



Short-, Medium-, and Long-Term Prediction of Carbon Dioxide Emissions using Wavelet-Enhanced Extreme Learning Machine

Mohamed Khalid AlOmar ^{1*}, Mohammed Majeed Hameed ^{1, 2}, Nadhir Al-Ansari ^{3*},
Siti Fatin Mohd Razali ², Mohammed Abdulhakim AlSaadi ⁴

¹ Department of Civil Engineering, Al-Maarif University College, 31001 Ramadi, Iraq.

² Department of Civil Engineering, Faculty of Engineering and Built Environment, Universiti Kebangsaan Malaysia (UKM), 43600 Bangi, Selangor, Malaysia.

³ Civil Engineering Department, Environmental and Natural Resources Engineering, Lulea University of Technology, 97187 Lulea, Sweden.

⁴ Natural and Medical Sciences Research Center (NMSRC), University of Nizwa Sultanate of Oman, Nizwa, Oman.

Received 07 July 2022; Revised 08 March 2023; Accepted 14 March 2023; Published 01 April 2023

Abstract

Carbon dioxide (CO₂) is the main greenhouse gas responsible for global warming. Early prediction of CO₂ is critical for developing strategies to mitigate the effects of climate change. A sophisticated version of the extreme learning machine (ELM), the wavelet enhanced extreme learning machine (W-EELM), is used to predict CO₂ on different time scales (weekly, monthly, and yearly). Data were collected from the Mauna Loa Observatory station in Hawaii, which is ideal for global air sampling. Instead of the traditional method (singular value decomposition), a complete orthogonal decomposition (COD) was used to accurately calculate the weights of the ELM output layers. Another contribution of this study is the removal of noise from the input signal using the wavelet transform technique. The results of the W-EELM model are compared with the results of the classical ELM. Various statistical metrics are used to evaluate the models, and the comparative figures confirm the superiority of the applied models over the ELM model. The proposed W-EELM model proves to be a robust and applicable computer-based technology for modeling CO₂ concentrations, which contributes to the fundamental knowledge of the environmental engineering perspective.

Keywords: Carbone Dioxide; Greenhouse Gas; Climate Change; Complete Orthogonal Decomposition.

1. Introduction

1.1. Background of the Study

Due to the increasing global warming in various spheres of life, the threat of climate change has been considered a serious environmental problem for the last twenty years. The intensive expansion of human and industrial activities in many countries has put enormous pressure on the environment by releasing large amounts of greenhouse gases into the atmosphere. The inefficient, systematic process of releasing gasses into the atmosphere eventually leads to the occurrence of 'climate extremism' and serious consequences that threaten human health, the economy, and human development and increase the concentration of pollutants in the atmosphere. For example, the global average surface temperature of the Earth has increased from 0.4 to 0.8°C since the end of the 19th century [1]. Carbon dioxide is the main cause of global warming on our planet [2–5] and thus plays an important role in maintaining the stability of the

* Corresponding author: mohd.alomar@yahoo.com; nadhir.alansari@ltu.se

 <http://dx.doi.org/10.28991/CEJ-2023-09-04-04>



© 2023 by the authors. Licensee C.E.J, Tehran, Iran. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).

climate system [6, 7]. Recently, the phenomenon of global warming has become the biggest environmental problem in human history [8, 9]. Although measures have been taken to reduce carbon emissions into the atmosphere, the concentration continues to increase. In this context, international organizations are making great efforts to reduce the negative effects of global warming by focusing on policies to reduce carbon dioxide emissions [10].

Based on the 1997 climate report, the Kyoto Protocol called for an 85% reduction in greenhouse gas emissions. In 2009, the United Nations Framework Convention on Climate Change (UNFCCC) took place. As a result of this convention, China promised to reduce CO₂ emissions by 40–45% by the end of 2020 [10]. Although CO₂ concentration set a record in May 2019, it reached 415 ppm at Mauna Loa Observatory Station. However, the International Energy Agency (IEA) reports that global CO₂ emissions from energy generation have decreased to 33 megatons, which is due to the reduction of CO₂ emissions from power plants in developed countries. The main reason for this decrease in CO₂ emissions in the energy sector is the expansion of the use of renewable energy and the reduction in the use of fossil fuels in power plants, as well as the switch to natural gas. It is worth noting that the global CO₂ concentration from the energy sector accounts for 41% of all CO₂ emissions, while the other CO₂ emission sources can be divided as follows: 20% from industry, 16% from road transport, 6% from other transport, and about 16% from various sectors and households [11]. Figure 1 shows the global CO₂ emissions of developed and developing countries based on IEA data [12].

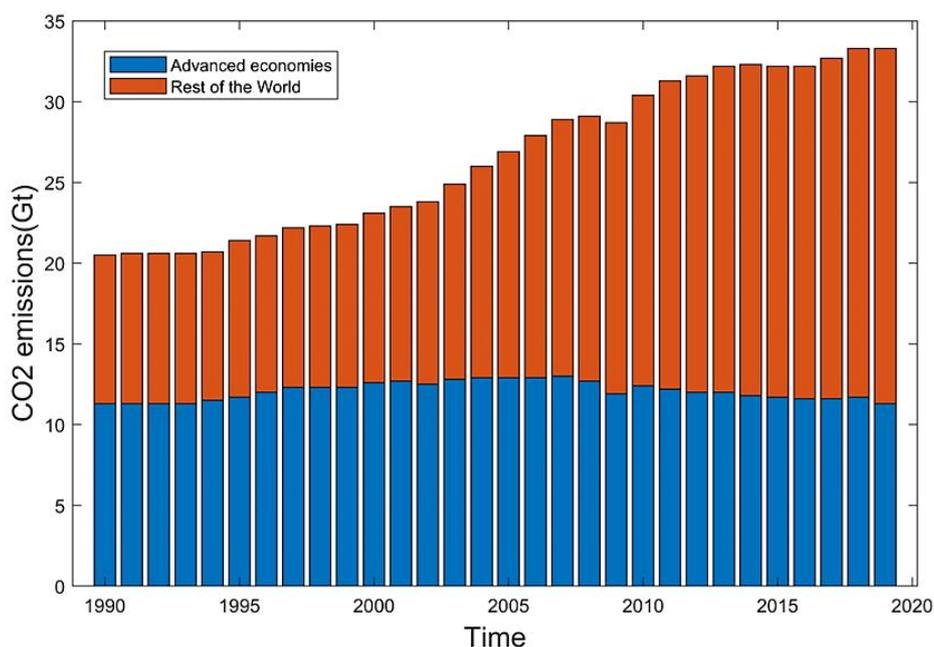


Figure 1. Records of CO₂ emissions over the last 30 years

Consequently, the main goal to preserve our planet is to keep the temperature increase on Earth below 2°C according to the recommendations of the Paris Agreement [13]. According to the recommendations of the Paris Agreement, developed countries must reduce CO₂ emissions from the energy industry. It is important to note that some of these countries have responded very well to the demands. For example, CO₂ emissions from the energy sector have decreased by 2.9% in the United States, 8% in Germany, 2% in the United Kingdom, and 4.3% in Japan. On the other hand, energy-related CO₂ emissions have increased in Asia due to the increased demand for coal for energy facilities [12].

Recently, the concentration of CO₂ emissions has increased dramatically, which puts enormous pressure on the entire ecosystem. CO₂ gas concentration has already exceeded the normal and safe level of about 350 ppm. Catastrophic weather events such as sea level rise and hurricanes can be avoided if the CO₂ concentration in the atmosphere is at 350 ppm. Moreover, the irreversible and dangerous consequences of climate change would be limited if the CO₂ level is at the safe level (350 ppm). It should be noted that the highest value determined in this study is 415.39 ppm.

1.2. Related Works and Research Gap

The prediction of CO₂ is an important issue in the environmental field because it significantly affects the temperature of the Earth's surface. Moreover, CO₂ concentration has increased exponentially in recent years and poses a significant threat to human life and the ecosystem. Accurate prediction of CO₂ concentration is not only important to provide important information to policy makers, but also can improve the quality of CO₂ emission management [14]. CO₂ prediction is crucial to monitor the changes of this gas over time and establish a reliable warning system; however, few studies have been conducted to predict this gas. Some studies have used artificial intelligence (AI) techniques to predict CO₂ emissions, including genetic algorithm [15, 16], artificial neural network (ANN) [17], Gaussian processes

regression method [18], and logarithmic mean Divisia index method [19, 20]. Mardani et al. [21] conducted a study using ANN and an adaptive neuro-fuzzy inference system (ANFIS) for CO₂ prediction. These models were validated against multilinear regression. The study found that the AI models had excellent prediction accuracy compared to the MLR model. In addition, Saleh et al. [22] used support vector regression to predict CO₂ emissions from energy consumption. As the hyperparameters of the applied model were calculated by trial-and-error method, the result of the study was satisfactory.

However, the standard ANN model has some problems, such as long training time and low generalization [22], while other models, such as Gaussian process regression, require accurate tuning of the hyperparameters of this model and appropriate selection of the kernel function. It is important to mention that the above studies could not consider the analysis of time series of CO₂. Moreover, the mentioned researchers obtained the required data from specific stations. In other words, these models can provide information on emissions at a regional scale, but not on global emissions. Detection of gas emissions at a geographically and climatically delineated location can provide an important indication of the increase in carbon dioxide concentration around the globe, which has not been studied before.

1.3. Research Significance

Accurate prediction of CO₂ gas is very important for establishing an advanced early warning system and can play a crucial role in evaluating the adopted global strategies to reduce the concentration of this gas in the air. The main challenge is that the atmospheric CO₂ concentration increases very rapidly in short time periods. In this study, the ability of a novel model called wavelet enhanced extreme learning machine (W-EELM) to predict CO₂ emissions is investigated. In a recent study [23], the classical ELM was improved by using a robust algorithm called the complete orthogonal decomposition algorithm (COD) instead of a classical algorithm (Singular Value Decomposition (SVD)) to optimally compute the output weights of ELM. Several studies have shown that the COD algorithm is much better than SVD through the calibration process.

Moreover, COD is more reliable, efficient and faster than SVD in solving technical problems [23, 24]. We integrate an enhanced extreme learning machine (EELM) with a discrete wavelet transform approach. The discrete wavelet transform technique removes the noise from the data set to obtain more accurate predictions. The proposed (W-EELM) was validated in predicting CO₂ emissions at different time scales (i.e., weekly, monthly, and yearly) against the standard ELM.

2. Methodology

2.1. Environmental Data and Site Description

This study uses CO₂ emission data collected from the Mauna Loa Observatory Station in the state of Hawaii, United States. This station measures environmental factors and emission gasses that contribute to global climate change. The observatory's location is ideal for collecting air samples because it is located in Hawaii on the side of the largest active volcano on Earth. Geographically, the station is about 3400 m above mean sea level, far enough away from pollution sources to facilitate scientists and researchers studying and analyzing air properties. The other geographic and hydrologic features of the observatory are listed in Table 1. In the 1950s, scientists studied atmospheric safety at Mauna Loa Observatory. This monitoring station can detect global climate changes by measuring various air gas concentrations. Many gas emissions are measured, including carbon dioxide, methane, sulfur dioxide, and nitrous oxide. Among these greenhouse gas emissions, CO₂ is the most critical pollutant that mainly causes global warming.

Table 1. General Information about Mauna Loa Observatory Station [25]

Geographical Information	
Latitude	19.54° N
Longitude	155.58° W
Altitude (meters above mean sea level)	3400 m
Hydrological Information	
Average Temperature	Day (9°C), Night (5°C)
Average Min Temperature	Day (4°C), Night (3°C)
Average Max Temperature	Day (12°C), Night (7°C)
Average Precipitation	1.5" per month
Average Wind Speed	5 m/sec
Max Average Wind Speed	20m/sec
Average Pressure	680 millibars

The CO₂ concentration data used in this study were collected by the National Oceanic and Atmospheric Administration of the U.S. Department of Commerce [26]. In addition, the data span multiple time scales (i.e., weekly, monthly, and yearly). The statistical characteristics of the data set with different time scales are shown in Table 2. The data set contains some missing values, which we fill by linear interpolation.

Table 2. statistical criteria

Time scale	Stage	Period	N	Rang	Mean	Standard deviation	CV
Weekly		19-5-1974 to 9-4-2000	1435	326.73 ~ 373.93	350.1961	12.3329	0.0352
Monthly	Training	Mar-1958 to May-1995	446	312.66 ~ 363.69	335.0417	13.9753	0.0417
Yearly		1957 to 1995	36	318.45 ~ 362.61	338.9078	14.3146	0.0422
Weekly		16-4-2000 to 15-3-2020	956	369.88 ~ 415.39	391.6733	11.8726	0.0303
Monthly	Testing	June-1996 to Feb-2020	297	358 ~ 414.66	384.9845	15.2106	0.0395
Yearly		1996 to 2019	23	363.73 ~ 411.44	386.3813	14.4219	0.0373

*N and CV respectively are the number of observations and variance coefficient.

