim, J. S. & Kamruzzaman, M. 2020 Deep learning-based maximum temperature forecasting assisted with meta-learning for hyperparameter optimization. *Atmosphere* **11** (5), 487.

Tian, C., Zhu, X., Hu, Z. & Ma, J. 2021 A transfer approach with attention reptile method and long-term generation mechanism for few-shot traffic prediction. *Neurocomputing* **452**, 15–27.

UNFCCC 2023 United Nations Framework Convention on Climate Change. Available from: https://unfccc.int/sites/default/files/NDC/2023–11/Brazil%20First%20NDC%202023%20adjustment.pdf.

Wang, X., Liu, J., Han, G., Wang, J. & Cui, J. 2024 End-to-End modulation recognition in underwater acoustic communications using temporal large kernel convolution with gated channel mixer. *IEEE Transactions on Vehicular Technology*.

Werner, D. & Lazaro, L. L. B. 2023 The policy dimension of energy transition: The Brazilian case in promoting renewable energies (2000–2022). *Energy Policy* **175**, 113480.

Wink Junior, M. V., dos Santos, L. G., Ribeiro, F. G. & da Trindade, C. S. 2023 Natural disasters and poverty: Evidence from a flash flood in Brazil. *Environment, Development and Sustainability* 1–22.

Wu, C. J., Raghavendra, R., Gupta, U., Acun, B., Ardalani, N., Maeng, K., Chang, G., Aga, F., Huang, J., Bai, C. & Gschwind, M. 2022 Sustainable ai: Environmental implications, challenges and opportunities. *Proceedings of Machine Learning and Systems* **4**, 795–813.

Zeng, A., Chen, M., Zhang, L. & Xu, Q. 2023 Are transformers effective for time series forecasting? In *AAAI Conference on Artificial Intelligence*, Vol. 37, No. 9.

Zhang, Y., Li, C., Chiew, F. H., Post, D. A., Zhang, X., Ma, N., Tian, J., Kong, D., Leung, L. R., Yu, Q. & Shi, J. 2023 Southern Hemisphere dominates recent decline in global water availability. *Science* **382** (6670), 579–584.

Zrira, N., Kamal-Idrissi, A., Farssi, R. & Khan, H. A. 2024 Time series prediction of sea surface temperature based on BiLSTM model with attention mechanism. *Journal of Sea Research* **198**, 102472.

First received 23 February 2024; accepted in revised form 10 July 2024. Available online 22 July 2024