Advanced Neural Network Models in Forecasting Carbon Emissions

Shengkai Xu*

Faculty of Science and Engineering, Nottingham University, Ningbo, China

*Corresponding author: ssysx2@nottingham.edu.cn

Abstract. The detrimental trend of the surging carbon emissions (CEs) and the subsequent issues of environmental problems have attracted global attention. Despite control measures that have been taken by governments worldwide to achieve carbon neutrality, further policymaking remains challenging. After browsing various research, forecasting CEs in advance is considered a proper measure to mitigate the trend by making the carbon capture data-oriented. Judging from influencing factors of concentrations, it is seen that a great number of researchers have focused on machine learning methods, especially for neural network models. Then neural network models implemented in three perspectives, including worldwide, local, and transportation, are analyzed in six metrics: Coefficient of determination (R2), root mean square error (RMSE), mean square error (MSE), mean absolute error (MAE), mean absolute percentage error (MAPE), Relative Error (RE). The results following these metrics suggest that the future of utilizing neural network models in forecasting CEs is promising in guiding future carbon neutrality works.

Keywords: Neural Network, Prediction Model, Carbon Emission.

1. Introduction

As CEs played a significant role in deteriorating the global climate, increasing the temperature via greenhouse effects. It is essential to forecast the emissions accurately and precisely in advance for better treatment. Among all the methods, researchers seemed to focus most on the neural network depending on its great potential to fit the actual curve.

With parallel connections comprised of straightforward adaptable units, the neural network simulates the interaction between a biological nervous system and the outside world. This artificial neural network (ANN) mimics the information processing carried out by the brain. The network solves a nonlinear problem based on the input information received from the neurons. Implementing neural networks for forecasting CEs involves first selecting training and verification samples from historic data. The network structure is established, parameters are set, and then network training is performed. Finally, once it has attained a sufficient level of accuracy, it can be applied as a prediction model [1].

The performance metrics are popular machine learning metrics standing for R^2, RMSE, MSE, MAE, MAPE, and RE. The equations for calculating these performance metrics are in equations (1)-(6) as below:

$$\mathrm{R}^2 = 1 - \frac{\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}{\sum_{i=1}^{n}(y_i-\bar{y})^2} \tag{1}$$

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

$$MSE = RMSE^2 \tag{3}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i-\hat{y}_i| \tag{4}$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{\bar{y}_i-y_i}{y_i}\right| \tag{5}$$

$$RE = \frac{Absolute\ Error}{True\ value} \tag{6}$$

After comparing distinct neural network structures and their performances, the evaluation would be made to distinguish the best model. Besides, whether the neural network is generally capable of accomplishing the task was suggested.

2. Neural Network (NN) Implementation

2.1. Worldwide CE Prediction

To predict worldwide carbon dioxide emissions, the Ray Optimization (RO) approach is combined with a multilayer perceptron neural network [2].

Population, gross domestic product (GDP), and oil and natural gas trade movements are all input factors for SEIs. According to the socioeconomic indicators (SEIs), the RO first specifies the worldwide fossil fuel and primary energy demand equations. The RO technique is then used to calculate CE on the basis of coal consumption, natural gas consumption (NG), crude oil consumption, and primary energy consumption.

Furthermore, the 1980–2006 data were used for both the implementation (1980–1999), and the testing (2000–2006) of the models. The CE for the world up to the year 2025 is then forecasted.
For each input variable (SEIs), the optimized network structure has been established.

1. Population: 6 neurons (1st layer); 1 neuron (2nd layer), having an MSE of 5.45E-8, R^2 of 99.99%.
2. GDP: 3 neurons (1st layer); 2 neurons (2nd layer), having an MSE of 3.29E-4, R^2 of 99.99%.
3. Oil trade movement: 7 neurons (1st layer); 2 neurons (2nd layer), having an MSE of 8.89E-4, R^2 of 99.96%.
4. NG trade movement: 2 neurons (1st layer); 1 neuron (2nd layer), having an MSE of 0.023, R^2 of 99.56%.

For the five parameters (Oil, NG, Coal, PE, and CO_2), the overall R^2 for the exponential developed models outweigh the linear ones, which are all over 99.50%.

The proposed model has a higher accuracy of CE prediction. Both the training error and testing error of distinct input variables in the established structure appear to be low. Additionally, the R^2 for different parameters in developed linear and exponential models is significantly high. To be reflected on CO2 emission prediction, it is spotted that the linear model possesses a lower relative error, which is only 0.0023905%.

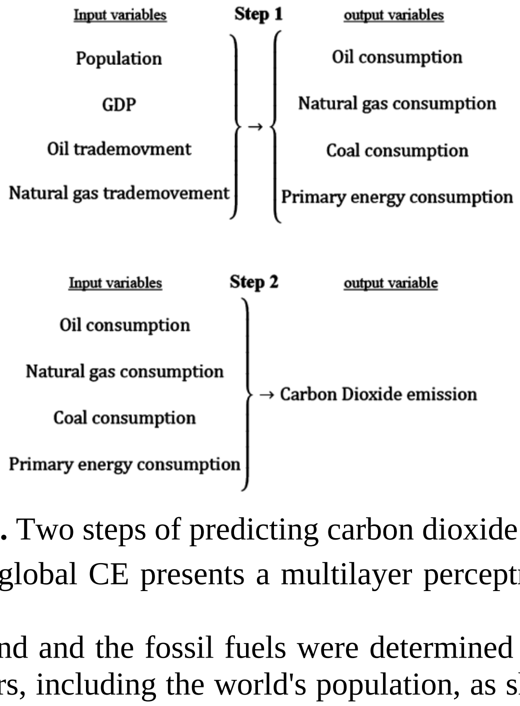

Figure 1. Two steps of predicting carbon dioxide emission

Another study estimating global CE presents a multilayer perceptron neural network integrated with the Bees Algorithm [3].