2.2. Extreme Learning Machine

Extreme learning machine (ELM) can be defined as a learning algorithm presented by Huang for training single hidden layer feedforward neural network (SLFN) [27]. Training classical SLFN can be performed using a backpropagation algorithm, which has several shortcomings, including time consumption, computational cost, and overfitting problems. Therefore, the ELM algorithm is proposed for training SLFN to achieve faster and more accurate modeling with better generalization [28-30]. The hidden layer of SLFN is the most important element in the structure of SLFN and significantly affects the efficiency of the model. According to Huang [27], if the transfer function of the hidden layer is infinitely differentiable in each interval and there are enough hidden nodes, it is not necessary to adjust all the weighting values of the network. In accordance with this fact, the ELM algorithm initializes the weight and bias values of the hidden layer randomly. For this reason, the ELM model requires less learning time than the classical neural network system. It is important to mention that the weights of the output layer can be calculated using the least squares method based on the Moore-Penrose generalized inverse function. Figure 2 shows the general structure of the modeling approach of ELM.

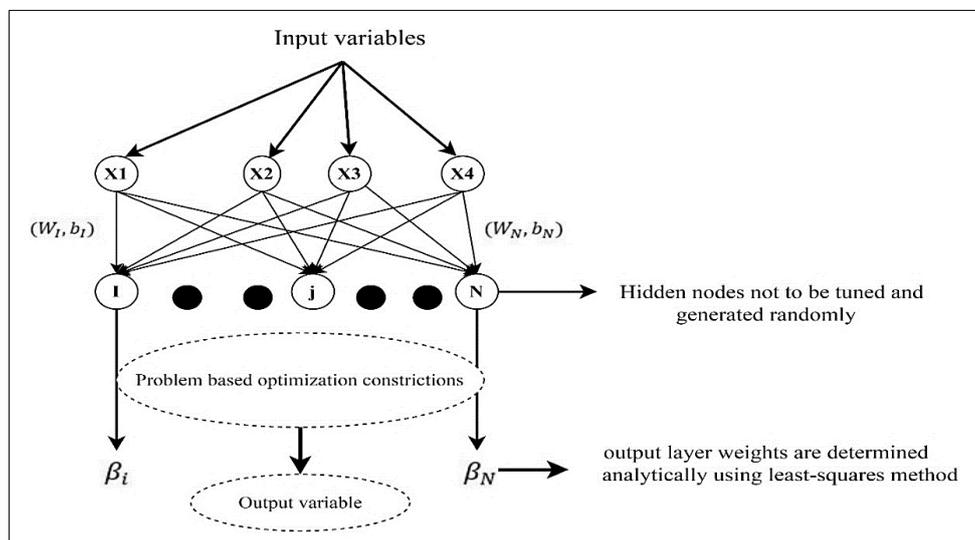


Figure 2. General structure of ELM modeling approach

The mathematical expression of the ELM model can be expressed by Equation 1:

$$\sum_{i=1}^L B_i g_i(a_i x_t + \beta_i) = z_t, i = 1, \dots, N \tag{1}$$

where L is the number of hidden nodes, $g_i(a_i x_t + \beta_i)$ is the output function of the hidden layer, a_i and β_i are the parameters (i.e., weights and biases) of the hidden nodes determined randomly, B_i refers to the weight values that map the i_{th} hidden nodes to the output node, and z_t is the output target of the ELM model. To achieve good generalization and more stable modeling, the required number of hidden nodes in the hidden layer should not exceed the number of input samples.

In the modeling approach of ELM, the parameters of the hidden nodes are determined randomly without iteratively tuning and adjusting their values, as is the case in various types of artificial intelligent modeling approaches [31]. The main matrix of ELM is the randomly generated values of the weights of the hidden layer. This can lead to zero error and provides an opportunity to analytically configure the weight values of the output neural network layer (B) for the training samples. Moreover, the parameter values of the internal activation function (a_i, β_i) are assigned according to the probability distribution. Finally, the general matrix describing the ELM model can be expressed in Equations 2 and 3 [32, 33]:

$$H(a, \beta, x) = \begin{bmatrix} g(x_1) \\ \vdots \\ g(x_n) \end{bmatrix} = \begin{bmatrix} g_1(a_1 \cdot x_1 + \beta_1) \dots g_L(a_L, \beta_L, x_1) \\ \vdots \\ g_1(a_1 \cdot x_n + \beta_1) \dots g_L(a_L, \beta_L, x_n) \end{bmatrix}_{N \times N} \tag{2}$$

and

$$B = \begin{bmatrix} B_1^T \\ \vdots \\ B_L^T \end{bmatrix}_{N \times 1} \quad \text{and} \quad Y = \begin{bmatrix} y_1^T \\ \vdots \\ y_L^T \end{bmatrix}_{N \times 1} \tag{3}$$

where H is the matrix of the hidden layer and T is the matrix transpose. The matrix of ELM can be simplified as Equation 4:

$$HB = Y \tag{4}$$

The least square method of Equation 3 can be expressed in Equation 5:

$$\hat{B} = H^\dagger Y \tag{5}$$

The H^\dagger is the Moore-Penrose inverse of Hussain (H) matrix. Conventionally, the singular value decomposition (SVD) method has been used primarily for ELM learning.

2.3. Enhanced Extreme Learning Machine (EELM) Model

The SVD approach is often used to solve linear problems using the least squares error method. However, the complete orthogonal decomposition (COD) algorithm can be used as an alternative to the traditional SVD method because it gives excellent results in solving linear problems with a simple and reliable computation [34]. In the present study, the COD algorithm is used to calculate the output weight values in order to improve the learning method and the efficiency of the model ELM. It should be noted that the COD method has special features that distinguish it from the usual SVD method. One of these advantages is that the COD algorithm can provide much more accurate results with a simpler computational process.

It is important to point out that the COD approach generally provides very efficient computational results regardless of the size of the H matrix. Moreover, the results of this algorithm are stable, and the required weights can be calculated using the least squares norm and in a shorter time. More details about the COD algorithm and its applications in tuning the ELM modeling technique can be found in Guo et al. [35].

2.4. Wavelet Transform

Wavelet transform (WT) is a mathematical approach used to remove noise and decompose data series. In WT analysis, a time series process includes two main components, low-frequency, and high-frequency components. The general characteristics of time series data include seasonal and cyclical trends, which are processed by the low-frequency components, while the chaotic and detailed elements are preserved in the high-frequency components. This strategy is similar to the variable separation approach in time series analysis, which can be used to identify inherent patterns in raw time series data.

The mathematical expression of the temporal wavelet transforms of a continuous time series, $x(t)$ can be defined as shown in Equation 6 [36, 37]:

$$W(m, n) = \frac{1}{\sqrt{m}} \int_{-\infty}^{+\infty} M^* \left(\frac{t-n}{m} \right) x(t) \cdot dt \tag{6}$$

Here, n describes the time shift of the function that supports a careful study of the signal; m represents the dilation, the symbol $*$ describes the complexity formula used for the conjugation, and $M(t)$ represents the mother wavelet. This transformation approach is primarily concerned with determining the time scale of the process. In practice, ecologists usually do not prefer the continuous wavelet transform because it generates redundant information that can affect the effectiveness of simulation models. Therefore, the discrete wavelet transform is often used instead of the continuous-time signal process. The formula of the discrete mother function can be expressed in Equation 7 [36]:

$$M_{a,b}(t) = \frac{1}{\sqrt{m^{a_o}}} M\left(\frac{t - bn_o m^{a_o}}{m^{a_o}}\right) \tag{7}$$

In this equation, a and b are integers. Wavelet dilation and wavelet translation are controlled by a and b , respectively. Normally, the parameters n_o and m^{a_o} are equal to 1 and 2, respectively. n_o describes the location parameter and its value should be greater than zero, while m^{a_o} indicates the step of refined dilation, which should be greater than one.

This logarithmic scaling for dilation and translation is called dyadic grid management. The dyadic wavelet function is presented in Equation 8 [36]:

$$M_{a,b}(t) = 2^{-a/2} \times M(2^{-a}t - b) \tag{8}$$

For a discrete-time series, x_i , the dyadic wavelet transform is represented in Equation 9 [38, 39]:

$$T_{a,b} = 2^{-a/2} \sum_{i=0}^{L-1} M(2^{-a}i - b)x_i \tag{9}$$

Thus, $T_{a,b}$ is the wavelet coefficient for the discrete wavelet with scale $m = 2^a$ and location $L = 2^a$ (where $i = 0, 1, 2, \dots, L - 1$). In addition, the smoothed signal component, which represents the overall trend of the time series, is considered as T . Thus, the discrete inverse transform function reconstructs the signal x_i as shown in Equation 10 [38, 39]:

$$x_i = T(t) + \sum_{a=1}^A T_{a,b}(t) \tag{10}$$

where $T(t)$ is the approximation sub-signal at level A , while $T_{a,b}$ is the details sub-signal at the level $a = 1, 2, \dots, A$ and the time dimension of t ($t = 1, 2, \dots, b$).

2.5. Modeling Development

In this study, the ability of the wavelet enhanced extreme learning machine (W-EELM) model with a single hidden layer to predict CO₂ emissions is investigated. The experiment of the modeling approach was conducted using MATLAB 2018b software environment. For this purpose, the experimental data set is divided into two stages; the first stage is used for training and constructing the proposed technique, while the second stage is used for testing. The assumed input combinations used to develop predictive models are expressed in Equations 11, 12, and 13: The input layer of W-EELM contains wavelet neurons (nodes) fed with subseries of CO₂ time series obtained by discrete wavelet transform (Equation 9).

The combination of artificial intelligence approaches and WT techniques provided a good level of accuracy in several disciplines [35-42].

$$m_1 = f(CO2^*_{t-1}) \tag{11}$$

$$m_2 = f(CO2^*_{t-1}, CO2^*_{t-2}) \tag{12}$$

$$m_3 = f(CO2^*_{t-1}, CO2^*_{t-2}, CO2^*_{t-2}) \tag{13}$$

where the symbol $*$ refers to the time scale of CO₂ concentration (i.e., weekly, monthly, and yearly).

Determining the hidden layer nodes in the middle layer of a neural network is critical to the development of the proposed model and has significant implications for model accuracy. For this reason, various combinations of hidden nodes ranging from 1 to 25 are used. Figure 3 clearly shows the methodology used to develop the proposed and comparable model to achieve the objective of this study.

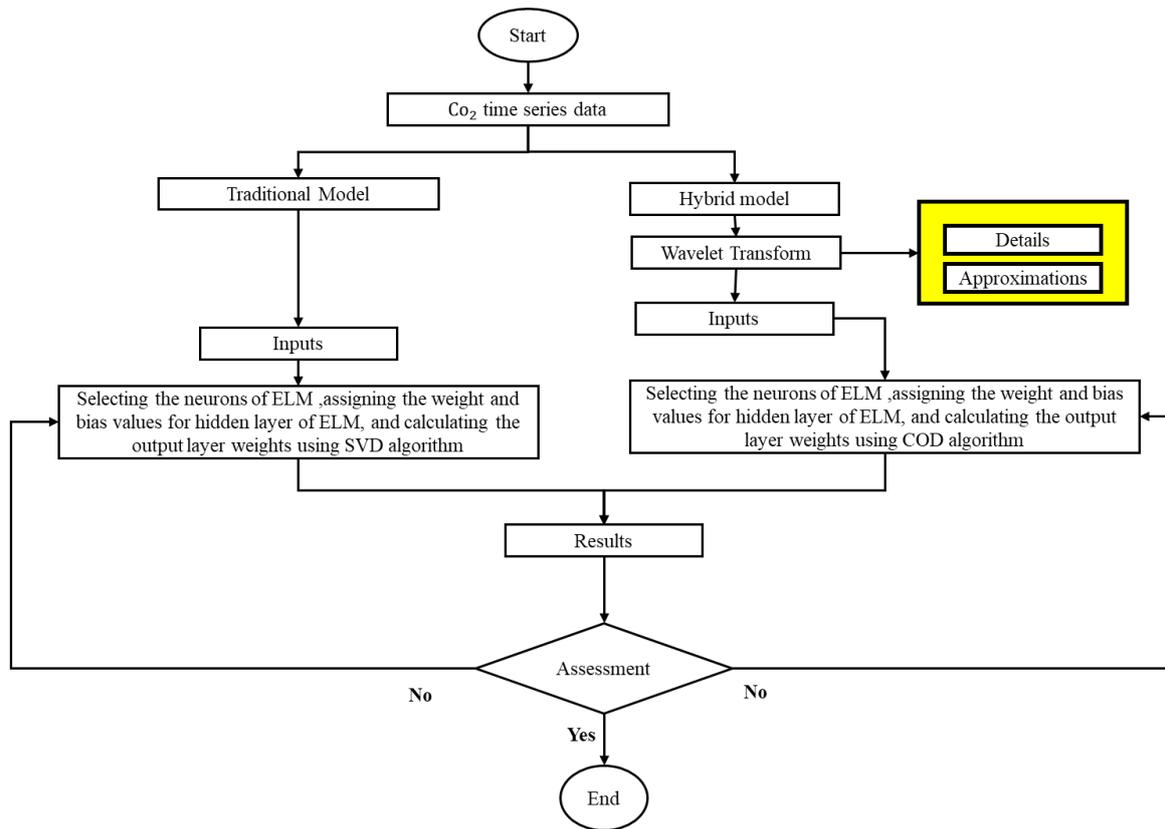


Figure 3. The methodology of the W-EELM models

The type of wavelet mother and the level of decomposition are crucial to increase the efficiency of the predictive model. In this study, the Daubechies wavelet was selected to decompose the raw time series data because it can extract useful information features from the data. Moreover, it has been widely used to deal with problems related to carbon dioxide emissions [40-43]. Moreover, the minimum decomposition level [44, 45] was calculated using $\log(N)$. Here, N denotes the number of original data used for the analysis of the five types of disappearance used in this study. The main steps of conduct the study can be summarized as follows:

- Preparation of the CO₂ time series data (weekly scale) and subsequent calculation of the monthly and yearly time scales.
- Select the possible input vectors for each time scale for the classical ELM model (Equations 11 to 13).
- Apply the WT approach to the selected inputs for the W-EELM model.
- Determine the number of hidden nodes and assign the ELM and W-EELM hidden layer weights and biases.
- Normalized the input and output variables (data training and data testing).
- Calculate the H matrix for both models according to Equation 2
- Activate the H matrix using the hyperbolic tangent sigmoid transfer function [46].
- For the classical ELM, compute the output weighting using SVD and the COD algorithm for W-EELM, and then perform the prediction.
- Denormalize the predicted and actual CO₂.
- Select the best models based on various statistical criteria, as shown in the following section and comparable plots.

