



Research paper

A comparative analysis to forecast carbon dioxide emissions

Md. Omer Faruque^a, Md. Afser Jani Rabby^a, Md. Alamgir Hossain^b, Md. Rashidul Islam^a,
Md Mamun Ur Rashid^c, S.M. Mueyeen^{d,*}^a Department of EEE, Rajshahi University of Engineering & Technology (RUET), Rajshahi, Bangladesh^b Queensland Micro-and Nano-technology Centre, Griffith University, Nathan, 4111, Brisbane, Australia^c Department of Electrical & Electronic Engineering, National Institute of Textile Engineering and Research (NITER), Dhaka 1350, Bangladesh^d Department of Electrical Engineering, Qatar University, Doha, 2713, Qatar

ARTICLE INFO

Article history:

Received 23 February 2022

Received in revised form 6 May 2022

Accepted 15 June 2022

Available online 28 June 2022

Keywords:

CO₂ emissions

Forecasting

Deep learning

FMOLS

CNN-LSTM

ABSTRACT

Despite the growing knowledge and commitment to climate change, carbon dioxide (CO₂) emissions continue to rise dramatically throughout the planet. In recent years, the consequences of climate change have become more catastrophic and have attracted widespread attention globally. CO₂ emissions from the energy industry have lately been highlighted as one of the world's most pressing concerns for all countries. This paper examines the relationships between CO₂ emissions, electrical energy consumption, and gross domestic product (GDP) in Bangladesh from 1972 to 2019 in the first section. In this purpose, we applied the fully modified ordinary least squares (FMOLS) approach. The findings indicate that CO₂ emissions, electrical energy consumption, and GDP have a statistically significant long-term cointegrating relationship. Developing an accurate CO₂ emissions forecasting model is crucial for tackling it safely. This leads to the second step, which involves formulating the multivariate time series CO₂ emissions forecasting challenges considering its influential factors. Based on multivariate time series prediction, four deep learning algorithms are analyzed in this work, those are convolution neural network (CNN), CNN long short-term memory (CNN-LSTM), long short-term memory (LSTM), and dense neural network (DNN). The root mean square error (RMSE), mean absolute error (MAE) and mean absolute percentage error (MAPE) are used to analyze and compare the performances of the predictive models. The prediction errors in MAPE of the CNN, CNN-LSTM, LSTM, and DNN are 15.043, 5.065, 5.377, and 3.678, respectively. After evaluating those deep learning models, a multivariate polynomial regression has also been employed to forecast CO₂ emissions. It seems to have nearly similar accuracy as the LSTM model, having a MAPE of 5.541.

© 2022 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

One of the most burning concerns confronting modern civilization is environmental issues. The principal cause underlying climate change is the production of greenhouse gases (GHGs), which mostly comprise carbon dioxide (CO₂) (Mitić et al., 2017). The effects of climate change can now be observed in every corner of the globe. As a consequence of global warming, Bangladesh is one of the most vulnerable countries to this climatic change (Sarkar et al., 2015), and due to the environmental concerns, Bangladesh is anticipated to decline by 3.4% of its 2015 gross domestic product (GDP) every year (Hasan and Chongbo, 2020). About 75% of the human-caused CO₂ in the last 20 years originated from fossil fuel combustion. However, fossil fuels comprise the dominant source of electricity, manufacturing activity, transportation, and consumption of goods and services, all of which

are directly linked to economic growth. As a result, it is vital to forecast CO₂ emissions from the perspective of energy consumption for climate policy and energy planning (Rayhan et al., 2018). As energy consumption and GDP have a strong connection with CO₂ emissions, those must be taken into account when predicting CO₂ emissions. Many scholars have analyzed and assessed the literature on the association between CO₂ emissions, environmental factors and other energy usages.

1.1. Literature review

CO₂ emission reduction is a worldwide issue that must be addressed for building sustainable society. The global carbon emission leads to livable issues of many countries and regions. Between 1970 and 2010, Bouznit and Pablo-Romero (2016) inspected the relationship between Algerian GDP and CO₂ emissions and the data show that GDP and CO₂ emissions have a substantial connection. According to Aftab et al. (2021), the

* Corresponding author.

E-mail address: sm.mueyeen@qu.edu.qa (S.M. Mueyeen).

autoregressive distributed lag model (ARDL) and Johansen cointegration were used to examine the relationship between CO₂ emissions, energy consumption, and economic progress both in bivariate and multivariate models for 1971–2019 in Pakistan. These findings indicate a positive and significant impact of a fast-growing economy and energy consumption on CO₂ emissions.

Khobai et al. (2017) reviewed the significant correlation between GDP and energy usage in South Africa. According to their document, there is a two-way causal determinant of economic progress and energy use. In Turkey, Gökmenoğlu and Taspınar (2016) discovered one-way statistical relationship between economic progress and energy usage. Ghosh et al. (2014) scrutinized the interconnection between economic growth, CO₂ emissions, and energy consumption in Bangladesh and found that energy consumption has a positive and significant influence on economic growth, but CO₂ emissions have a negative and minor impact. Using fully modified ordinary least squares (FMOLS) and Johansen cointegration approach, Mirza and Kanwal (2017) found population density and energy usages had a substantial influence on environmental degradation in Pakistan. To check the robustness of these literatures exhibit that Pakistan's population density and energy consumption contribute to CO₂ emissions. For the country of Bangladesh, Sarkar et al. (2018) investigated the trends in energy consumption and CO₂ emissions. It has been proven that the rising rate of CO₂ emissions in Bangladesh is higher than the increasing rate of GDP and energy consumption. Cai et al. (2018) utilized the ARDL limit test, to determine the relationship between renewable energy usages, economic growth, and CO₂ emissions for a group of seven (G7) countries. When real GDP per capita and CO₂ emissions are employed as dependent variables, it shows that cointegration exists in G7 countries. From 1997 to 2016, Salari et al. (2021) inspected the relationship between CO₂ emissions, energy consumption, and GDP at the state level in the United States. For both static and dynamic models, the results show a long-run relationship between various types of energy consumption and CO₂ emissions. Wasti and Zaidi (2020) examined the affiliation between CO₂ emissions, energy consumption, GDP, and trade liberalization in their research of Kuwait. This statistic reveals unidirectional causation between GDP and CO₂ emissions and between energy usage and trade liberalization. A study by Valadkhani et al. (2019) examined the contribution of several main energy sources (oil, coal, natural gas, hydroelectricity, and other renewables) to global CO₂ emissions. In 11 Asian states from 1960 to 2014, Rahman (2017) analyzed the correlation between CO₂ emissions, energy consumption, economic development, exports, and population density to evaluate the long-term consequences on CO₂ emissions by FMOLS and dynamic ordinary least squares (DOLS) techniques. These five factors appear to be cointegrated throughout time, based on the test findings. The research by Shahbaz et al. (2013) in Indonesia, Ang (2008) in France and Malaysia have exposed that economic expansion affects energy consumption and CO₂ emissions. Much recent researches, such as Saint Akadiri et al. (2019), You and Lv (2018) have used the KOF index to investigate the effect of globalization on CO₂ emissions. In addition to these, a summary of some recent studies on CO₂ emission is shown in the Table 1 below:

On the basis of these literature, CO₂ emissions can be influenced by a variety of factors, and it is important to forecast CO₂ emissions. As a result, various studies have recently been developed new models to predict CO₂ emissions, and some of them are on the way. According to Li et al. (2020), China has paid close attention to model and forecast China's CO₂ emissions since the Copenhagen climate summit. Various viewpoints and forecast outcomes have emerged on China's ability to meet its pledge to reduce CO₂ intensity. Yuan et al. (2012) prediction model showed

that CO₂ emissions would be reduced by 45% if China maintains annual economic growth rates of 7% and 6% in the 12th and 13th five year plans (FYP) periods, respectively. Using the logarithmic mean division index (LMDI) decomposition method, Islam et al. (2021) predicted that CO₂ emissions in Bangladesh would peak at 58.97 Mtoe by 2040. Heydari et al. (2019) used a general regression neural network (GRNN) and grey wolf optimization (GWO) to assess the trend of CO₂ emissions in Iran, Canada, and Italy, and the findings show that the suggested method is more accurate in long-term CO₂ emissions forecasting. Fang et al. (2018) utilized an enhanced Gaussian processes regression model to forecast CO₂ emissions, and he discovered that China's overall CO₂ emissions would continue to rise, but at a slower rate, and that the US and Japan will have a good handle on their CO₂ emissions soon. Hosseini et al. (2019) used time series and regression analysis to forecast Iran's CO₂ emissions in 2030. According to their findings, Iran is unlikely to keep its goal to the Paris agreement based on the business as usual (BAU's) assumptions. Moreover, Ofosu-Adarkwa et al. (2020) projected CO₂ emissions from the Chinese cement sector, and their proposed system (hybrid Verhulst-GM (1, N)) can forecast emissions with a 97% precision. According to Ameyaw et al. (2019), the long short-term memory (LSTM) based CO₂ from combustion projection in China depicts a declining trend up to 2030. Their analysis also stated that, if renewable energy investments are not increased, countries' intended nationally determined contributions (INDCs) will be jeopardized. Li (2020) utilized LSTM, support vector machines (SVM), and kernel least squares (KLS) models to forecast CO₂ emissions in China, with the findings demonstrating that KLS was more accurate than other current techniques. For the first time in China, Huang et al. (2019) used LSTM to undertake grey relationship analysis, principal component analysis for CO₂ emissions forecasting.

Amarpuri et al. (2019) in India, a deep learning hybrid method was used to predict CO₂ emissions. A deep learning hybrid model convolutional neural network-long short-term memory (CNN-LSTM) was utilized to make the prediction. To anticipate CO₂ emissions, Fatima et al. (2019) utilized simple exponential smoothing (SES) and autoregressive integrated moving average (ARIMA) models. They discovered that the ARIMA model is appropriate since it has the lowest fractional mean absolute error (FMAE) value.

1.2. Research gaps and contributions

Numerous studies have investigated the CO₂ emissions from a variety of perspectives, such as a causal relationship of CO₂ emissions and the factors that may or may not have influences on it. However, there are still some rooms to work on it, and some factors which motivate to do this work, e.g. the data samples which were used before being too small, not taking into consideration of some extremely crucial influencing factors, lack of focus on Bangladesh. It is worth mentioning that, Bangladesh badly needs the investigation of CO₂ forecasting because, in 2019, it has produced around 0.66 tonnes of CO₂ per capita, which was only 0.05 tons per capita in 1970, that indicates an annual average climbing rate of 5.48% (Anon, 2022). Taking into account all those influential factors on CO₂ emission, the problem formulation leads to multivariate time series forecasting methodologies, which are difficult to execute but provide more accurate results. Time series forecasting issues have been explicitly addressed by machine learning and regression techniques in the literature, however, to the best of the authors' knowledge, machine learning or deep learning approaches have not extensively been used to anticipate CO₂ emissions considering those influential factors. To fulfill the research gaps, this paper have evaluated several deep learning algorithms' efficacy, along with a multivariate polynomial regression. The study will have a significant

Table 1
Summary of recent studies.

