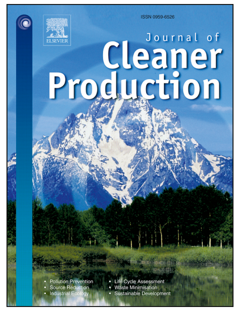
Acheampong, Alex O.; Boateng, Emmanuel B.. "Modelling carbon emission intensity: application of artificial neural network". Published in Journal of Cleaner Production Vol. 225, Issue 10 July 2019, p. 833-856 (2019).

**Available from**: http://dx.doi.org/10.1016/j.jclepro.2019.03.352

# Accepted Manuscript




Modelling carbon emission intensity: Application of artificial neural network

Alex O. Acheampong, Emmanuel B. Boateng

PII: S0959-6526(19)31068-6

DOI: https://doi.org/10.1016/j.jclepro.2019.03.352

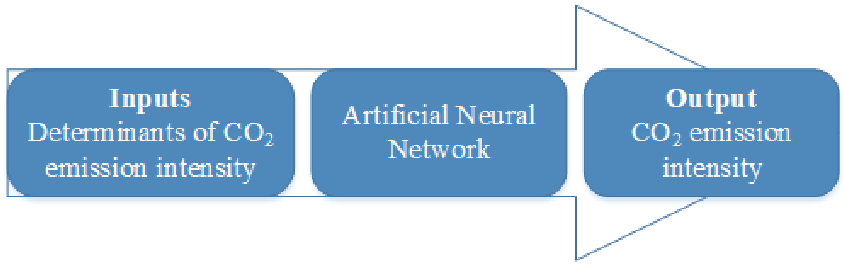
Reference: JCLP 16356

To appear in: *Journal of Cleaner Production*

Received Date: 1 December 2018

Revised Date: 11 February 2019

Accepted Date: 31 March 2019


**Inputs**
Determinants of $CO_2$ emission intensity
Artificial Neural Network
**Output**
$CO_2$ emission intensity

# Modelling Carbon Emission Intensity: Application of Artificial Neural Network

**Alex O. Acheampong[1,*], Emmanuel B. Boateng[2]**

[1] *Newcastle Business School, University of Newcastle, Australia*
[2] *School of Health Sciences, University of Newcastle, Australia*

**Abstract**

This study applies an artificial neural network (ANN) to develop models for forecasting carbon emission intensity for Australia, Brazil, China, India, and USA. Nine parameters that play an essential role in contributing to carbon emissions intensity were selected as input variables. The input parameters are economic growth, energy consumption, R&D, financial development, foreign direct investment, trade openness, industrialisation, and urbanisation. The study used quarterly data which span over the period 1980Q1-2015Q4 to develop, train and validate the models. To ensure the reproducibility of the results, twenty simulations were performed for each country. After numerous iterations, the optimal models for each country were selected based on predefined criteria. A 9-5-1 multi-layer perceptron with back-propagation algorithm was sufficient in building the models which have been trained and validated. Results from the validated models show that the predicted versus actual values indicate approximately zero errors along with higher coefficients of determination ($R^2$) of 0.80 for Australia, 0.91 for Brazil, 0.95 for China, 0.99 for India and 0.87 for USA. The Partial Rank Correlation Coefficient (PRCC) results reveal that for Australia, R&D has the highest sensitivity weight while for Brazil and the USA, urbanisation has the highest sensitivity weight. For China, population size has the highest sensitivity weight while energy consumption has the highest sensitivity weight in India. The ANN models presented in this study have been validated and reliable to predict the growth of $CO_2$ emission intensity for Australia, Brazil, China, India, and USA with negligible forecasting errors. The models developed from this study could serve as tools for international organizations and environmental policymakers to forecast and help in climate change policy decision-making.

**Keywords:** Carbon emissions; Artificial neural network; Forecasting; Sensitivity

[*] Corresponding Author: The University of Newcastle, Newcastle Business School, 409 Hunter Street, Newcastle, NSW 2300, Australia or by via email at alexopokuacheampong@gmail.com/alex.acheampong@uon.edu.au

## 1. Introduction

This study sought to use Artificial Neural Network (ANN) to develop models for forecasting carbon emission intensity for Australia, Brazil, China, India, and the USA. Global warming has been the most challenging environmental issue in the history of humanity (Acheampong, 2018). Thus, the increasing concentration of greenhouse gases in the atmosphere leading to global warming has severe implications for both economic and human development. Carbon dioxide is the primary greenhouse gas behind global warming. Therefore, efforts by international organisations to mitigate the adverse effect of global warming have been a focus on policies to reduce carbon emissions (Tamazian, Chousa, & Vadlamannati, 2009). Global carbon emissions have been increasing despite the global effort to reduce it. According to the International Energy Agency (2018) report, global energy-related carbon emissions increased by 1.4% in 2017. This represents an absolute increase of 460 million tons (Mt) reaching a historic high of 32.5 gigatons (Gt) for the past three years after remaining flat. This astronomic increase in carbon emissions conflicts with the Paris agreement on climate change to reduce carbon emissions.

While the global economy has witnessed an increase in carbon emissions, countries such as the USA, UK, Mexico, and Japan have experienced a sharp reduction in carbon emissions in 2017. For instance, carbon emissions dropped by 0.5% representing 25Mt to 4810 Mt in the USA (International Energy Agency, 2018). On the other hand, the role of the Asia economies in carbon emissions cannot be underestimated. Two-third of global carbon emissions comes from Asian countries. Specifically, China and India are major players contributing to the increase in global carbon emissions. In the recent report by International Energy Agency (2018), carbon emissions increase from China increased by 9.1 gigatons in 2017 which is 1% higher than the level of emissions in 2014. Additionally, India experienced per-capital emissions of 1.7t $CO_2$. Southeast Asian has also contributed significantly to global emissions, with Indonesia playing a major in this region.

Understanding the future trend of carbon emissions at the global, regional, and national level could provide insight for developing appropriate environmental policies and strategies to mitigate carbon emissions. Thus, developing a reliable model for predicting the growth of carbon emissions could serve as the tool for international organisations and environmental policymakers to design and implement appropriate environmental policies and strategies to control environmental problems. Over the decades, some researchers have used classical statistical and econometric approaches to model or forecast the growth of carbon emissions. For instance, regression analysis has been the most popular estimation technique to study the causal relationship between carbon emissions and other independent variables such as economic growth, population, energy consumption, technology, globalisation and among others (see for example, Ahmad et al., 2017; Ahmed, Rehman, & Ozturk, 2017; Al-Mulali, Ozturk, & Solarin, 2016; Almeida, Cruz, Barata, & García-Sánchez, 2017; Grossman & Krueger, 1995; Köne & Büke, 2010). However, the effectiveness of regression depends on the reliability and availability of independent variables (Zhou, Ang, & Poh, 2006). Additionally, given that variables for modelling carbon emissions are chaotic, non-stationary, and non-linear, the classical statistical and econometric approaches are not suitable for modelling such a complex behaviour (Gallo, Contò, & Fiore, 2014; Hussain & Reynolds, 1975; Stanley, 1997).

In addition to the statistical and regression approaches, some scholars have also employed time series models such as Box-Jenkins Autoregressive Integrated Moving Average (ARIMA) and Autoregressive Moving Average (ARMA) to forecast emissions. For accurate forecasting using ARIMA and ARMA models, a large number of historical observation for the variable of interest is required (Pao, Fu, & Tseng, 2012; Zhou et al., 2006). Other researchers have also employed the Grey Model (GM) prediction especially, GM (1, 1) to forecast carbon emissions (see Ding, Dang, Li, Wang, & Zhao, 2017; Lin, Liou, & Huang, 2011; Pao & Tsai, 2011; Majeed Safa, Nejat, Nuthall, & Greig, 2016; Wu, Liu, Liu, Fang, & Xu, 2015). Generally, GM performs best with limited data (Yin & Tang, 2013). However, the forecasting accuracy of the GM (1, 1) has been questioned (see Zhou et al., 2006). Additionally, comparative studies have shown that ANN produces superior forecasting results relative to ARIMA, ARMA, classical statistical and regression approaches (Falat & Pancikova, 2015; Kohzadi, Boyd, Kermanshahi, & Kaastra, 1996; Prybutok, Yi, & Mitchell, 2000; Stamenković, Antanasijević, Ristić, Perić-Grujić, & Pocajt, 2015; Valipour, Banihabib, & Behbahani, 2013).

The forecasting ability of ANN has made it received a widespread application in the field of engineering (Ahmadi, 2011; Ahmadi, Soleimani, Lee, Kashiwao, & Bahadori, 2015; Valipour et al., 2013), agriculture (Khoshroo, Emrouznejad, Ghaffarizadeh, Kasraei, & Omid, 2018; Majeed Safa et al., 2016; M. Safa & Samarasinghe, 2011; Soltanali, Nikkhah, & Rohani, 2017), energy (Deb, Zhang, Yang, Lee, & Shah, 2017; Debnath & Mourshed, 2018; Jebaraj & Iniyan, 2006), and finance (Kara, Acar Boyacioglu, & Baykan, 2011; Moghaddam, Moghaddam, & Esfandyari, 2016). One of the main advantages of ANN is its ability to use prior information to model a complex non-linear system, and its forecast results are robust since it can approximate non-linear input-output relationship to any degree of accuracy in an iterative manner (M. Safa & Samarasinghe, 2011; Sözen, 2009). Also, ANN can handle noisy data, accommodating multiple variables with non-linear, linear, and unknown interactions and make a good generalisation (Colwell, 1994; Hagan, Demuth, Beale, & De Jesús, 1996; M. Safa & Samarasinghe, 2011). Despite the forecasting ability of ANN, its application for forecasting carbon emissions intensity is still limited (see Zhao, et al., 2018). Therefore, this study utilised ANN to develop models for forecasting carbon emissions intensity for

Australia, Brazil, China, India, and USA. These countries are studied because they are among the top carbon-emitting countries.

In this direction, this study makes several distinct contributions to a new body of knowledge: Firstly, unlike the previous studies which have forecasted carbon emissions using only economic growth and population as input variables (see Pao et al., 2012; Pao & Tsai, 2011; Zhao & Du, 2015), this study incorporates other variables such as energy consumption, R&D, financial development, FDI, trade openness, industrialisation, and urbanisation, which play important role in contributing to carbon emissions, in our model to prevent underestimation of carbon emissions intensity. Secondly, unlike the previous forecasting studies on emissions, this study utilises the Partial Rank Correlation Coefficient (PRCC) to conduct sensitivity analysis to determine the input variable that is most influential in contributing to carbon emissions for the respective countries. Khoshroo et al. (2018), Marino, Hogue, Ray, and Kirschner (2008) and Saltelli and Marivoet (1990) argue that PRCC is the most reliable and efficient method for sensitivity analysis. Additionally, unlike previous studies, this study uses high-frequency data to provide accurate forecasting models. Finally, given that this study focuses on the major carbon-emitting countries, the models that will be developed from this study would help environmental planners in climate change policy decision-making.

The remaining sections are organised as follows. Section 2 provides a literature review while section 3 provides an overview of the research methodology and data, followed by results and discussions in section 4. Section 5 also presents the proposed closed-form formula for forecasting carbon emissions intensity while section 6 presents the sensitivity analysis. Conclusions and policy implications are presented in section 7.

**Nomenclature**

| | |
|---|---|
| AD | Absolute deviation |
| ANN | Artificial neural network |
| ARIMA | Autoregressive Integrated Moving Average |
| ARMA | Autoregressive Moving Average |
| $bias_k$ | biases of the *k*th hidden neuron |
| BP | back-propagation |
| CI | confidence interval |
| $CO_2$ | Carbon emission intensity |
| $e^{-y}$ | exponential function |
| ENER | Energy consumption |
| FD | Financial development |
| FDI | Foreign direct investment |
| FFMLP | feedforward multilayer perceptron |
| GDP | Gross Domestic Product |
| GM | Grey Model |
| Gt | gigatons |
| IMF | International Monetary Fund |
| INDUS | Industrialisation |
| MAD | mean absolute deviation |
| $\max(y)$ | maximum observation |
| $ME_y$ | mean error on prediction |
| $\min(y)$ | minimum observation |
| MLP | multi-layer perceptron |
| MSE | mean squared error |
| $MSE_{test}$ | mean squared error on test dataset |
| $MSE_{train}$ | mean squared error on training dataset |
| Mt | million tons |
| $N^h$ | number of neurons in the hidden layer |
| $N^i$ | number of input parameters |
| $N^o$ | number of output parameters |
| $N^{tr}$ | number of training samples |
| $O_1$ | Output |
| $P_{mean}$ | population mean |
| POP | Population |
| PRCC | Partial Rank Correlation Coefficient |
| R&D | Research and Development |
| $R^2$ | Coefficients of determination |
| ReLU | rectifier function |
| RNN | recurrent neural network |
| SD | standard deviation |
| SE | standard error |
| SVR | support vector regression |
| TRAD | Trade Openness |
| URB | Urbanisation |
| $x_{stand}$ | standardization |
| $y$ | observation for the output parameter |
| $\bar{y}$ | mean of the $CO_2$ intensities |
| $y_a$ | actual $CO_2$ intensity |
| $y_{norm}$ | normalization |
| $y_p$ | predicted $CO_2$ intensity |

*Greek letters*

| | |
|---|---|
| $\theta_r(x)$ | rectifier function |
| $\theta_s(y)$ | sigmoid function |
| $\mu$ | mean of the observations |
| σ | standard deviation of the observations |

*Subscript*

| | |
|---|---|
| $i$ | input data (0, 1, 2, 3, 4,….n) |
| $j$ | $j$th input parameter |

*Superscript*

| | |
|---|---|
| $bias_o$ | bias of the output neuron |
| $H_k$ | *k*th hidden neurons |
| $I_j$ | Input value of $j$th input parameter |
| $n$ | number of input data |
| $q$ | numbers of input parameters (closed-form formula) |
| $r$ | number of hidden neurons (closed-form formula) |
| $w_{j,k}^{ih}$ | weight of the link between $I_j$ and $H_k$ |
| $w_{k,1}^{ho}$ | weight of the link between $H_k$ and $O_1$ |

## 2. Determinants of carbon emissions

In this section, we provide a brief overview of the literature that supports the variables that were used as inputs for the modelling. Following the existing literature, we explore the impact of economic growth, energy consumption, population size, R&D, urbanisation, globalisation (FDI and trade openness) and industrialisation on carbon emissions. For clarity purpose, the literature is divided into the following segments: economic growth–carbon emissions nexus, energy consumption–carbon emissions nexus, population size–carbon emissions nexus, R&D–carbon emissions nexus, urbanisation–carbon emissions nexus, financial development–carbon emissions nexus, industrialisation–carbon emissions nexus and globalisation (FDI and trade openness)–carbon emissions nexus.

*2.1. Economic growth-carbon emissions nexus*

Economic growth is argued to be the primary force behind the persistent increase in global environmental pollution (carbon emissions). The nexus between economic growth and carbon emissions have been widely studied. Some scholars are of the view that economic growth has an adverse effect on the environment by increasing carbon emissions (Grove, 1992) while others contend that economic growth is necessary to improve the quality of the environment (Meadows, Randers, & Meadows, 1992). Majority of the studies on the nexus between economic growth and carbon emissions is much rooted in the Environmental Kuznets Curve (EKC) hypothesis[1]. The EKC hypothesis assumes that an inverted U-shaped relationship exists between economic growth and carbon emissions. Thus, at the early stages of economic growth, carbon emissions increase, but beyond a certain level of economic growth, carbon emissions reduces (Grossman & Helpman, 1991; Grossman & Krueger, 1995; Stern, 2004). Findings from the empirical studies on the impact of economic growth on

[1] For literature on the nexus between carbon emissions and economic growth see the extensive literature survey of (Dinda, 2004).

carbon emissions remain highly contentious. For instance, Ahmad et al. (2017) studied the nexus between economic growth and carbon emissions in Croatia using ARDL and the results reveal that inverted U-shape relation between carbon emissions and economic growth in the long run and this supports the EKC hypothesis. The Granger causality based on the VECM approach shows that bi-directional causality exists between carbon emissions and economic growth in the short run and unidirectional causality from economic growth to carbon emissions in long run.

For the case of Asia economies, Apergis and Ozturk (2015) employed the GMM to examine the nexus between economic growth and carbon emissions. The results confirm the validity of the EKC hypothesis. Similarly, Narayan and Narayan (2010) used panel cointegration technique to investigate the relationship between economic growth and carbon emissions for 43 developing countries and their findings confirm the EKC hypothesis in Middle Eastern and South Asian countries while the EKC hypothesis was not confirmed in Africa, East Asia and Latin America. Using GMM, Tamazian and Bhaskara Rao (2010) found that the EKC hypothesis exists in transitional economies. In another study, Tamazian, Chousa, and Vadlamannati (2009) found that economic growth degrades the environment by increasing carbon emissions. Additionally, Stern and Common (2001) and Stern (2004) found that carbon emissions monotonically increases with economic growth, which does not confirm the EKC hypothesis. Similarly, using GMM-PVAR, the empirical findings of Acheampong (2018) revealed that economic growth reduces carbon emissions at the global level and Caribbean-Latin America countries while it has an insignificant effect on carbon emissions for countries in sub-Saharan Africa, Asia-Pacific and the Middle East and North Africa (MENA). Using data from Malaysia, Saboori, Sulaiman, and Mohd (2012) revealed that the EKC hypothesis exists. In China, Liu and Bae (2018) found that economic growth increases carbon emissions.