2.6. Performance Measures

Each predictive model developed in this study was evaluated using several performance measures (see Equations 14 to 21). The statistical measures are correlation coefficient (R), Nash-Sutcliffe efficiency (NSE), mean absolute error (MAE), root mean square error (RMSE), mean absolute percentage error (MAPE), relative error (RE), mean absolute relative error (MARE), and residual. The mathematical expressions of these metrics can be written as follows [47-50]:

$$R = \frac{\sum_{i=1}^n (CO_{2iObs} - \bar{CO}_{2iObs})(CO_{2iPre} - \bar{CO}_{2iPre})}{\sqrt{\sum_{i=1}^n (CO_{2iObs} - \bar{CO}_{2iObs})^2} \sqrt{\sum_{i=1}^n (CO_{2iPre} - \bar{CO}_{2iPre})^2}} \quad (14)$$

$$MSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (CO_{2iObs} - CO_{2iPre})^2} \quad (15)$$

$$MAE = \frac{1}{N} \sum_{i=1}^n |(CO_{2iObs} - CO_{2iPre})| \quad (16)$$

$$MAPE (\%) = \frac{1}{N} \sum_{i=1}^n \left| \frac{(CO_{2iObs} - CO_{2iPre})}{CO_{2iObs}} \right| \times 100 \quad (17)$$

$$NE = 1 - \left[\frac{\sum_{i=1}^n (CO_{2iObs} - CO_{2iPre})^2}{\sum_{i=1}^n (CO_{2iObs} - \overline{CO_{2iObs}})^2} \right] \quad (18)$$

$$RE (\%) = \frac{CO_{2iObs} - CO_{2iPre}}{CO_{2iObs}} \times 100 \quad (19)$$

$$MARE (\%) = \frac{1}{N} \sum_{i=1}^n \left| \frac{CO_{2iObs} - CO_{2iPre}}{CO_{2iObs}} \right| \times 100 \quad (20)$$

$$Residual = CO_{2iObs} - CO_{2iPre} \quad (21)$$

where CO_{2iObs} and CO_{2iPre} are the observed and predicted i_{th} values of CO_2 , $\overline{CO_{2iObs}}$ and $\overline{CO_{2iPre}}$ are the average observed and predicted CO_2 values in a trained and tested sample, n denotes the total number of data points.

3. Result and Discussion

The ability of the wavelet-enhanced extreme learning machine (W-EELM) and the classical extreme learning machine (ELM) modeling approaches to predict CO_2 emissions over Mauna Loa Observatory Station in the United States is presented. In this study, three different time scales (weekly, monthly, and yearly) were selected to predict CO_2 gas emissions one step ahead based on previous data points. The short-term (weekly) prediction is important for developing advanced early warning systems, while the medium-term (monthly) prediction is critical for evaluating overall mitigation strategies and reducing CO_2 concentrations. In addition, the long-term (yearly) data can help with global management, policies, and strategies for countries exporting more CO_2 , as well as monitoring the improvements achieved.

The reason for modeling short- and long-term predictions is to intensively study the behavior of the proposed models in capturing the patterns of CO_2 gas emissions over different time series. In general, the analysis of time series data becomes more complex over time as some of the key data features are lost due to averaging. Therefore, the qualifications of a reliable model are effectively considered in short- and long-term predictions. In this study, three different input combinations (m1, m2, and m3) were introduced in the previous section (see Equations 11 to 13) and used to develop the models for one-step ahead prediction of CO_2 . For the W-EELM models, the input data were denoised using the wavelet transform (WT) to obtain clean and stationary data. These data were fed into the input nodes of the EELM approach. It is worth noting that Daubechies (db) was used as the wavelet mother and that two different decomposition levels were used for each time scale (weekly, monthly, and yearly).

In this study, 60% of the data set was used to build the model, and the rest was used for testing purposes. Various statistical matrices and graphical representations were used to evaluate the predictive quality of each model. Table 3 lists the statistical evaluation matrices for each predictive model in the training and testing phases. According to the quantitative analysis, several prediction models show good performance in predicting the weekly CO_2 concentrations in the training phase. The statistical measures such as RMSE, MAE, MAPE, and NE ranged from 0.4575 to 0.9893 (ppm), 0.3699 to 0.8074 (ppm), 0.1058% to 0.2325%, and 0.9936 to 0.9986, respectively. The Wm2-WEELM³ model with only two input variables (m2) and three decomposition levels showed the highest prediction accuracy compared to other similar models. However, the training phase cannot provide a meaningful impression of the most reliable modeling approaches because the trained models are based on label input data. Therefore, the testing phase is crucial for the selection of a reliable predictive model because unlabelled input data is fed to the model, which can accurately determine the generalizability of the model at this stage. The superiority of the Wm2-WEELM³ model was shown in comparison to other comparable models, which had the lowest values for RMSE (0.5899 ppm), MAE (0.4745 ppm), MAPE (0.1206%), and the highest value for NE (0.9975). In addition, the accuracy of predictive models decreases when the number of input variables is increased because unnecessary information is added, which can hinder the training process of the model and decrease the generalization efficiency. This phenomenon can be clearly demonstrated by the Wm3-WEELM⁴ model, where 15 input variables are introduced into the system, which complicates the mathematical operations of the model and reduces the prediction accuracy to the lowest level.

Table 3. Performance prediction capabilities for predictive models in training and testing phases

Time Scale	Model*	Training Set				Testing Set			
		RMSE	MAE	MAPE (%)	NE	RMSE	MAE	MAPE (%)	NE
Weekly	Wm1-ELM	0.5228	0.4119	0.1178	0.9982	0.6103	0.4747	0.1211	0.9974
	Wm1-WEELM ³	0.4634	0.3759	0.1076	0.9986	0.5924	0.4766	0.1212	0.9975
	Wm1-WEELM ⁴	0.9893	0.8074	0.2325	0.9936	1.0053	0.8017	0.2042	0.9928
	Wm2-ELM	0.5221	0.4113	0.1176	0.9982	0.6097	0.4741	0.1209	0.9974
	Wm2-WEELM³	0.4615	0.3741	0.1071	0.9986	0.5899	0.4745	0.1206	0.9975
	Wm2-WEELM ⁴	0.8869	0.7188	0.2066	0.9948	1.0313	0.8167	0.2079	0.9924
	Wm3-ELM	0.7557	0.6243	0.1786	0.9962	0.84	0.6859	0.175	0.995
	Wm3-WEELM ³	0.4575	0.3699	0.1058	0.9986	0.6101	0.4914	0.1253	0.9974
Monthly	Wm3-WEELM ⁴	0.8293	0.6696	0.1922	0.9955	1.0277	0.8135	0.2071	0.9925
	Mm1-ELM	1.206	1.0382	0.3106	0.9925	1.3213	1.1848	0.3081	0.9924
	Mm1-WEELM²	1.0967	0.8738	0.2618	0.9938	1.1501	0.929	0.2421	0.9943
	Mm1-WEELM ³	1.1919	0.9736	0.2912	0.9927	1.3323	1.1061	0.2875	0.9923
	Mm2-ELM	1.1896	1.0507	0.3137	0.9928	1.302	1.1563	0.3005	0.9926
	Mm2-WEELM ²	1.1458	0.914	0.2743	0.9933	1.2275	0.9807	0.2559	0.9934
	Mm2-WEELM ³	2.587	2.1455	0.6428	0.9658	2.485	2.0939	0.5457	0.9731
	Mm3-ELM	1.2116	1.0837	0.3237	0.9925	1.3285	1.1944	0.3107	0.9923
Yearly	Mm3-WEELM ²	1.1006	0.8681	0.2601	0.9938	1.1601	0.9297	0.2424	0.9941
	Mm3-WEELM ³	1.2514	1.0157	0.305	0.992	1.3302	1.0556	0.2759	0.9923
	Ym1-ELM	0.3342	0.2506	0.0748	0.9994	0.5743	0.446	0.1162	0.9985
	Ym1-WEELM¹	0.2747	0.1975	0.0585	0.9996	0.4674	0.3783	0.0983	0.999
	Ym1-WEELM ²	0.5039	0.3936	0.1192	0.9986	0.7902	0.6663	0.1725	0.9971
	Ym2-ELM	0.3961	0.3062	0.0908	0.9992	0.5513	0.3797	0.0978	0.9985
	Ym2-WEELM ¹	0.2996	0.2389	0.0706	0.9995	0.4953	0.4023	0.1042	0.9988
	Ym2-WEELM ²	0.4158	0.3339	0.0986	0.9991	0.524	0.3932	0.1017	0.9986
Yearly	Ym3-ELM	0.4075	0.3224	0.0959	0.9992	0.5413	0.3791	0.0973	0.9985
	Ym3-WEELM ¹	0.2991	0.2359	0.0695	0.9995	0.5259	0.4265	0.1104	0.9986
	Ym3-WEELM ²	0.4323	0.3513	0.1042	0.999	0.536	0.4003	0.1038	0.9986

Note: The bold fonts represent the best model accuracy; the symbol * represents the wavelet decomposition level; m1, m2 and m3 are the input variables as shown in the methodology section; and W, M, and Y refer to the time scale (i.e., weekly, monthly, and yearly).

Regarding the prediction of the monthly time scale of CO₂, the accuracy of the predictions varies from one model to another. The statistical metrics are tabulated in Table 3, and the values of RMSE, MAE, MAPE, and NE ranged from 1.0967 to 2.587 (ppm), 0.8681 to 2.1455 (ppm), 0.2601% to 0.6428%, and 0.95658 to 0.9938 separately. In this phase, the Mm1-WEELM² model had the highest accuracy, while the Mm2-WEELM³ model (with eight input variables) had the lowest prediction accuracy compared with the other eight similar models. However, in the phase test, the Mm1-WEELM² maintained its excellent prediction performance and provided the lowest RMSE of 1.501 ppm, MAE of 0.929 ppm, MAPE of 0.2421%, and the highest value of NE (0.9943).

From Tables 3 and 4, it can be seen that both models are sensitive to the number of hidden nodes and input combinations. For the weekly scale, the best model is Wm2-WEELM³. The model requires two raw input variables to achieve the best accuracy in this case. However, for the monthly and yearly time scales, the model needs only one raw input variable. Another important observation is that the models for predicting weekly CO₂ generally require a higher number of hidden models than for the monthly and yearly time scales. This could be related to the nature of CO₂ records, where data are fluctuating and scattered on a short time scale, while fluctuating on the medium and long-term scales.

Table 4. The parameters of the applied models at different time scale

Time scale	Model	Hidden nodes
Weekly	Wm1-ELM	7
	Wm1-WEELM ³	12
	Wm1-WEELM ⁴	4
	Wm2-ELM	7
	Wm2-WEELM³	25
	Wm2-WEELM ⁴	8
	Wm3-ELM	4
	Wm3-WEELM ³	6
	Wm3-WEELM ⁴	11
Monthly	Mm1-ELM	5
	Mm1-WEELM²	3
	Mm1-WEELM ³	2
	Mm2-ELM	3
	Mm2-WEELM ²	1
	Mm2-WEELM ³	8
	Mm3-ELM	6
	Mm3-WEELM ²	3
	Mm3-WEELM ³	5
Yearly	Ym1-ELM	21
	Ym1-WEELM¹	13
	Ym1-WEELM ²	14
	Ym2-ELM	10
	Ym2-WEELM ¹	10
	Ym2-WEELM ²	2
	Ym3-ELM	10
	Ym3-WEELM ¹	7
	Ym3-WEELM ²	2

Note: Bold font is the best model.