The primary energy demand and the fossil fuels were determined with support from a series of social and economic indicators, including the world's population, as shown in Figure 1. Then, there were two steps following these input indicators.

And the errors in the output of step 1 are shown below in Table 1.

Table 1. Actual data of the output variables (in exajoules) compared with the modelling result from 2000 to 2006

Relative error (%)	2000	2001	2002	2003	2004	2005	2006	Average
$Oil_{exponential}$	-0.433	-0.12	-0.158	-0.14	-1.13	-0.281	0.826	0.441
Oil_{linear}	-0.524	-0.654	-0.029	0.465	-0.713	-0.107	0.043	0.362
$NG_{exponential}$	0.283	0.791	-0.158	0.059	-0.147	-0.098	0.46	0.285
NG_{linear}	-0.211	0.696	-0.278	0.033	-0.355	-0.37	0.063	0.287
$Coal_{exponential}$	2.497	1.63	1.936	-3.022	-3.707	-3.958	-4.041	2.97
$Coal_{linear}$	0.496	2.273	1.293	-1.621	-2.734	-4.151	-7.645	2.888
$PE_{exponential}$	0.7	0.724	0.436	-0.324	-0.845	-0.286	0.699	0.573
PE_{linear}	-0.33	0.434	0.557	0.576	0.006	0.126	-0.28	0.33

For the installation period (1980 - 1999), the following data of R^2 is shown:

$$Oil_{linear} = 99.95\%,\ Oil_{exponential} = 99{:}31\%$$

$$NG_{linear} = 99.94\%,\ NG_{exponential} = 99.97\%$$

$$Coal_{linear} = 98.07\%,\ Coal_{exponential} = 98.95\%$$

$$PE_{linear} = 99.72\%,\ PE_{exponential} = 99.45\%$$

$$CO_{2linear} = 99.98\%,\ CO_{2exponential} = 99.99\%$$

According to the data, CO_2 emission has been forecasted with the linear model best performed to have a relative error of 0.457% compared with a 0.765% of the exponential model.

2.2. Local CE Prediction

To exemplify, using a neural network to predict the CE in different cities of Turkey, the proposed model shows good performance. The predicted data were compared with actual values in 2014-2017 with performance metrics: R^2 0.976, MSE 3.198, RMSE 1.788 (round to 3 decimal places). Besides, the RReliefF algorithm was used to find parameters having weighted effects on the prediction of CE [4].

To train the neural network, the base model was used along with input data in the .tif format, which corresponded to emissions from 81 provinces in Turkey for 2015. In accordance with the results, the ANN model containing 24 neurons, a single hidden layer, and a tangent-sigmoid activation function exhibited the highest R^2, MSE, and RMSE values. The data will be used to train 70% of the system, validate 15% of the system, and test 15% of the system. As the learning algorithm, Levenberg-Marquardt was selected since it is generally suited for curve fitting and nonlinear regression modeling. As can be seen from Figure 2, this neural network performs well overall [5].

There are two layers in the optimized multilayer perceptron (MLP), a hidden layer, ten hidden layer neurons, and a single output layer. In the MLP, 1000 iterations are performed. Both inputs and outputs have tangent activation functions.

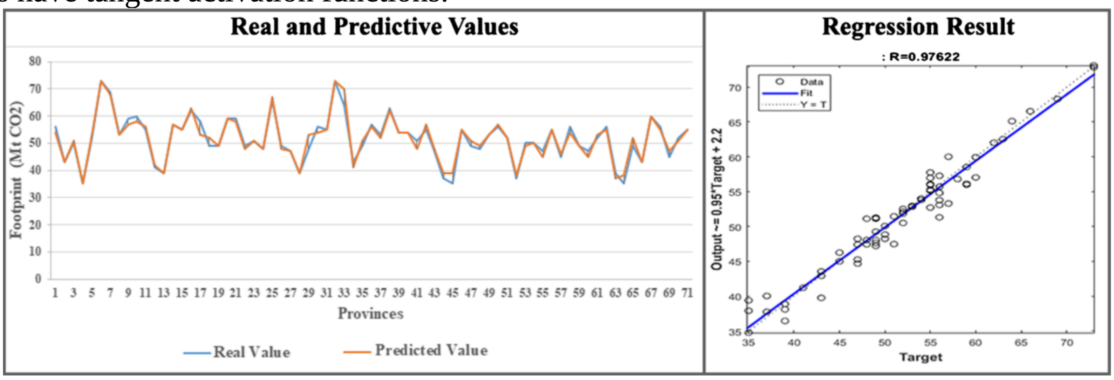

Figure 2. Comparing real values with predictions (2015)

As shown in the figure, this optimized model with one hidden layer, ten neurons, tangent activation function, and 1000 iterations has an R^2 of 0.97622.