Investigator(s)	Period	Nation(s)	Variables	Methods	Results
Rahman and Kashem (2017)	1972–2011	Bangladesh	CO ₂ , EN, IG	ARDL, Granger causality	EC → CO ₂ (+) IG → CO ₂ (+)
Muhammad et al. (2020)	1990–2014	Pakistan	CO ₂ , PD, EN	FMOLS, DOLS,	PD → CO ₂ (+) EC → CO ₂ (+)
Chontanawat (2020)	1971–2015	ASEAN	CO ₂ , EG, EN	Cointegra- tion, Causality	EG → CO ₂ (+) EC → CO ₂ (+)
Uddin et al. (2016)	1971–2006	Srilanka	CO ₂ , EG, EN	Cointegra- tion, Causality	EG → CO ₂ (+) EN → CO ₂ (+)
Khan et al. (2019)	1971–2016	Pakistan	CO ₂ , EF, EN	ARDL	EF → CO ₂ (-) EN → CO ₂ (+)
Li et al. (2018)	1993–2013	US and China	HS, GCF, CO ₂	Input Output analysis	HS → CO ₂ (+) GCF → CO ₂ (+)
Bekun et al. (2020)	1971–2015	Nigeria	GDP, FDI, EN, CO ₂	ARDL	GDP ² → CO ₂ (-) GDP → CO ₂ (+) EN → CO ₂ (+) FDI → CO ₂ (-)
Etokakpan et al. (2020)	1980–2014	Malaysia	GDP, GCF, GL, EN, CO ₂	Combined cointegra- tion test, ARDL	GDP → CO ₂ (+) GCF → CO ₂ (+) GL → CO ₂ (+) EN → CO ₂ (+)
Adebayo (2020)	1980–2018	Mint Nations	GDP, TO, URB, EN, CO ₂	Westerlund Coint, PMG, CDS, Panel causality	GDP ≠ CO ₂ TO → CO ₂ (-) URB → CO ₂ (+) EN → CO ₂ (+)
Arango Miranda et al. (2020)	1990–2016	Mexico, US, Canada	GDP, GDP ² , TO, HDI, CO ₂	OLS, VECM, Causality	For US and Mexico GDP ² → CO ₂ (-) For Canada GDP ² ≠ CO ₂
Leal and Marques (2020)	1990–2014	OECD countries	GDP, GDP ² , TO, CO ₂	FMOLS, DOLS, AMG, CDS	GDP ² ≠ CO ₂ REN → CO ₂ (+) OP → CO ₂ (+)
Dogan and Inglesi-Lotz (2020)	1980–2014	European countries	ENE, GDP, GDP ² , CO ₂	Panel OLS, FMOLS, FE	GDP ² ≠ CO ₂ GDP → CO ₂ (-)
Bekhet et al. (2017)	1980–2011	GCC countries	GDP, FD, EN, CO ₂	ARDL, VECM	FD → CO ₂ (-) EN → CO ₂ (+) GDP → CO ₂ (+)
Koc and Bulus (2020)	1971–2017	South Korea	GDP, GDP ² , EN, REN, CO ₂	ARDL	GDP ² ≠ CO ₂ GDP → CO ₂ (+) EN → CO ₂ (+) REN → CO ₂ (-)

Note: →: Unidirectional causality, ≠: not equal causality, GDP: real growth, TO: trade openness, EN: energy consumption, CO₂: Carbon dioxide emissions, FD: financial development, FDI: foreign direct investment, REN: renewable energy, HDI: human development index, IG: industrial growth, PD: population density, EG: economic growth, EF: economic factor, MINT: Economies of Mexico, Indonesia, Nigeria and Turkey, OECD: The Organisation for Economic Co-operation and Development, GCC: The Gulf Cooperation Council, (-) and (+): negative and positive relationship.

impact on CO₂ emission policies since it is being undertaken in the world's eighth most populous nation along with a number of environmental issues taken into consideration. This study has the following contributions to the existing knowledge:

- Bangladesh suffers from a lack of environmental awareness, notably in terms of CO₂ emissions. This study examines the dynamics and causal linkages among the level of CO₂ emissions, energy consumption and GDP in Bangladesh

for the period of 1972–2019 applying FMOLS cointegration technique.

- Despite the fact that similar indicators have been studied in the past, there have been very few studies conducted on Bangladeshi data. More crucially, Bangladesh has received very little attention from scholars utilizing the FMOLS approach to analyze similar variables. This research adds value into the current literature by incorporating deep learning to forecast CO₂ emissions, especially in the case of Bangladesh.

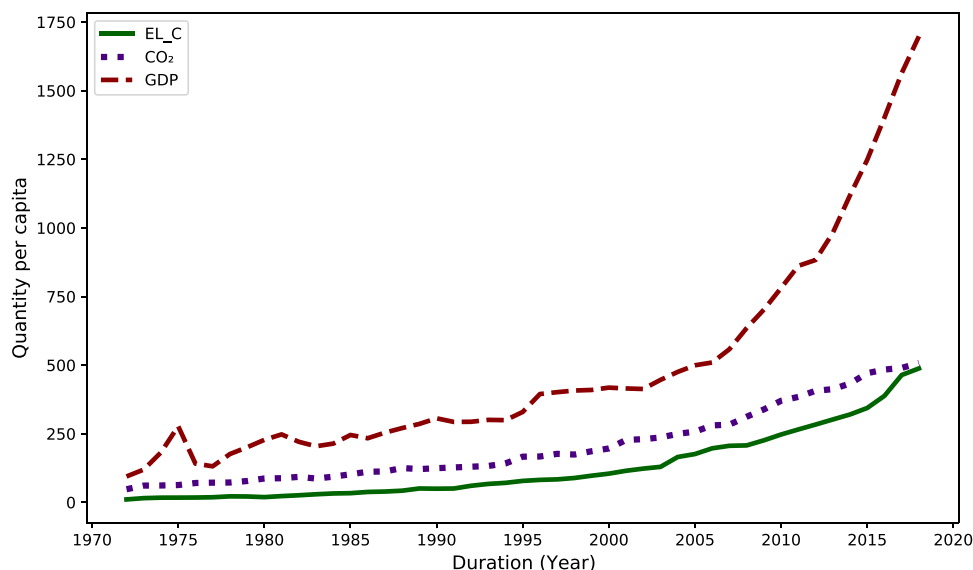


Fig. 1. Variable data graph.

- Forecasting CO₂ emissions using deep learning models while accounting for GDP and per capita energy use, i.e. multivariate time series forecasting. This is an important consideration of this research owing to the higher accuracy from practical perspective.
- Analyzing the forecasting result using a multivariate polynomial regression is performed. The empirical and forecasted findings of this research will offer policymakers with a better knowledge of the relationship between energy consumption, CO₂ emissions, and economic growth allowing them to develop energy and climate policies to improve the environment by reducing CO₂ emissions.

The paper is organized as follows. The data and methodology part is shown in Section 2. Section 3 explains the experimental setup, data preprocessing, model configuration, accuracy metrics, and the experimental results. Single-step ahead forecasting of CO₂ emissions is carried out in Section 5. Section 6 is reserved for a discussion on the findings as well as some perspective research directions. Finally, Section 7 outlines the conclusions and policy implications.

2. Data and methodology

2.1. Variables

Global warming has now become a major environmental issue, and it is mostly caused by greenhouse gas emissions, particularly CO₂ emissions, which have increased dramatically in recent years. CO₂ emissions are related to both energy consumption and national economic growth. Bangladesh has employed a variety of energies, but for this study, we will only consider electrical energy, which we have referred to as energy consumption. So this paper examines the relationship between CO₂ emissions, GDP, and energy consumption, where CO₂ is employed as a dependent variable. GDP and electrical consumption are used as independent variables to derive the long-run relationship between these variables. This paper has been carried out by utilizing Bangladesh's yearly time series data from 1972 to 2019, and all the necessary data for this work has been collected from the world development indicators (WDI)¹ (Anon, 0000).

Fig. 1 shows the graph that indicates both the independent and dependent variables that have been examined in this paper, where EL_C stands for electrical energy consumption (KWh) per capita; meanwhile, GDP and CO₂ emissions are the other two indicators. Compared to the other metrics, GDP exhibits to be a rising exponential trend. The data of CO₂ emissions ranges from 0.047 to 0.598 tons per capita, which are fraction values, according to the data source the values of GDP and electrical energy consumptions are in the range of two to four digit decimal numbers. Nonetheless, to deal appropriately with the other two components, the unit of CO₂ emissions has been changed from metric tons per capita to kg per capita for convenience.

2.2. Unit root testing

Prior to assessing the long-run interconnection of multiple variables via cointegration in applied time series econometrics, stationary data must be gathered and evaluated (Yuping et al., 2021). To establish a long-term relationship between variables, they must be stationary at the first difference. The unit root test is performed to determine whether a variable is stationary or not. In this work the augmented Dickey–Fuller (ADF) and Philips-Perron (PP) tests are used to carry out the unit root test. Table 2 shows the outcomes of the ADF and PP tests.

From Table 2, the ADF and PP test findings have shown that all the variables, i.e., GDP, CO₂ and electrical energy consumption are non-stationary initially. After encountering the first difference, those variables have become stationary since the related variables' probability (P) values are less than 5%. There are several techniques available for estimating cointegrating relationships. Therefore, the following sub-section explains the state-of-the-art methodologies employed in this empirical investigation.

2.3. Johansen cointegration test

Since the variables are stationary at the first difference, the unit root test allows us to conduct the Johansen cointegration test. The purpose of this test is to observe the long-term relationship between them. We have assumed that if no cointegration exists between the dependent and the independent variables is treated as the null hypothesis. In contrast, if there is cointegration will be considered as the alternative hypothesis.

¹ The authors are happy to share the data set upon request.

Table 2
Stationary test.

Variables	ADF test		PP test	
	Intercept and trend	Intercept and trend	Intercept and trend	Intercept and trend
	Level	1st difference	Level	1st difference
CO ₂ emissions (kg per capita)	1	0.0029	1	0.0029
GDP per capita (current US\$)	1	0.0073	1	0.0091
Electric power consumption (kWh per capita)	1	0.001	1	0.0011

Table 3
Johansen cointegration test.

Hypothesized no.of CE (s)	Unrestricted cointegration rank test (Trace)			
	Eigenvalue	Trace statistic	5% Critical value	Probability
None*	0.458446	44.82479	35.0109	0.0033
At most 1	0.2869	17.22577	18.39771	0.0723
At most 2	0.043678	2.009745	3.841465	0.1563
Hypothesized no.of CE (s)	Unrestricted cointegration rank test (Maximum eigenvalue)			
	Eigenvalue	Max-eigen statistic	5% Critical value	Probability
None*	0.458446	27.59902	24.25202	0.0174
At most 1	0.2869	15.21602	17.14769	0.0935
At most 2	0.043678	2.009745	3.841465	0.1563

2.3.1. Cointegration test results

Before performing Johansen cointegration test, it is necessary to calculate the optimum lag length. Akaike information criterion (AIC) and Bayesian information criteria are used to determine the optimal order of lag in the model, as described by Pesaran et al. (2001). Results of Johansen cointegration are reported in Table 3. Hence, according to the *P*-value, there is one cointegration among the selected set of variables at a 5% level of significance.