Finally, Table 3 lists the quantitative assessments of nine predictive models used to predict one year ahead. The statistical parameters of all predictive models for the training phase ranged from 0.2747 to 0.5039 ppm, 0.1975 to 0.3936 ppm, 0.0585% to 0.1192%, and 0.9986 to 0.9969 for RMSE, MAE, MAPE, and NE, respectively. The superiority of the Ym1-WEELM¹ model in predicting CO₂ yearly emissions over the other eight predictive models was clearly demonstrated in both the training and testing sets. Moreover, with only two input variables, this model achieved lower values for RMSE (0.4674 ppm), MAE (0.3783 ppm), and MAPE (0.0983%) and the highest value for NE (0.999) during the testing set. The superiority of the proposed modeling approach was demonstrated by the reduction of RMSE criteria during the training and testing phases (see Figure 4). Thus, the proposed model has shown a significant improvement in prediction compared to the classical model in the yearly time interval with 13.65% during the testing phase, as shown in Figure 4.

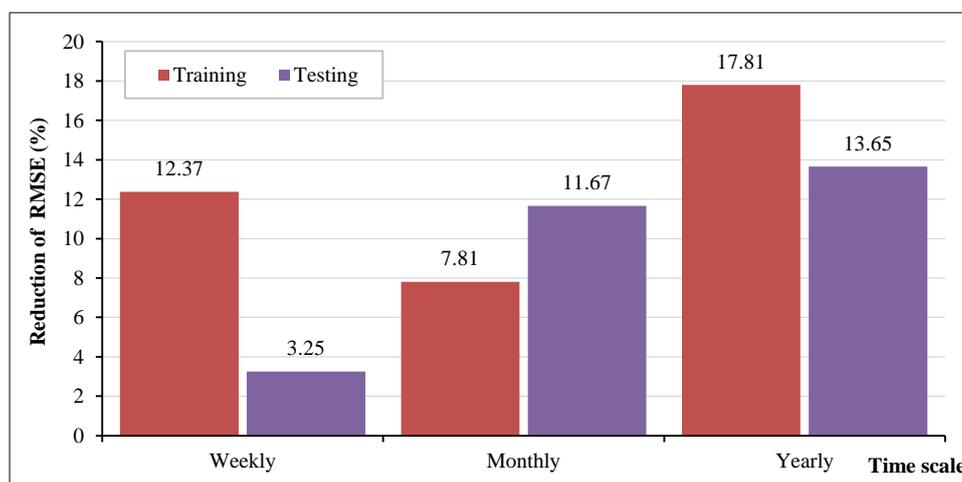


Figure 4. Superiority of W-EELM models over classical ones in reduction RMSE values

However, the percentage performance improvement varies from 3.25% to 11.67% for the weekly and monthly time scales, respectively. The main observation is that the highest improvement occurs when predicting CO₂ yearly, followed by the monthly and weekly time scales. The explanation for this phenomenon is that the accuracy of long-term predictions using conventional models (e.g., ELM) decreases significantly because these approaches cannot effectively capture the dynamics of the environmental system. In addition, the dynamic trend of the time series data is lost when larger time scales are used and the observed data points are averaged.

To further investigate the adequacy of each predictive model, various graphical representations were created, including scatter plots, relative error counts, and boxplots. The scatter plots for the training and testing phases, shown in Figures 5 through 10, serve as a better means for comparative evaluation of the classical and hybrid modeling approaches. In addition to the correlation coefficient (*R*), these figures show a more meaningful visualization of the variance between the observed and predicted values of CO₂ gas emissions. Regarding the weekly time scale, the graphical representations (Figures 5 and 6) showed that the hybrid models with two decomposition levels achieved the best prediction accuracy compared to the classical models. Moreover, these figures provided a deep understanding of how each model predicted the scattering values around the ideal line. For example, the Wm2-WEELM³ exhibited lower scatter and achieved the highest prediction efficiency with *R* of 0.994 and 0.991 for the training and testing phases separately.

Similarly, the standard models based on Figures 7 and 8 for the monthly time scale had lower prediction accuracy and relatively more scattered points than the hybrid models. It is obvious that the Mm1-WEELM² model is superior to all other predictive models and achieves perfect accuracy with *R* between 0.9969 and 0.9972 for the training and testing phases. Finally, the general situation regarding the yearly predictions for CO₂ emissions has not changed significantly, as the hybrid models show great ingenuity and high prediction accuracy with the highest *R* values for both phases compared to the classical models (see Figures 9 and 10). The Ym1-WEELM¹ model exhibited lower scatter and deviation from the ideal line, resulting in a higher *R*-value (0.9998 to 0.9996) for the training and testing series, respectively, compared to other predictive models. It is worth noting that the complex models, especially the hybrid models that require many inputs, did not predict the weekly, monthly, and yearly CO₂ gas emissions very well.

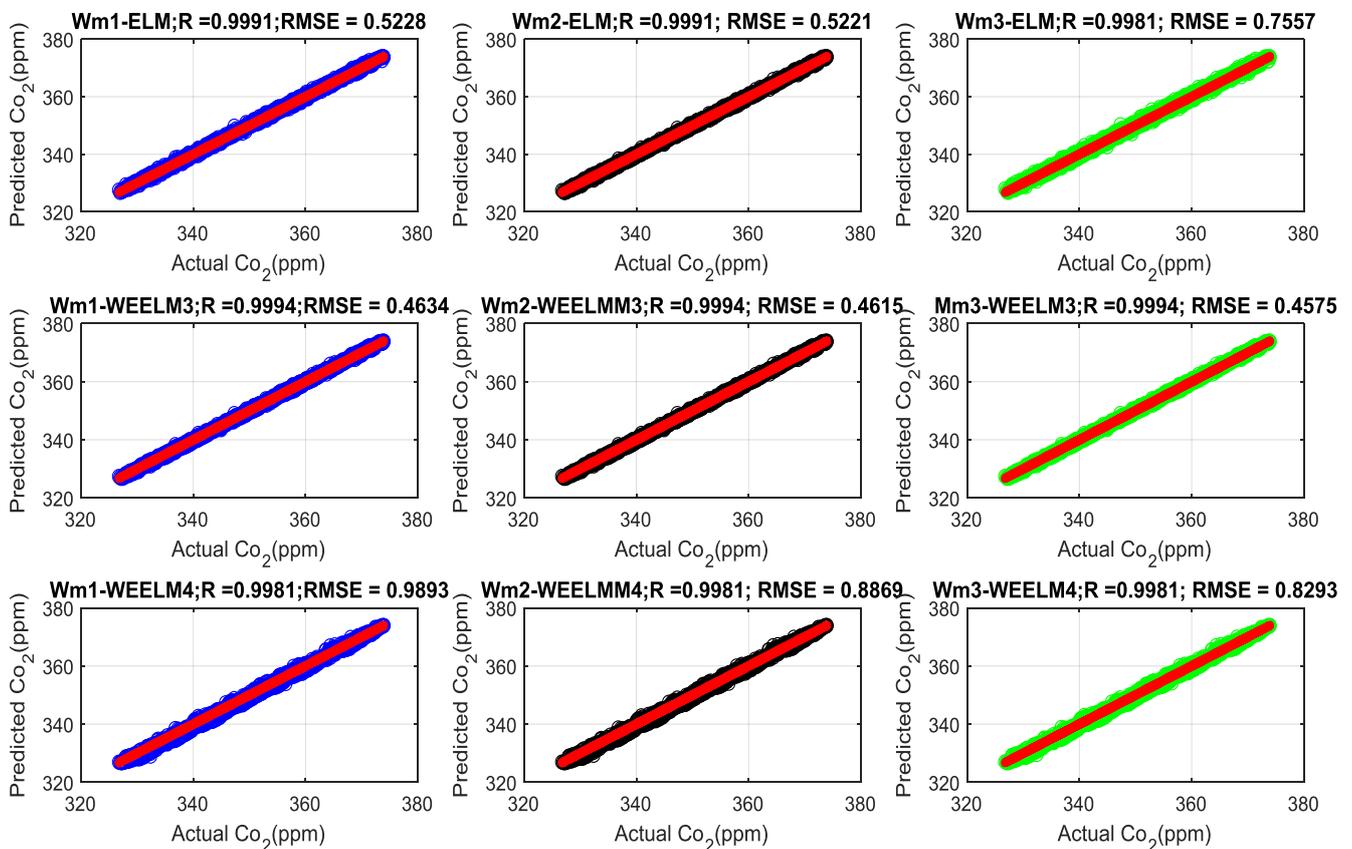


Figure 5. Weekly time scale of the training phase: classical ELM and W-EELM models

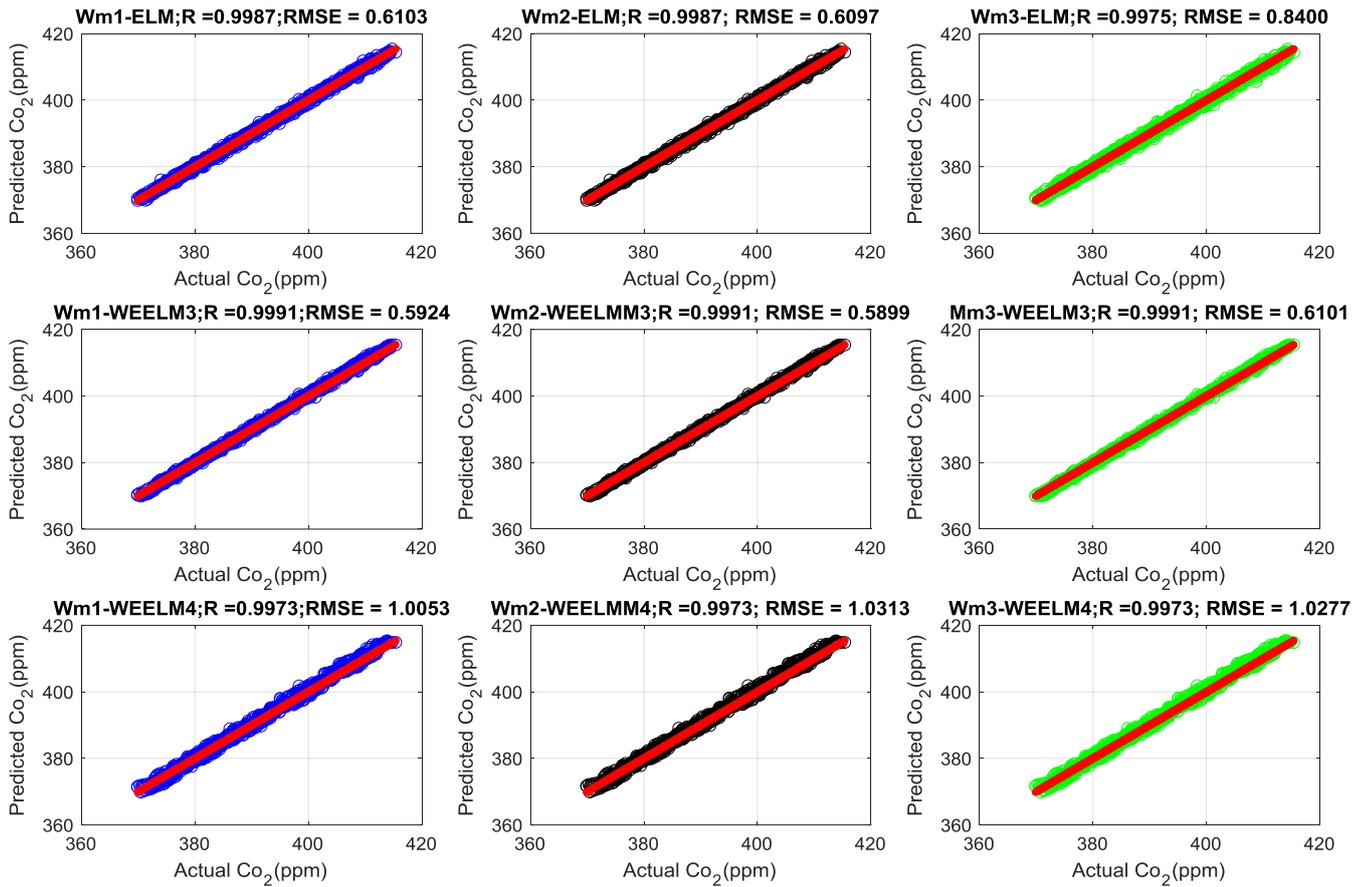


Figure 6. Weekly time scale of the testing phase: classical ELM and W-EELM models