The artificial neural network was trained with nine parameters from another study that provides predictions of the intensity of CE in five countries (Australia, Brazil, China, India, and the United States). In the study, nine parameters were taken into account: urbanization, financial development, research & development, energy consumption, economic growth, trade openness, population, foreign direct investment, and industrialization [6].

It was necessary to include quarterly data from 1980 Q1 to 2015 Q4 aiming to train the model to optimize its performance. All five countries were then validated using the optimal models of 9-5-1 MLP supported by a backpropagation algorithm after twenty simulations. In Australia, the R^2 of the model was 0.80. In Brazil, it was 0.91; in China, it was 0.95; in India, it was 0.99; and in the United States, it was 0.87.

Figure 3 below illustrates the training process.

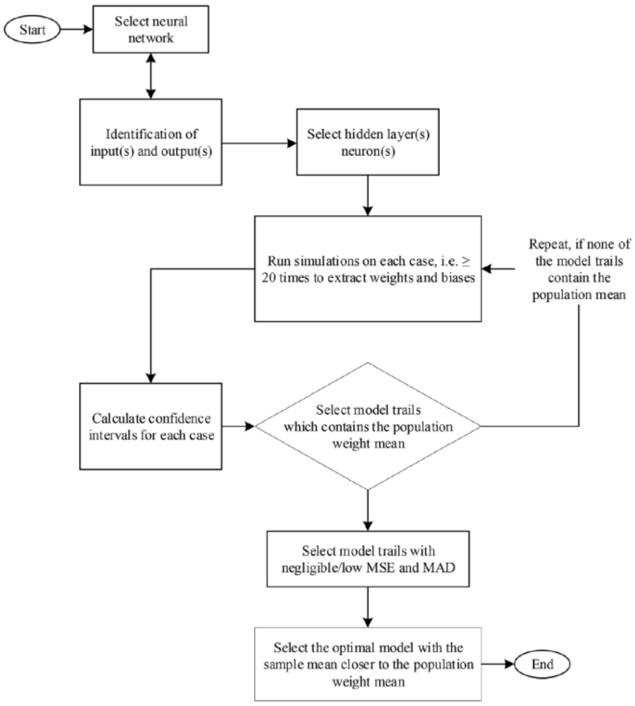

Figure 3. Model selection flow chart

The performance of the model is determined by the selection of the hidden layer neurons. An optimized neuron number of 5 in the hidden layer is calculated based on the trial-and-error approach. The neural work is configured with the above nine input parameters as the input layer, 5 hidden layers of neurons, and CE intensity as the output layer. The CO_2 intensities are shown in Figure 4.

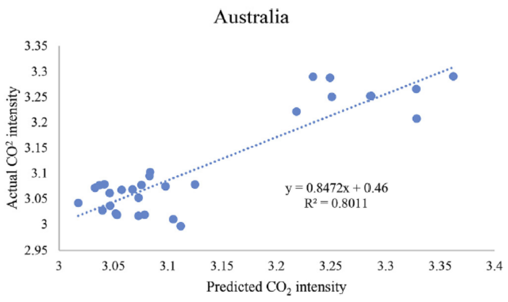

Fig. 4a. Scatter chart of actual and predicted CO_2 intensities.

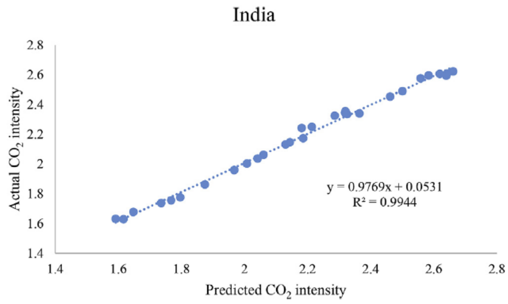

Fig. 4d. Scatter chart of actual and predicted CO_2 intensities.

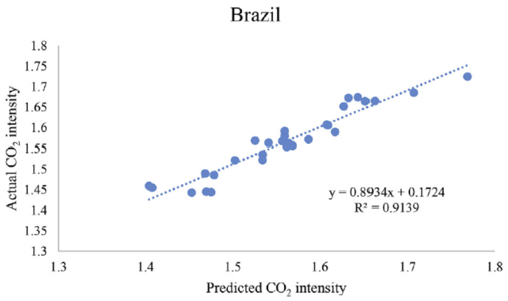

Fig. 4b. Scatter chart of actual and predicted CO_2 intensities.

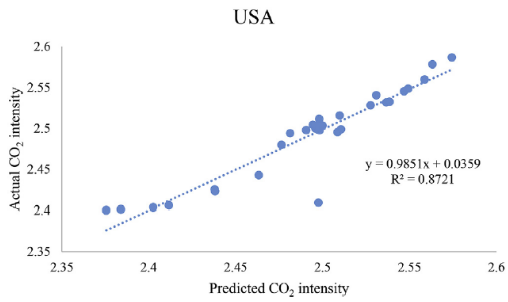

Fig. 4e. Scatter chart of actual and predicted CO_2 intensities.