From Table 3, a single cointegrating vector is observed by trace statistics. We may reject the null hypothesis since the trace value 44.8247 of rank zero is greater than the 5 percent critical value of 35.0109. Since the maximal eigenvalue of 27.5990 for rank zero is higher than the 5 percent critical value of 24.2520, a similar result has been observed in eigenvalue statistics. The Johansen outcome indicates a long-term connection between dependent and independent variables. That means long-run associations exist among them and a high impact on CO₂ emissions of GDP and electrical energy consumption per capita. To assess the impact on CO₂ emissions of explanatory variables, i.e., electric energy consumption per capita and GDP, the FMOLS test is applied in the following subsection.

2.4. FMOLS estimation

The FMOLS technique gives credible estimations for small sample sizes and provides a robustness test on the results. The FMOLS technique, established and developed by Philips and Hansen (1990), is used to estimate one cointegrating relationship that contains a combination of I(1) (Bashier and Siam, 2014). After establishing a cointegrating relationship between the variables, it is necessary to estimate the long-term dynamics among the variables. The FMOLS approach would be utilized in the following step to estimate long-term elasticities. The FMOLS assesses the impact of GDP and electric energy consumption on CO₂ emissions after quantifying the long-run association among the variables in the previous section.

The outcomes suggest that electric energy consumption has a positive and remarkable contribution to increase CO₂ emissions. Nevertheless, the coefficient observes that there is an adverse consequence of GDP on CO₂ discharge shown in Table 4. Following the results of the data, it appears that a decrease in GDP increases CO₂ emissions in a very slow and gradual manner.

Table 4
FMOLS results.

Variables	Coefficient	Std. Error	t-Statistic	Probability
EL_C	1.526638	0.16593	9.200493	0
GDP	−0.142563	0.052762	−2.70198	0.0098
C	83.97603	7.650033	10.97721	0

According to the result of FMOLS, one unit increase in electric energy consumption causes a 1.526 unit increase in CO₂ emissions. Again, one unit expansion in GDP causes a 0.142 unit decline in CO₂ outflow. The *P*-value for GDP and the electric energy consumption is substantial at the 5% level, and both factors are significant for CO₂ emissions.

The previous subsection detailed the estimation for the cointegration models of per capita CO₂ emissions for Bangladesh. The next part will elaborate the future projection of CO₂ emissions using multivariate time series, which will consider the GDP and electric energy usage. They have a significant impact on CO₂ discharge over a period which is proved by the FMOLS result. To correctly estimate CO₂ emissions, we have used deep learning algorithms, and the analysis has been undertaken using a variety of accuracy metrics.

2.5. Forecasting problem formulation

A time-series approach is used in this paper to forecast CO₂ emissions because conventional methods have several limitations (Hossain et al., 2021a). The time-series approach ($X = x_1, x_2, \dots, x_t$) is a sequence of observations in a time frame, where x_t refers to the observation at time t and X is the total number of observations. A challenge may have two or more parallel input time series and an output time series that depends on the input time series. Here, GDP and energy consumption are input time series, whereas CO₂ emissions are considered as output. Using the first three-time steps of each parallel input time series as input features, the models then relate this value to the fourth time step of output time series as target variable. As our goal is to single-step ahead forecasting we took the fourth observation as our output feature. The input data can be of any previous data (such as 3, 5, 10), but we have found better results in the three input previous data. If it is multi-step forecasting then

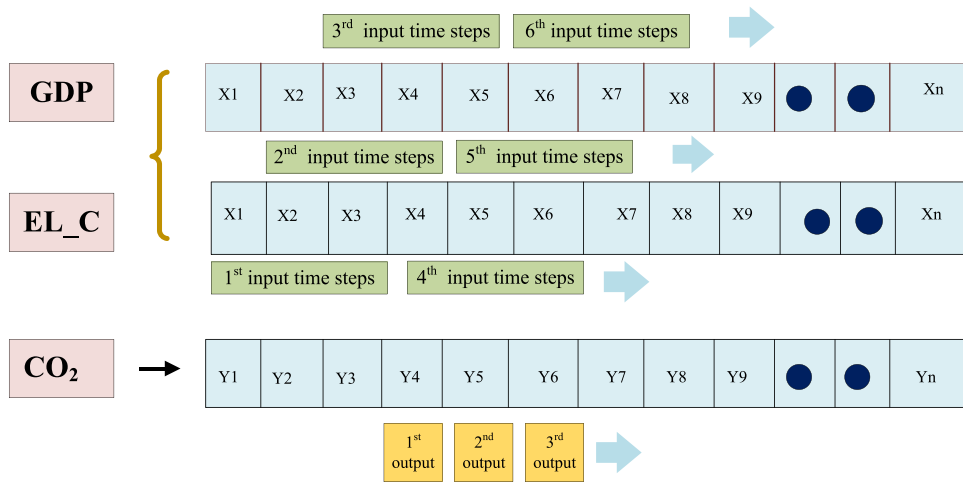


Fig. 2. Input-output dataset for training the network.

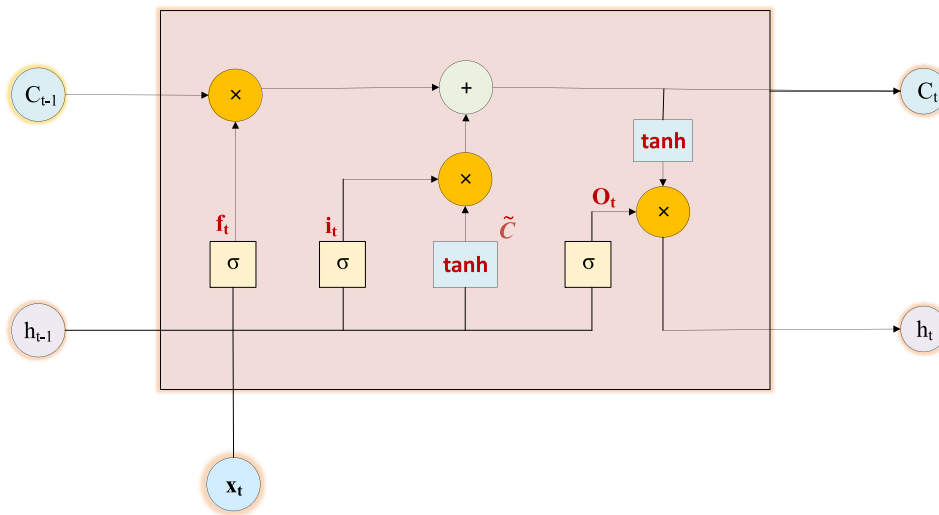


Fig. 3. Architecture of LSTM network.

additional time steps beyond four will be considered. This system works as a sliding window approach. We have discarded some values for which we did not have values in the input time series at previous time steps during the output time series. The Fig. 2 depicts a graphical representation of the problem formulation for the time series forecasting algorithm.

3. Forecasting models

3.1. Long short-term memory (LSTM) neural network

The LSTM is an advanced gated memory unit that eliminates the vanishing gradient problems that restrict the effectiveness of a basic RNN (Li et al., 2021; Zeroual et al., 2020). Owing to its high long-term memory function, this neural network is capable of exploring in depth the long-term relationships and trends of limited data samples (Hossain et al., 2020). As an interpretation from the completion of the sequence, an LSTM layer must be fed a three-dimensional input, so that input data must be reshaped before being fed into it.

Fig. 3 illustrates the basic LSTM network construction, where the current variable vector x_t , the previous output h_{t-1} , and the previous cell state C_{t-1} are inputs to the LSTM cell. As seen in Fig. 3, the input gate (i_t), forget gate (f_t), output gate (o_t),

and memory cell (\tilde{C}_t) are depicted by small boxes. They can be computed by using the following equations (Li et al., 2021; Zhou and Chen, 2021):

$$f_t = \sigma g (W_f \cdot x_t + U_f \cdot h_{t-1} - 1 + b_f) \quad (1)$$

$$i_t = \sigma g (W_i \cdot x_t + U_i \cdot h_{t-1} - 1 + b_i) \quad (2)$$

$$o_t = \sigma g (W_o \cdot x_t + U_o \cdot h_{t-1} - 1 + b_o) \quad (3)$$

$$\tilde{C}_t = \tanh (W_C \cdot x_t + U_C \cdot h_{t-1} - 1 + b_C) \quad (4)$$

where \cdot denotes the matrix multiplication operation, b_f , b_i , b_o , and b_C are four bias vectors, the weight matrices U_f , U_i , U_o , and U_C connect the previous output to the three gates and the memory cell. The W_f , W_i , W_o , and W_C are the weight matrices, the $\sigma g()$ represents the gate activation function, which here is a sigmoid function, and $\tanh()$ is the hyperbolic tangent function, as seen in the equations above. The cell output state C_t and the layer output h_t can be determined as follows:

$$C_t = (f_t \oplus C_{t-1}) + i_t \oplus \tilde{C}_t \quad (5)$$

$$h_t = o_t \oplus \tanh (C_t) \quad (6)$$

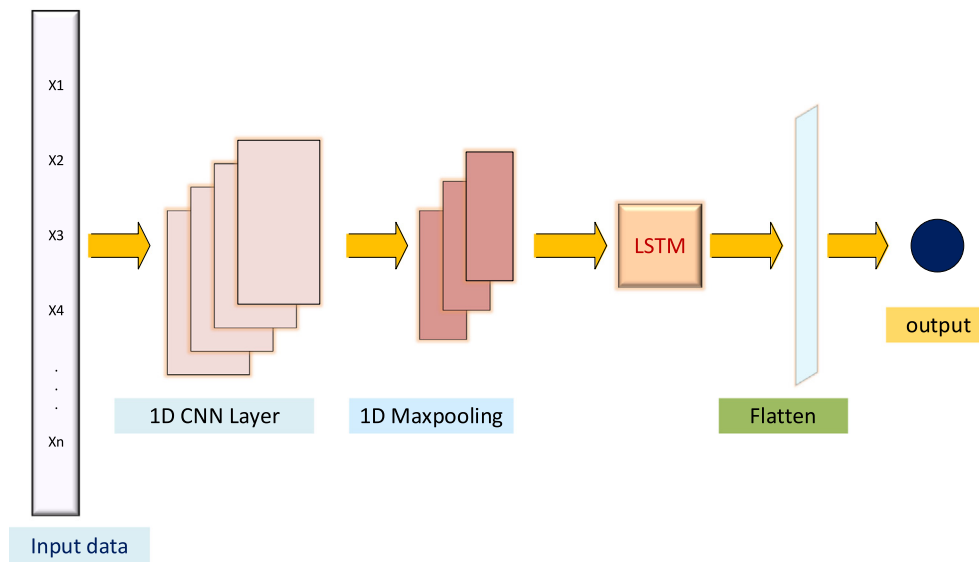


Fig. 4. Architecture of CNN-LSTM network.

where \oplus denotes the element-wise matrix/vector, multiplication operator.