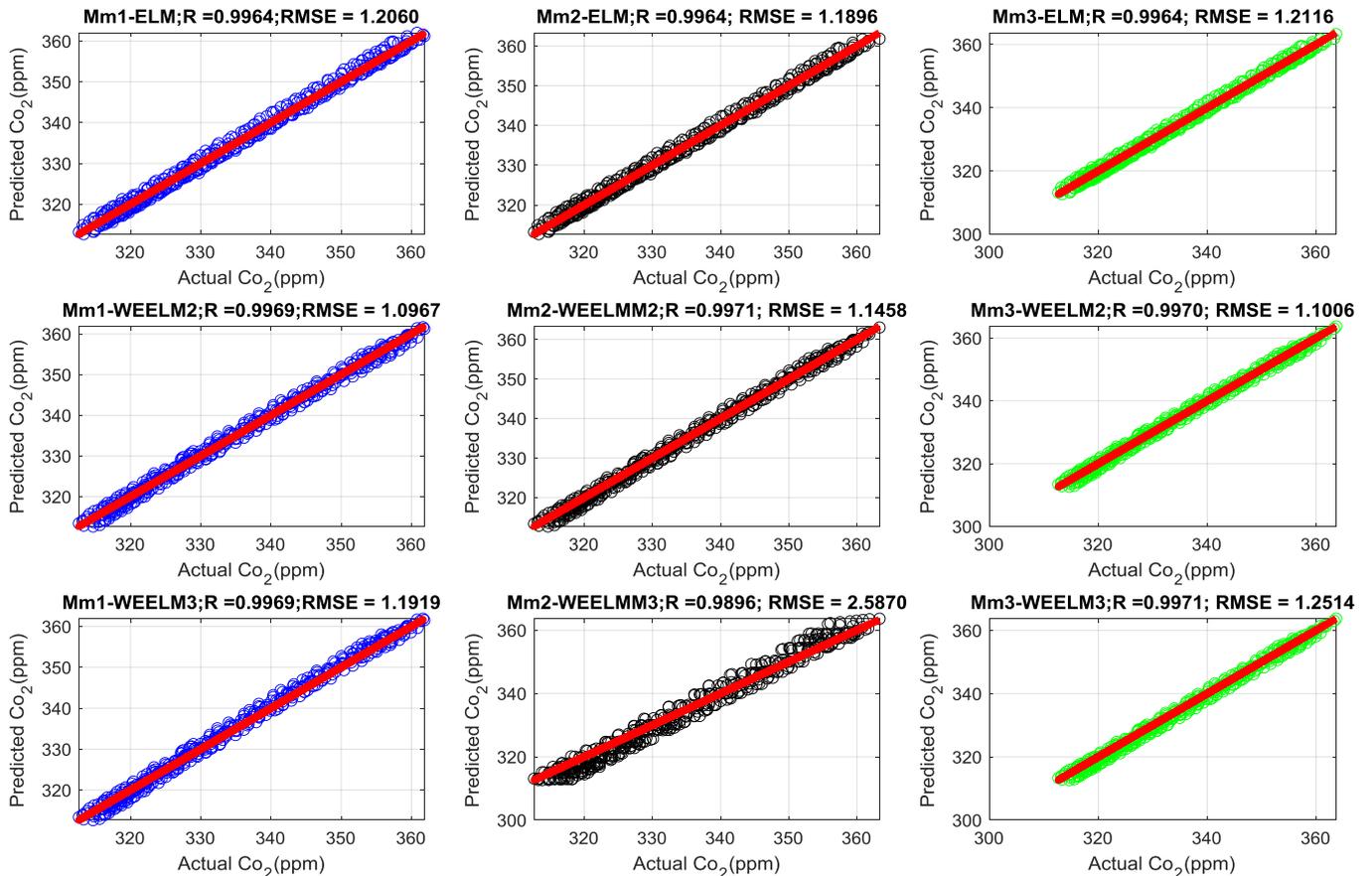


Figure 7. Monthly time scale of the training phase: classical ELM and W-EELM models

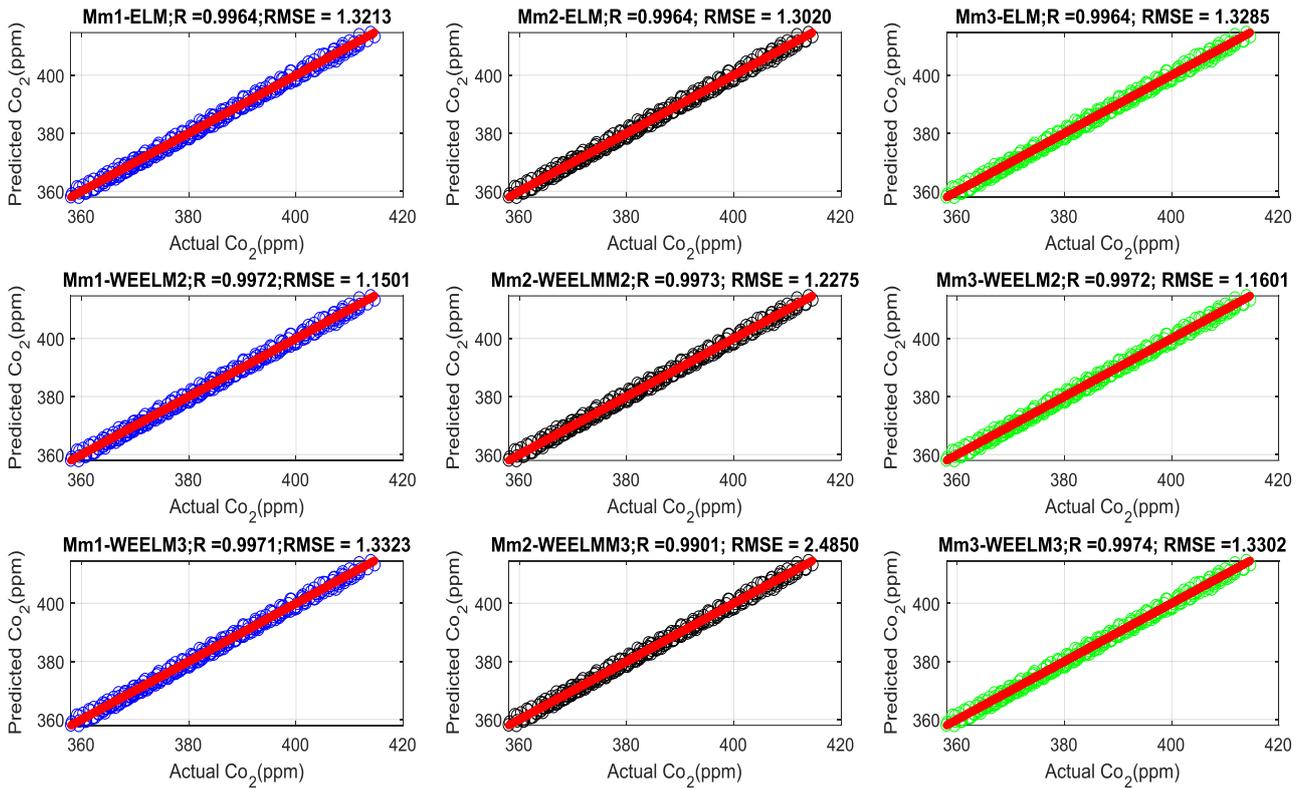


Figure 8. Monthly time scale of the testing phase: classical ELM and W-EELM models

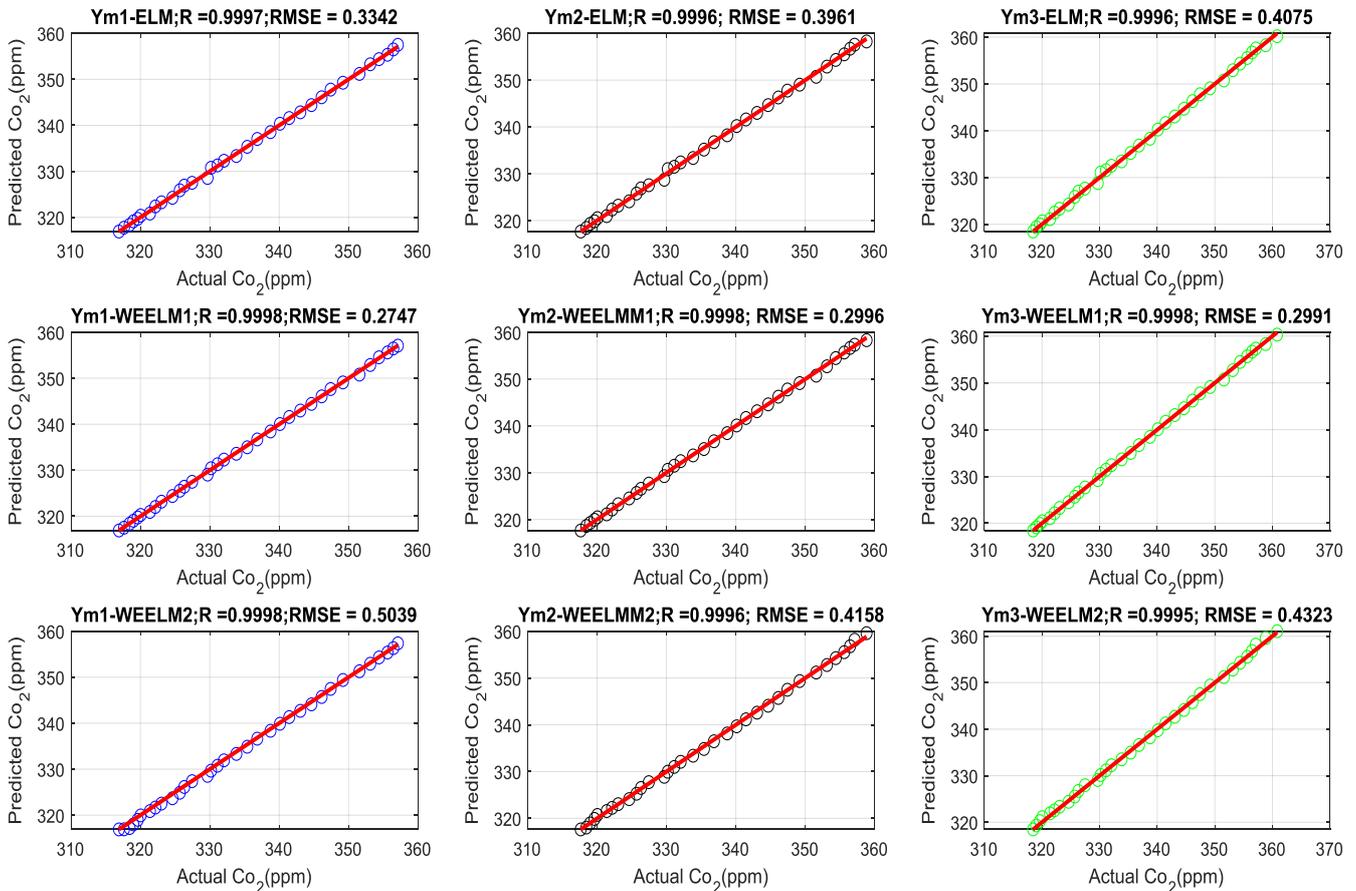


Figure 9. Yearly time scale of the training phase: classical ELM and W-EELM models