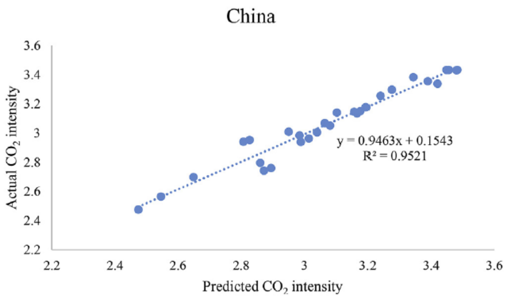

Fig. 4c. Scatter chart of actual and predicted CO_2 intensities.

Figure 4. CO_2 intensities scattering diagram (X-axis: predicted value, Y-axis: real value)

The R^2 value of the five selected countries follows an order of Australia (0.8011) < USA (0.8721) < Brazil (0.91) < China (0.95) < India (0.9944), indicating that the model offers the best fitting result for the data in India. The results of the sensitivity analysis also show that Australia has the highest sensitivity weight for research and development. As for the sensitivity weight for urbanization, Brazil and the United States are the most sensitive. India's energy consumption has the highest sensitivity weight, while China's population size has the largest sensitivity weight.

Another research collected CE data between 1999 and 2019 from different Chinese provinces and analyzed them, and the emissions of 10,000 residents' direct consumption were determined based on population data for all provinces. To offer a scientific foundation for CE management and planning, the trend and characteristics of changes in CE between 2027 and 2032 are studied and anticipated based on data on direct CEs of people of Chinese provinces from 1999 to 2019. The weighted factors included as input data were residential energy consumption types, CO_2 emission coefficients, electrical CE coefficients, and populations (total, urban, and rural) [7].

Within this paper, three types of artificial neural networks were discussed (Backpropagation Neural Network (BPNN), Radial Basis Function (RBF), and Elman neural network). After considering prediction accuracy, the complexity of construction, robustness, and fault tolerance, the Elman neural network seems to be the best one (Figure 5).

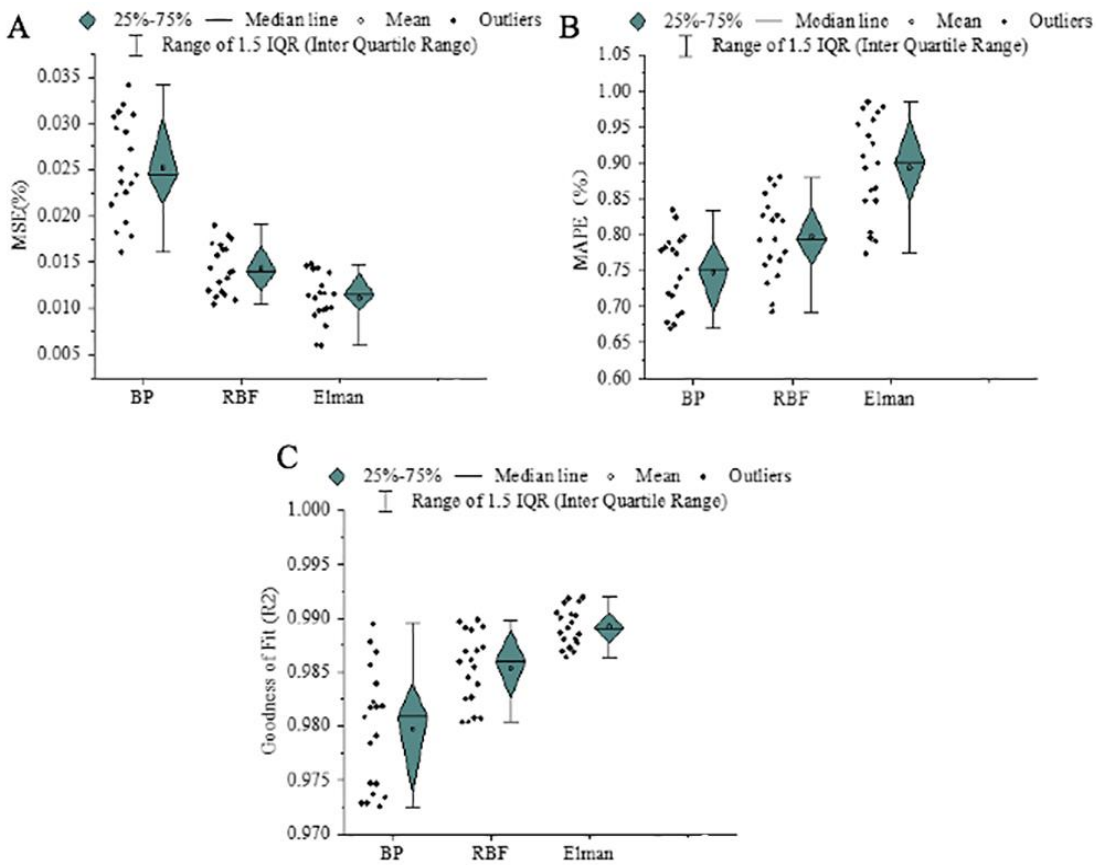

Figure 5. Analyzing the performance of various prediction models

The overall performances of the three models are shown above, suggesting Elman is the best with MSE of 0.005-0.015% and R^2 of 0.985-0.995.

The Yellow River Basin occupies a significant position in China, both as an ecological barrier and as an area of rapid economic development. It is crucial to understand the factors influencing CEs in the Yellow River Basin to achieve carbon peaking in China. Thus, a research study was conducted to forecast CEs in the Yellow River Basin. The value of the index was predicted based on five weighted factors (the population, the GDP, the industrial structure, the urbanization rate, and the energy intensity) [8].