*2.2. Energy consumption-carbon emissions nexus*

The Kaya identity shows that one of the key factors that influence the evolution of carbon emissions in the intensity of energy consumption (see Acheampong, 2018). Additionally, the IEA (2018) report further suggests that energy intensity is one of the two drivers of carbon emissions, the other being carbon intensity. While global carbon intensity declined less in 2017 than in 2016, the rate remains similar to the average rate of improvement in 2014-2016 – partly driven by the increasing expansion of renewables. However, the slower improvement in the energy intensity of energy demand in 2017 was not sufficient to counteract the effect of higher economic growth, leading to the increase in global energy-related carbon emissions in 2017 (IEA, 2018). The empirical literature suggests that energy consumption has an important impact on carbon emissions. For instance, using data from China, Zhang and Cheng (2009) reported that energy consumption increases carbon emissions. Similarly, Shahbaz, Hye, Tiwari, and Leitão (2013) reported that energy consumption increases carbon emissions in Indonesia. Also, using data from Kuwait, Salahuddin, Alam, Ozturk, and Sohag (2018), revealed that energy consumption increases carbon emissions. Similarly, using data from 14 MENA countries, Omri (2013) reported that energy consumption increases carbon emissions. Focusing on Bangladesh, Jahangir Alam, Ara Begum, Buysse, and Van Huylenbroeck (2012) reported that energy consumption increases carbon emissions. Similarly, Halicioglu (2009) reported that energy consumption stimulates carbon emissions in Turkey. The results of Begum, Sohag, Abdullah, and Jaafar (2015) also revealed that energy consumption stimulates carbon emissions in Malaysia. Using data from China, the empirical findings of Wu, Shen, Zhang, Skitmore, and Lu (2016) revealed that energy consumption increases carbon emissions.

*2.3. Population-carbon emissions nexus*

Population size, which refers to the total number of people living in a particular country, has an important effect on carbon emissions. In almost all climate models, population size is the only demographic variable considered (Zhu & Peng, 2012). In a classical study, Birdsall (1992) argue that population size affects carbon emissions through energy use and deforestation. Additionally, population size could influence the scale and structure of consumption and production, thereby increasing carbon emissions (Zhu et al., 2012). The empirical findings of Zhu et al. (2012) revealed that population size has no significant impact on carbon emissions but it is population structure, population age and household size that matter. Using data from 93 countries, the findings of Shi (2003) revealed that population size is proportionally related to the growth of carbon emissions. Using data from Europe, Weber and Sciubba (2018) reported that the population growth rate has a considerable effect on carbon emissions is Western Europe but has a negligible effect on carbon emissions in Eastern Europe. Focusing on Malaysia, the findings of Begum et al. (2015) revealed that the population growth rate has no effect on carbon emissions. Using 128 countries, Dong et al. (2018) reported that population size contributes significantly to the growth of carbon emissions.

### *2.4. R&D-carbon emissions nexus*

Technological innovation is another important variable that influences the environment (carbon emissions). Technological innovation could be helpful in switching to more sustainable sources of energy including renewables which could reduce carbon emissions (Shahbaz, Nasir, & Roubaud, 2018). Shahbaz et al. (2018) further argue that innovation related to energy is more prone to influence energy consumption and, hence, carbon emissions, specifically energy innovations which are intuitively more relevant and important for environmental quality. However, investment in research and development (R&D) is crucial for facilitating the promotion of technological progress, which could lead to greater efficiency in energy and the use of natural resources, thereby reducing carbon emissions (Churchill, Inekwe, Smyth, & Zhang, 2019). On the other hand, the positive effect of R&D on economic growth and trade could increase carbon emissions through the scale effect of larger production associated with economic growth and trade liberalisation (Churchill et al., 2019). In an empirical study, Tamazian et al. (2009) reported that R&D contributes to the mitigation of carbon emissions in BRICS. The study of Churchill et al. (2019) also revealed that R&D reduces carbon emissions in G7 countries. Using data from France, Shahbaz et al. (2018) reported that R&D contributes to the reduction of carbon emissions. Fernández-Fernández et al. (2018) also used data from the European Union (15), the United States and China; their results revealed that R&D contributes to the reduction of carbon emissions. Using data from China, Zhang, Peng, Ma, and Shen (2017), R&D is conducive for reducing carbon emission. On the other hand, Jiao, Jiang, and Yang (2018) found that R&D generally increases carbon emissions; however, considering regional effects, R&D reduces carbon emissions.

*2.5. Urbanisation-carbon emissions nexus*

Urbanization may also play important role in the evolution of carbon emissions. The theoretical linkage between urbanization and environmental quality has been discussed in the work of (Poumanyvong & Kaneko, 2010) and (Sadorsky, 2014). Ecological modernization theory, urban environmental transition theory and compact city theory are the major theories for explaining the impact of urbanization on the environment (Poumanyvong et al., 2010; Sadorsky, 2014). The ecological modernization theory discusses the impact of urbanization on the environment at the national level while the latter theories focus the impact at the city level (Poumanyvong et al., 2010). The ecological modernization theory argues that environmental problems increase as society becomes modernize and thereafter, seek to address environmental problems at the advanced stage of economic development. Thus, as a society makes transition from low-level of economic development to an intermediate level of development, environmental problems increases; however, at the advanced stage of economic development with efficient technology, urban agglomeration and knowledge spillover effect, societies seek to minimize environmental problems such as mitigating carbon emissions (Gouldson & Murphy, 1997; Mol & Spaargaren, 2000; Poumanyvong et al., 2010).

Similar to the ecological modernization theory, the urban environmental transition theory also argues that environmental problems differ across different stages of economic development at the city level (McGranahan, 2010). Thus, as cities become prosperous by increasing production, environmental problems also increase; however, as cities become wealthier or at the advanced stage of development, environmental problems reduce as results of improvement in environmental regulation, technological progress and structural change in the economy (Poumanyvong et al., 2010; Sadorsky, 2014). As ecological modernization and the urban environmental transition theories argue for both negative and positive effect of urbanization on the environment, the net effect of urbanization on the environment is indeterminate (Sadorsky, 2014). On the other hand, the compact city theory focuses on the positive externality of urbanization on the environment. Thus, rapid urbanization help cities to facilitate economies of scale for urban infrastructure and these economies of scale reduces environmental pollution (Poumanyvong et al., 2010). Thus, high urban density helps to reduce travel distance, car dependency, energy consumption and carbon emissions (Burton, 2000; Capello & Camagni, 2000). However, some scholar argues that increasing urbanization could result in traffic congestions and overcrowding which will consequently increase energy consumption and carbon emissions (Breheny, 2001; Poumanyvong et al., 2010; Rudlin & Falk, 1999). Empirically, Poumanyvong et al. (2010) found that urbanization increases carbon emission. Using data from emerging countries, Sadorsky (2014) found that urbanization could either increase or reduce carbon emissions depending on the estimator.

Using data from China, Wang, et al. (2016) reported that urbanization contributes to the growth of carbon emissions. Similarly, Zhang and Lin (2012) reported that urbanization increases carbon emissions in China. Additionally, using data from China, the findings of Wu et al. (2016) revealed that urbanization increases carbon emissions. Using data from 23 European countries, Al-Mulali, Ozturk, and Lean (2015) found that urbanisation increases carbon emissions. Liu and Bae (2018) further found that urbanisation contributes to the increased carbon emissions in China. Contrarily, Bekhet and Othman (2017) found that urbanization contributes to carbon emissions. They further found that an inverted U-shaped relationship exists between urbanization and carbon emissions.

*2.6. Financial development-carbon emissions nexus*

Recently, research on the nexus between financial development and carbon emissions has received interest among energy and environmental economists. It is argued that financial development could reduce carbon emissions, as it attracts foreign direct investment and further promotes research and development, which in turn enhance the quality of the environment (Tamazian et al., 2009). On the other hand, financial development could worsen the quality of the environment by increasing carbon emissions. (Sadorsky, 2010, 2011) argues that a developed financial system makes it easy for economic agents to have access to cheap credits to purchase big-ticket items and expand their existing plants, which increase energy consumption, thereby increasing carbon emissions. While it is argued theoretically that the impact of financial development could either improve or worsens the environment, the empirical findings remain ambiguous. For instance, one category of empirical findings report that financial development reduces carbon emissions (see Al-Mulali, Tang, & Ozturk, 2015; Tamazian & Bhaskara Rao, 2010; Tamazian et al., 2009) while the second category report that financial development simulates the growth of carbon emissions (see Boutabba, 2014; Sehrawat, Giri, & Mohapatra, 2015; Shahbaz, Shahzad, Ahmad, & Alam, 2016). The third category of the empirical literature also suggests that financial development has no relationship with carbon emissions (see Dogan & Turkekul, 2016; Maji, Habibullah, & Saari, 2017; Omri, Daly, Rault, & Chaibi, 2015).

*2.7. Industrialisation-carbon emissions nexus*

Industrialisation, which is a critical path to economic and social modernization, has a significant impact on the environment. Industrialization refers to an increase in industrial activity, and that rapid industrialization leads to higher energy usage because higher value-added manufacturing uses more energy than does traditional agriculture or basic manufacturing (Sadorsky, 2013). In other words, industrialisation promotes the rapid growth of fossil fuel consumption and produces significant amounts of carbon dioxide and other greenhouse gas emissions (Li & Lin, 2015). Using data from China, Wang, Shi, Li, and Wang (2011) reported that industrialisation increases carbon emissions. Similarly, Liu and Bae (2018) found that industrialisation increases the intensity of carbon emissions in China. Using data from MENA countries, the empirical results of Al-Mulali and Ozturk (2015) revealed that industrialisation contributes to the increase in carbon emissions. Using data from China, Zhou, Zhang, and Li (2013) found that industrialisation reduces carbon emissions. The empirical results of Li and Lin (2015) revealed that across all income groups, industrialisation fuel the growth of carbon emissions.

*2.8. Globalisation-carbon emissions nexus*

The role of foreign direct investment (FDI) and trade openness on the environment has been highly debated in the literature. Trade openness impact on the environment through the scale effect, technique effect and composition effect (Antweiler, Copeland, & Taylor, 2001; Ghani, 2012). The scale effect of trade openness on the environment occurs through the growth of the economy. Thus, the scale effect suggests that trade openness facilitate economic growth which in turn result in higher carbon emissions. Additionally, the technique effect suggests that trade openness promotes the transfer of environmentally friendly technologies which could result in reducing carbon emissions. According to the composition effect, trade openness could affect the environment by changing the structure of the economy. In addition, FDI could improve or worsen environmental quality. Studies on the impact of FDI on the environment (carbon emissions) are deeply rooted in the Pollution-haven hypothesis, Pollution-halo hypothesis and scale effect hypothesis (Pao & Tsai, 2011). According to the pollution haven, FDI degrades the quality of the environment by increasing carbon emissions. Thus, weak environmental regulation in a host country could attract the inflow of FDI by multinational companies that are pollution intensive, thereby increasing carbon emissions (Shahbaz et al., 2018). Like the pollution-haven hypothesis, the scale effect hypothesis also suggests that the inflow of FDI could contribute significantly to a host countries' economic output, which in turn, increase carbon emissions (Pao & Tsai, 2011; Shahbaz et al., 2018). The pollution-halo effect hypothesis also suggests that FDI could reduce carbon emissions by increasing the spread the environmentally friendly technologies.

Empirically, Acheampong (2018) found that trade openness decreases carbon emissions at the global level, Asia-Pacific, MENA and Sub-Saharan Africa countries. Similarly, Shahbaz, Kumar Tiwari, and Nasir (2013) found that trade openness improves environmental quality by reducing carbon emissions in South Africa. Antweiler et al. (2001) further reported that trade is important for improving the quality of the environment by reducing carbon emissions. Contrarily, Ren, Yuan, Ma, and Chen (2014) found that trade increases carbon emissions in China. Using data from five South Asian countries, Ahmed, Rehman, and Ozturk (2017) found that trade openness increases carbon emissions. Focusing on the impact of FDI on the environment, Shahbaz, Nasreen, Abbas, and Anis (2015) found that at the global level, FDI increases carbon emissions. However, they concluded that the impact of FDI on carbon emissions is sensitive to income and regional groups. In France, Shahbaz et al. (2018) found that FDI increases carbon emissions. Similarly, Ren et al. (2014) found that FDI increases carbon emissions in China. Contrarily, using 19 of the G20 countries, Lee (2013) reported that FDI contributes to the reduction in carbon emissions.

## 3. Data and methodology

### *3.1. Dataset*

The study used time series data which spans between 1980-2015. However, to develop accurate models, the study follows Sbia, Shahbaz, and Hamdi (2014) and Shahbaz, Hoang, Mahalik, and Roubaud (2017) to use quadratic-sum approach to convert the annual data from low-frequency data to high-frequency data. Therefore, quarterly data between 1980Q1-2015Q4 was used for the study. This period represented 144 quarters. Table 1 presents the proxies for the variables and the justification for selecting the input variables used for the modelling. In selecting the input variables, the study follows the literature on carbon emissions to select the fundamental variables that influence carbon emissions intensity. Except for financial development, all the remaining variables were sourced from World Bank (2016). The financial development index was obtained from the International Monetary Fund (IMF)[2]. Table 2 also presents the descriptive statistics for variables.

[2] http://data.imf.org/?sk=F8032E80-B36C-43B1-AC26-493C5B1CD33B

Table 1. Variables for the study.

| Variable | Code | Proxies | Definitions | Reference |
|---|---|---|---|---|
| Carbon emissions intensity | *$CO_2$* | $CO_2$ intensity (kg per kg of oil equivalent energy use) | Carbon emissions intensity is the volume of carbon emissions due to economic activity/economic growth. It is also defined as carbon emissions emitted per unit of energy consumed. | |
| Energy consumption | *ENER* | Energy use (kg of oil equivalent per capita) | Energy use refers to use of primary energy before transformation to other end-use fuels. | Destek and Sarkodie (2019); Sarkodie and Strezov (2018). |
| Financial development | *FD* | The financial development index is a broad-based measure which comprises bank-based and market-based indicators of financial development. | Financial development refers to the increased flow of foreign direct investment, banking and stock market activities. | Shahbaz, Kumar Tiwari, and Nasir, (2013); Tamazian and Bhaskara Rao, (2010); Tamazian et al. (2009). |
| Foreign direct investment | *FDI* | Foreign direct investment, net inflows (% of GDP) | Foreign direct investment is the net inflows of investment to acquire a lasting management interest in an enterprise operating in an economy other than that of the investor. | Ren, Yuan, Ma and Chen (2014); Sarkodie et al. (2019); Zhang and Zhou (2016). |
| Economic growth | *GDP* | GDP per capita (constant 2010 US$) | GDP per capita is gross domestic product divided by midyear population. It is the sum of gross value added by all resident producers in the economy plus any product taxes and minus any subsidies not included in the value of the products. | Ben Jebli, Ben Youssef and Ozturk, (2016); Grossman and Krueger, (1995); Saboori, Sulaiman and Mohd (2012). |
| Industrialisation | *INDUS* | Industry, value added (% of GDP) | Industrialization refers to an increase in industrial activity. It comprises value added in mining, manufacturing, construction, electricity, water, and gas. | Wang, Shi, Li and Wang (2011); |
| R&D | *R&D* | Trademark applications, total | The R&D covers basic research, applied research, and experimental development. | Jiao, Jiang and Yang (2018); Shahbaz, Nasir and Roubaud, (2018). |
| Population | *POP* | Population, total | Total population refers to the total number of people living in a particular geographical area. It is based on the de facto definition of population, which counts all residents regardless of legal status or citizenship. | Zhu and Peng (2012) |
| Trade Openness | *TRAD* | Trade (% of GDP) | Trade is the sum of exports and imports of goods and services measured as a share of the gross domestic product. | Acheampong (2018); Ren et al., (2014) |
| Urbanisation | *URB* | Urban population (% of total) | Urban population refers to people living in urban areas as defined by national statistical offices. | Poumanyvong and Kaneko (2010); Sadorsky (2014). |

Table 2. Descriptive statistics.