3.2. Convolution neural network (CNN)

The CNN has been widely used in a variety of fields, including pattern classification, image processing, radiology, and model identification (Yu et al., 2020; Yao et al., 2021). In this paper, the one-dimensional CNN is utilized to extract characteristics of the complex interaction between nonlinear input and output data. Instead of conventional matrix multiplication, the convolution layer uses a mathematical convolution operator and cross-correlation. The convolution operation can be described in the following way (Hossain et al., 2021b):

$$c_i^l = f_1 \left(\sum_{i=1}^n x_i * w_i^l + b_j \right) \quad (7)$$

In the convolution layer, c_i^l is the i_{th} element of l_{th} feature, w_i^l is the i_{th} element of the l_{th} kernel, b_j is bias, $*$ is convolution operation and f_1 denotes the activation function. x_i is the i_{th} element of the input, e.g. GDP and energy consumption here. Various pooling approaches are available on the CNN architecture, however, max-pooling is the most commonly employed after CNN layers. Using a pooling layer after convolution reduces the data dimension by combining the output of neuron clusters.

3.3. CNN-Long short-term memory networks (CNN-LSTM)

The CNN-LSTM architecture uses CNN layers for feature extraction on input data and combines them with LSTM to improve sequence prediction (Rajagukguk et al., 2020; Guo et al., 2020). In this paper, the 1D CNN-LSTM network has been employed to forecast CO₂ emissions. According to Fig. 4, there are five layers in the 1D CNN-LSTM network architecture. Firstly, the original signals are transferred into the first 1D convolution layer for feature extraction and feature selection. The second layer is the 1D maximum pooling layer, followed by the 1D convolution layer, and this is followed by stacking the one LSTM layer. Finally, a fully connected layer with the relu activation function is placed at the end to predict.

3.4. Artificial neural networks (ANN)

The ANNs are data-driven, flexible models capable of approximating a vast class of nonlinear problems to any desired level of accuracy. A wide range of ANN models have been developed and widely used in different applications (Hossain et al., 2021b; Guo et al., 2021). The multilayer perceptron is one of the most extensively used ANN models in time series forecasting (Panigrahi and Behera, 2017). This paper's model architecture comprises an input layer, two hidden layers, and one output layer.

In a forecasting problem, the number of inputs and neurons in the hidden layer is configurable. In contrast, the output layer contains only one neuron with auxiliary components, such as weight, bias, activation function so on. With a multilayer perceptron, the fundamental neural network design is shown in Fig. 5. The input layer accepts input values, while the hidden layer analyzes those input values. The output layer collects the data from the hidden layer and decides the final output. The training procedure will be repeated until the difference between the neural network output and the supervisor comes within an acceptable range (Hameed et al., 2019). The mathematical formula for ANN can be stated as follows (Guo et al., 2021):

$$A_n = \sum_{j=1}^n (w_j \cdot I_j) + b \quad (8)$$

where n denotes the number of input, w_j , b indicates the weight and bias, respectively. I_j is the input, and A_n refers to the output of the ANN.

4. Experimental results

The experiments are carried out in google colab (Colab) using python 3.0 with open source libraries like Tensorflow, Pandas, Numpy, and Keras. The experimental setup is based on a working environment having Intel(R) Core (TM) i5-7400 CPU @ 2.5 GHz with 4 GB RAM under 64-bit Windows 10 Pro Operating system. The entire dataset was analyzed before feeding to the model. Time series forecasting of the CO₂ emissions dataset is modeled using four deep learning models mentioned in the below section. The optimum results are considered from each model's eleven unique runs. The working and parameters optimization of all four versions are explored in sub-sections.

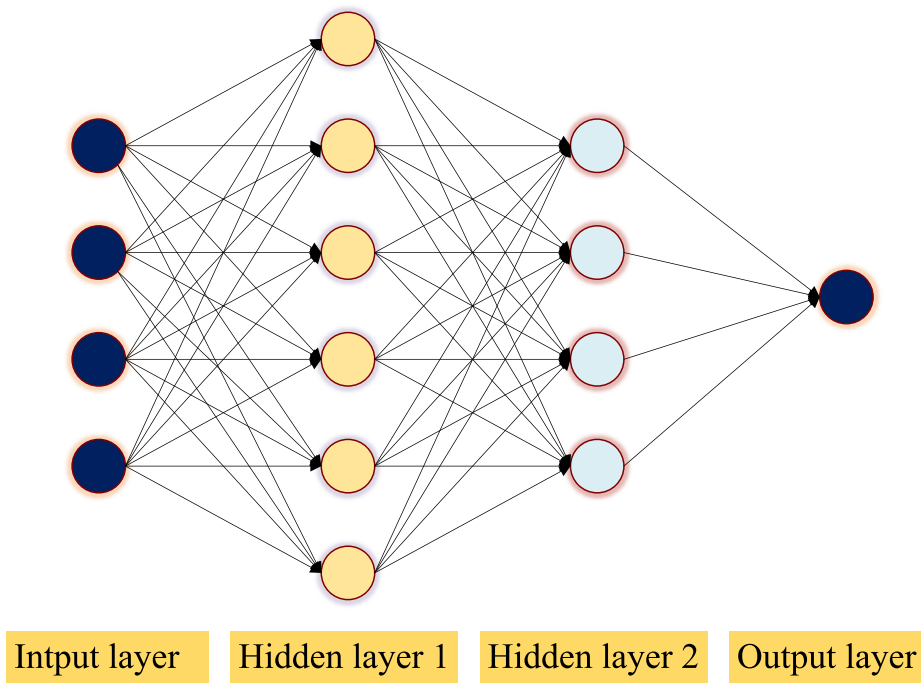


Fig. 5. Architecture of artificial neural networks.

4.1. Data preprocessing

To forecast CO₂ emissions, the CO₂ data from the dataset has been transformed from metric tons per capita to kg per capita, which is previously mentioned. Electric energy consumption (kWh per capita) and GDP per capita (current US\$) are the other two variables. The dataset only comprised 48 samples in yearly format, which is insufficient for a deep learning model to be appropriately trained. The Pandas library has a `resample()` function that can resample a series or a data frame. Down-sampling is a technique for grouping observations and making the number of samples less than the original. In contrast, creating room for additional observations is known as up-sampling which can be accomplished using the resampling technique. We have changed our annual datasets into monthly datasets by invoking up-sampling and setting "M" as desired in the respective function, where M is denoted by month. The `interpolation()` method of the Pandas library is used to interpolate missing values, and there is a great range of simple and more complicated interpolation techniques available. However, we used the linear approach to fill the intermediate interpolated missing data, where a total of 565 data points were generated throughout the process.

- **Missing values:** As mentioned earlier Bangladesh-specific data was obtained from the WDI, and there were no missing figures.
- **Scaling:** The literature suggests that normalizing approaches substantially influence the performance of a model. As the data of raw time series ranges widely, the optimization methods used to derive the objective functions in some machine learning and deep learning models will not function properly unless the data is normalized. So that the normalization technique should be selected based on the problem and model at hand. Normalization is a technique for scaling (also called min–max scaling) numerical data that involves scaling each input variable separately to the range 0–1, the most precise range for floating-point data.

Table 5

Data Features.

Feature	CO ₂	EL_C	GDP
Count	565	565	565
min	47.78	10.65	97.38
max	598.74	510	1856
std	140.54	130.69	406.14
25%	92.53	31.68	237.7
50%	166.8	80.08	362.07

Numerical input variables are normalized using MinMax scaler API to acceptable ranges for better performance. The data characteristics are given in Table 5 after the raw annual data has been transformed into monthly data using the up-sampling method.

4.2. Model configuration

As a consequence of over-fitting and under-fitting, the number of neuron layers in deep learning models affect predicting accuracy. In deep learning, it is not true that adding layers would enhance prediction accuracy (Hossain et al., 2021b). Due to over-fitting, prediction models learn well during training but have greater mistakes during prediction. On the other hand, a limited number of layers might cause uncertainty related to poor fitting owing to the inability to map input–output relationships. The performance of a hybrid deep learning model must be improved by combining network layers in an appropriate manner (Hossain et al., 2021b). The configurations we have used are discussed below:

1. **CNN–LSTM:** One 1D convolutional layer, one LSTM layer, and a fully connected dense layer make up this model. The root mean square loss function is used in conjunction with the adam optimizer. The batch size is 32, with a learning rate of 0.0001.
2. **CNN:** This model comprises one 1D convolutional layer and a fully connected dense layer. In combination with the adam optimizer, the root mean square loss function is

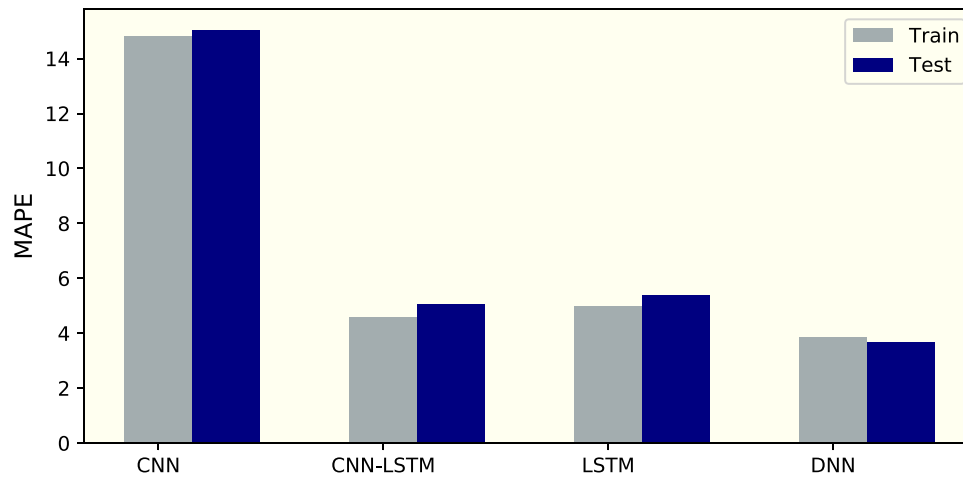


Fig. 6. MAPE values of CNN, CNN-LSTM, LSTM and DNN model.

employed. A learning rate of 0.0001 is likewise employed here, with a batch size of 32. Sixty-four filters and three kernels are used in the first layer of the convolutional neural network. Both layers are activated by relu's activation function.

3. LSTM: This model is a combination of one fully interconnected LSTM layer and one dense layer. The only difference between this model and CNN is the batch size, which is 8 in this case. This model performs best with batch sizes of 8 in this work.
4. DNN: This model comprises three dense layers, having 32 neurons, 8 neurons, and one fully connected layer, respectively. All hidden layers use the relu activation function with batch size of 8, which outperforms all the models.