An LSTM (Long short-term memory, a type of artificial neural network) - optimized sparrow search algorithm was used to forecast CEs in the Yellow River Basin. The Yellow River Basin is experiencing an increase in CEs, with significant differences between provinces. Based on a comparison with 2000, the Yellow River Basin's CE intensity decreased gradually from 5.187 t/10,000 RMB to 1.672 t/10,000 RMB in 2019. In contrast to Shandong, which emits the most carbon dioxide, Qinghai emits less than a tenth as much. GDP per capita was the most significant contributor to CEs in the Yellow River Basin between 2000 and 2010, followed by population after 2010.

To specify, in traditional recurrent neural network (RNN) models, problems including gradient disappearance and explosion are the results of error transmission. An improved RNN model, the LSTM model, addresses this problem. Input gates, forgetting gates, and output gates are introduced in the LSTM model to control information transmission. In addition, the model achieves effective processing of long time series information by avoiding the two problems mentioned above, which are common when handling time series data. Besides, the sparrow search algorithm (SSA) is proposed as a new swarm optimization algorithm. Sparrows are observed to perform both anti-predatory and foraging behaviors. This inspired the development of the SSA, and the flowchart of SSA-LSTM is shown in Figure 6 [9].

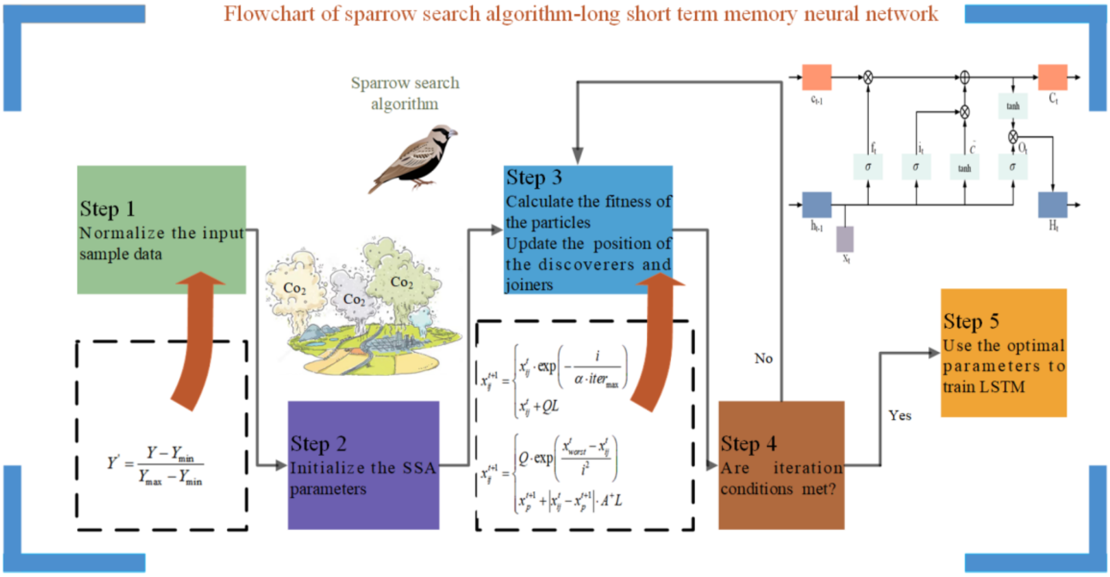

Figure 6. The flowchart of SSA-LSTM

After acknowledging the overall theory of SSA-LSTM, the optimized parameter was determined through training, testing, and validating, the detailed ranges of parameters are shown in Table 2, and the prediction result of different models are shown in Figure 7.

Table 2. Parameter optimization range of LSTM

Parameter	Range
Learning rate	$(1*10^{-3}, 1*10^{-2})$
Number of iterations	(50, 200)
Number of neurons	(1, 200)