| | Count | Mean | SD | Min | Max |
|---|---|---|---|---|---|
| **Australia** | | | | | |
| ENER | 144 | 5268.8320 | 423.9819 | 4533.5690 | 5971.2290 |
| FD | 144 | 0.6808 | 0.2168 | 0.2731 | 0.9657 |
| FDI | 144 | 2.4624 | 1.7497 | -4.3786 | 7.8503 |
| GDP | 144 | 41787.9100 | 8398.9680 | 29725.5000 | 55179.3500 |
| INDUS | 144 | 25.5816 | 1.1295 | 22.3563 | 29.3606 |
| R&D | 144 | 37201.0600 | 18086.5500 | 12935.0600 | 73741.6600 |
| POP | 144 | 18800000.0000 | 2610000.0000 | 14600000.0000 | 24000000.0000 |
| TRAD | 144 | 37.3866 | 4.7907 | 28.1058 | 46.2559 |
| URB | 144 | 16400000.0000 | 2540000.0000 | 12500000.0000 | 21500000.0000 |
| $CO_2$ | 144 | 3.1120 | 0.1100 | 2.8758 | 3.3627 |
| **Brazil** | | | | | |
| ENER | 144 | 1084.3660 | 166.5520 | 870.0903 | 1494.8180 |
| FD | 144 | 0.4030 | 0.1575 | 0.1713 | 0.6273 |
| FDI | 144 | 2.0784 | 1.4981 | 0.0883 | 5.1136 |
| GDP | 144 | 9109.4750 | 1362.5440 | 7192.9950 | 11961.0800 |
| INDUS | 144 | 29.1012 | 8.1757 | 18.8637 | 43.1403 |
| R&D | 144 | 88282.5300 | 37664.9800 | 27442.0600 | 164319.7000 |
| POP | 144 | 167000000.0000 | 25600000.0000 | 120000000.0000 | 207000000.0000 |
| TRAD | 144 | 21.2789 | 4.5309 | 14.2097 | 29.8731 |
| URB | 144 | 132000000.0000 | 29700000.0000 | 78200000.0000 | 177000000.0000 |
| $CO_2$ | 144 | 1.5872 | 0.0987 | 1.3841 | 1.7763 |
| **China** | | | | | |
| ENER | 144 | 1103.4160 | 518.0744 | 596.2954 | 2238.4880 |
| FD | 144 | 0.4056 | 0.1301 | -0.0559 | 0.6458 |
| FDI | 144 | 2.9237 | 1.6337 | 0.2046 | 6.7016 |
| GDP | 144 | 2161.0150 | 1822.3990 | 345.7698 | 6642.6710 |
| INDUS | 144 | 44.9490 | 1.9043 | 39.8404 | 49.0031 |
| R&D | 144 | 430376.7000 | 543678.5000 | 18218.9100 | 2206486.0000 |
| POP | 144 | 1210000000.0000 | 120000000.0000 | 977000000.0000 | 1370000000.0000 |
| TRAD | 144 | 36.8526 | 14.2636 | 12.0060 | 64.6305 |
| URB | 144 | 437000000.0000 | 174000000.0000 | 186000000.0000 | 770000000.0000 |
| $CO_2$ | 144 | 3.0654 | 0.2692 | 2.4261 | 3.4830 |
| **India** | | | | | |
| ENER | 144 | 414.3776 | 95.8066 | 283.4874 | 654.1635 |
| FD | 144 | 0.3230 | 0.0974 | 0.1857 | 0.4696 |
| FDI | 144 | 0.8444 | 0.9051 | -0.0012 | 3.7726 |

| | | | | | |
|---|---|---|---|---|---|
| GDP | 144 | 818.1663 | 392.8301 | 382.6152 | 1805.4140 |
| INDUS | 144 | 28.7170 | 1.4005 | 25.9435 | 31.7937 |
| R&D | 144 | 80968.1400 | 70535.7400 | 14350.7800 | 300555.9000 |
| POP | 144 | 1010000000.0000 | 186000000.0000 | 691000000.0000 | 1310000000.0000 |
| TRAD | 144 | 28.9482 | 14.8495 | 12.2991 | 56.5884 |
| URB | 144 | 282000000.0000 | 79600000.0000 | 159000000.0000 | 432000000.0000 |
| $CO_2$ | 144 | 2.1817 | 0.3280 | 1.5424 | 2.7770 |
| **USA** | | | | | |
| ENER | 144 | 7562.8570 | 351.4622 | 6692.1100 | 8074.1540 |
| FD | 144 | 0.7201 | 0.1951 | 0.2862 | 0.8938 |
| FDI | 144 | 1.2952 | 0.7862 | 0.2494 | 3.7214 |
| GDP | 144 | 41135.7900 | 7427.4450 | 28281.2400 | 52366.9900 |
| INDUS | 144 | 20.9829 | 0.7909 | 18.9394 | 23.4137 |
| R&D | 144 | 195258.0000 | 99127.9600 | 45927.1300 | 389988.5000 |
| POP | 144 | 274000000.0000 | 29400000.0000 | 226000000.0000 | 322000000.0000 |
| TRAD | 144 | 22.9258 | 4.3746 | 16.4905 | 31.2702 |
| URB | 144 | 214000000.0000 | 30000000.0000 | 167000000.0000 | 263000000.0000 |
| $CO_2$ | 144 | 2.4977 | 0.0563 | 2.3631 | 2.6313 |

## *3.2. Methodology*

### *3.2.1. Artificial neural networks*

The study aims to develop models for predicting/forecasting carbon emissions intensity for high carbon-emitting countries such as Australia, Brazil, China, India, and USA. ANN is employed and incorporated in the proposed theoretical framework of the model as shown in Fig. 1a. The diagram (Fig. 1a) generally depicts the possible relationship connecting nine (9) determinants (inputs) of $CO_2$ emission intensity and $CO_2$ emission intensity (output) for the selected countries.

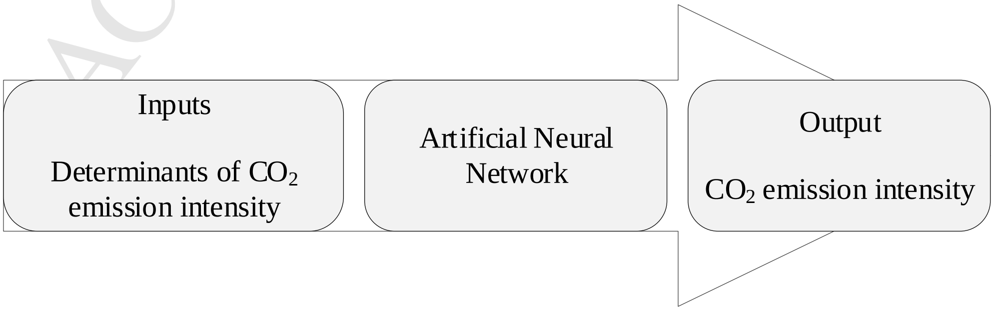



Fig. 1a. A predictive model of $CO_2$ emission intensity.

Artificial neural networks (ANNs) are data processing systems that mimic the way data is processed in the human brain (Boateng, Pillay, & Davis, 2019). An ANN is made up of input, hidden, and output layers which consist of numerous processing components called neurons (Boateng et al., 2019). The neurons process the data and feed forward to the subsequent layer. These neurons are connected by corresponding links between layers. On each connected link is a numeric weight. ANNs can automatically adjust their weights to enhance their behaviour, unlike statistical models (Boussabaine, 1996). The approach adopted in ANN does not require prior expertise in computer programming to develop and compute solutions as required in other numerical solutions (Ghritlahre & Prasad, 2018). A problematic issue in statistical model development is multicollinearity, i.e., the high degree of correlation among independent variables, which is much better dealt with in ANN because the assumption of independent variables being uncorrelated is not made (Detienne, Detienne, & Joshi, 2003).

Moreover, statistical tools cannot deal effectively with nonlinearity while ANNs are inherently nonlinear nonparametric models that can deal with indefinable nonlinearity in a straightforward manner (Detienne et al., 2003). Also, ANNs are especially suitable to find solutions for problems that have fuzzy information and are highly complex where individuals usually make decisions on an intuitional basis (Ghritlahre & Prasad, 2018). Besides, unlike most statistical approaches, ANNs do not need predefined mathematical equations of the relationship between the model inputs and corresponding outputs (Shahin & Elchalakani, 2008). These enable ANNs to overcome the limitations of existing modelling methods. Despite the differences between ANNs and statistical approaches, both techniques can be combined into a solid and powerful methodological platform (Karlaftis & Vlahogianni, 2011). This is because ANN is like a 'black box' and hence lacks self-explanation. As expressed by Alaka et al. (2018 p. 173), 'the more accurate the tool, the less transparent the result.' Consequently, statistical approaches such as descriptive statistics are often incorporated to produce explanatory results that can easily be interpreted and understood.

ANN has become a popular and useful tool for modelling accurate predictions to solve complex and nonlinear problems in diverse industrial domains. Many researchers have used ANN in the field of energy utilization and conversion systems for performance predictions (Kalogirou, 2000), solar radiations predictions (Yadav & Chandel, 2014), length of stay predictions on post-coronary care units (Mobley, Leasure, & Davidson, 1995), bankruptcy predictions (Adnan Aziz & Dar, 2006), and performance prediction of solid desiccant dehumidifier cooling methods (Jani, Mishra, & Sahoo, 2017). However, the application of ANN in environmental economics is quite rare.

### *3.2.2. Optimal model selection*

The performance of a neural network model primarily depends on the architecture of the network and the tuning of various parameters. Fig. 1b. illustrates a robust process used in selecting the optimal predictive models for the five countries. Details of the whole process are explained in the proceeding sections.

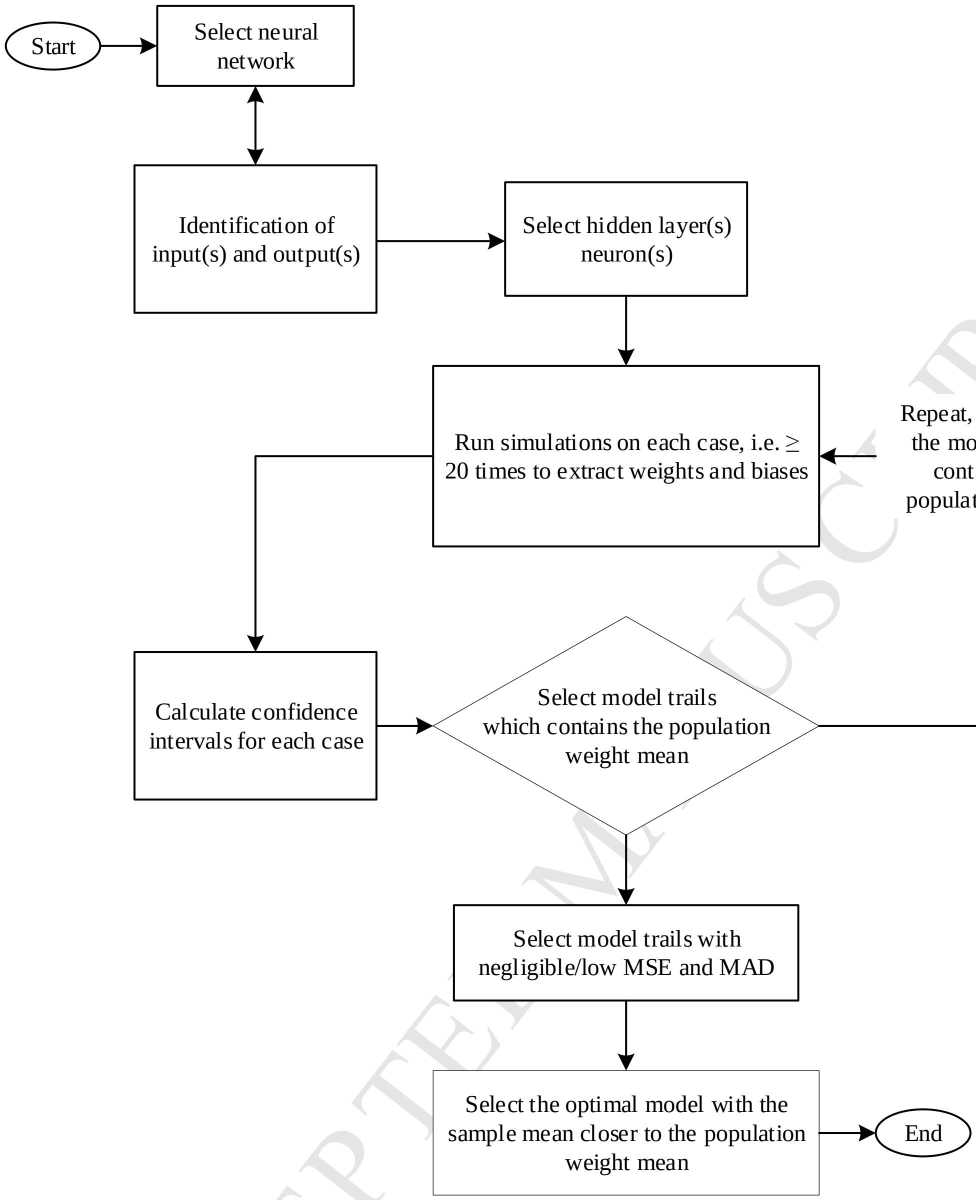



Fig. 1b. Flowchart of optimal model selection.
Source: Authors' construct.

*3.2.3. Development of ANN models*

Due to the severe computations on the high dimensional data when training the ANN, features of the data set are scaled using standardization ($x_{\mathrm{stand}}$) and normalization ($y_{\mathrm{norm}}$) on the inputs and output data respectively (Boateng et al., 2019). In normalization, the observations range from $0 \leq y_{norm} \leq 1$ to increase the rate of training the network. Hence, the output variable ($CO_2$ intensity) is normalized for each country, as it further facilitates the use of the sigmoid function for the output layer. Normalization is expressed in Eq. (1):

$$y_{norm} = \frac{y - \min(y)}{\max(y) - \min(y)} \tag{1}$$

Where $y$ is the observation for the parameter, min($y$) and max(y) is the minimum and maximum observations respectively. Standardisation is expressed in Eq. (2):

$$x_{stand} = \frac{x - \mu}{\sigma} \tag{2}$$

Where $\mu$ is the mean of the observations for the parameter and σ is the standard deviation of the observations for the parameter. Standardisation tends to centre the input values towards zero (0). Standardising the input data into a lesser array of variability would likely aid the effective learning of the neural network while improving the numerical state of the optimisation problem (StatSoft Inc., 2008). Thus, scaling the data eliminates any instances of one variable dominating the other (Boateng et al., 2019). The 144 observations for individual countries (Australia, Brazil, China, India, and USA) were randomly split into training and test sets to ensure the reliability of the results over time. 80% (115 observations) of the data set were used for training the model, while the remaining 20% (29 observations) were used to validate the model. Similar data ratio has been commonly used in previous studies (see Abidoye & Chan, 2018; Lam, Yu, & Lam, 2008; Morano, Tajani, & Torre, 2015).

Afterwards, a multi-layer perceptron (MLP) with back-propagation (BP) is selected to achieve the optimal performance of the ANN model. MLP is the most commonly used neural network (Ghaedi & Vafaei, 2017; Pérez-Sánchez, Fontenla-Romero, & Guijarro-Berdiñas, 2016). The BP algorithm was selected as the learning algorithm. The BP algorithm is often used to iteratively minimize the cost function concerning the interconnection weight and neurons thresholds (El Kadi, 2006; Kartalopoulos & Kartakapoulos, 1997). Therefore, the MLP with a BP algorithm can approximate any continuous function to meet the desired accuracy (Patel & Jha, 2015). The sigmoid function, which ranges from 0 to 1, was selected as the activation function for the output layer while the rectifier function (ReLU) was used as the activation function for the hidden layer to perform efficient computations. Currently, the ReLU is the most popular activation function for deep neural networks (LeCun, Bengio, & Hinton, 2015) while the sigmoid function is the most popular activation function for ANNs (Alvanitopoulos, Andreadis, & Elenas, 2010). The sigmoid and rectifier functions are defined as Eqs. (3) and (4) respectively;

$$\theta_s(y) = \frac{1}{1+e^{-y}} \tag{3}$$

$$\theta_r(x) = max(0, x) \tag{4}$$

Where $\theta_s(y)$ is the sigmoid function, $\theta_r(x)$ is the rectifier function, and $e^{-y}$ is the exponential function.