4.3. Parameters tuning

We have taken substantial amount of tuning, a very time-consuming method, for all of the models because the performance of all of them are heavily dependent on parameter adjustment (Hossain et al., 2021b). Special efforts to calibrate the parameters of the CNN-LSTM and CNN forecasting models have taken because their structures are more complex than that of other deep learning models (Yao et al., 2021). All hyperparameters are manually adjusted through trial and error. Our models have some limitations as we could not able to employ optimization algorithms such as grid search and, random search. However, in the CNN model, the tuning for the kernel an essential task to design convolutional layers and is defined as an operator that transforms the information in the entire data set. The selection of kernel 3 has produced the best result out of eleven unique runs, followed by kernels 5, 6, and 7. Moreover, the best result are obtained while utilizing 64 filters. For the best outcomes, the number of hidden layers and neurons are also optimized for all deep learning models. In all models, a learning rate of 0.0001 provides a better level of accuracy. An early stopping approach has been utilized to minimize overfitting and optimize the efficiency of the models.

4.4. Forecasting evaluation metrics

The forecasting accuracy metric indicates how effectively a model predicts the future value. This endeavor strengthen its prediction powers that will benefit overall planning and make predictions far more agile in changing circumstances. Let us take a look at the most well-known forecast KPIs one by one. The

data are separated into training and test sets. All models have been trained using training sets (80% of the data set), however, the RMSE, MAE, and MAPE values for the training and test sets have been evaluated and compared for each model. The accuracy metrics we have used are explained below:

1. **MAPE:** MAPE measures the accuracy of a forecasting system and is determined as the average absolute percent error for each period minus actual values divided by real values, and it is expressed as a percentage. It can be written as follows (Li et al., 2020):

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left[\frac{A(t) - F(t)}{A(t)} \right] \quad (9)$$

In Eq. (9) A(t) and F(t) indicate the actual and predicted values, respectively. The number of samples is denoted by the letter n . Fig. 6 illustrates the MAPE of all the deep learning models, where DNN has the lowest MAPE value of 3.853 and 3.678 for the test and training set, respectively. For the CNN model, the maximum error is recorded, and it is clearly shown that the testing errors are slightly larger than the training set errors for all models except DNN.

2. **RMSE:** RMSE is a common tool for analyzing the error of a model when predicting the quantitative data. The standard deviation of the residuals is denoted by RMSE. Residues are a measurement of how far data points depart from the regression line in regression analysis. It can be expressed as follows (Zhou and Chen, 2021):

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (f(i) - o(i))^2} \quad (10)$$

In Eq. (10) $f(i)$ is the forecasted value and $o(i)$ is the observed value. The other term, n , stands for the number of samples. Fig. 7 shows the RMSE for each model which we have used to estimate CO₂ emissions for both train and test data. According to the results, the DNN has the lowest RMSE value of 8.393 for the train data and 8.099 for the test data, respectively. CNN, on the other hand, has the largest error of any the projected model, which is 21.80 for the training set and 21.96 for the testing set.

3. **MAE:** The MAE is a statistical error that evaluates the average magnitude of error in a set of forecasts without taking into account the direction of the projections. It is used to find out how accurate continuous variables are.

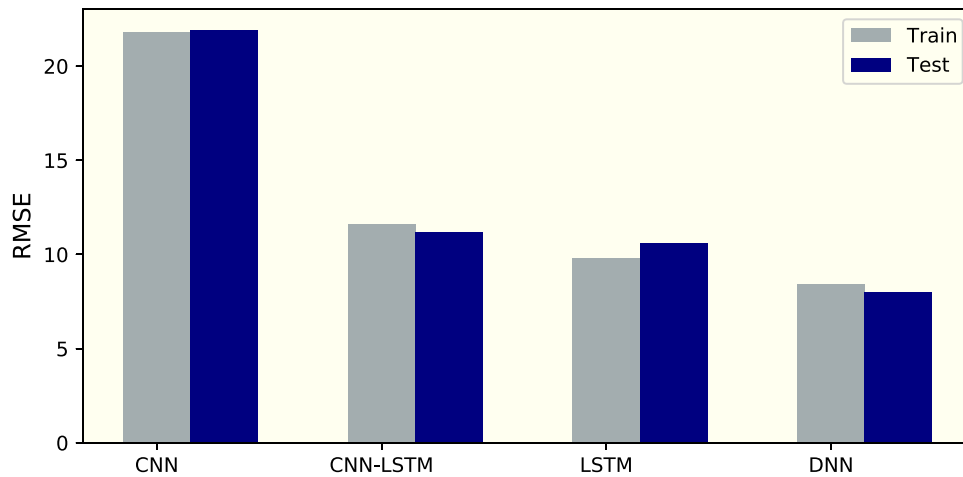


Fig. 7. RMSE values of CNN, CNN-LSTM, LSTM and DNN model.

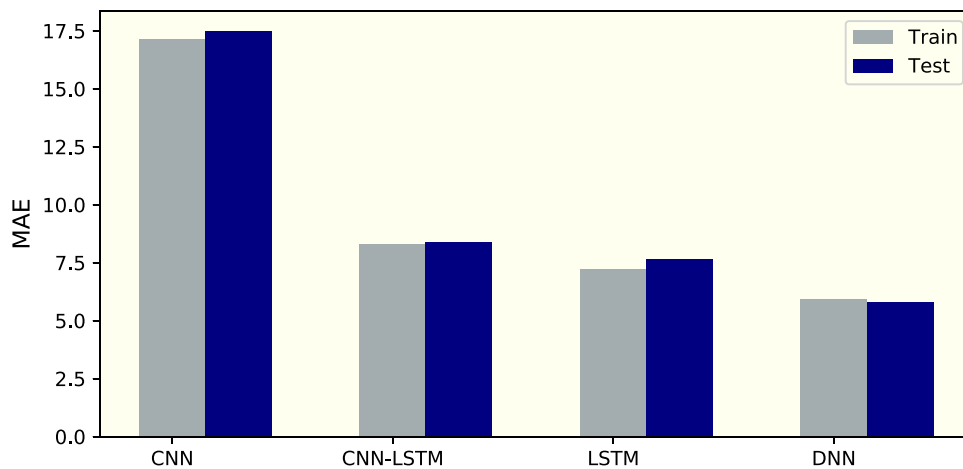


Fig. 8. MAE values of CNN, CNN-LSTM, LSTM, DNN model.

In other words, over the verification sample, the MAE is the average of the absolute values of the differences between the forecast and the actual observation. It can be expressed as (Heydari et al., 2019):

$$MAE = \frac{1}{n} \sum_{i=1}^n (y(i) - x(i)) \tag{11}$$

where $y(i)$ denotes prediction value, $x(i)$ denotes true value and n is the total number of data points. Fig. 8, displays the MAE of all the developed models. The bar chart illustrates a gradual decrease in MAE values starting with CNN and ending with the DNN, which has the lowest error of 5.933 for the training set and 5.820 for the test set, respectively. The training and testing errors for the CNN model are 17.164 and 17.469, respectively.

4.5. Forecast accuracy metrics for multivariate polynomial regression

Simple polynomial regression is a form of regression that is applied to a single regressor, whereas multivariate polynomial regression involves multiple regressors. The later one is used to predict values when several variables are involved (Sinha, 2013). We have intended to develop polynomial regression after studying the CO₂ emissions curve, and the result seems excellent. Two steps are required to accomplish polynomial regression. Using the polynomial features function from sklearn, we first transform

the inputs into a polynomial. Characteristics that are generated by exponentiating the existing features are known as polynomial features. This model has been tested optimistically from second to fifth order. However, it performs the best in the third order. It is worth mentioning that the complexity of the model rises as the degree of the polynomial increases. As a result, the order of polynomial must be chosen very carefully. After performing the transformation, the input features are extended into several new terms. Moreover, the problem formulation method is different from those deep learning models used in this study. Taking the first time step of GDP and energy consumption as input and the second time step of CO₂ emissions as the target variable, have made the system reliable to one step ahead forecasting. The general equation of third-order polynomial with two variables can be expressed as (Sinha, 2013):

$$p(x, y) = a_{3,0}x^3 + a_{2,1}x^2y + a_{1,2}xy^2 + a_{0,3}y^3 + a_{2,0}x^2 + a_{1,1}xy + a_{0,2}y^2 + a_{1,0}x + a_{0,1}y + a_{0,0} \tag{12}$$

where, $a_{3,0}, a_{2,1}, a_{1,2}, a_{0,3}, a_{2,0}, a_{1,1}, a_{0,2}, a_{1,0}, a_{0,1}, a_{0,0} \in R$, as well $a_{3,0}, a_{2,1}, a_{1,2}, a_{0,3}, a_{2,0}, a_{1,1}$, and $a_{0,2}$ cannot be equal 0.

In our model, x stands for GDP, and y represents energy consumption. Subsequently the linear regression is used to fit the parameters after converting the data into polynomial features. The Fig. 9 portrays the pipeline that can be used to explain the complete procedure.

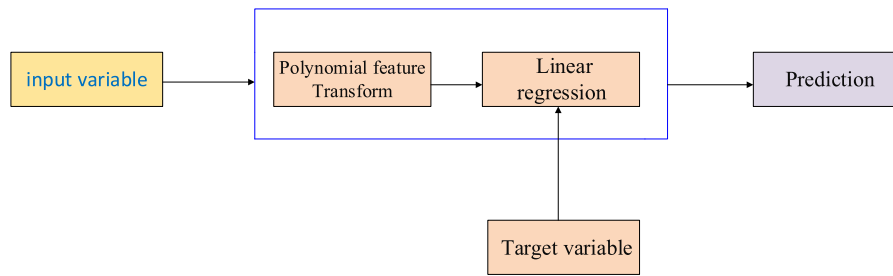


Fig. 9. The pipeline of polynomial regression.

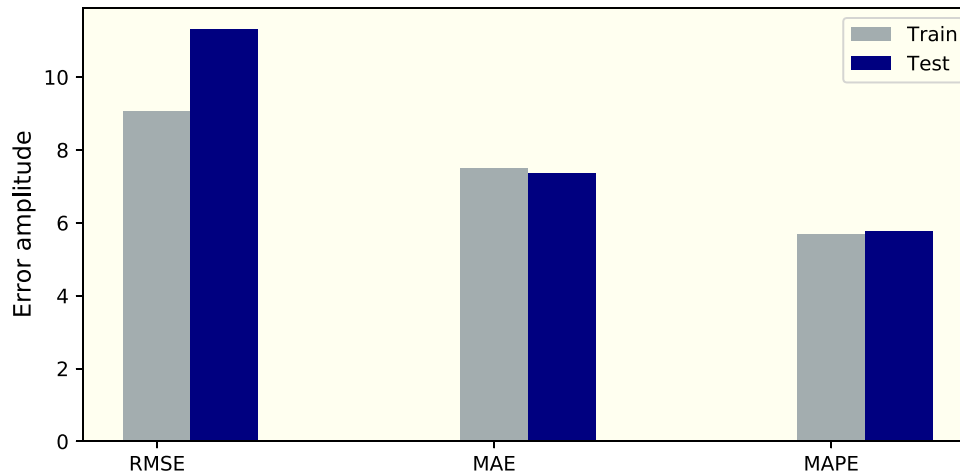


Fig. 10. RMSE, MAE, and MAPE values for multivariate polynomial regression model.

Data are split into training and test sets according to 80% and 20% before applying linear regression. Both the training and test sets are used to assess the accuracy of the system.