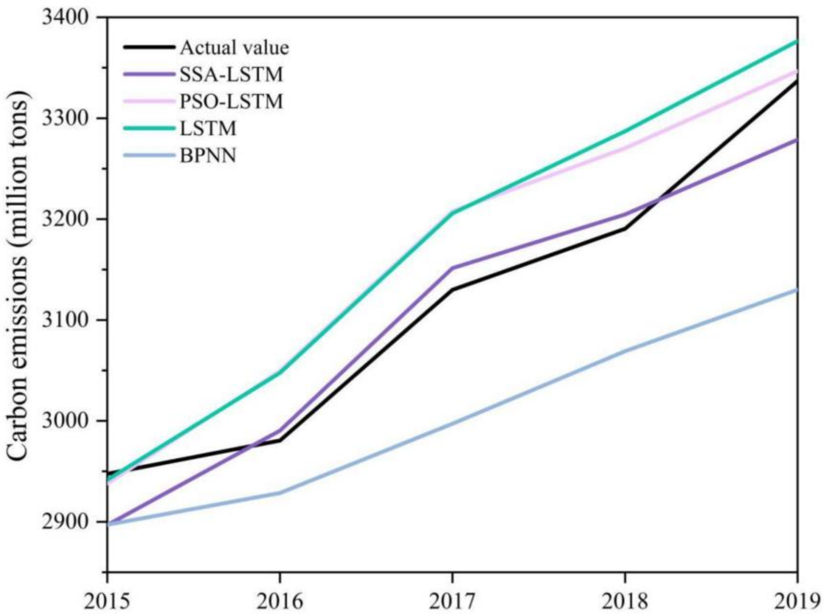

Figure 7. Comparison of the model prediction result

Table 3. Prediction model evaluation index results

	SSA-LSTM	PSO-LSTM	LSTM	BPNN
MAE	30.90	49.29	57.01	126.77
RMSE	36.67	59.04	65.11	112.64
MAPE	0.0099	0.0155	0.0178	0.0370

This optimized SSA-LSTM was set to have a specific range of 3 parameters. It was further proved to best fit the actual value of CEs with errors of 30.90 (MAE), 36.67 (RMSE), and 0.0099 (MAPE), as illustrated in Table 3.

To estimate China's CEs differently, the author identified 16 potential influencing factors (including population, railway operating mileage, freight volume, integrated circuit, and urbanization) and used gray relational analysis to distinguish the factors that are closely correlated with CEs. For example, population, railway operating mileage, freight volume, and urbanization have a grey relational degree over 0.99. On the other side, the integrated circuit only has a degree of 0.516. To extract the four principal components from the input data, principal component analysis (PCA) was used. This reduced the amount of redundancy in the data [10].

As means of predicting CEs in China, the LSTM method has been developed. Later, the gaussian process regression (GPR) and back propagation neural networks (BPNN) were compared with the LSTM method. All in all, LSTM was more effective at predicting CEs based on simulation results than BPNN and GPR, indicating that LSTM is effective at predicting CEs.

The overall MAPE of LSTM has been calculated to be at 1.983%, which is much lower than GPR's 7.101% and BPNN's 7.186%. For its further development, improved factors were implemented to reach a MAPE of only 0.436%.

The refinement can be directly spotted on the line chart, as shown in Figure 8.

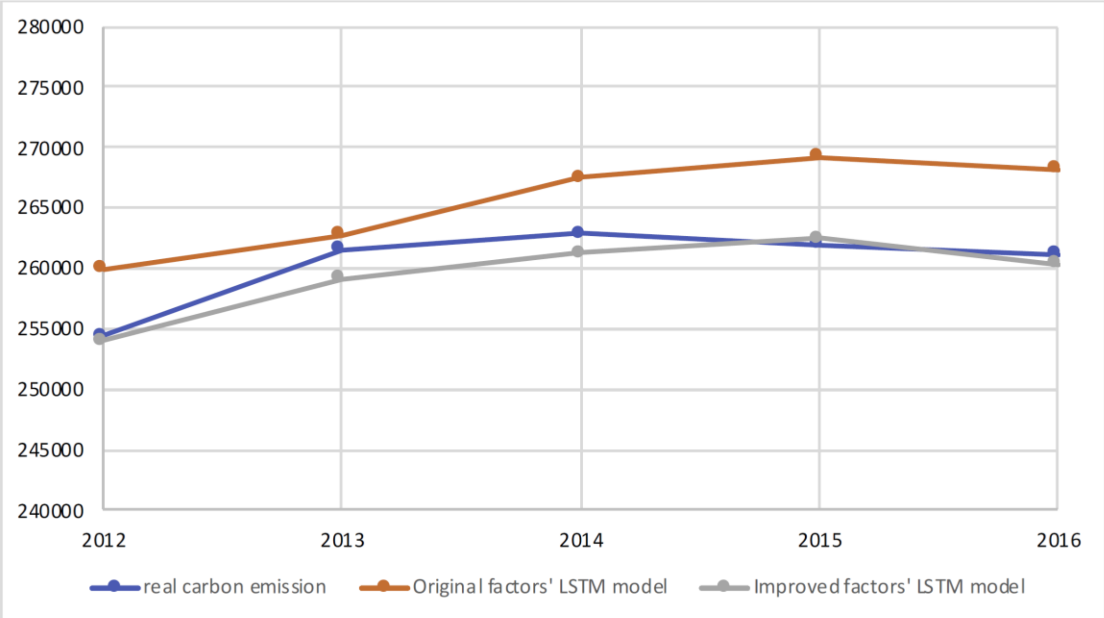

Figure 8. Prediction of CEs based on the improved factors

2.3. Transportation CE Prediction

Transportation CEs contribute significantly to the increase of greenhouse gases, which threaten human health and climate change. Therefore, it is imperative to be able to master real-time information about transportation CEs. However, it seems to be impossible for traditional methods to make it.

This study was based on spatiotemporal datasets of data collected from a variety of sources in the city, including taxi GPS data, transportation CE data, road networks, points of interest (POI), and meteorological data. In this paper, a three-layer perceptron neural network (3-layer PNN) is proposed as a tool for forecasting real-time and finely-grained information regarding transportation CE across the city in real-time [11].

Five real data sources were analyzed in Zhuhai, China, in order to evaluate the PNN. It was found that the PNN was superior to three well-known machine learning approaches (Gaussian Naive Bayes, Logistic Regression, and Linear Regression) and two deep learning approaches (Deep Belief Networks and Stacked Denoising Autoencoders).

Six data sources in Zhuhai were recorded during 2015/08/01-2015/10/31 with about 2,200 hours of time span.