## 4. Results and discussion

### *4.1. Training of models*

In training the neural network, the selection of the hidden layer neurons is crucial to the performance of the model (Boateng et al., 2019). The optimum number of hidden layer neurons generally has to be found using a trial and error approach (Maier & Dandy, 2001). However, some general guidelines may be followed. Hecht-Nielsen (1987) suggests the following upper limit for the number of hidden layer nodes in order to ensure that the neural network can approximate any continuous function. The upper limit of the number of hidden layer nodes is calculated using Eq. (5):

$$N^h \leq 2N^i + 1 \tag{5}$$

Where $N^h$ is the number of neurons in the hidden layer and $N^i$ is the number of input parameters. For this present study, with nine (9) input parameters, the upper limit of the number of neurons in the hidden layer is given in Eq. (6) and (7):

$$N^h \leq 2(9) + 1 \tag{6}$$

$$N^h \leq 19$$

(7)

From this, the number of hidden layer neurons should not be more than 19. However, in order to ensure that the network does not overfit the training data, the relationship between the number of training samples and network size also needs to be considered (Maier & Dandy, 2001). Overfitting is where the model performs well on the training data but poorly on the test/validation data, and underfitting is where the model performs well on the test data and poorly on the training data. Rogers and Dowla (1994) recommend the following upper limit for the number of hidden layer nodes to satisfy the above criteria using Eq. (8):

$$N^h \leq \frac{N^{tr}}{N^i+1} \tag{8}$$

Where $N^{tr}$ is the number of training samples. Consequently, the upper limit for the number of hidden layer neurons may be taken as the smaller of the values for $N^h$ obtained from the two formulas. For this present study, with 115 training samples, the upper limit of the number of hidden layer is given in Eq. (9) and (10);

$$N^h \leq \frac{115}{9+1} \tag{9}$$

$$N^h \leq 11.5{\sim}12 \tag{10}$$

From this, the number of neurons in the hidden layer should not be more than 12. Consideration should be given to the selection of the hidden layer neurons since it affects the architecture of the network as well as the accuracy. If the network architecture is too complex, overfitting may occur, and if the architecture is too simple, the preferred estimate skill may not be achieved (Hippert, Pedreira, & Souza, 2001). From experimentations and practice, the problem could be argued as trivial. Therefore, the hidden layer neurons required in an MLP with a single layer could further be determined using a simplified formula, Eq. (11):

$$N^h = \frac{N^i + N^o}{2} \tag{11}$$

$$N^h = \frac{9+1}{2}. \tag{12}$$

$$N^h = 5 \text{ Hidden layer neurons.} \tag{13}$$

Where $N^o$ is the number of output parameters. From the computation, five (5) neurons in the hidden layer were deemed optimal in configuring the neural networks for this study. As a result, a 9-5-1 MLP with BP was sufficient to perform the necessary predictive capabilities with a minimal/negligible error. Fig. 2 illustrates the configuration of the three-layer feed-forward MLP.

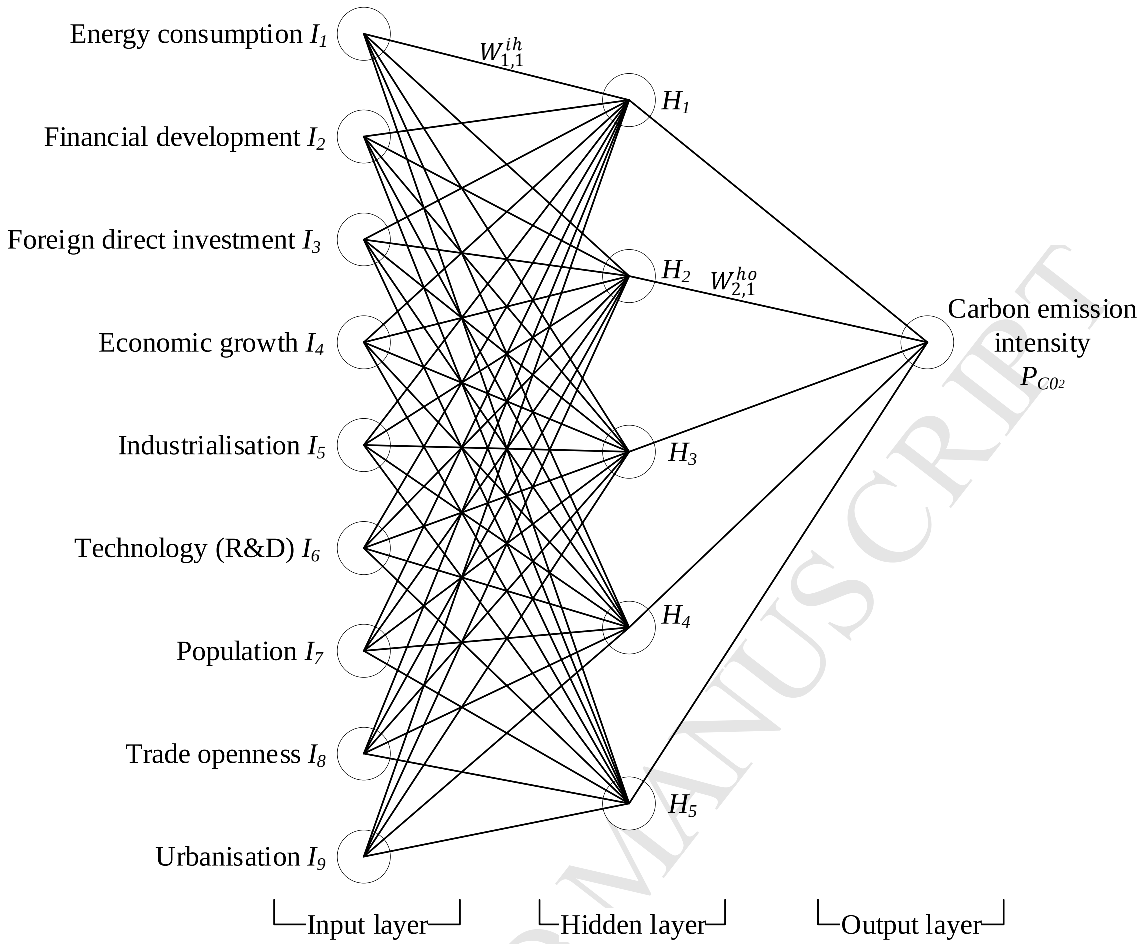



Fig. 2. Configuration of the developed neural network (9-5-1).

At this stage, a stochastic gradient descent batching is applied to the entire neural network to find the optimal weights. The 9-5-1 MLP for each country is trained a number of times to update its weights after every 5 observations. The stochastic gradient descent is initialized to improve the accuracy and minimize the loss over the various rounds (Boateng et al., 2019). During the training of the neural network, there is some randomness involved because at the start of training the weights would be randomly initialized. This sort of randomization results in different results at various iterations. To ensure stable results (reproducibility) each time the weights are initialized, a robust approach is employed by conducting repeated evaluation experiments (Boateng et al., 2019). In this approach, each case is run at least 20 times with different random weights at the start and then the mean is taken to calculate confidence intervals (CIs). The accuracies, means, standard deviations (SDs), standard errors (SEs), and intervals are evaluated to estimate the skill of the stochastic model at a 95% confidence interval while simultaneously checking the mean squared errors (MSEs) on both the training and test sets. The MSE metric is used since it is very closely related to the forecast accuracy. The MSEs are determined using Eq. (14):

$$MSE = \sum_{i=1}^{n}(y_a - y_p)^2 \quad (14)$$

Where n is the number of input data (i = 0, 1, 2, 3, 4,….n), $y_a$ and $y_p$ are the actual and predicted $CO_2$ intensities respectively. For each country, the computed MSE on both the training and test sets are presented in Appendix Table Ia - Ie. Coefficient of determination ($R^2$) was further computed for each trial on the actual (independent) and predicted (dependent) $CO_2$ intensities for each country. The $R^2$ denotes the ratio of the change in the output parameter that is predictable from the input parameter. The $R^2$ coefficient ranges from 0 to 1, and a coefficient close to 1 depicts an excellent performance. $R^2$ is determined using Eq. (15):

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_a - y_p)^2}{\sum_{i=1}^{n}(y_a - \bar{y})^2} \quad (15)$$

Where $\bar{y}$ is the mean of the $CO_2$ intensities. $R^2$ for each trial in each country is presented in Appendix Table Ia - Ie. After 20 simulated model trials for each country, the CIs are presented in Fig. 3a – 3e. For Australia, except the 10$^{th}$ and 16$^{th}$ runs, the remaining runs contain the population weights mean (see Fig. 3a) while for Brazil, except the 3$^{rd}$ and 4$^{th}$ runs, the remaining runs contain the population weights mean (see Fig. 3b). For China, except the 1$^{st}$, 13$^{th}$, and 19$^{th}$ runs, the remaining runs contain the population weights mean (see Fig. 3c). For India, except the 1$^{st}$, 2$^{nd}$, 3$^{rd}$, and 13$^{th}$ runs, the remaining runs contain the population weights mean (see Fig. 3d). Lastly, for USA, except the 3rd and 4th runs, the remaining runs contain the population weights mean (see Fig. 3e). By eliminating the runs that could not contain the population weights mean, random effects are avoided to ensure reproducibility of results.

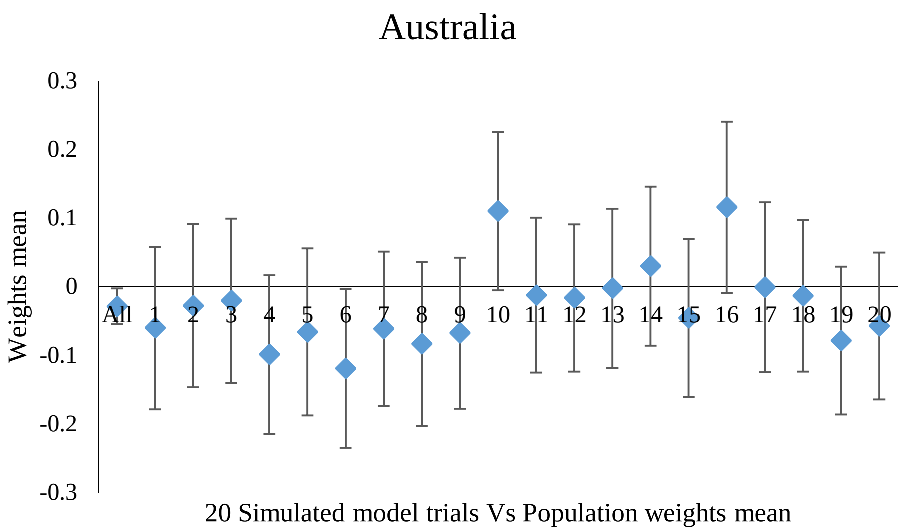



Fig. 3a. Point and interval estimates for 20 trials.

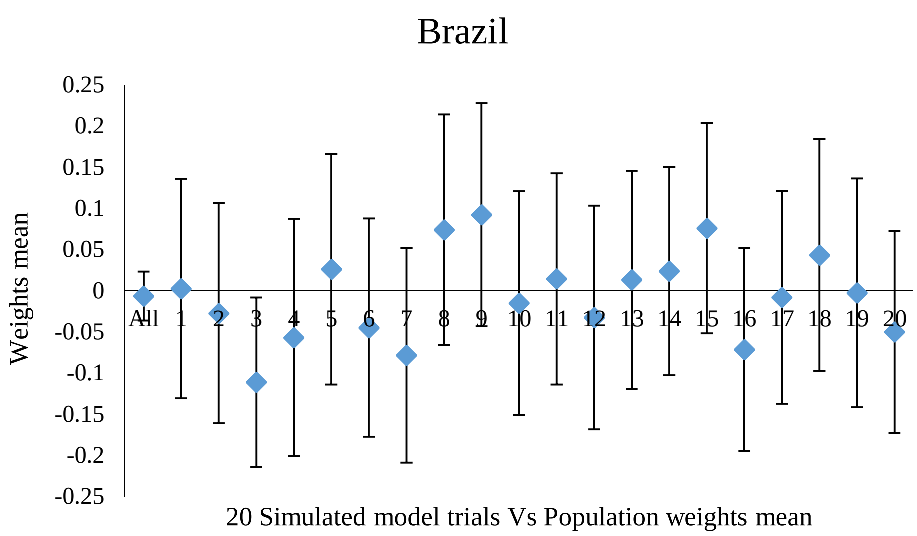



Fig. 3b. Point and interval estimates for 20 trials.

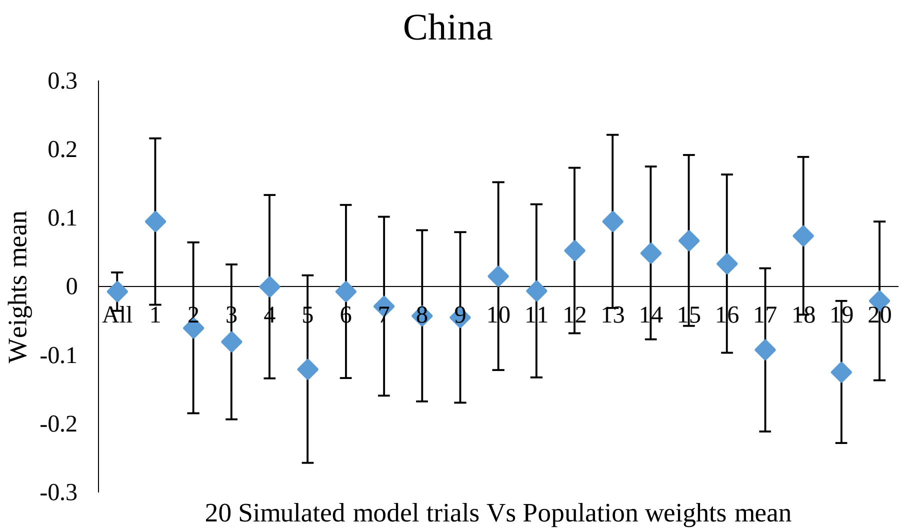



Fig. 3c. Point and interval estimates for 20 trials.

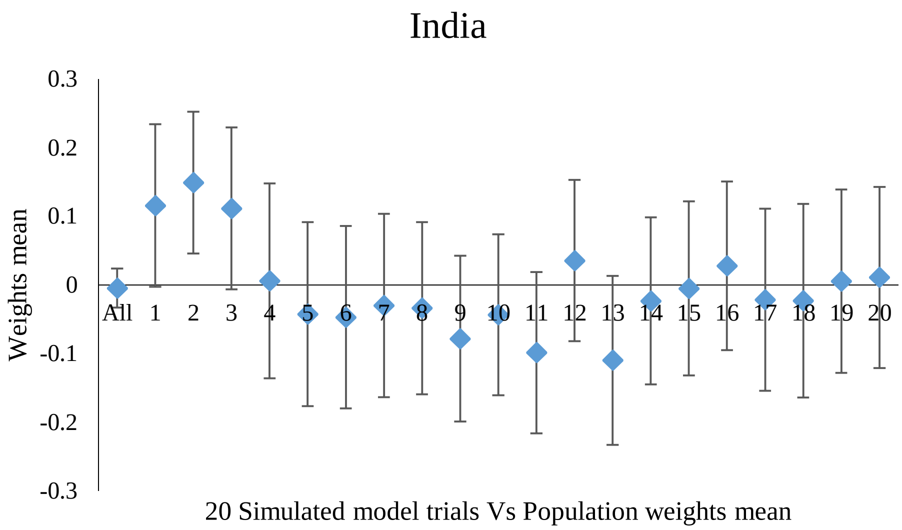



Fig. 3d. Point and interval estimates for 20 trials.

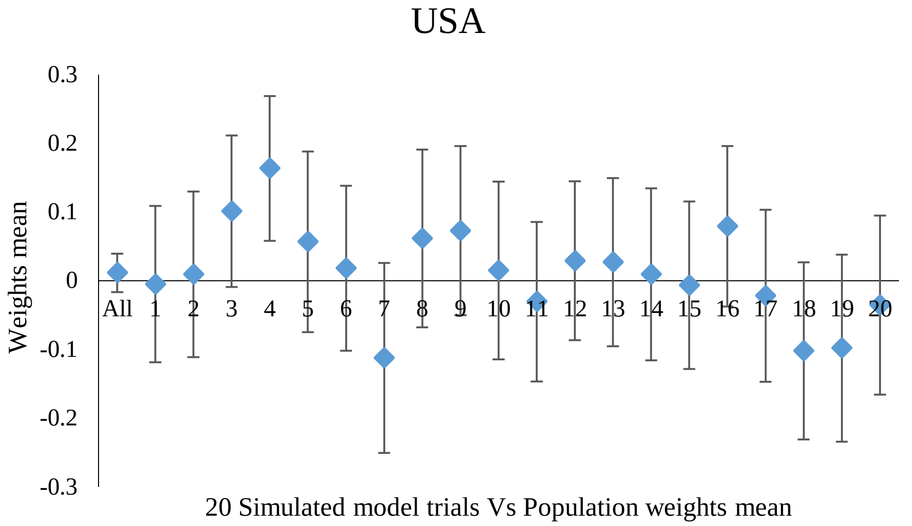



Fig. 3e. Point and interval estimates for 20 trials.

The $CO_2$ intensities resulted from the various iterations for each country is denormalized to obtain the original and predicted $CO_2$ intensities. Denormalisation is computed using Eq. (16);

$$y = y_{norm}\big(\max(y) - \min(y)\big) + \min(y) \qquad (16)$$

Further, for all countries mean absolute deviations (MADs) were computed after denormalization. Trials with negligible/low MSEs and MADs were selected for further evaluation. The MADs are presented in Appendix Table Ia - Ie. The MADs for Australia ranges between 0.01054 to 0.01649; 0.01227 to 0.02305 for Brazil; 0.01536 to 0.03093 for China; 0.00919 to 0.03652 for India and 0.00458 to 0.03976 for USA. The MADs are determined using Eq. (17):

$$MAD = \frac{1}{n}\sum_{i=1}^{n} |y_a - y_p| \qquad (17)$$

The optimal models for each country were finally selected based on the run with the sample weight mean closer to the population weights mean with negligible MAD and MSE.