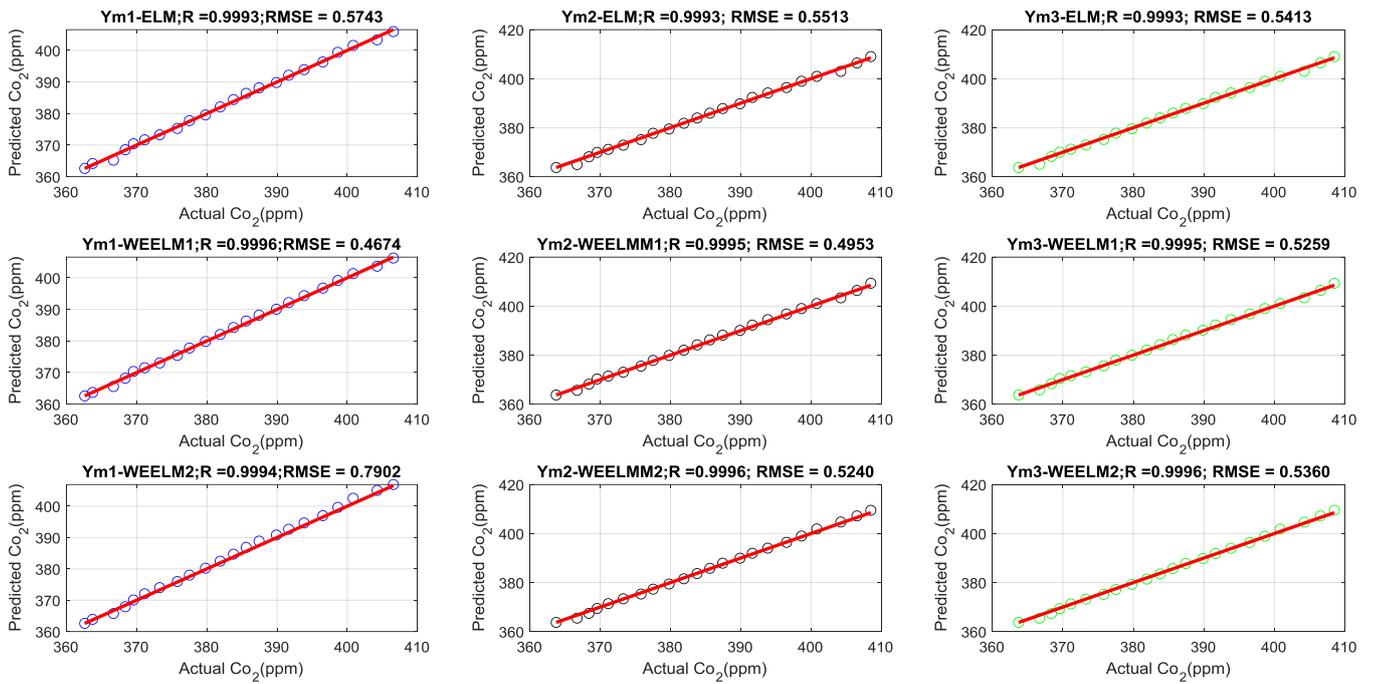


Figure 10. Yearly time scale of the testing phase: classical ELM and W-EELM models

The RE value is calculated for each observation during the testing phase to provide a much more meaningful graphical assessment of the prediction error. In addition, the RE value can provide a meaningful and clear explanation of the ability of the predictive models to predict CO₂ gas emissions on different time scales. The results presented in Figures 11 to 13 show that for each time scale (i.e., weekly, monthly, and yearly), the hybrid models provide less RE and MARE than the classical prediction approaches, except for the models that require the highest decomposition level. For example, the weekly CO₂ prediction models provided the lowest value of RE, with only one observation just above 0.5%. The classical models, on the other hand, provided approximately larger values of RE and a number of observations above the benchmark of 0.5%, ranging from 3 to 10.

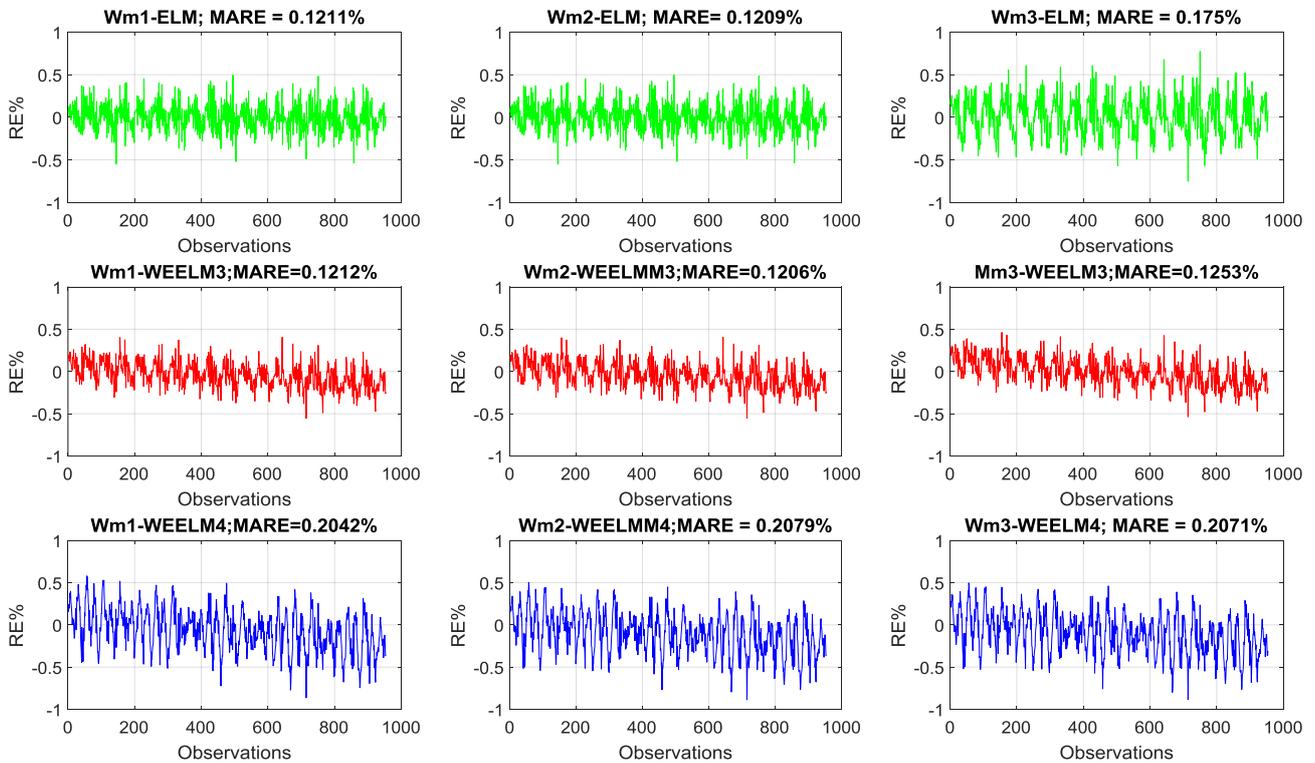


Figure 11. Distribution of relative error for each predictive model during the testing phase: weekly time scale

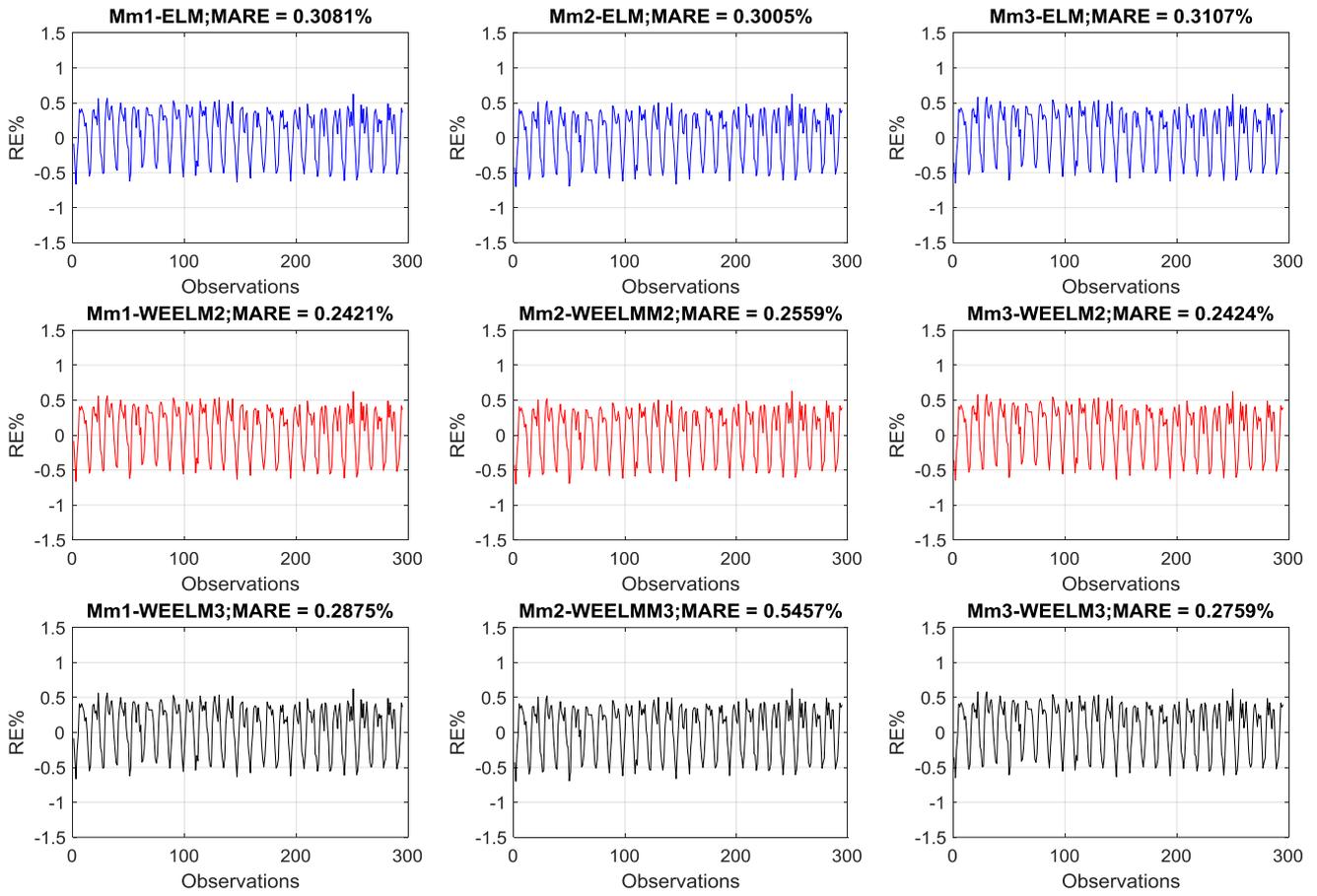


Figure 12. Distribution of relative error for each predictive model during the testing phase: monthly time scale

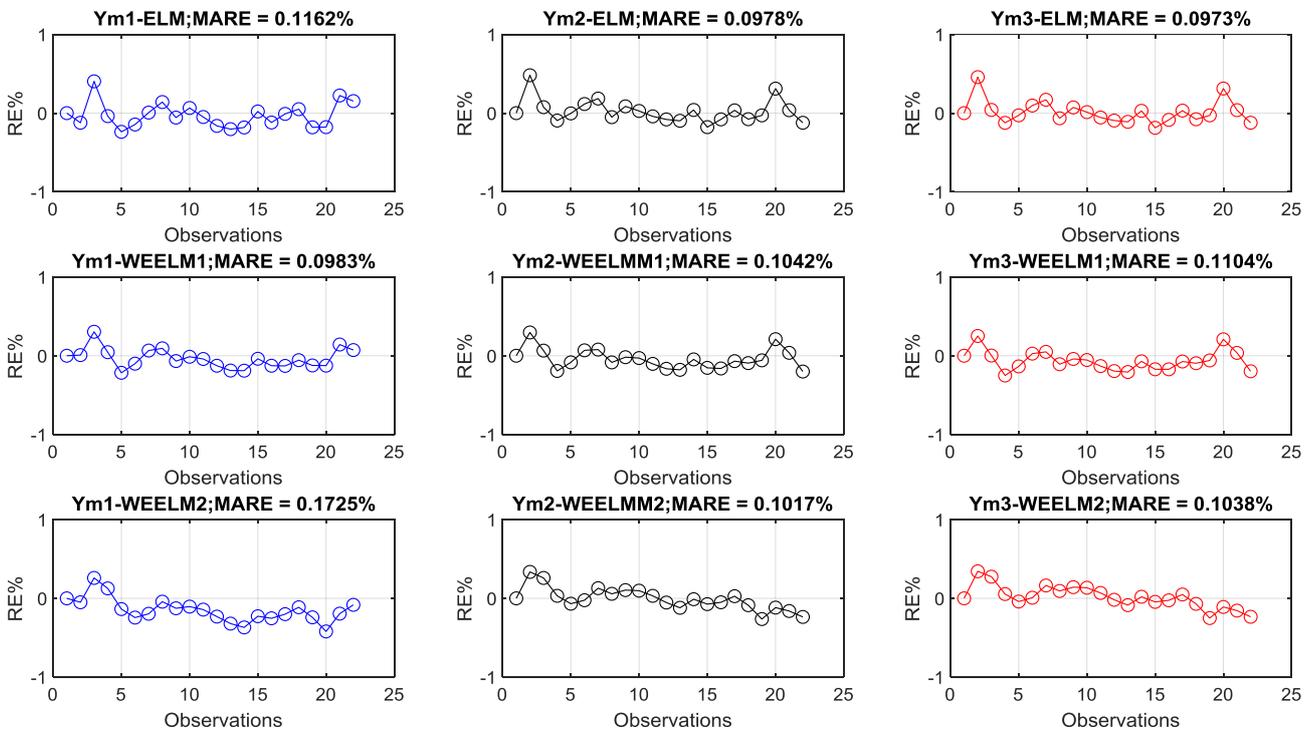


Figure 13. Distribution of relative error for each predictive model during the testing phase: yearly time scale

In addition, the Mm1-WEELM² and Ym1-WEELM¹ models have much lower relative error in the other predictions at monthly and yearly time scales compared to the traditional models. Finally, the best MARE criterion of all proposed models was found to be 0.1206%, 0.2421%, and 0.0983% for weekly, monthly, and yearly time series forecasts,

respectively. The boxplot was created for all the models used in this study to give a more informative and clear overview of the outliers and robustness of the models during the testing phase [51].