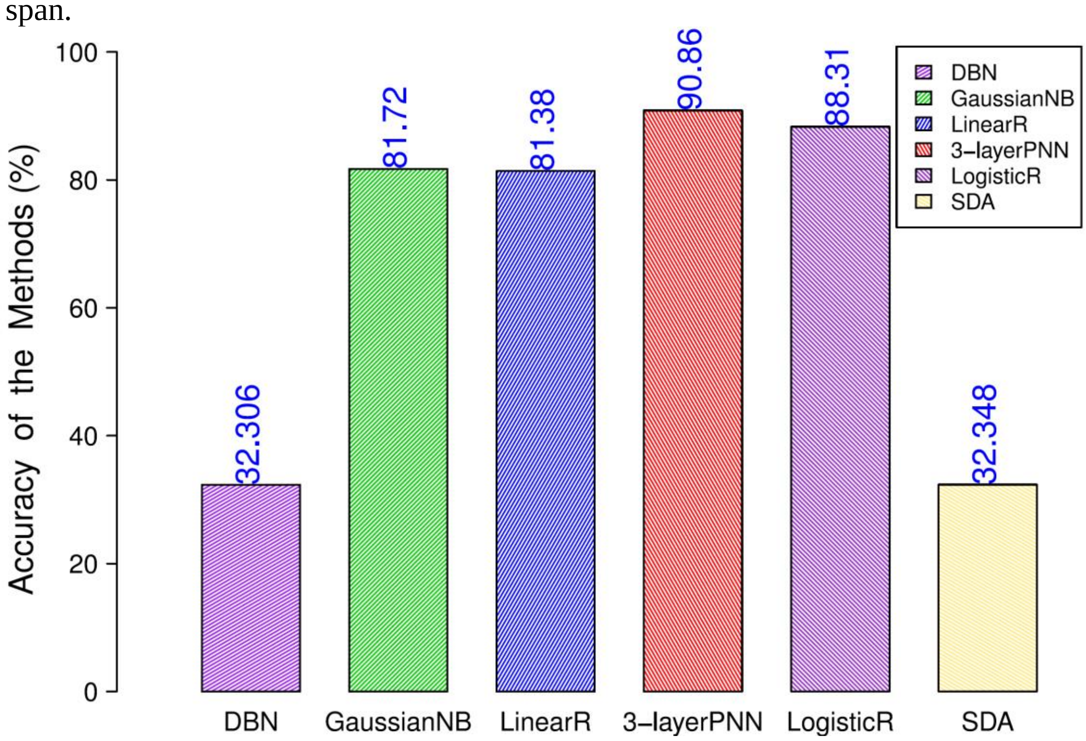

Figure 9. Overall results comparing different methods

The detailed datasets and high accuracy (90.86%) are shown in Figure 9.

According to another paper, the CEs from the aviation industry in EU countries were estimated using artificial neural networks. Four factors are important to consider in this data: type of flight, age of the fleet, number of flights, and number of passengers. The most suitable ANN network for domestic flights is one with two hidden layers and 20 neurons. In contrast, ANN networks with two hidden layers and 15 neurons are most suitable for international flights. For domestic flights, the MAPE error was 7.8%, while for international flights, it was 6.7%. There are 98.1% correlation coefficients for domestic flights and 99.0% correlation coefficients for international flights [12].

Besides carbon dioxide (CO_2), a study was conducted on carbon monoxide (CO). An integrated model combining neural networks, data mining, and geographic information systems was developed to predict emissions on the New Klang Valley Expressway in Malaysia. A field survey and open-source data are employed to generate prediction maps in small urban areas through correlation-based feature selection (CFS). To assist in the survey, 15-minute intervals were recorded four times a day on weekends and weekdays for the measurement of traffic CO concentrations [13].

A grid search procedure and a CFS model were utilized in the development of the model in order to optimize the network architecture, hyperparameters, and predictors. Therefore, six predictors of traffic CO were used: vehicle number, heavy truck number, motorbike number, temperature, wind speed, and digital surface model. The detailed architecture of the proposed neural network is shown in Figure 10.