Using this criteria, the 18th run for Australia (SE = 0.05641, MAD = 0.01193, $MSE_{train}$ = 0.01090, $MSE_{test}$ = 0.01020), the 11th run for Brazil (SE = 0.06544, MAD = 0.01345, $MSE_{train}$ = 0.00570, $MSE_{test}$ = 0.00420), the 11th run for China (SE = 0.06433, MAD = 0.01536, $MSE_{train}$ = 0.00330, $MSE_{test}$ = 0.00330), the 15th run for India (SE = 0.06480, MAD = 0.00919, $MSE_{train}$ = 0.00045, $MSE_{test}$ = 0.00042), and the 6th run for USA (SE = 0.06124, MAD = 0.00458, $MSE_{train}$ = 0.00440, $MSE_{test}$ = 0.00530) were selected (see Appendix Table Ia – Ie). This indicates how precise and close the weight points tend to approach/converge at the true population weights mean with the given data (Boateng et al., 2019). Table 3 presents the selected models and their confidence limits.

Table 3. Selected models and their confidence limits on extracted weights.

| Country | Run | Lower limit* | Upper limit* | Mean | $P_{mean}$ |
|---|---|---|---|---|---|
| Australia | 18th | -0.124120166 | 0.097003346 | -0.01355841 | -0.029038403 |
| Brazil | 11th | -0.114179821 | 0.142364429 | 0.014092304 | -0.006875717 |
| China | 11th | -0.132262709 | 0.119915949 | -0.00617338 | -0.007379938 |
| India | 15th | -0.132083021 | 0.121917025 | -0.005082998 | -0.004772717 |
| USA | 6th | -0.101997647 | 0.138054655 | 0.018028504 | 0.011375343 |

*95% confidence interval, $P_{mean}$ = population mean

The $R^2$ for each selected model in individual countries are shown in Fig. 4a – 4e. Australia, Brazil, China, India, and USA achieved coefficients of 0.8011, 0.9139, 0.9521, 0.9944, and 0.8721 respectively. The high $R^2$ values indicate how well the 9-5-1 MLP with BP models fit the data. Therefore, there are strong relationships between the developed models and the output variables for each country.

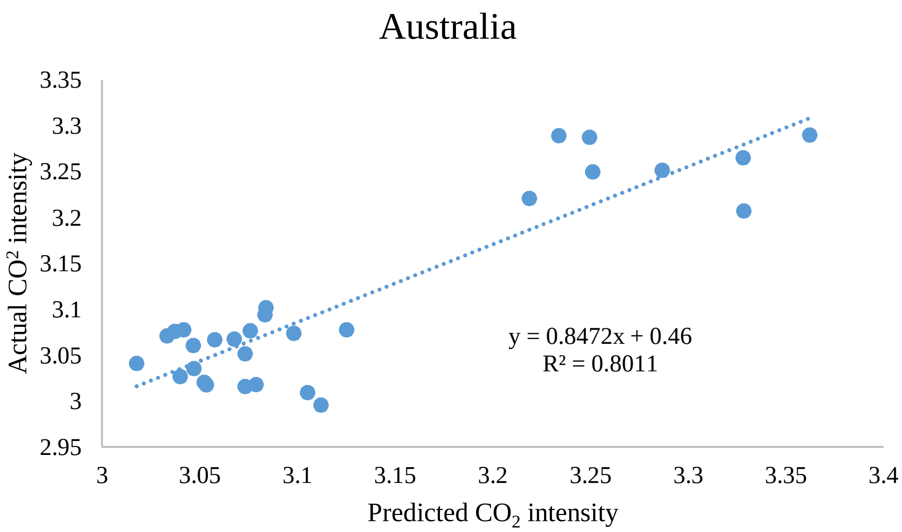



Fig. 4a. Scatter chart of actual and predicted $CO_2$ intensities.

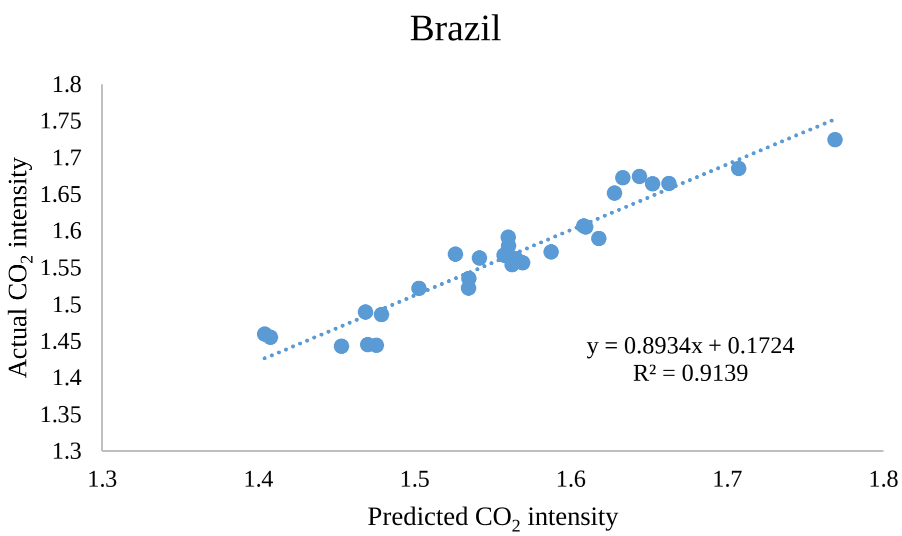



Fig. 4b. Scatter chart of actual and predicted $CO_2$ intensities.

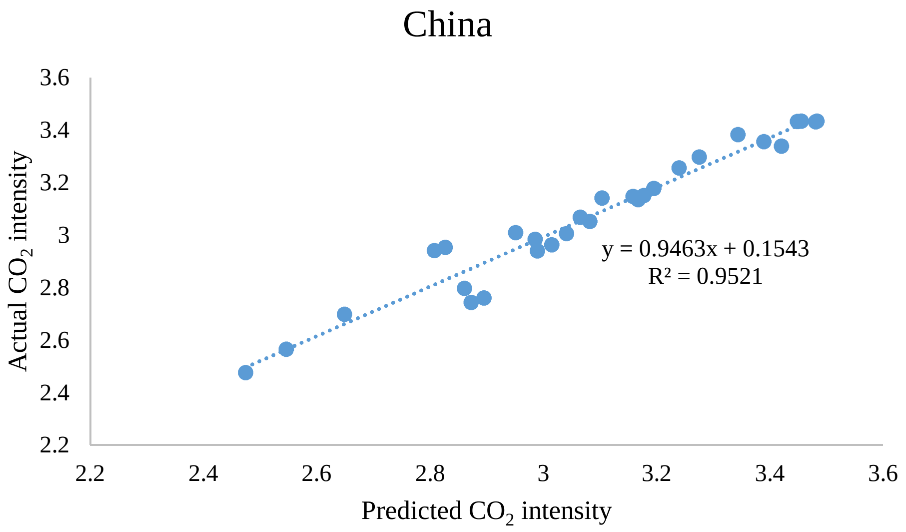



Fig. 4c. Scatter chart of actual and predicted $CO_2$ intensities.

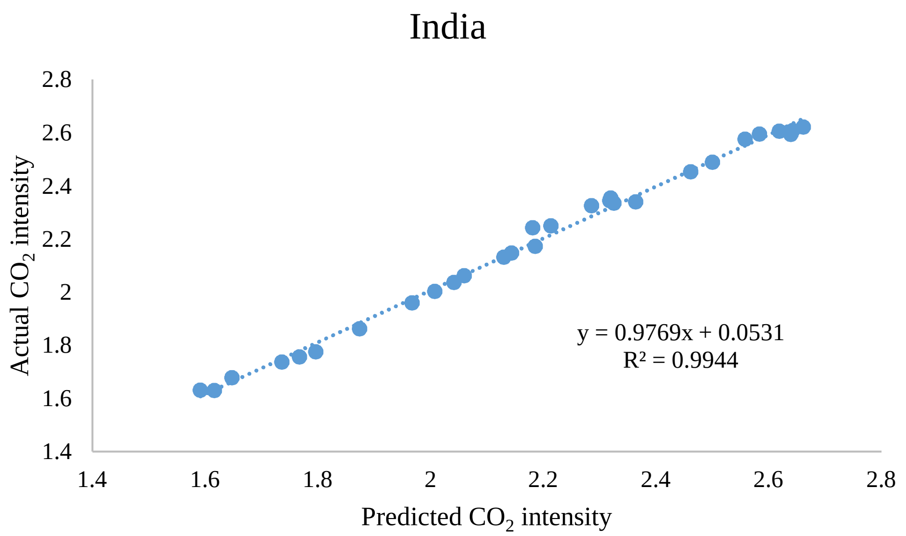



Fig. 4d. Scatter chart of actual and predicted $CO_2$ intensities.

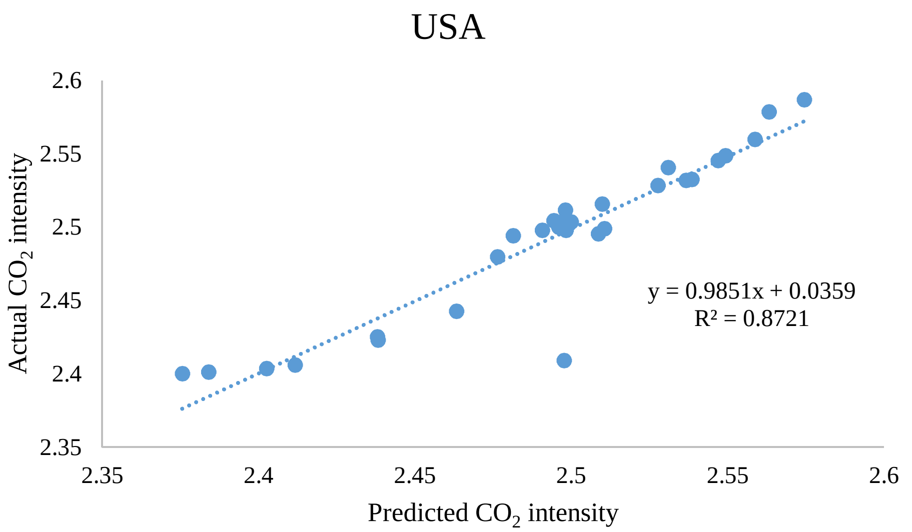



Fig. 4e. Scatter chart of actual and predicted $CO_2$ intensities.

The MSE values on both the training and test sets are illustrated in Fig. 5. The MSE values for each case are approximately Zero (0). This affirms that the developed models are sufficient to perform the necessary computations with minimal or negligible forecasting error.

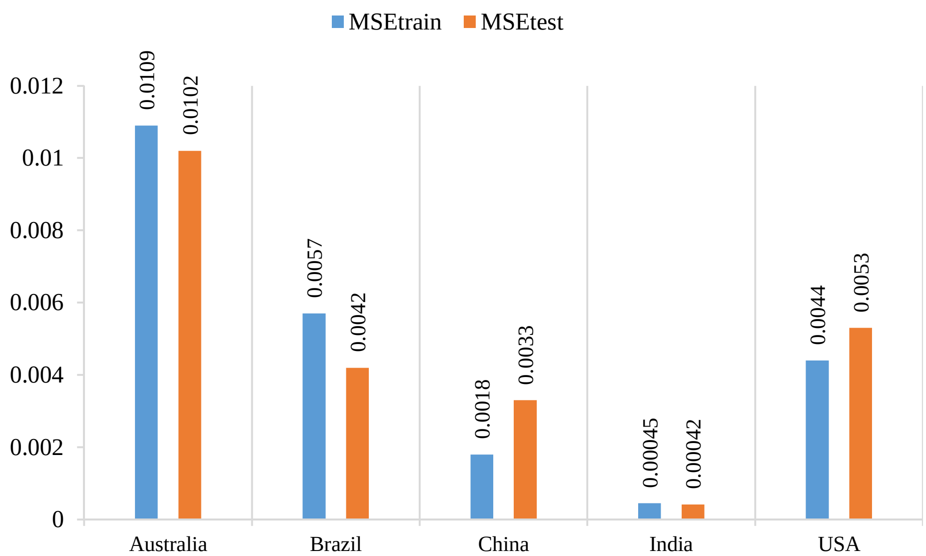



Fig. 5. Regression metric scores.

*4.2. Validation of the ANN models*

With 29 test samples, the 9-5-1 MLPs were employed to predict the $CO_2$ emission intensities from the 9 input parameters. Absolute deviations (ADs) were computed to validate the models. The AD is equal to the positive proportion of the difference between the actual ($y_a$) and predicted ($y_p$) observations to the actual observation ($y_a$) of the model (Patel & Jha, 2016). AD for Australia ranged from 0.000128408 to 0.037267504, 0.000544838 to 0.039480646 for Brazil, 0.000551518 to 0.047491399 for China, 0.000404588 to 0.028095168 for India, and 0.00024287 to 0.03549516 for USA (see Appendix Table IIa – IIc). The range of ADs shows that the trained model is capable of forecasting the intensity of $CO_2$ emissions for each country. Fig. 6a – 6e show the actual $CO_2$ emission intensities versus the predicted $CO_2$ emission intensities from the 29 test samples (quarters) for each country.

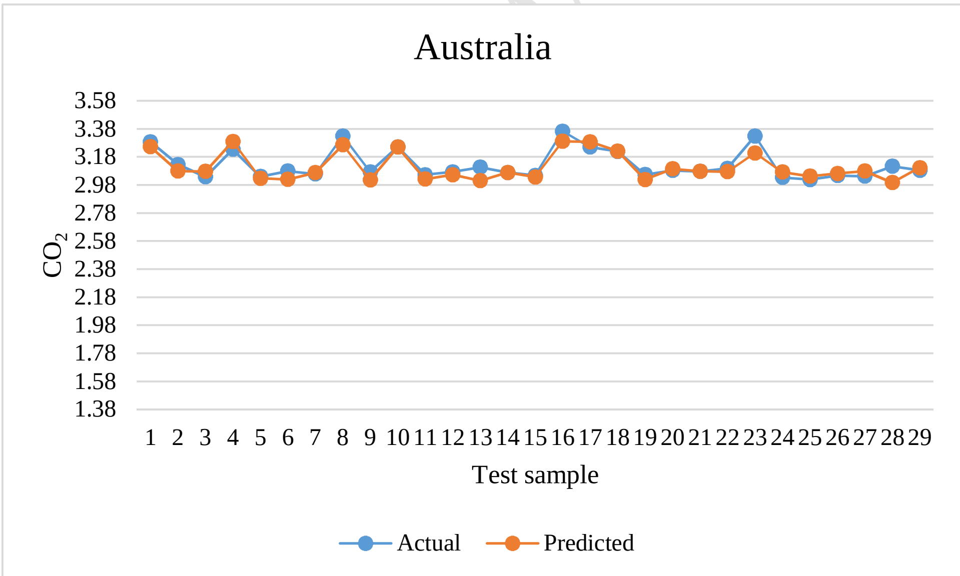



Fig. 6a. Actual versus predicted $CO_2$ intensities.

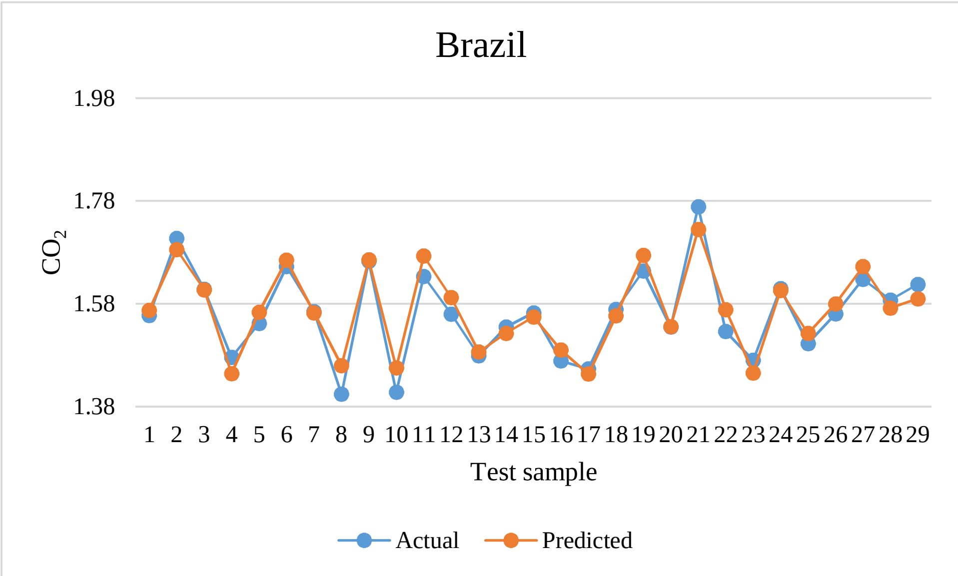



Fig. 6b. Actual versus predicted $CO_2$ intensities.