Fig. 10 shows the MAE, RMSE, and MAPE of polynomial regression. The bar chart depicts training errors of 9.58, 7.33, and 5.04 for RMSE, MAE, and MAPE, respectively. Testing data shows errors of 11.03, 7.86, and 5.54 in the same order. The errors are pretty similar to the LSTM model. Because of the limited amount of data points available, there have been some overfitting problems. As we have used the annual data to predict the following year, the training range for the model seems to be very narrow, resulting in slightly overfitting problems. It can be resolved by collecting more data in the future.

4.6. Validation losses

A validation error is a vehicle for assessing the generalization of a model that measures how well a model fits new data. Typically, the validation loss is more significant than the training loss, but this problem has overcome with the early stopping approach in our study. Consequently, the training and testing errors in the models have employed considerably close in magnitude.

Fig. 11 illustrates the validation losses which have been computed by the testing data set, which is used as a validation set. The box plot enables us to observe the distributional characteristics of validation losses of eleven unique runs. The CNN model has the lowest performance metric, followed by the LSTM, CNN-LSTM and DNN models. Fig. 11 demonstrates an outlier of the CNN-LSTM model, where the DNN model, on the other hand, outperforms the others, as the median line of this model lies at the bottom.

4.7. Forecasting the CO₂ emissions

This section shows the original and predicted CO₂ emissions results for all of the models in Fig. 12. After training the models,

it is necessary to assess the performance to predict CO₂ emissions. Regarding this purpose the training data set has been utilized to measure the accuracy metrics displayed in the previous section.

The whole data set has been predicted in this section, allowing us to compare all the models in terms of actual and predicted graphs. According to Fig. 12(d), the CNN model shows the inferior result of all the models examined in this paper to predict CO₂ emissions. Fundamentally, CNN exhibits a wide range of hyperparameters and various specific architectures, which are often complex and provide difficulties in selecting the most appropriate value among many possible combinations. Aside from that, CNN is very sensitive to set its hyperparameters, which have a significant effect on the efficiency and behavior of its architecture. Fig. 12(d) demonstrates the projected value is far away from the actual value at the starting region, indicating poor result.

It was expected that the hybrid deep learning model would produce the best outcomes, it has not happened in this study. According to the data in the testing set, CNN-LSTM has slightly higher dispersion from its real value in the peak region, as shown in Fig. 12(b). To enhance the performance of the hybrid deep learning model, it is also necessary to fine-tune the hyperparameters, such as the number of layers, neurons, and so on. However, the CNN-LSTM hybrid model has substantially more complexity (Guo et al., 2020). To get a better result from this model, an experimental approach has been conducted to fine-tune its parameters. The best result has provided the DNN model and is much more generalized in terms of effectiveness which can be seen in Fig. 12(a).

4.8. Forecasting one step ahead

For univariate time series, a single step ahead forecast is easier than multivariate forecasting. Whenever a multivariate forecasting problem is encountered, it is complicated as the output series

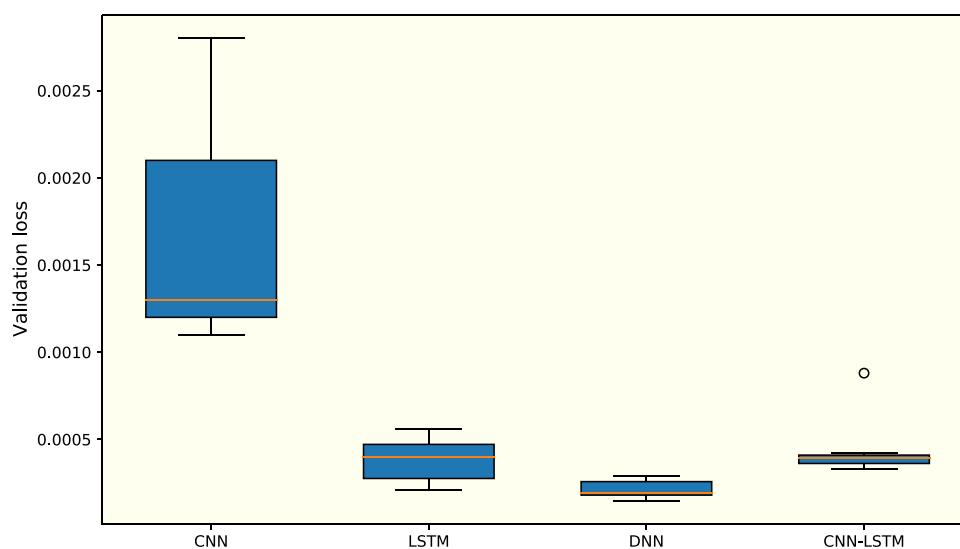


Fig. 11. Validation loss for CNN–LSTM, DNN, LSTM, and CNN model.

is completely dependent on one or more input time series that also need to be forecasted. We have performed our analysis using data from 1972 to 2019 and proceeded to estimate CO₂ emissions for the years 2020 and 2021. To forecast the CO₂ emissions in 2021, we must need the data from 2020 as input, but the actual electrical energy consumptions per capita data is not available on WDI. That is why we have to take assistance from other sources. However, the per capita electrical energy generation was 510 KWh in 2019, which we have considered as consumption per capita (Ebn Sharif, 2020). According to the problem formulation section, we can only forecast CO₂ emissions for the year 2020. As a result, forecasting for 2021 in the next phase, we have included the forecasted value of CO₂ emissions for 2020 in the original data set. As shown in Fig. 13, all the deep learning models except CNN predict a rising trend in CO₂ emissions. The DNN, LSTM, CNN, and CNN–LSTM predict CO₂ emissions of 665, 650.53, 575, and 612.50 kg per capita, respectively, for the year 2020, which was 598 kg per capita in 2019.

Despite the fact that both GDP and energy consumption increased from 2018 to 2019, according to the FMOLS results, GDP is less significant than energy consumption. This supports the increasing trend of CO₂ emissions as energy consumption has a positive impact on it. Later, the rising rate of GDP outpaces the increasing rate of energy consumption, resulting in reduction in the trend by 2021. For the year 2021, the models provide the forecasted values of 645, 668, 580, and 619 kg per capita for DNN, LSTM, CNN, and CNN–LSTM, respectively. The prediction of polynomial regression is quite close to LSTM, which forecasts the CO₂ emissions of 654 and 665.5 kg per capita in 2020 and 2021, respectively.

It is conceivable to expect a better outcome from these models if we could access to the actual data in some of the situations mentioned earlier. However, the outcomes are utterly reliable. Incorporating more explanatory factors into the model may lead to a more accurate estimation of CO₂ emissions. From the result, we may conclude that energy consumption is one of the vital factors for rising emissions. Switching to clean energy and proper formulation, implementation of environmental laws and policies should be targeted to minimize emissions in future (Rahman and Kashem, 2017).

5. Discussion

This long-term forecasting study has revealed the best outcomes with the assistance of the DNN technique. Compared to the

other approaches, the DNN algorithm has shown better performance in terms of evaluation metrics accuracy. The effectiveness of this forecasting model has been assessed by comparing it to other deep learning models such as CNN, LSTM, and a hybrid model that combines CNN and LSTM as well as a multivariate polynomial regression. The MAPE, MAE, and RMSE are used to compare these models. The accuracy metrics of multivariate polynomial regression and LSTM models are more likely to be similar. For the years 2020 and 2021, the polynomial regression provides the predicted output that is nearly similar to the LSTM model, which is a very satisfactory result in a single step forward prediction. During this work, the road-maps of all models have been filled with information obtained throughout trial and error. Several problems occurred during this procedure, particularly in data processing and choosing the best forecasting model. This procedure is repeated until the models get the desired outcomes. It might be conceivable that the results of this analysis would have been more accurate if we had used the hyperparameter optimization technique. As a result, selecting the most appropriate set of parameters will be a significant concern of practitioners for future studies. However, the greenhouse effect is well known to have several severe consequences, including an increase in global pests and diseases, sea-level rise, a shift in temperature, and desertification. Nonetheless, Bangladesh should pay greater attention to CO₂ emissions, make necessary changes to climate and energy policies, and collaboratively address climate change problems, regardless of whether CO₂ emissions are increasing or decreasing. Even though Bangladesh is an overpopulated and developing nation, industrial output continues to increase in the country. These factors need to be considered if we want to forecast CO₂ emissions correctly in future. Many additional techniques, such as DOLS, ARDL, etc., may be used to derive empirical cointegration, long and short-run dynamics, and causal connections (Bastola and Sapkota, 2015; Rahman, 2017). Finally, to maintain the economy and public health over the time, Bangladesh must implement environmental protection laws.

6. Conclusion and policy implications

In this paper, the FMOLS approach has been used to analyze the impact of energy consumption and GDP on CO₂ emissions in Bangladesh from 1972 to 2019. The ADF and PP unit root tests are used to check the stationarity of the variables. The estimated outcomes reveal the evidence of cointegration and a