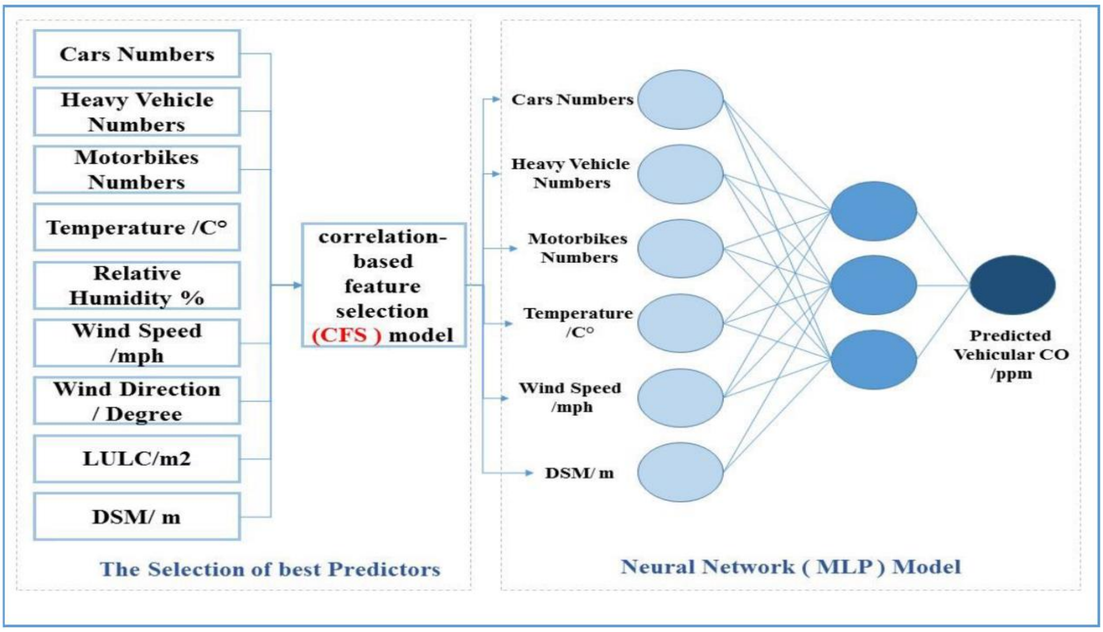

Figure 10. The architecture of the proposed neural network for traffic CO prediction

In total, 352 instances were utilized, of which 247 (70%) were training instances, and 105 (30%) were testing instances. After the analysis, parameters were best chosen to refine the training process.

The selected MLP & RBF have the number of hidden units in the range of 3-40 using BFGS & RBFT as the training algorithm. The hidden & output activation is identity, logistic, Tanh, and Exponential. Finally, the learning rate is (0.1, 0.5) with a momentum of (0.1-0.9).

After optimization and comparison between different model structures, the CFS-MLP model was proven to be the best, with its structure and results being displayed as shown in Table 4.

Table 4. The best structure of the CFS-MLP model with its results

Best structure	6-3-1
Correlation coefficient	0.980
MAE (ppm)	0.8925
RMSE (ppm)	1.2736
Total number of instances	247

To summarize, this CFS-MLP model with 6 input layers, 3 hidden layers, and 1 output layer has an R^2 of 0.9604 (deducing from correlation coefficient).

3. Conclusion

In summary, several typical models have been compared and analyzed in this paper. The assessed models involve multiple aspects of CEs prediction, which could be applied to large ranges as the worldwide prediction, or more specifically be applied to a local region or even the transportation area. The ANN has been integrated with several other methods, such as RO or Bees algorithm, to provide more accurate predictions. The choosing of the influencing factors and the number of hidden layers and neurons would effectively affect the accuracy of the prediction, which has been proved by a comparison of multiple reported training processes in this paper. More carefully designing the training set is necessary to acquire precise prediction results. In addition, according to the research, the LMST and Elman NN are promising neural networks in further research. According to the marvelous performance of these models, it can be concluded that neural network models within different structures are pragmatic tools to use when forecasting CEs.

References

[1] Li X, Zhou W, Liu Z, et al. Research Overview on Carbon Emission Models[C]//2021 16th International Conference on Computer Science & Education (ICCSE). IEEE, 2021: 172-176.

[2] Assareh E, Nedaei M. A metaheuristic approach to forecast the global carbon dioxide emissions[J]. International Journal of Environmental Studies, 2018, 75(1): 99-120.

[3] Behrang M A, Assareh E, Assari M R, et al. Using bees algorithm and artificial neural network to forecast world carbon dioxide emission[J]. Energy Sources, Part A: Recovery, Utilization, and Environmental Effects, 2011, 33(19): 1747-1759.

[4] Bozdağ A. Local-based mapping of carbon footprint variation in Turkey using artificial neural networks[J]. Arabian Journal of Geosciences, 2021, 14(6): 1-15.

[5] Moran D, Kanemoto K, Jiborn M, et al. Carbon footprints of 13 000 cities[J]. Environmental Research Letters, 2018, 13(6): 064041.

[6] Acheampong A O, Boateng E B. Modelling carbon emission intensity: Application of artificial neural network[J]. Journal of Cleaner Production, 2019, 225: 833-856.

[7] Jin H. Prediction of direct carbon emissions of Chinese provinces using artificial neural networks[J]. Plos one, 2021, 16(5): e0236685.

[8] Zhao J, Kou L, Wang H, et al. Carbon Emission Prediction Model and Analysis in the Yellow River Basin Based on a Machine Learning Method[J]. Sustainability, 2022, 14(10): 6153.

[9] Xue J, Shen B. A novel swarm intelligence optimization approach: sparrow search algorithm[J]. Systems Science & Control Engineering, 2020, 8(1): 22-34.

[10] Huang Y, Shen L, Liu H. Grey relational analysis, principal component analysis and forecasting of carbon emissions based on long short-term memory in China[J]. Journal of Cleaner Production, 2019, 209: 415-423.

[11] Lu X, Ota K, Dong M, et al. Predicting transportation carbon emission with urban big data[J]. IEEE Transactions on Sustainable Computing, 2017, 2(4): 333-344.

[12] Demir A S, Eminler Ö E. Estimation of CO2 emissions from air transportation in EU countries by artificial neural networks[J]. International Journal of Global Warming, 2020, 21(3): 234-243.

[13] Azeez O S, Pradhan B, Shafri H Z M, et al. Modeling of CO emissions from traffic vehicles using artificial neural networks[J]. Applied Sciences, 2019, 9(2): 313.