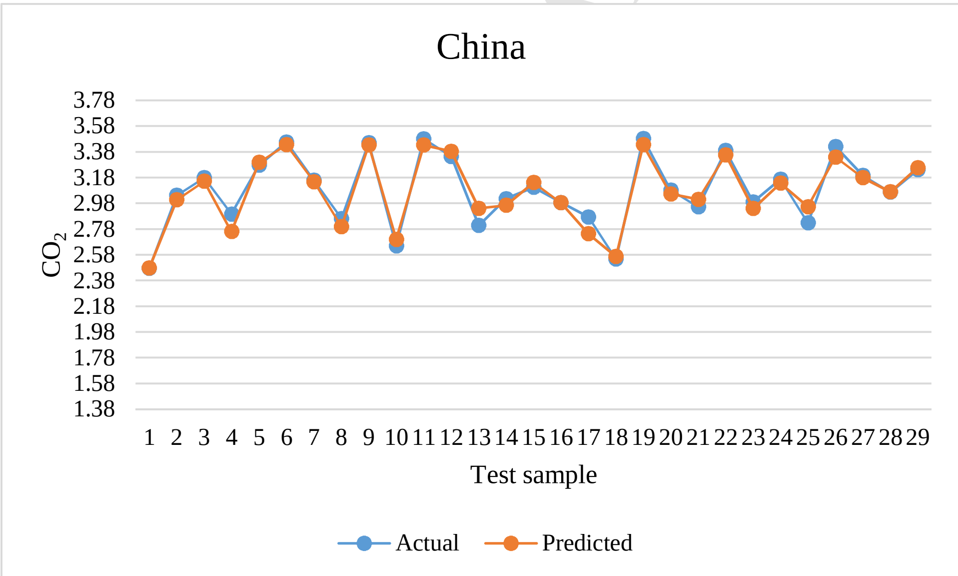



Fig. 6c. Actual versus predicted $CO_2$ intensities.

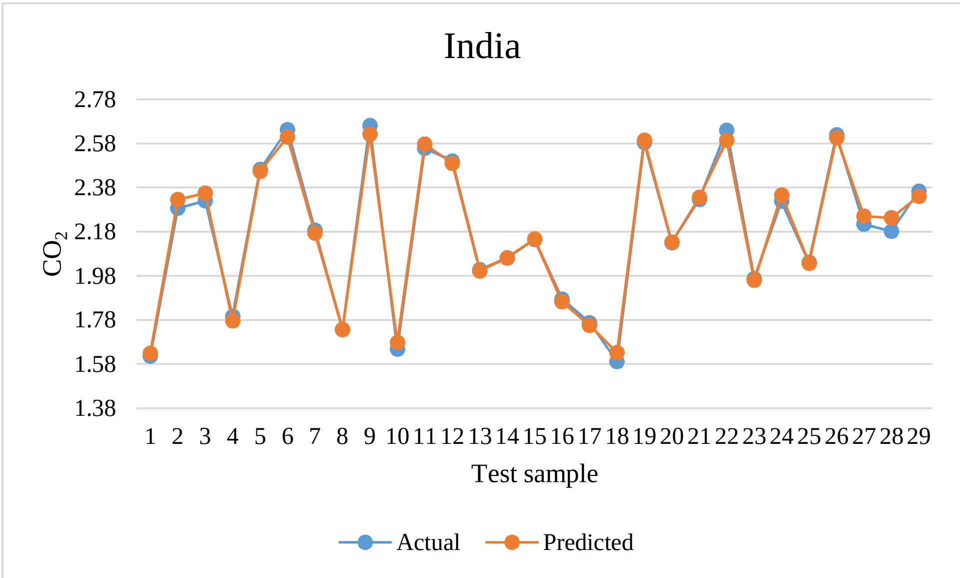



Fig. 6d. Actual versus predicted $CO_2$ intensities.

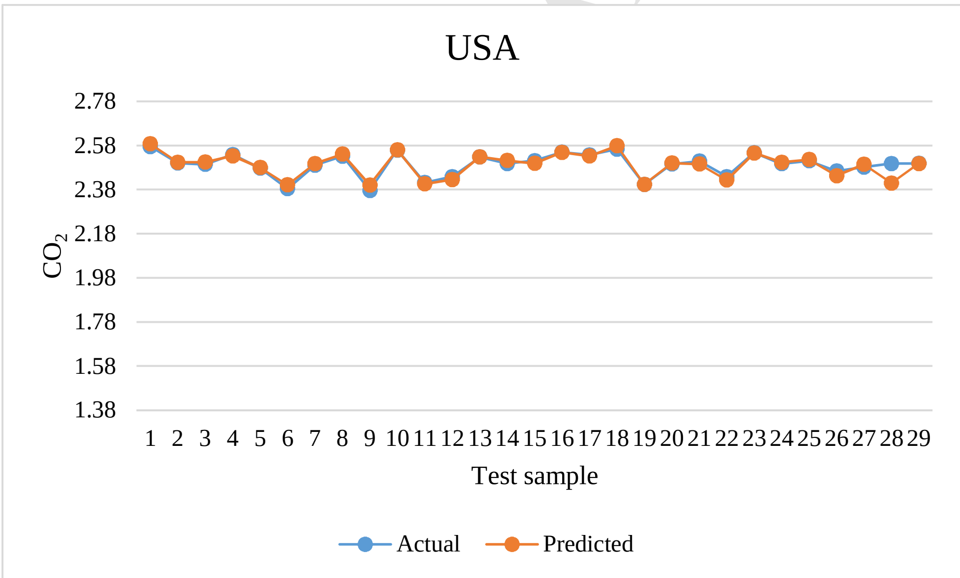



Fig. 6e. Actual versus predicted $CO_2$ intensities.

## 5. Closed-form formula for predicting $CO_2$ emission intensity

For the purpose of environmental policymakers and consultants, a simplified closed-form Eq. (18) and (19) can be used for predicting the intensity of $CO_2$ emission. The closed-form equations need the values of the inputs, weights of the links between the neurons in different layers, and the biases of the output and input neurons (Patel & Jha, 2015; Tadesse, Patel, Chaudhary, & Nagpal, 2012). The closed-form formula presented in this study is suitable for use where the activation function for the output layer is the sigmoid function. The output $O_1$ from Fig. 2 can be obtained from computing Eq. (18) and (19). Where Eq. (18) is the formula for predicting the output (carbon emission intensity).

$$O_1 = \frac{1}{1+e^{-\left(bias_o+\sum_{k=1}^{r}\frac{w_{k,1}^{ho}}{1+e^{-H_k}}\right)}} \tag{18}$$

Where $r$ is the number of hidden neurons respectively; $bias_o$ is the bias of the output layer neuron; $w_{k,1}^{ho}$ is the weight of the link between $H_k$ and $O_1$. Eq. (19) is used for calculating the *kth* hidden layer neuron.

$$H_k = \sum_{j=1}^{q} w_{j,k}^{ih} \times I_j + bias_k \tag{19}$$

Where $q$ is the number of the input parameters; $bias_k$ is the bias of the *kth* hidden layer neuron ($H_k$); $w_{j,k}^{ih}$ the weight of the link between $I_j$ and $H_k$. For each country, the weights and biases are presented in Tables 4a – 4e.

The closed-form expression can be used to predict $CO_2$ emission intensity based on the previous input values. An illustrative example is demonstrated afterwards.

Table 4a. Weight values and biases for neural network (Australia).

| Link | Weight/bias | Number of hidden layer neuron (*k*) | | | | |
|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 |
| Input to hidden layer | $w_{1,k}^{ih}$ | 0.480 | -0.677 | -0.371 | 0.582 | -0.750 |
| | $w_{2,k}^{ih}$ | -0.026 | 0.326 | 0.018 | -0.312 | -0.252 |
| | $w_{3,k}^{ih}$ | -0.216 | 0.228 | 0.126 | -0.265 | 0.180 |
| | $w_{4,k}^{ih}$ | -0.183 | 0.529 | -0.179 | -0.385 | 0.457 |
| | $w_{5,k}^{ih}$ | 0.262 | 0.220 | -0.314 | 0.401 | -0.089 |
| | $w_{6,k}^{ih}$ | -0.500 | 0.486 | -0.419 | -0.229 | 0.532 |
| | $w_{7,k}^{ih}$ | -0.304 | 0.597 | -0.397 | 0.190 | -0.088 |
| | $w_{8,k}^{ih}$ | 0.039 | 0.171 | -0.210 | 0.527 | -0.194 |
| | $w_{9,k}^{ih}$ | 0.029 | -0.347 | -0.225 | 0.314 | -0.619 |
| | $bias_k$ | -0.036 | 0.298 | 0.002 | -0.195 | 0.022 |

| Link | Weight/bias | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| Hidden layer to output | $w^{ho}_{k,1}$ | -0.971 | -0.363 | 0.452 | 0.276 | 0.786 |
| | $bias_o$ | -0.161 | | | | |

Table 4b. Weight values and biases for neural network (Brazil).

| Link | Weight/bias | Number of hidden layer neuron ($k$) | | | | |
|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 |
| Input to hidden layer | $w^{ih}_{1,k}$ | 0.237 | 0.538 | -0.358 | -0.035 | 0.549 |
| | $w^{ih}_{2,k}$ | -0.459 | -0.651 | 0.299 | -0.376 | 0.210 |
| | $w^{ih}_{3,k}$ | -0.330 | 0.848 | 0.196 | -0.739 | -0.592 |
| | $w^{ih}_{4,k}$ | 0.197 | 0.074 | -0.873 | 0.250 | -0.022 |
| | $w^{ih}_{5,k}$ | 0.208 | 0.067 | -0.346 | 0.650 | -0.307 |
| | $w^{ih}_{6,k}$ | 0.046 | 0.313 | 0.165 | -0.303 | -0.557 |
| | $w^{ih}_{7,k}$ | 0.033 | -0.466 | 0.181 | -0.338 | -0.505 |
| | $w^{ih}_{8,k}$ | 0.178 | 0.238 | -0.234 | 0.033 | -0.254 |
| | $w^{ih}_{9,k}$ | 0.538 | -0.315 | 0.054 | 0.030 | -0.145 |
| | $bias_k$ | -0.352 | 0.032 | 0.268 | 0.070 | 0.024 |
| Hidden layer to output | $w^{ho}_{k,1}$ | 1.062 | 1.148 | 1.091 | -0.552 | 0.028 |
| | $bias_o$ | -0.236 | | | | |

Table 4c. Weight values and biases for neural network (China).

| Link | Weight/bias | Number of hidden layer neuron ($k$) | | | | |
|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 |
| Input to hidden layer | $w^{ih}_{1,k}$ | 0.240 | -0.147 | -0.552 | 0.645 | 0.384 |
| | $w^{ih}_{2,k}$ | 0.265 | -0.026 | 0.316 | 0.491 | 0.498 |
| | $w^{ih}_{3,k}$ | 0.299 | -0.266 | -0.657 | -0.462 | -0.549 |
| | $w^{ih}_{4,k}$ | 0.349 | -0.484 | -0.777 | -0.649 | 0.522 |
| | $w^{ih}_{5,k}$ | 0.847 | 0.285 | -0.183 | 0.724 | 0.196 |
| | $w^{ih}_{6,k}$ | 0.294 | 0.359 | 0.501 | -0.101 | 0.200 |
| | $w^{ih}_{7,k}$ | 0.128 | -0.636 | 0.177 | -0.348 | -0.403 |
| | $w^{ih}_{8,k}$ | -0.172 | -0.959 | 0.373 | 0.505 | -0.301 |
| | $w^{ih}_{9,k}$ | -0.167 | -0.751 | -0.273 | -0.383 | 0.507 |
| | $bias_k$ | 0.158 | -0.222 | 0.097 | -0.324 | -0.190 |
| Hidden layer to output | $w^{ho}_{k,1}$ | 1.064 | -0.727 | 0.867 | -0.484 | 0.190 |
| | $bias_o$ | 0.044 | | | | |

Table 4d. Weight values and biases for neural network (India).

| Link | Weight/bias | Number of hidden layer neuron ($k$) | | | | |
|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 |
| Input to hidden layer | $w^{ih}_{1,k}$ | -0.497 | 0.383 | -0.468 | 0.626 | -0.253 |
| | $w^{ih}_{2,k}$ | -0.508 | 0.130 | 0.180 | -0.467 | 0.775 |
| | $w^{ih}_{3,k}$ | 0.237 | 0.063 | 0.794 | 0.218 | 0.410 |
| | $w^{ih}_{4,k}$ | 0.132 | 0.460 | 0.011 | 0.234 | -0.154 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | $w^{ih}_{5,k}$ | -0.264 | 0.359 | -0.503 | -0.316 | 0.355 |
| | $w^{ih}_{6,k}$ | 0.016 | 0.284 | 0.228 | -0.541 | -0.139 |
| | $w^{ih}_{7,k}$ | -0.326 | 0.229 | -0.919 | -0.329 | 0.240 |
| | $w^{ih}_{8,k}$ | 0.344 | -0.142 | -0.167 | 0.238 | -0.709 |
| | $w^{ih}_{9,k}$ | -0.772 | -0.250 | 0.061 | 0.381 | -0.125 |
| | $bias_k$ | -0.169 | 0.134 | -0.307 | 0.267 | 0.427 |
| Hidden layer to output | $w^{ho}_{k,1}$ | -1.091 | 0.149 | -0.490 | 1.296 | 0.342 |
| | $bias_o$ | 0.122 | | | | |

Table 4e. Weight values and biases for neural network (USA).

| Link | Weight/bias | Number of hidden layer neuron ($k$) | | | | |
|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 |
| Input to hidden layer | $w^{ih}_{1,k}$ | 0.299 | 0.671 | -0.756 | 0.075 | -0.057 |
| | $w^{ih}_{2,k}$ | -0.168 | -0.370 | -0.374 | -0.255 | 0.183 |
| | $w^{ih}_{3,k}$ | -0.107 | -0.646 | -0.148 | 0.657 | 0.563 |
| | $w^{ih}_{4,k}$ | -0.436 | -0.319 | 0.305 | -0.864 | -0.037 |
| | $w^{ih}_{5,k}$ | 0.348 | 0.161 | -0.263 | -0.547 | -0.244 |
| | $w^{ih}_{6,k}$ | -0.155 | 0.461 | -0.231 | -0.066 | -0.098 |
| | $w^{ih}_{7,k}$ | -0.548 | 0.712 | 0.486 | -0.585 | 0.253 |
| | $w^{ih}_{8,k}$ | 0.574 | -0.587 | 0.211 | -0.080 | -0.238 |
| | $w^{ih}_{9,k}$ | 0.189 | 0.582 | 0.709 | -0.060 | -0.034 |
| | $bias_k$ | -0.036 | 0.298 | 0.002 | -0.195 | 0.022 |
| Hidden layer to output | $w^{ho}_{k,1}$ | -0.427 | 0.359 | -0.004 | -0.380 | -0.214 |
| | $bias_o$ | -0.305 | | | | |

*Illustrative example*

For practical purpose, we demonstrate how to use the above closed-form formula for predicting/forecasting carbon emission intensity[3]. Using India as an illustration, consider the nine input parameters ($I_1$ to $I_9$) for determining the $CO_2$ emission intensity for 2011Q4 (Table 5). The $CO_2$ emission intensity for the next quarter (2012Q1) $O_1$ may be obtained by the following steps:

Table 5. Input values and the output value for 2011Q4, India.

| Period | 2011Q4 (destandardised) | 2011Q4 (standardised) |
|---|---|---|
| Energy consumption ($I_1$) | 586.5490144 | 1.797073524 |
| Financial development ($I_2$) | 0.405568259 | 0.847736897 |
| Economic growth ($I_3$) | 1.905736931 | 1.583596755 |

[3] Following the steps used for the illustration, one can predict for the remaining countries (Australia, Brazil, China and the USA) by substituting the weights, biases and values of the input variables for respective countries into the closed-form formula.

| | | |
|---|---|---|
| Foreign direct investment ($I_4$) | 1440.250772 | 1.172634771 |
| Industrialisation ($I_5$) | 30.00708424 | 0.921162953 |
| Technology ($I_6$) | 198210.5313 | 1.662170004 |
| Population ($I_7$) | 1253238699 | 1.307734943 |
| Trade openness ($I_8$) | 56.58837488 | 1.861353516 |
| Urbanisation ($I_9$) | 393609283.8 | 1.402126680 |
| $CO_2$ emission intensity ($O$) | 2.583927266 | 0.843589044* |
| $CO_2$ emission intensity for 2012Q1 | $y_p$ | $O_1$? |

Note: *normalised, destandardisation: $x = (x_{stand} \times \sigma) + \mu$

Step 1. Insert the standardised values of the input parameters (Table 5) and weights and biases of input to the hidden layer (Table 4d) in Eq. (19) to compute $H_1$ to $H_5$ as given in Eq. (20) - (24).

$$H_1 = (-0.497I_1 - 0.508I_2 - 0.237I_3 + 0.132I_4 - 0.264I_5 + 0.016I_6 - 0.326I_7 + 0.344I_8 - 0.772I_9) - 0.169 \quad (20)$$

$$H_2 = (0.383I_1 + 0.130I_2 + 0.063I_3 + 0.460I_4 + 0.359I_5 + 0.284I_6 + 0.229I_7 - 0.142I_8 - 0.250I_9) + 0.134 \quad (21)$$

$$H_3 = (-0.468I_1 + 0.180I_2 + 0.794I_3 + 0.011I_4 - 0.503I_5 + 0.228I_6 - 0.919I_7 - 0.167I_8 + 0.061I_9) - 0.307 \quad (22)$$

$$H_4 = (0.626I_1 - 0.467I_2 + 0.218I_3 + 0.234I_4 - 0.316I_5 - 0.541I_6 - 0.329I_7 + 0.238I_8 + 0.381I_9) + 0.267 \quad (23)$$

$$H_5 = (-0.253I_1 + 0.775I_2 + 0.410I_3 - 0.154I_4 + 0.355I_5 - 0.139I_6 + 0.240I_7 - 0.709I_8 - 0.125I_9) + 0.427 \quad (24)$$

The values of $H_1$ to $H_5$ are obtained as -2.047745743, 2.05904476, -1.236657889, -0.972341664, and 0.012887535, respectively.