Figures 14 to 16 show the boxplots of the residuals between the actual and predicted CO₂ levels for all models. As Figure 14 shows, all predictive models for weekly scale had higher outliers than Wm1-WEELM³, Wm2-WEELM³ Wm3-WEELM³ models. However, the Wm3-WEELM³ model shows a tendency toward the third quartile, while the other models show smaller interquartile range (IQR) errors. As for the monthly time scale, only three models (Mm1-WEELM², Mm2-WEELM² and Mm3-WEELM²) had smaller residual error among all nine models, as shown in Figure 15. In addition, Mm2-WEELM² model showed a slight trend toward the third quartile. Finally, for the yearly estimates, the following models (Ym1-WEELM¹, Ym2-WEELM¹, and Ym3-WEELM¹) performed best among all nine models, as shown in Figure 16. The Ym2-WEELM¹ and Ym3-WEELM¹ models had the highest range of suspected outliers (1.5 IQR).

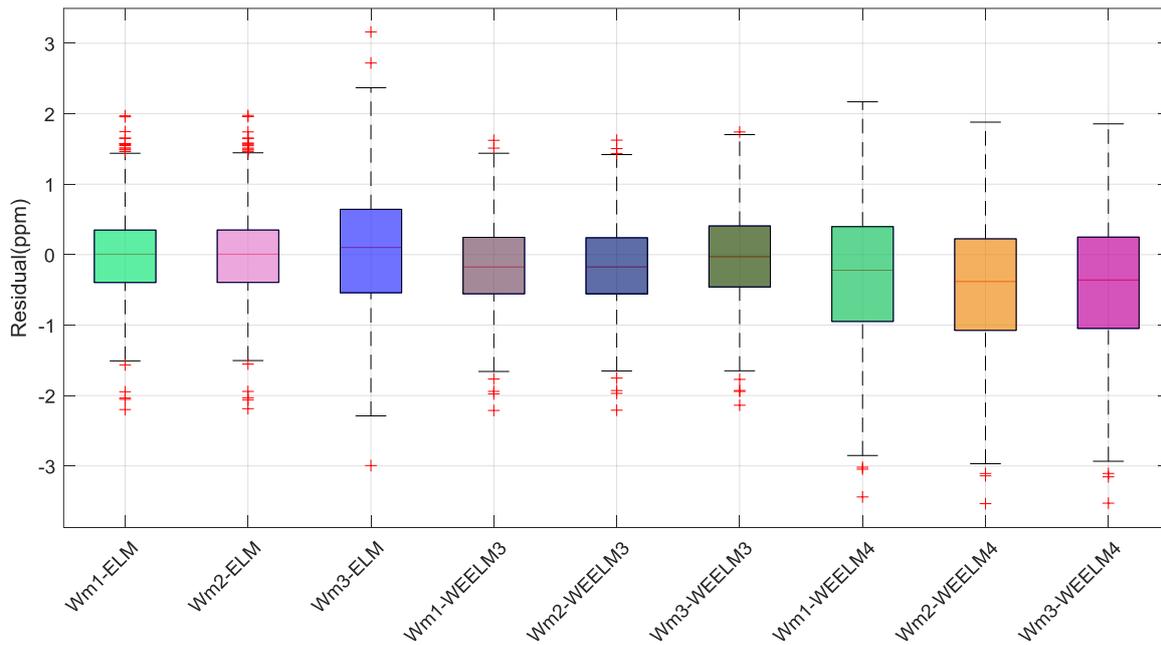


Figure 14. The box plots graphical presentation over the testing phases for all applied predictive models: weekly

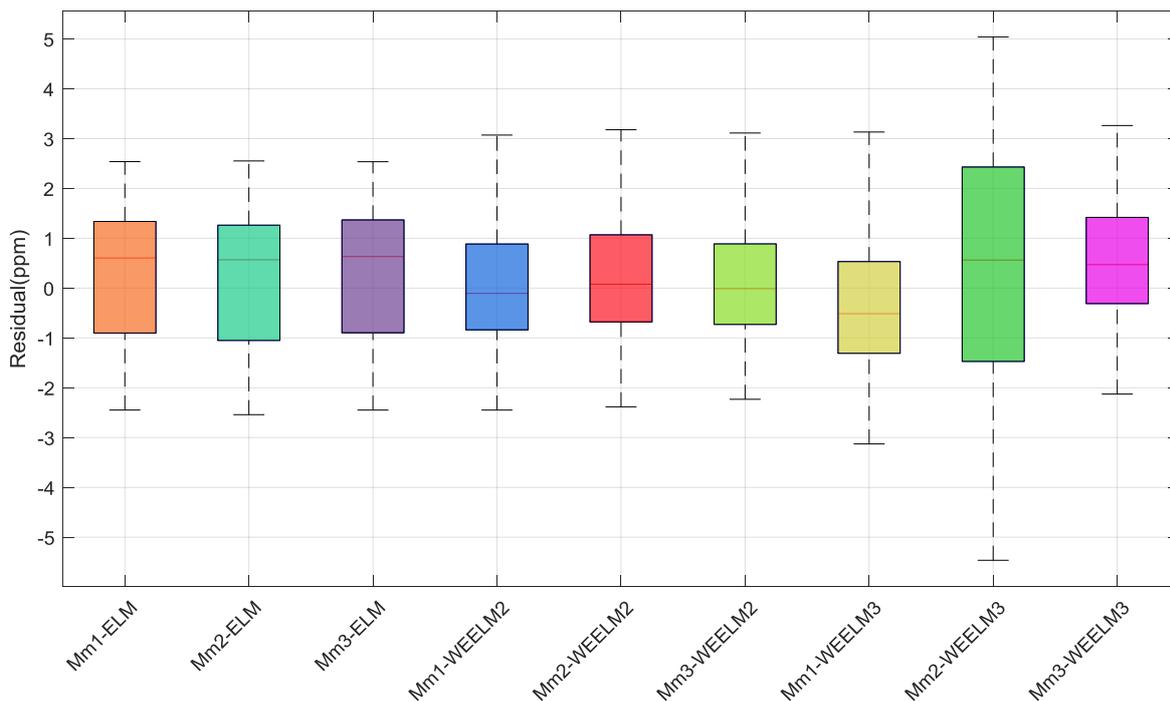


Figure 15. The boxplots for the testing phases for all applied predictive models: weekly scale

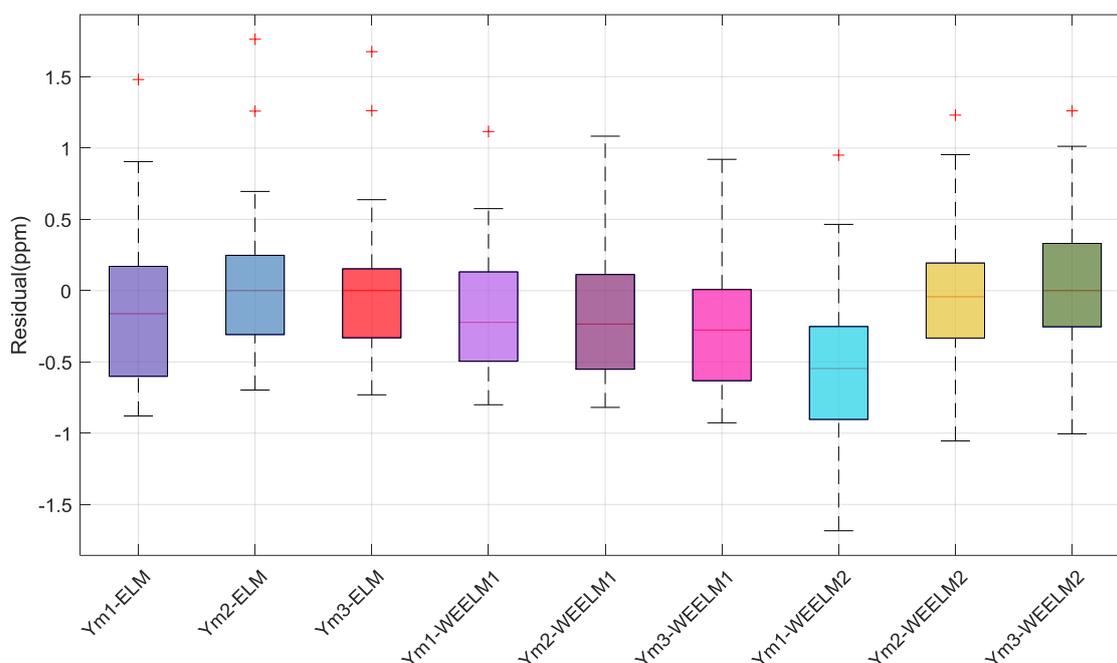


Figure 16. The box plots for the testing phases for all applied predictive models: yearly scale

4. Conclusion

Reliable predictions can help decision makers to take effective actions to reduce CO₂ concentration. In this work, the efficiency of W-EELM and the classical model ELM in predicting CO₂ emissions at different time scales (i.e., short-, medium-, and long-term) is investigated. The main contribution of this study is to improve the predictive capacity and stability of the classical model ELM by denoising the input data using the technique WT and using the algorithm COD instead of SVD to calculate the weights of the hidden layer of ELM more accurately. The CO₂ data were collected from Mauna Loa Observatory Station, 60% of the data sets were used for training and calibration, and the rest (end of the time series) was used for model accuracy testing. The predictive capacity of the two models was evaluated using various statistical metrics. The superiority of the proposed model capacities over the classical models was clearly shown in the reduction of the value of the RMSE criterion in the testing phases. The greatest improvement was obtained in the prediction on the yearly time scale (16.65%), followed by the monthly (11.67%) and weekly (3.25%).

5. Declarations

5.1. Author Contributions

Conceptualization: M.M.H., S.F.M.R., and M.K.A.; methodology: M.M.H. and M.A.A.; software: M.M.H. and M.K.A.; validation: M.M.H., S.F.M.R., and M.K.A.; formal analysis: M.M.H., and S.F.M.R.; investigation: M.A.A.; resources, M.M.H. and M.K.A.; data curation: M.M.H.; writing—original draft preparation: M.M.H., M.K.A., S.F.M.R., and M.A.A.; writing—review and editing, M.M.H., M.K.A., and S.F.M.R.; visualization, M.M.H.; supervision, N.A.; project administration: M.M.H. All authors have read and agree to the published version of the manuscript.

5.2. Data Availability Statement

The data used in this study is available online (<https://www.esrl.noaa.gov/gmd/>).

5.3. Funding and Acknowledgements

The authors would like to thank Al-Maarif University College for funding this research.

5.4. Conflicts of Interest

The authors declare no conflict of interest.

6. References

- [1] Qiao, W., Lu, H., Zhou, G., Azimi, M., Yang, Q., & Tian, W. (2020). A hybrid algorithm for carbon dioxide emissions forecasting based on improved lion swarm optimizer. *Journal of Cleaner Production*, 244, 118612. doi:10.1016/j.jclepro.2019.118612.