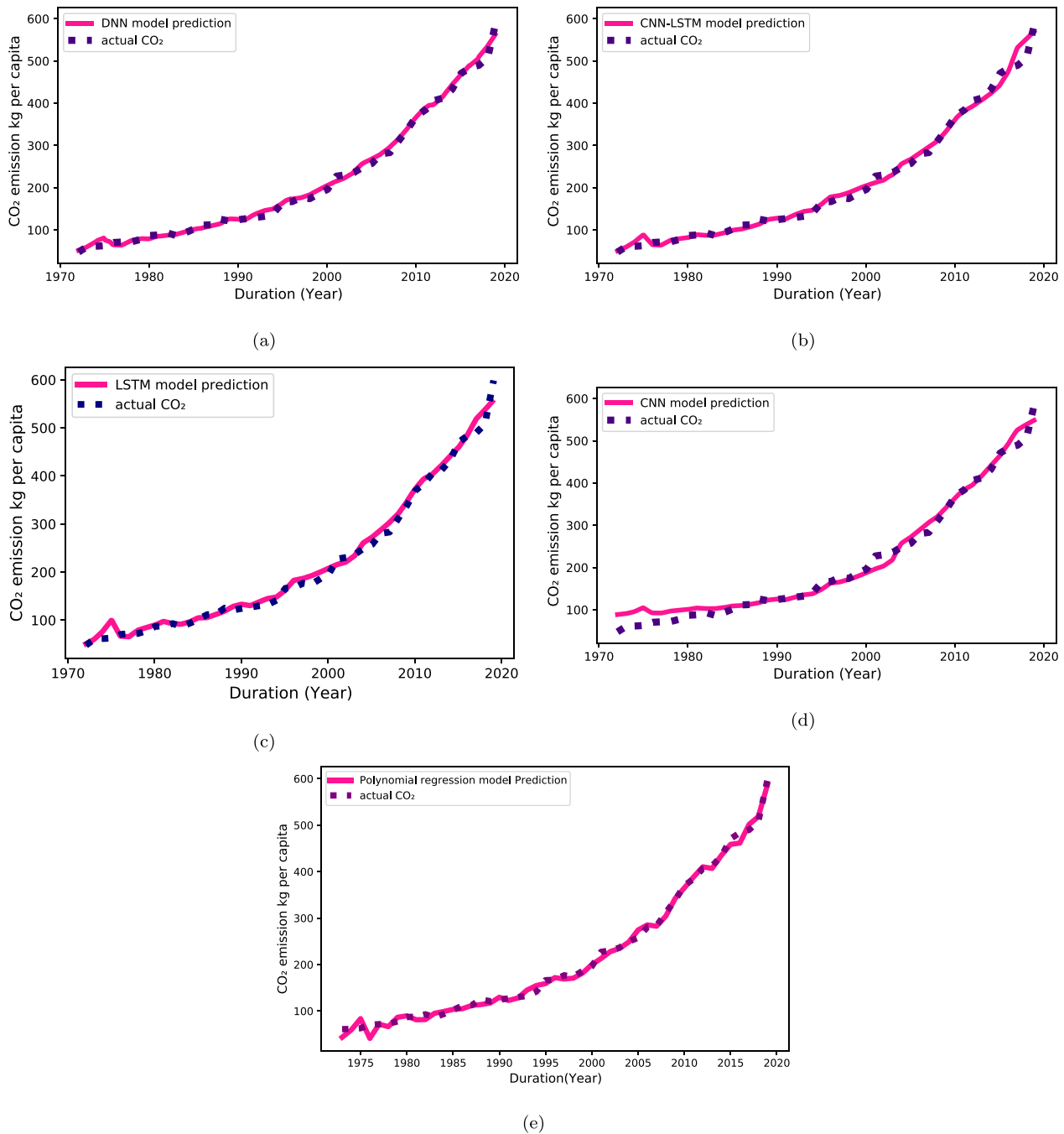


Fig. 12. Forecasting CO₂ emissions by (a) DNN, (b) CNN–LSTM, (c) LSTM, (d) CNN, and (e) polynomial regression model.

long-term connection between GDP, energy consumption, and CO₂ emissions. As previously stated, energy consumption positively impacts CO₂ emissions, while GDP has a negative impact in Bangladesh. However, the nonlinear relationships between CO₂ emissions and their influencing factors are terribly complicated, making CO₂ forecasting a difficult job. A large number of researchers have made significant contributions to forecast CO₂ emissions. The accuracy of these approaches seem to be poor because of the complicated nonlinear relationship between CO₂ emissions and other factors, which is very difficult to model. Furthermore, the majority of these techniques do not have the capability of predicting CO₂ emissions levels. They can only anticipate it if they have obtained all the necessary variables in advance. The second section of this study explores the potential

of employing the state-of-the-art of several deep learning techniques for forecasting CO₂ emissions in Bangladesh. In addition to that, a simple multivariate polynomial regression has also been conducted to forecast CO₂ emissions, which comprises a quite similar result with one of the deep learning models named LSTM. In future, this work will be extended to develop a multi-step prediction model of CO₂ emissions that takes into account a wider range of variables at the same time. It is expected that by gathering a large amount of data, the forecasting outcome can be further improved.

It is undeniable that climate change is unfolding, and deep learning is playing an important role in this catastrophe. Fortunately, several recent research has begun in details the environmental costs of their revolutionary deep learning approaches, sometimes even including CO₂ emissions as an objective to be

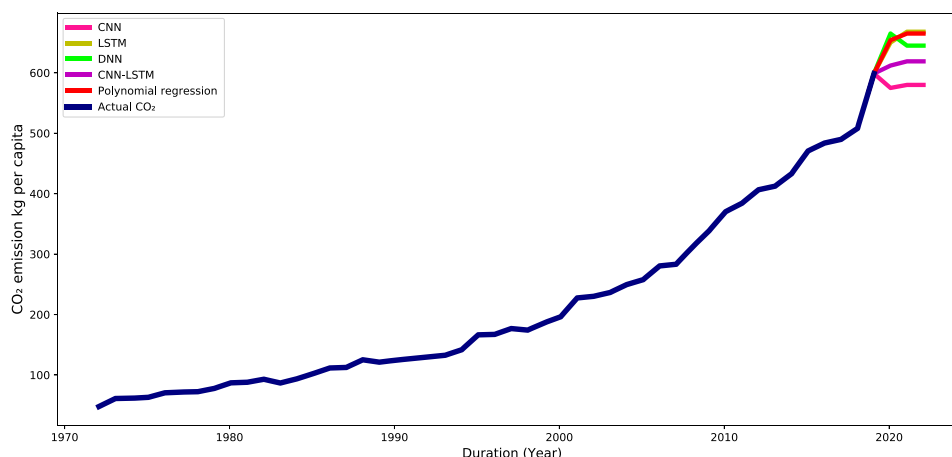


Fig. 13. Single step ahead prediction by CNN, LSTM, DNN, CNN–LSTM, polynomial regression model.

minimized. Considering Bangladesh is a densely populated country in the early stages of development, the study's findings indicate that energy consumption has an impact on CO₂ emissions. In this context, the immediate import and installation of energy-efficient technologies may result in long-term CO₂ reduction.

Although the focus of this study is on Bangladesh, the findings can be used to provide a viable alternative for energy usage in the future, reducing CO₂ emissions globally. To reduce CO₂ emissions, Bangladesh must implement energy conservation and environmental protection policies for encouraging local and international investors in an environmental friendly resources. Additionally, the government should impose emission restrictions on enterprises and factories that release CO₂ particularly for the coal based power plants which are rising in Bangladesh. The country's energy policy should emphasize on research and investment in clean energy. In this aspect, technological advancement through research and development is essential despite the fact that carbon capture and storage (CCS) technology is an effective way to capture carbon but it is costly. Furthermore, the authors propose that readily available renewable resources with intelligent control can be a viable solution for effectively reducing CO₂ emissions.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The publication of this article was funded by Qatar National Library. All authors contributed equally to the manuscript preparation.

References

Adebayo, Tomiwa Sunday, 2020. New insights into export-growth nexus: wavelet and causality approaches. *Asian J. Econ. Bus. Account.* 32–44.

Aftab, Samina, Ahmed, Aftab, Chandio, Abbas Ali, Korankye, Bright Asiamah, Ali, Aamir, Fang, Wang, 2021. Modeling the nexus between carbon emissions, energy consumption, and economic progress in Pakistan: Evidence from cointegration and causality analysis. *Energy Rep.* 7, 4642–4658.

Amarpuri, Lakshay, Yadav, Navdeep, Kumar, Girish, Agrawal, S., 2019. Prediction of CO₂ emissions using deep learning hybrid approach: A case study in Indian context. In: 2019 Twelfth International Conference on Contemporary Computing. IC3, pp. 1–6.

Ameyaw, Bismark, Yao, Li, Oppong, Amos, Agyeman, Joy Korang, 2019. Investigating, forecasting and proposing emission mitigation pathways for CO₂ emissions from fossil fuel combustion only: A case study of selected countries. *Energy Policy* 130, 7–21.

Ang, James B., 2008. Economic development, pollutant emissions and energy consumption in Malaysia. *J. Policy Model.* 30 (2), 271–278.

Anon, 0000. <https://databank.worldbank.org/source/world-development-indicators>.

Anon, 2022. Knoema <https://knoema.com/atlas/{Bangladesh}/{CO2-emissions-per-capita}>.

Arango Miranda, Raul, Hausler, Robert, Romero Lopez, Rabindranath, Glaus, Mathias, Pasillas-Diaz, Jose Ramon, 2020. Testing the environmental kuznets curve hypothesis in north America's free trade agreement (NAFTA) countries. *Energies* 13 (12), 3104.

Bashier, Al-Abdulrazag, Siam, Amani Jaser, 2014. Immigration and economic growth in Jordan: FMOLS approach. *Int. J. Human. Soc. Sci. Edu.(IJHSSE)* 1 (9), 85–92.

Bastola, Umesh, Sapkota, Pratikshya, 2015. Relationships among energy consumption, pollution emission, and economic growth in Nepal. *Energy* 80, 254–262.

Bekhet, Hussain Ali, Matar, Ali, Yasmin, Tahira, 2017. CO₂ emissions, energy consumption, economic growth, and financial development in GCC countries: Dynamic simultaneous equation models. *Renew. Sustain. Energy Rev.* 70, 117–132.

Bekun, Festus Victor, Agboola, Mary Oluwatoyin, Joshua, Udi, 2020. Fresh insight into the EKC hypothesis in Nigeria: accounting for total natural resources rent. In: *Econometrics of Green Energy Handbook*. Springer, pp. 221–243.

Bouznit, Mohammed, Pablo-Romero, María del P., 2016. CO₂ emission and economic growth in Algeria. *Energy Policy* 96, 93–104.

Cai, Yifei, Sam, Chung Yan, Chang, Tsangyao, 2018. Nexus between clean energy consumption, economic growth and CO₂ emissions. *J. Clean. Prod.* 182, 1001–1011.

Chontanawat, Jaruwat, 2020. Relationship between energy consumption, CO₂ emission and economic growth in ASEAN: Cointegration and causality model. *Energy Rep.* 6, 660–665.

Dogan, Eyup, Inglesi-Lotz, Roula, 2020. The impact of economic structure to the environmental Kuznets curve (EKC) hypothesis: evidence from European countries. *Environ. Sci. Pollut. Res.* 27 (11), 12717–12724.

Ebn Sharif, Saleh, 2020. Bangladesh power sector review for 2020. <http://dx.doi.org/10.13140/RG.2.2.33692.51845>.

Etokakpan, Mfonobong Udom, Solarin, Sakiru Adebola, Yorucu, Vedat, Bekun, Festus Victor, Sarkodie, Samuel Asumadu, 2020. Modeling natural gas consumption, capital formation, globalization, CO₂ emissions and economic growth nexus in Malaysia: Fresh evidence from combined cointegration and causality analysis. *Energy Strat. Rev.* 31, 100526.

Fang, Debin, Zhang, Xiaoling, Yu, Qian, Jin, Trenton Chen, Tian, Luan, 2018. A novel method for carbon dioxide emission forecasting based on improved Gaussian processes regression. *J. Clean. Prod.* 173, 143–150.

Fatima, Samreen, Ali, Sayed Saad, Zia, Syeda Sadia, Hussain, Ehtesham, Fraz, Tayyab Raza, Khan, Mehwish Shafi, 2019. Forecasting carbon dioxide emission of Asian countries using ARIMA and simple exponential smoothing models. *Int. J. Econ. Environ. Geol.* 10 (1), 64–69.

Ghosh, Bikash Chandra, Alam, Khandakar Jahangir, Osmani, Md Ataul Gani, 2014. Economic growth, CO₂ emissions and energy consumption: The case of Bangladesh. *Int. J. Bus. Econ. Res.* 3 (6), 220–227.

Gökmenoğlu, Korhan, Taspınar, Nigar, 2016. The relationship between CO₂ emissions, energy consumption, economic growth and FDI: the case of Turkey. *J. Int. Trade Econ. Dev.* 25 (5), 706–723.