Step 2. Insert the values of $H_1$ to $H_5$ and the weights and biases of the hidden to output layer (Table 4d) in Eq. (18) as given in Eq. (25). The value of the predicted output $O_1$ is 0.756145926.

$$O_1 = \frac{1}{1+e^{-\left(0.122 - \frac{1.091}{1+e^{-H_1}} + \frac{0.149}{1+e^{-H_2}} - \frac{0.490}{1+e^{-H_3}} + \frac{1.296}{1+e^{-H_4}} + \frac{0.342}{1+e^{-H_5}}\right)}} \quad (25)$$

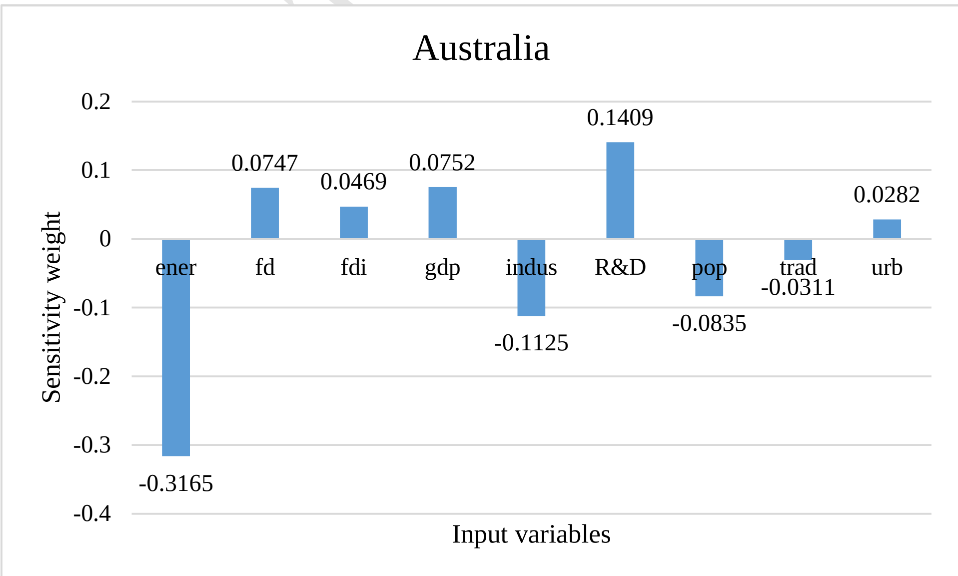
Step 3. The value obtained from Eq. (25) is the normalized value ($y_{norm}$). Eq. (16) is used to denormalize $O_1$ as given in Eq. (26):

$$y_p = 0.756145926(2.777034539 - 1.542419792) + 1.542419792 \quad (26)$$

The actual $CO_2$ intensity ($y_a$) for 2012Q1 is 2.6430 and the predicted $CO_2$ intensity ($y_p$) for 2012Q1 is 2.4760. The AD for this forecast is 0.0632.

## 6. Sensitivity analysis

Sensitivity analysis is conducted to identify the extent to which each input variable contributes to the intensity of carbon emissions in Australia, Brazil, China, India, and the USA. To conduct the sensitivity analysis, the Partial Rank Correlation Coefficient (PRCC) between carbon emission intensity and each input variable is calculated for each country. Fig 7a-7d depicts the normalised sensitivity weight of each input variable for each country. Fig. 7a shows that in Australia, R&D has the highest sensitivity weight, followed by economic growth, financial development, foreign direct investment and urbanisation. As depicted in Fig 7a, the PRCC results show that R&D (0.1409), economic growth (0.0752), financial development (0.0747), foreign direct investment (0.0469) and urbanisation (0.0282) increase carbon emissions intensity while energy consumption (-0.3165), industrialisation (-0.1125), population (-0.0835) and trade openness (-0.0311) reduce carbon emission intensity in




Australia.

Fig. 7a. Sensitivity analysis of $CO_2$ emission intensity determinants.

Fig 7b indicates that in Brazil, urbanisation has the highest sensitivity weight followed by R&D, energy consumption and financial development. PRCC results show that in Brazil (see Fig.7b), urbanisation (0.4791), R&D (0.3686), energy consumption (0.2526) and financial development (0.1491) increases carbon emissions intensity while population (-0.4956), trade openness (-0.3145), industrialisation (-0.3094), economic growth (-0.2176), population (-0.0835) and foreign direct investment (-0.0991) contribute to reduction in carbon emission intensity.

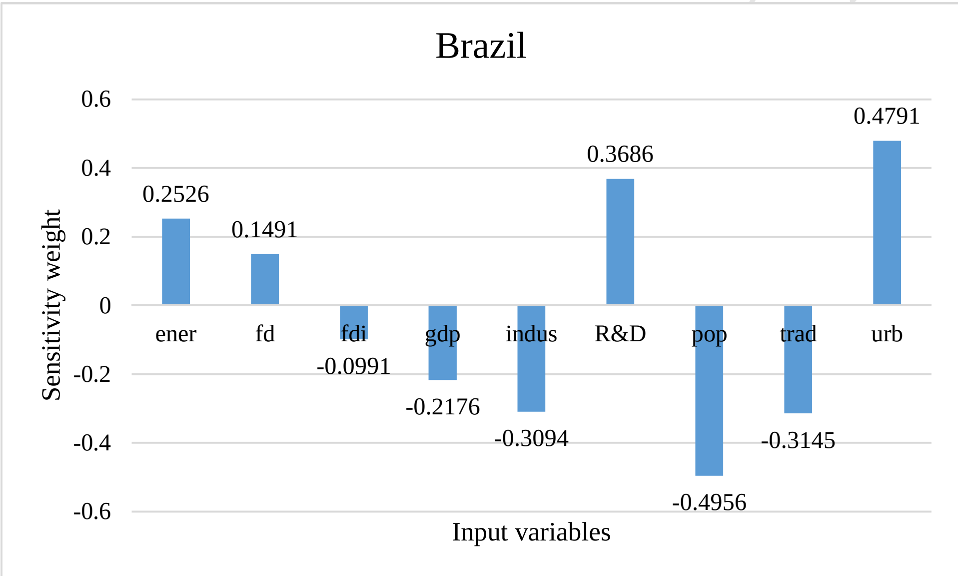



Fig. 7b. Sensitivity analysis of $CO_2$ emission intensity determinants.

In China, as depicted Fig 7c, population size has the highest sensitivity weight followed by economic growth, energy consumption, trade openness, industrialisation, and financial development. The PRCC results as shown in Fig. 7c shows that population (0.7053), economic growth (0.6047), energy consumption (0.5822), trade openness (0.4625), industrialisation (0.276) and financial development (0.2055) are the forces behind carbon emissions in China while urbanisation (-0.6716), R&D (-0.3021) and foreign direct investment (-0.3094) reduce the intensity of carbon emissions.

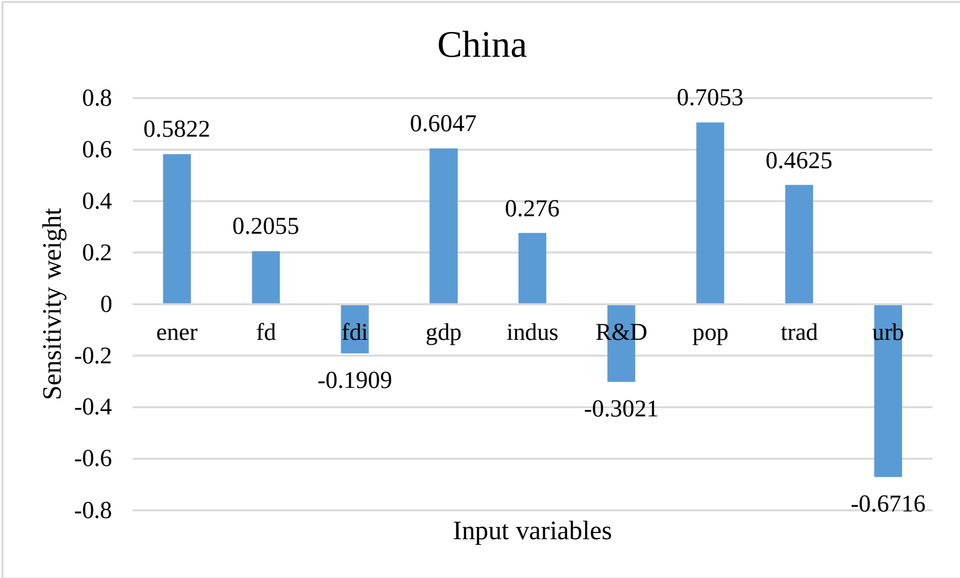



Fig. 7c. Sensitivity analysis of $CO_2$ emission intensity determinants.

For India, Fig 7d shows that energy consumption has the highest sensitivity weight followed by foreign direct investment, population, and economic growth. As depicted in Fig. 7d, the factors responsible for the growth of carbon emission intensity include energy consumption (0.8558), foreign direct investment (0.6256), population (0.4636) and economic growth (0.1177) while financial development (-0.5202), R&D (-0.3814), industrialisation (-0.3526), urbanisation (-0.334) and trade openness (-0.1274) contribute to the reduction in carbon emission intensity.

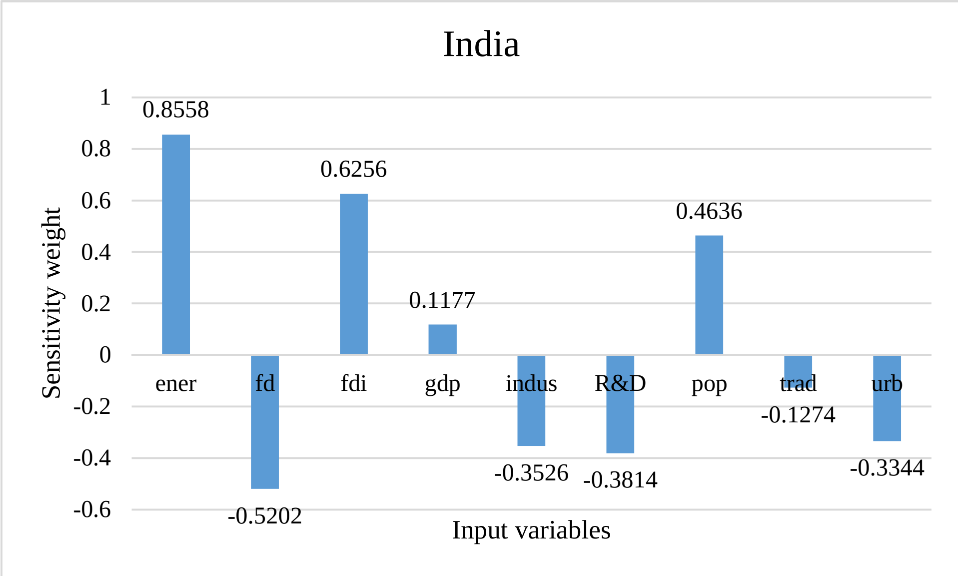



Fig. 7d. Sensitivity analysis of $CO_2$ emission intensity determinants.

For USA, urbanisation has the highest sensitivity weight, followed by foreign direct investment, economic growth and R&D (see Fig 7e). Fig. 7e shows that the factors contributing to the rise of carbon emission intensity in USA include urbanisation (0.4541) energy consumption (0.3211), foreign direct investment (0.2044), economic growth (0.1504) and R&D (0.0667) while trade openness (-0.4261), population (-0.4025), financial development (-0.3599) and industrialisation (-0.0616) reduces carbon emission intensity.

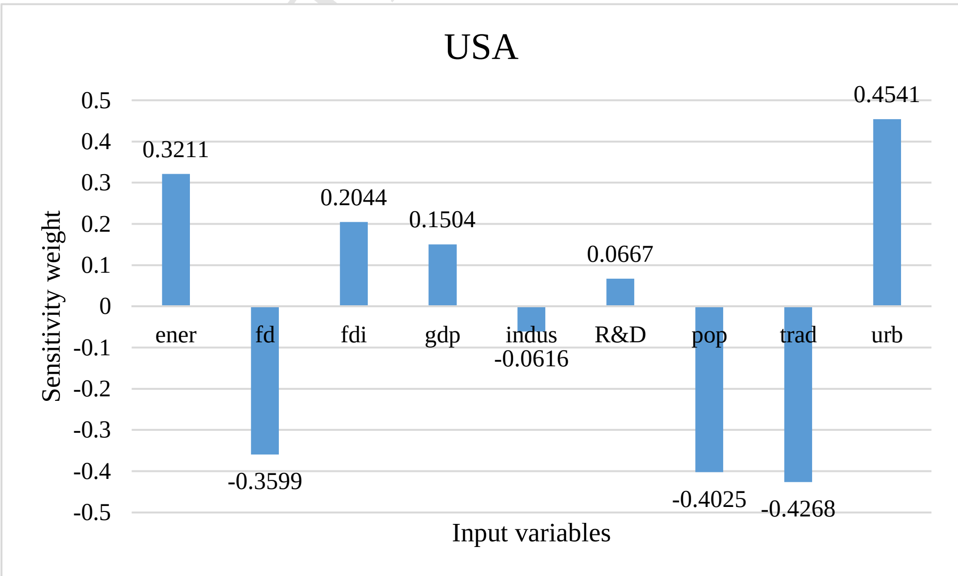



Fig. 7e. Sensitivity analysis of $CO_2$ emission intensity determinants.

The results from the sensitivity analysis revealed that each of the input has an important but different influence on the intensity of carbon emissions of the countries considered. Therefore, carbon emissions models that tend to ignore these variables could result in the underestimation of the actual carbon emission intensity.

## 7. Conclusion and policy implications

The applicability of the artificial neural network (ANN) technique for predicting $CO_2$ emission intensity was evaluated for Australia, Brazil, China, India, and USA. The type of neural network used for each country was the feed forward multi-layer perceptron (FFMLP). A stochastic gradient descent with the backpropagation algorithm was employed to train the networks over several iterations. The 9-5-1 FFMLPs take into account energy consumption, financial development, foreign direct investment, economic growth, industrialisation, technology, population, trade openness, and urbanization scores of the selected countries. Five ANN models were developed with 115 quarters and validated on 29 quarters for the prediction of $CO_2$ emission intensity for the five countries. The results of this study are very promising and showed good generalization. The predicted versus actual values indicate negligible or approximately zero errors for the AD, MAD, MSE, SD, SE, and ME along with higher coefficients of determination ($R^2$) of 0.80 for Australia, 0.91 for Brazil, 0.95 for China, 0.99 for India, and 0.87 for USA. This study does not only proposes a novel ANN technique for predicting $CO_2$ emission intensity but also presents a closed-form solution for predicting $CO_2$ emission intensity for Australia, Brazil, China, India, and USA with insignificant forecasting deviations. Software developers could also use the closed-form solution, the model architecture, and the extracted weights and biases of each parameter to develop a $CO_2$ emission intensity application on various platforms for the five countries using any programming language.

As a machine learning (ML) technique, the developed ANN models overcome the limitations of statistical approaches and are very practical to use. Due to the stochastic nature of neural networks, this study further proposes a robust methodology in selecting an optimal model for stable reproducibility of results. The ANN models presented in this study have been validated and reliable to predict the growth of $CO_2$ emission intensity for Australia, Brazil, China, India, and USA with negligible forecasting errors. Additionally, the results from the sensitivity analysis revealed that for Australia, R&D has the highest sensitivity weight while for Brazil and the USA, urbanisation has the highest sensitivity weight. For China, population size has the highest sensitivity weight while energy consumption has the highest sensitivity weight in India. The implication from the sensitivity results is that environmental policymakers in each respective country should prioritise these variables when designing and implementing environmental (climate change) policies. Additionally, the models developed from this study could serve as tools for international organizations and environmental policymakers to design and implement environmental policies and strategies to monitor and control environmental problems.