- [2] Acheampong, A. O., & Boateng, E. B. (2019). Modelling carbon emission intensity: Application of artificial neural network. *Journal of Cleaner Production*, 225, 833–856. doi:10.1016/j.jclepro.2019.03.352.
- [3] Li, F., Zhang, J., Oko, E., & Wang, M. (2017). Modelling of a post-combustion CO₂ capture process using extreme learning machine. *International Journal of Coal Science and Technology*, 4(1), 33–40. doi:10.1007/s40789-017-0158-1.
- [4] Alias, N. I., Jayakumar, J. K., & Zain, S. M. (2018). Characterization of waste cooking oil for biodiesel production. *Jurnal Kejuruteraan*, 1(2), 79–83. doi:10.17576/jkukm-2018-si1(2)-10.
- [5] Ridzuan, N. H. A. M., Marwan, N. F., Khalid, N., Ali, M. H., & Tseng, M. L. (2020). Effects of agriculture, renewable energy, and economic growth on carbon dioxide emissions: Evidence of the environmental Kuznets curve. *Resources, Conservation and Recycling*, 160, 104879. doi:10.1016/j.resconrec.2020.104879.
- [6] Marjanović, V., Milovančević, M., & Mladenović, I. (2016). Prediction of GDP growth rate based on carbon dioxide (CO₂) emissions. *Journal of CO₂ Utilization*, 16, 212–217. doi:10.1016/j.jcou.2016.07.009.
- [7] Yatim, P., Mamat, M. N., Mohamad-Zailani, S. H., & Ramlee, S. (2016). Energy policy shifts towards sustainable energy future for Malaysia. *Clean Technologies and Environmental Policy*, 18(6), 1685–1695. doi:10.1007/s10098-016-1151-x.
- [8] Acheampong, A. O. (2018). Economic growth, CO₂ emissions and energy consumption: What causes what and where? *Energy Economics*, 74, 677–692. doi:10.1016/j.eneco.2018.07.022.
- [9] Yaacob, N. F. F., Yazid, M. R. M., Maulud, K. N. A., & Basri, N. E. A. (2020). A review of the measurement method, analysis and implementation policy of carbon dioxide emission from transportation. *Sustainability (Switzerland)*, 12(14), 5873. doi:10.3390/su12145873.
- [10] An, Q., Wen, Y., Xiong, B., Yang, M., & Chen, X. (2017). Allocation of carbon dioxide emission permits with the minimum cost for Chinese provinces in big data environment. *Journal of Cleaner Production*, 142, 886–893. doi:10.1016/j.jclepro.2016.07.072.
- [11] Foster, V., & Bedrosyan, D. (2014). Understanding CO₂ emissions from the global energy sector. *Live Wire*, 2014/5, World Bank, Washington, United States.
- [12] IEA. (2020). Global CO₂ emissions in 2019. International Energy Agency (IEA), Paris, France. Available online: <https://www.iea.org/articles/global-co2-emissions-in-2019> (accessed on March 2023).
- [13] Gielen, D., Boshell, F., Saygin, D., Bazilian, M. D., Wagner, N., & Gorini, R. (2019). The role of renewable energy in the global energy transformation. *Energy Strategy Reviews*, 24, 38–50. doi:10.1016/j.esr.2019.01.006.
- [14] Lu, H., Huang, K., Azimi, M., & Guo, L. (2019). Blockchain technology in the oil and gas industry: A review of applications, opportunities, challenges, and risks. *IEEE Access*, 7, 41426–41444. doi:10.1109/ACCESS.2019.2907695.
- [15] Reijnders, L., & Huijbregts, M. A. J. (2008). Palm oil and the emission of carbon-based greenhouse gases. *Journal of Cleaner Production*, 16(4), 477–482. doi:10.1016/j.jclepro.2006.07.054.
- [16] Zhang, Y., & Fu, B. (2023). Impact of China's establishment of ecological civilization pilot zones on carbon dioxide emissions. *Journal of Environmental Management*, 325, 116652. doi:10.1016/j.jenvman.2022.116652.
- [17] Baareh, A. K. (2013). Solving the Carbon Dioxide Emission Estimation Problem: An Artificial Neural Network Model. *Journal of Software Engineering and Applications*, 06(07), 338–342. doi:10.4236/jsea.2013.67042.
- [18] Fang, D., Zhang, X., Yu, Q., Jin, T. C., & Tian, L. (2018). A novel method for carbon dioxide emission forecasting based on improved Gaussian processes regression. *Journal of Cleaner Production*, 173, 143–150. doi:10.1016/j.jclepro.2017.05.102.
- [19] Lin, B., & Ahmad, I. (2017). Analysis of energy related carbon dioxide emission and reduction potential in Pakistan. *Journal of Cleaner Production*, 143, 278–287. doi:10.1016/j.jclepro.2016.12.113.
- [20] Oh, I., Wehrmeyer, W., & Mulugetta, Y. (2010). Decomposition analysis and mitigation strategies of CO₂ emissions from energy consumption in South Korea. *Energy Policy*, 38(1), 364–377. doi:10.1016/j.enpol.2009.09.027.
- [21] Mardani, A., Liao, H., Nilashi, M., Alrasheedi, M., & Cavallaro, F. (2020). A multi-stage method to predict carbon dioxide emissions using dimensionality reduction, clustering, and machine learning techniques. *Journal of Cleaner Production*, 275, 122942. doi:10.1016/j.jclepro.2020.122942.
- [22] Saleh, C., Dzakiyullah, N. R., & Nugroho, J. B. (2016). Carbon dioxide emission prediction using support vector machine. *IOP Conference Series: Materials Science and Engineering*, 114, 012148. doi:10.1088/1757-899x/114/1/012148.
- [23] Yaseen, Z. M., Sulaiman, S. O., Deo, R. C., & Chau, K. W. (2019). An enhanced extreme learning machine model for river flow forecasting: State-of-the-art, practical applications in water resource engineering area and future research direction. *Journal of Hydrology*, 569, 387–408. doi:10.1016/j.jhydrol.2018.11.069.

- [24] Shao, Z., Zhai, Q., Wu, J., & Guan, X. (2021). Data Based Linear Power Flow Model: Investigation of a Least-Squares Based Approximation. *IEEE Transactions on Power Systems*, 36(5), 4246–4258. doi:10.1109/TPWRS.2021.3062359.
- [25] Global Monitoring Laboratory. (2023). U.S. Department of Commerce, National Oceanic & Atmospheric Administration, NOAA Research, Boulder, United States. Available online: <https://gml.noaa.gov/obop/mlo/aboutus/siteInformation/mlosite.html> (accessed on March 2023).
- [26] G.M. Laboratory (2020). Earth System Research Laboratories. U.S. Department of Commerce, National Oceanic & Atmospheric Administration, NOAA Research, Boulder, United States. Available online: <https://gml.noaa.gov/> (accessed on March 2023).
- [27] Huang, G.-B., Zhu, Q.-Y., & Siew, C.-K. (2006). Extreme learning machine: Theory and applications. *Neurocomputing*, 70(1–3), 489–501. doi:10.1016/j.neucom.2005.12.126.
- [28] Wang, W., Peng, W., Xu, J., Zhang, R., & Zhao, Y. (2018). Driving factor analysis and forecasting of CO₂ emissions from power output in China using scenario analysis and CSCWOA-ELM method. *Energies*, 11(10), 2709. doi:10.3390/en11102709.
- [29] Masmoudi, S., Elghazel, H., Taieb, D., Yazar, O., & Kallel, A. (2020). A machine-learning framework for predicting multiple air pollutants' concentrations via multi-target regression and feature selection. *Science of the Total Environment*, 715, 136991. doi:10.1016/j.scitotenv.2020.136991.
- [30] Wei, S., Yuwei, W., & Chongchong, Z. (2018). Forecasting CO₂ emissions in Hebei, China, through moth-flame optimization based on the random forest and extreme learning machine. *Environmental Science and Pollution Research*, 25(29), 28985–28997. doi:10.1007/s11356-018-2738-z.
- [31] Hameed, M. M., Alomar, M. K., Khaleel, F., & Al-Ansari, N. (2021). An Extra Tree Regression Model for Discharge Coefficient Prediction: Novel, Practical Applications in the Hydraulic Sector and Future Research Directions. *Mathematical Problems in Engineering*, 2021, 7001710. doi:10.1155/2021/7001710.
- [32] Hameed, M. M., Alomar, M. K., Mohd Razali, S. F., Kareem Khalaf, M. A., Baniya, W. J., Sharafati, A., & Alsaadi, M. A. (2021). Application of Artificial Intelligence Models for Evapotranspiration Prediction along the Southern Coast of Turkey. *Complexity*, 2021, 8850243. doi:10.1155/2021/8850243.
- [33] Hameed, M. M., AlOmar, M. K., Baniya, W. J., & AlSaadi, M. A. (2022). Prediction of high-strength concrete: high-order response surface methodology modeling approach. *Engineering with Computers*, 38(2), 1655–1668. doi:10.1007/s00366-021-01284-z.
- [34] Hough, P. D., & Vavasis, S. A. (1997). Complete orthogonal decomposition for weighted least squares. *SIAM Journal on Matrix Analysis and Applications*, 18(2), 369–392. doi:10.1137/S089547989528079X.
- [35] Guo, W., Xu, T., & Lu, Z. (2016). An integrated chaotic time series prediction model based on efficient extreme learning machine and differential evolution. *Neural Computing and Applications*, 27(4), 883–898. doi:10.1007/s00521-015-1903-2.
- [36] Addison, P. S., Murray, K. B., & Watson, J. N. (2001). Wavelet Transform Analysis of Open Channel Wake Flows. *Journal of Engineering Mechanics*, 127(1), 58–70. doi:10.1061/(asce)0733-9399(2001)127:1(58).
- [37] Zhu, H., Hu, W., Jing, Y., Cao, Y., Bi, R., & Yang, W. (2018). Soil organic carbon prediction based on scale-specific relationships with environmental factors by discrete wavelet transform. *Geoderma*, 330, 9–18. doi:10.1016/j.geoderma.2018.05.029.
- [38] Nourani, V., Hosseini Baghanam, A., Adamowski, J., & Kisi, O. (2014). Applications of hybrid wavelet-Artificial Intelligence models in hydrology: A review. *Journal of Hydrology*, 514, 358–377. doi:10.1016/j.jhydrol.2014.03.057.
- [39] Sharghi, E., Nourani, V., Najafi, H., & Molajou, A. (2018). Emotional ANN (EANN) and Wavelet-ANN (WANN) Approaches for Markovian and Seasonal Based Modeling of Rainfall-Runoff Process. *Water Resources Management*, 32(10), 3441–3456. doi:10.1007/s11269-018-2000-y.
- [40] Fosten, J. (2019). CO₂ emissions and economic activity: A short-to-medium run perspective. *Energy Economics*, 83, 415–429. doi:10.1016/j.eneco.2019.07.015.
- [41] Pla, C., Cuezva, S., Garcia-Anton, E., Fernandez-Cortes, A., Cañaveras, J. C., Sanchez-Moral, S., & Benavente, D. (2016). Changes in the CO₂ dynamics in near-surface cavities under a future warming scenario: Factors and evidence from the field and experimental findings. *Science of the Total Environment*, 565, 1151–1164. doi:10.1016/j.scitotenv.2016.05.160.
- [42] Ma, D., Tan, W., Zhang, Z., Wang, X., Xia, F., & Hu, J. (2018). Recognition of leak CO₂ with wavelet analysis based on correlation monitoring between CO₂ and O₂ in atmosphere. *Process Safety and Environmental Protection*, 114, 64–78. doi:10.1016/j.psep.2017.12.009.
- [43] Oh, Y. Y., Yun, S. T., Yu, S., Kim, H. J., & Jun, S. C. (2018). Characterization of Environmental Drivers Controlling the Baseline of Soil Surface CO₂ Flux using Wavelet-based Multiresolution State-Space Model and Wavelet Denoising. *Energy Procedia*, 154, 157–162. doi:10.1016/j.egypro.2018.11.026.

- [44] Ebrahimi, H., & Rajaei, T. (2017). Simulation of groundwater level variations using wavelet combined with neural network, linear regression and support vector machine. *Global and Planetary Change*, 148, 181–191. doi:10.1016/j.gloplacha.2016.11.014.
- [45] Nury, A. H., Hasan, K., & Alam, Md. J. B. (2017). Comparative study of wavelet-ARIMA and wavelet-ANN models for temperature time series data in northeastern Bangladesh. *Journal of King Saud University - Science*, 29(1), 47–61. doi:10.1016/j.jksus.2015.12.002.
- [46] AlOmar, M. K., Hameed, M. M., & AlSaadi, M. A. (2020). Multi hours ahead prediction of surface ozone gas concentration: Robust artificial intelligence approach. *Atmospheric Pollution Research*, 11(9), 1572–1587. doi:10.1016/j.apr.2020.06.024.
- [47] Yan, H., Mou, Y., Xu, X., Du, J., Wang, R., & Liu, P. (2018). A comparative assessment of predicting CH₄ adsorption on different activated carbons using generalized regression neural network (GRNN), and adaptive network-based fuzzy inference system (ANFIS). *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*, 41(16), 1983–1992. doi:10.1080/15567036.2018.1548527.
- [48] Metia, S., Ha, Q. P., Duc, H. N., & Scorgie, Y. (2020). Urban air pollution estimation using unscented Kalman filtered inverse modeling with scaled monitoring data. *Sustainable Cities and Society*, 54, 101970. doi:10.1016/j.scs.2019.101970.
- [49] Hameed, M., Sharqi, S. S., Yaseen, Z. M., Afan, H. A., Hussain, A., & Elshafie, A. (2017). Application of artificial intelligence (AI) techniques in water quality index prediction: a case study in tropical region, Malaysia. *Neural Computing and Applications*, 28(1), 893–905. doi:10.1007/s00521-016-2404-7.
- [50] Ali, A., Abdulrahman, A., Garg, S., Maqsood, K., & Murshid, G. (2019). Application of artificial neural networks (ANN) for vapor-liquid-solid equilibrium prediction for CH₄-CO₂ binary mixture. *Greenhouse Gases: Science and Technology*, 9(1), 67–78. doi:10.1002/ghg.1833.
- [51] Alomar, M. K., Hameed, M. M., Al-Ansari, N., & Alsaadi, M. A. (2020). Data-Driven Model for the Prediction of Total Dissolved Gas: Robust Artificial Intelligence Approach. *Advances in Civil Engineering*, 2020, 6618842. doi:10.1155/2020/6618842.