- Guo, Li-Na, She, Chen, Kong, De-Bin, Yan, Shuai-Ling, Xu, Yi-Peng, Khayat-nezhad, Majid, Gholinia, Fatemeh, 2021. Prediction of the effects of climate change on hydroelectric generation, electricity demand, and emissions of greenhouse gases under climatic scenarios and optimized ANN model. *Energy Rep.* 7, 5431–5445.
- Guo, Xifeng, Zhao, Qiannan, Zheng, Di, Ning, Yi, Gao, Ye, 2020. A short-term load forecasting model of multi-scale CNN-LSTM hybrid neural network considering the real-time electricity price. *Energy Rep.* 6, 1046–1053.
- Hameed, Waleed I, Sawadi, Baha A, Al-Kamil, Safa J, Al-Radhi, Mohammed S, Al-Yasir, Yasir IA, Saleh, Ameer L, Abd-Alhameed, Raed A, 2019. Prediction of solar irradiance based on artificial neural networks. *Inventions* 4 (3), 45.
- Hasan, Mohammad Maruf, Chongbo, Wu, 2020. Estimating energy-related CO2 emission growth in Bangladesh: The LMDI decomposition method approach. *Energy Strat. Rev.* 32, 100565.
- Heydari, Azim, Garcia, Davide Astiaso, Keynia, Farshid, Bisegna, Fabio, De Santoli, Livio, 2019. Renewable energies generation and carbon dioxide emission forecasting in microgrids and national grids using GRNN-GWO methodology. *Energy Procedia* 159, 154–159.
- Hossain, Md Alamgir, Chakraborty, Ripon K, Elsawah, Sondoss, Gray, Evan Mac A, Ryan, Michael J, 2021a. Predicting wind power generation using hybrid deep learning with optimization. *IEEE Trans. Appl. Supercond.*
- Hossain, Md Alamgir, Chakraborty, Ripon K, Elsawah, Sondoss, Ryan, Michael J, 2020. Hybrid deep learning model for ultra-short-term wind power forecasting. In: 2020 IEEE International Conference on Applied Superconductivity and Electromagnetic Devices. ASEM, IEEE, pp. 1–2.
- Hossain, Md Alamgir, Chakraborty, Ripon K, Elsawah, Sondoss, Ryan, Michael J, 2021b. Very short-term forecasting of wind power generation using hybrid deep learning model. *J. Clean. Prod.* 296, 126564.
- Hosseini, Seyed Mohsen, Saifoddin, Amirali, Shirmohammadi, Reza, Aslani, Alireza, 2019. Forecasting of CO2 emissions in Iran based on time series and regression analysis. *Energy Rep.* 5, 619–631.
- Huang, Yuansheng, Shen, Lei, Liu, Hui, 2019. Grey relational analysis, principal component analysis and forecasting of carbon emissions based on long short-term memory in China. *J. Clean. Prod.* 209, 415–423.
- Islam, Md Monirul, Khan, Muhammad Kamran, Tareque, Mohammad, Jehan, Noor, Dagar, Vishal, 2021. Impact of globalization, foreign direct investment, and energy consumption on CO2 emissions in Bangladesh: Does institutional quality matter? *Environ. Sci. Pollut. Res.* 1–21.
- Khan, Muhammad Kamran, Teng, Jian-Zhou, Khan, Muhammad Imran, Khan, Muhammad Owais, 2019. Impact of globalization, economic factors and energy consumption on CO2 emissions in Pakistan. *Sci. Total Environ.* 688, 424–436.
- Khobai, Hlalefang, Le Roux, Pierre, et al., 2017. The relationship between energy consumption, economic growth and carbon dioxide emission: The case of South Africa. *Int. J. Energy Econ. Policy* 7 (3), 102–109.
- Koc, Suleyman, Bulus, Gokay Canberk, 2020. Testing validity of the EKC hypothesis in South Korea: role of renewable energy and trade openness. *Environ. Sci. Pollut. Res.* 27 (23), 29043–29054.
- Leal, Patrícia Hipólito, Marques, António Cardoso, 2020. Rediscovering the EKC hypothesis for the 20 highest CO2 emitters among OECD countries by level of globalization. *Int. Econ.* 164, 36–47.
- Li, Yan, 2020. Forecasting Chinese carbon emissions based on a novel time series prediction method. *Energy Sci. Eng.* 8 (7), 2274–2285.
- Li, Yunlu, Ma, Guiqing, Yang, Junyou, Wang, Haixin, Feng, Jiawei, Ma, Yihua, 2021. Dynamic equivalent modeling for power converter based on LSTM neural network in wide operating range. *Energy Rep.* 7, 477–484.
- Li, Yan, Wei, Yigang, Dong, Zhang, 2020. Will China achieve its ambitious goal?—forecasting the CO2 emission intensity of china towards 2030. *Energies* 13 (11), 2924.
- Li, Yaoguang, Zhang, Yan, Zhao, Xinyue, Tian, Xin, 2018. The influence of US and China's CO2 transfer embodied in final consumption on global emission. *Energy Procedia* 152, 835–840.
- Mirza, Faisal Mehmood, Kanwal, Afra, 2017. Energy consumption, carbon emissions and economic growth in Pakistan: Dynamic causality analysis. *Renew. Sustain. Energy Rev.* 72, 1233–1240.
- Mitić, Petar, Munitlak Ivanović, Olja, Zdravković, Aleksandar, 2017. A cointegration analysis of real GDP and CO2 emissions in transitional countries. *Sustainability* 9 (4), 568.
- Muhammad, Faqeer, Karim, Rehmat, Muhammad, Khair, Asghar, Amna, 2020. Population density, CO2 emission and energy consumption in Pakistan: A multivariate analysis. *Int. J. Energy Econ. Policy* 10 (6), 250.
- Ofose-Adarkwa, Jeffrey, Xie, Naiming, Javed, Saad Ahmed, 2020. Forecasting CO2 emissions of China's cement industry using a hybrid Verhulst-GM (1, N) model and emissions' technical conversion. *Renew. Sustain. Energy Rev.* 130, 109945.
- Panigrahi, Sibarama, Behera, Himansu Sekhar, 2017. A hybrid ETS-ANN model for time series forecasting. *Eng. Appl. Artif. Intell.* 66, 49–59.
- Pesaran, M. Hashem, Shin, Yongcheol, Smith, Richard J., 2001. Bounds testing approaches to the analysis of level relationships. *J. Appl. Econometrics* 16 (3), 289–326.
- Rahman, Mohammad Mafizur, 2017. Do population density, economic growth, energy use and exports adversely affect environmental quality in Asian populous countries? *Renew. Sustain. Energy Rev.* 77, 506–514.
- Rahman, Mohammad Mafizur, Kashem, Mohammad Abul, 2017. Carbon emissions, energy consumption and industrial growth in Bangladesh: Empirical evidence from ARDL cointegration and Granger causality analysis. *Energy Policy* 110, 600–608.
- Rajagukguk, Rial A, Ramadhan, Raden A.A., Lee, Hyun-Jin, 2020. A review on deep learning models for forecasting time series data of solar irradiance and photovoltaic power. *Energies* 13 (24), 6623.
- Rayhan, Istihak, Akter, Khaleda, Islam, Mohammad Safiqul, Hossain, Mohammad Amzad, 2018. Impact of urbanization and energy consumption on CO2 emissions in Bangladesh: an ARDL bounds test approach. *Int. J. Sci. Eng. Res.* 9 (6), 838–843.
- Saint Akadiri, Seyi, Alola, Andrew Adewale, Akadiri, Ada Chigozie, 2019. The role of globalization, real income, tourism in environmental sustainability target. Evidence from Turkey. *Sci. Total Environ.* 687, 423–432.
- Salari, Mahmoud, Javid, Roxana J., Noghanibehambari, Hamid, 2021. The nexus between CO2 emissions, energy consumption, and economic growth in the US. *Econ. Anal. Policy* 69, 182–194.
- Sarkar, SKM, Sadeka, S., Sikdar, MHM, 2015. Badiuzzaman (2015) energy consumption and CO2 emission in Bangladesh: trends and policy implications. *Asia Pac. J. Energy Environ.* 2 (3), 175–182.
- Sarkar, Md Sujahangir Kabir, Sadeka, Sumaiya, Sikdar, Md Mehedi Hasan, et al., 2018. Energy consumption and CO2 emission in Bangladesh: trends and policy implications. *Asia Pac. J. Energy Environ.* 5 (1), 41–48.
- Shahbaz, Muhammad, Hye, Qazi Muhammad Adnan, Tiwari, Aviral Kumar, Leitão, Nuno Carlos, 2013. Economic growth, energy consumption, financial development, international trade and CO2 emissions in Indonesia. *Renew. Sustain. Energy Rev.* 25, 109–121.
- Sinha, Priyanka, 2013. Multivariate polynomial regression in data mining: methodology, problems and solutions. *Int. J. Sci. Eng. Res.* 4 (12), 962–965.
- Uddin, Md Gazi Salah, Bidisha, Sayema Haque, Ozturk, Ilhan, 2016. Carbon emissions, energy consumption, and economic growth relationship in Sri Lanka. *Energy Sources, Part B: Econ. Plan. Policy* 11 (3), 282–287.
- Valadkhani, Abbas, Smyth, Russell, Nguyen, Jeremy, 2019. Effects of primary energy consumption on CO2 emissions under optimal thresholds: Evidence from sixty countries over the last half century. *Energy Econ.* 80, 680–690.
- Wasti, Syed Khurram Arslan, Zaidi, Shumaila Waqar, 2020. An empirical investigation between CO2 emission, energy consumption, trade liberalization and economic growth: A case of Kuwait. *J. Build. Eng.* 28, 101104.
- Yao, Shunyu, Xu, Yi-Peng, Ramezani, Ehsan, 2021. Optimal long-term prediction of Taiwan's transport energy by convolutional neural network and wildebeest herd optimizer. *Energy Rep.* 7, 218–227.
- You, Wanhai, Lv, Zhike, 2018. Spillover effects of economic globalization on CO2 emissions: a spatial panel approach. *Energy Econ.* 73, 248–257.
- Yu, Yixiao, Han, Xueshan, Yang, Ming, Yang, Jiajun, 2020. Probabilistic prediction of regional wind power based on spatiotemporal quantile regression. *IEEE Trans. Ind. Appl.* 56 (6), 6117–6127.
- Yuan, Jiahai, Hou, Yong, Xu, Ming, 2012. China's 2020 carbon intensity target: consistency, implementations, and policy implications. *Renew. Sustain. Energy Rev.* 16 (7), 4970–4981.
- Yuping, Li, Ramzan, Muhammad, Xincheng, Li, Murshed, Muntasir, Awosusi, Abraham Ayobamiji, BAH, Sununu Ibrahim, Adebayo, Tomiwa Sunday, 2021. Determinants of carbon emissions in Argentina: The roles of renewable energy consumption and globalization. *Energy Rep.* 7, 4747–4760.
- Zeroual, Abdelhafid, Harrou, Fouzi, Dairi, Abdelkader, Sun, Ying, 2020. Deep learning methods for forecasting COVID-19 time-series data: A comparative study. *Chaos Solitons Fractals* 140, 110121.
- Zhou, Cheng, Chen, Xiyang, 2021. Predicting China's energy consumption: Combining machine learning with three-layer decomposition approach. *Energy Rep.* 7, 5086–5099.