Future studies could consider performance evaluation of ANN models for prediction of $CO_2$ emission intensity with other ML approaches such as support vector regression (SVR) and recurrent neural network (RNN). This sort of evaluation would offer a dais for the methodological rigour in the selection of other ML tools that may give predictions that are more accurate. Other high $CO_2$ emitting countries such as Russia, Japan, and Germany could adopt the flow process of this study's methodology to develop robust predictive ANN models for guiding decision making when drafting environmental (climate change) policies.

## Appendix

Appendix Ia. Comparison of selected optimal model and other ANN models (Australia).

| Model | $R^2$ | $MSE_{train}$ | $MSE_{test}$ | SD | SE | MAD | $ME_y$ |
|---|---|---|---|---|---|---|---|
| 1 | 0.75280 | 0.01290 | 0.01290 | 0.42690 | 0.06037 | 0.01410 | 0.02076 |
| 2 | 0.64220 | 0.01180 | 0.01830 | 0.42849 | 0.06060 | 0.01646 | 0.01560 |
| 3 | 0.76190 | 0.01240 | 0.01260 | 0.43203 | 0.06110 | 0.01410 | 0.02152 |
| 4 | 0.82680 | 0.01180 | 0.00830 | 0.41794 | 0.05910 | 0.01054 | 0.01273 |
| 5 | 0.79290 | 0.01140 | 0.00930 | 0.43954 | 0.06216 | 0.01112 | 0.00656 |
| 6 | 0.74940 | 0.01120 | 0.01430 | 0.41651 | 0.05890 | 0.01456 | 0.02506 |
| 7 | 0.68230 | 0.01310 | 0.01470 | 0.40492 | 0.05726 | 0.01447 | 0.01224 |
| 8 | 0.80390 | 0.00750 | 0.01080 | 0.43129 | 0.06099 | 0.01235 | 0.02030 |
| 9 | 0.75600 | 0.01190 | 0.01360 | 0.39712 | 0.05616 | 0.01322 | 0.02507 |
| 10 | 0.69430 | 0.01390 | 0.01840 | 0.41533 | 0.05874 | 0.01649 | 0.03371 |
| 11 | 0.67830 | 0.01180 | 0.01580 | 0.40661 | 0.05750 | 0.01509 | 0.01525 |
| 12 | 0.74020 | 0.01320 | 0.01180 | 0.38728 | 0.05477 | 0.01280 | 0.00765 |
| 13 | 0.78040 | 0.01150 | 0.01080 | 0.41848 | 0.05918 | 0.01265 | 0.01598 |
| 14 | 0.64300 | 0.01040 | 0.01740 | 0.41846 | 0.05918 | 0.01573 | 0.01162 |
| 15 | 0.78840 | 0.01070 | 0.01020 | 0.41634 | 0.05888 | 0.01234 | 0.01297 |
| 16 | 0.70310 | 0.00880 | 0.01380 | 0.45046 | 0.06370 | 0.01378 | 0.00786 |
| 17 | 0.76070 | 0.00910 | 0.01110 | 0.44654 | 0.06315 | 0.01210 | 0.01064 |
| **18** | **0.80110** | **0.01090** | **0.01020** | **0.39887** | **0.05641** | **0.01193** | **0.01778** |
| 19 | 0.71020 | 0.01320 | 0.01310 | 0.38845 | 0.05494 | 0.01276 | 0.00600 |
| 20 | 0.77700 | 0.01290 | 0.01050 | 0.38632 | 0.05463 | 0.01224 | 0.01251 |

Note: MSE = mean square error, SD = standard deviation, SE = standard error, MAD = mean absolute deviation, $ME_y$ = mean error on prediction

Appendix Ib. Comparison of selected optimal model and other ANN models (Brazil).

| Model | $R^2$ | $MSE_{train}$ | $MSE_{test}$ | SD | SE | MAD | $ME_y$ |
|---|---|---|---|---|---|---|---|
| 1 | 0.89430 | 0.00650 | 0.00540 | 0.48114 | 0.06804 | 0.01479 | -0.00897 |
| 2 | 0.88680 | 0.00740 | 0.00650 | 0.48239 | 0.06822 | 0.01622 | -0.01403 |
| 3 | 0.70370 | 0.01800 | 0.01450 | 0.37132 | 0.05251 | 0.02305 | -0.01125 |
| 4 | 0.87870 | 0.00690 | 0.00610 | 0.51981 | 0.07351 | 0.01589 | -0.00825 |
| 5 | 0.81460 | 0.00880 | 0.00900 | 0.50558 | 0.07150 | 0.02011 | -0.00877 |
| 6 | 0.75540 | 0.01580 | 0.01250 | 0.47784 | 0.06758 | 0.02040 | -0.01319 |
| 7 | 0.81900 | 0.01230 | 0.00860 | 0.47053 | 0.06654 | 0.01820 | -0.00641 |
| 8 | 0.88110 | 0.00540 | 0.00620 | 0.50556 | 0.07150 | 0.01600 | -0.00920 |
| 9 | 0.91020 | 0.00480 | 0.00420 | 0.48854 | 0.06909 | 0.01227 | -0.00312 |
| 10 | 0.89100 | 0.00650 | 0.00610 | 0.48999 | 0.06929 | 0.01591 | -0.01232 |
| **11** | **0.91390** | **0.00570** | **0.00420** | **0.46277** | **0.06544** | **0.01345** | **-0.00597** |
| 12 | 0.79740 | 0.01240 | 0.01050 | 0.49058 | 0.06938 | 0.02174 | -0.01332 |
| 13 | 0.88450 | 0.00740 | 0.00640 | 0.47772 | 0.06756 | 0.01633 | -0.01286 |
| 14 | 0.80050 | 0.01290 | 0.00980 | 0.45652 | 0.06456 | 0.01977 | -0.01001 |
| 15 | 0.90760 | 0.00550 | 0.00500 | 0.46081 | 0.06517 | 0.01460 | -0.01031 |
| 16 | 0.74890 | 0.01450 | 0.01290 | 0.44537 | 0.06298 | 0.02216 | -0.01377 |
| 17 | 0.87870 | 0.00700 | 0.00650 | 0.46594 | 0.06589 | 0.01613 | -0.01168 |
| 18 | 0.85470 | 0.00890 | 0.00730 | 0.50782 | 0.07182 | 0.01763 | -0.00943 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 19 | 0.84500 | 0.01160 | 0.00770 | 0.50151 | 0.07092 | 0.01719 | -0.00882 |
| 20 | 0.86580 | 0.00800 | 0.00700 | 0.44253 | 0.06258 | 0.01736 | -0.01128 |

Note: MSE = mean square error, SD = standard deviation, SE = standard error, MAD = mean absolute deviation, $ME_y$ = mean error on prediction

Appendix Ic. Comparison of selected optimal model and other ANN models (China).

| Model | $R^2$ | $MSE_{train}$ | $MSE_{test}$ | SD | SE | MAD | $ME_y$ |
|---|---|---|---|---|---|---|---|
| 1 | 0.90790 | 0.00490 | 0.00650 | 0.43691 | 0.06179 | 0.02537 | 0.02060 |
| 2 | 0.93890 | 0.00370 | 0.00400 | 0.44958 | 0.06358 | 0.01795 | 0.00483 |
| 3 | 0.92060 | 0.00480 | 0.00560 | 0.40740 | 0.05762 | 0.02277 | 0.01653 |
| 4 | 0.93380 | 0.00310 | 0.00450 | 0.48241 | 0.06822 | 0.01990 | 0.01332 |
| 5 | 0.84030 | 0.00900 | 0.01420 | 0.49285 | 0.06970 | 0.03087 | 0.05190 |
| 6 | 0.83180 | 0.01430 | 0.01170 | 0.45484 | 0.06432 | 0.02740 | -0.01303 |
| 7 | 0.92000 | 0.00350 | 0.00530 | 0.47023 | 0.06650 | 0.02045 | 0.00511 |
| 8 | 0.95550 | 0.00230 | 0.00320 | 0.44949 | 0.06357 | 0.01794 | 0.01729 |
| 9 | 0.91480 | 0.00380 | 0.00560 | 0.44856 | 0.06344 | 0.02265 | 0.00120 |
| 10 | 0.91250 | 0.00290 | 0.00580 | 0.49364 | 0.06981 | 0.02065 | 0.01016 |
| **11** | **0.95210** | **0.00180** | **0.00330** | **0.45489** | **0.06433** | **0.01536** | **0.01128** |
| 12 | 0.92050 | 0.00360 | 0.00530 | 0.43513 | 0.06154 | 0.02127 | 0.01063 |
| 13 | 0.94290 | 0.00300 | 0.00390 | 0.45553 | 0.06442 | 0.01840 | 0.01299 |
| 14 | 0.93830 | 0.00340 | 0.00430 | 0.45426 | 0.06424 | 0.01920 | 0.01384 |
| 15 | 0.92940 | 0.00330 | 0.00470 | 0.44908 | 0.06351 | 0.01993 | 0.00545 |
| 16 | 0.92370 | 0.00290 | 0.00510 | 0.46805 | 0.06619 | 0.02045 | 0.00669 |
| 17 | 0.92870 | 0.00350 | 0.00480 | 0.42924 | 0.06070 | 0.01980 | 0.01139 |
| 18 | 0.93450 | 0.00410 | 0.00460 | 0.41507 | 0.05870 | 0.02044 | 0.01741 |
| 19 | 0.83630 | 0.01030 | 0.01480 | 0.37316 | 0.05277 | 0.03093 | 0.05035 |
| 20 | 0.93890 | 0.00240 | 0.00420 | 0.41736 | 0.05902 | 0.01935 | 0.01190 |

Note: MSE = mean square error, SD = standard deviation, SE = standard error, MAD = mean absolute deviation, $ME_y$ = mean error on prediction

Appendix Id. Comparison of selected optimal model and other ANN models (India).

| Models | $R^2$ | $MSE_{train}$ | $MSE_{test}$ | SD | SE | MAD | $ME_y$ |
|---|---|---|---|---|---|---|---|
| 1 | 0.96870 | 0.00500 | 0.00360 | 0.42662 | 0.06033 | 0.02529 | -0.02168 |
| 2 | 0.93530 | 0.00820 | 0.00640 | 0.37290 | 0.05274 | 0.03652 | -0.03435 |
| 3 | 0.93600 | 0.00730 | 0.00590 | 0.42554 | 0.06018 | 0.03484 | -0.02437 |
| 4 | 0.98260 | 0.00140 | 0.00140 | 0.51222 | 0.07244 | 0.01750 | 0.00730 |
| 5 | 0.98460 | 0.00140 | 0.00110 | 0.48376 | 0.06841 | 0.01670 | 0.00599 |
| 6 | 0.98250 | 0.00093 | 0.00130 | 0.47938 | 0.06779 | 0.01500 | -0.00433 |
| 7 | 0.98680 | 0.00086 | 0.00099 | 0.48183 | 0.06814 | 0.01510 | 0.00561 |
| 8 | 0.96480 | 0.00210 | 0.00260 | 0.45240 | 0.06398 | 0.01906 | -0.01174 |
| 9 | 0.95520 | 0.00260 | 0.00320 | 0.43564 | 0.06161 | 0.02040 | 0.00338 |
| 10 | 0.93130 | 0.00420 | 0.00480 | 0.42340 | 0.05988 | 0.02908 | -0.00064 |
| 11 | 0.97320 | 0.00170 | 0.00190 | 0.42415 | 0.05998 | 0.02126 | 0.00353 |
| 12 | 0.94780 | 0.00290 | 0.00370 | 0.42418 | 0.05999 | 0.02982 | -0.00639 |
| 13 | 0.98740 | 0.00088 | 0.00093 | 0.44412 | 0.06281 | 0.01388 | -0.00667 |
| 14 | 0.96830 | 0.00200 | 0.00220 | 0.43985 | 0.06220 | 0.02085 | 0.00087 |
| **15** | **0.99440** | **0.00045** | **0.00042** | **0.45818** | **0.06480** | **0.00919** | **-0.00280** |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 16 | 0.99060 | 0.00069 | 0.00068 | 0.44326 | 0.06269 | 0.01164 | -0.00406 |
| 17 | 0.98220 | 0.00160 | 0.00130 | 0.47885 | 0.06772 | 0.01675 | 0.00521 |
| 18 | 0.96940 | 0.00200 | 0.00220 | 0.50918 | 0.07201 | 0.02186 | -0.00113 |
| 19 | 0.98490 | 0.00120 | 0.00120 | 0.48186 | 0.06815 | 0.01722 | 0.00927 |
| 20 | 0.95850 | 0.00250 | 0.00290 | 0.47591 | 0.06730 | 0.02446 | 0.00143 |

Note: MSE = mean square error, SD = standard deviation, SE = standard error, MAD = mean absolute deviation, $ME_y$ = mean error on prediction

Appendix Ie. Comparison of selected optimal model and other ANN models (USA).

| Model | $R^2$ | $MSE_{train}$ | $MSE_{test}$ | SD | SE | MAD | $ME_y$ |
|---|---|---|---|---|---|---|---|
| 1 | 0.84680 | 0.00550 | 0.00690 | 0.41045 | 0.05805 | 0.00533 | -0.00062 |
| 2 | 0.81800 | 0.00810 | 0.00720 | 0.43459 | 0.06146 | 0.00564 | -0.00472 |
| 3 | 0.87480 | 0.00420 | 0.00580 | 0.39801 | 0.05629 | 0.00537 | 0.00254 |
| 4 | 0.85110 | 0.00480 | 0.00650 | 0.38049 | 0.05381 | 0.00567 | 0.00141 |
| 5 | 0.88900 | 0.00370 | 0.00530 | 0.47439 | 0.06709 | 0.00499 | 0.00169 |
| **6** | **0.87210** | **0.00440** | **0.00530** | **0.43302** | **0.06124** | **0.00458** | **0.00118** |
| 7 | 0.89690 | 0.00280 | 0.00550 | 0.49859 | 0.07051 | 0.00486 | 0.00291 |
| 8 | 0.87300 | 0.00410 | 0.00560 | 0.46706 | 0.06605 | 0.00521 | 0.00124 |
| 9 | 0.86730 | 0.00610 | 0.00550 | 0.44389 | 0.06278 | 0.00512 | -0.00027 |
| 10 | 0.86030 | 0.00340 | 0.00730 | 0.46693 | 0.06603 | 0.00636 | 0.00096 |
| 11 | 0.89030 | 0.00300 | 0.00540 | 0.41896 | 0.05925 | 0.00466 | 0.00251 |
| 12 | 0.84240 | 0.00540 | 0.00630 | 0.41731 | 0.05902 | 0.00532 | -0.00034 |
| 13 | 0.83300 | 0.00800 | 0.00680 | 0.44197 | 0.06250 | 0.00518 | -0.00490 |
| 14 | 0.88570 | 0.00280 | 0.00590 | 0.45145 | 0.06384 | 0.00518 | 0.00326 |
| 15 | 0.88220 | 0.00480 | 0.00620 | 0.44029 | 0.06227 | 0.00578 | -0.00075 |
| 16 | 0.84810 | 0.00550 | 0.00600 | 0.42097 | 0.05953 | 0.00448 | -0.00038 |
| 17 | 0.87900 | 0.00320 | 0.00620 | 0.45185 | 0.06390 | 0.00479 | 0.00358 |
| 18 | 0.88500 | 0.00230 | 0.00600 | -0.10243 | 0.46541 | 0.00539 | 0.00368 |
| 19 | 0.85260 | 0.00400 | 0.00690 | 0.49119 | 0.06946 | 0.00544 | 0.00191 |
| 20 | 0.78400 | 0.00300 | 0.00670 | 0.46985 | 0.06645 | 0.03976 | -0.00843 |

Note: MSE = mean square error, SD = standard deviation, SE = standard error, MAD = mean absolute deviation, $ME_y$ = mean error on prediction

1 Appendix IIa. Results of evaluating predictions for $CO_2$ emission intensities on selected optimal models.

| Test samples | Australia | | | | Brazil | | | |
|---|---|---|---|---|---|---|---|---|
| | $y_a$ | $y_p$ | $AD_{9-5-1}$ | $E_y$ | $y_a$ | $y_p$ | $AD_{9-5-1}$ | $E_y$ |
| 1 | 3.286642486 | 3.251830856 | 0.010591852 | 0.034811631 | 1.557095671 | 1.567267068 | 0.006532288 | -0.010171397 |
| 2 | 3.125154889 | 3.077880797 | 0.01512696 | 0.047274092 | 1.707167999 | 1.685288219 | 0.012816418 | 0.021879779 |
| 3 | 3.037133734 | 3.076276916 | 0.012888198 | -0.039143182 | 1.608201304 | 1.60699964 | 0.00074721 | 0.001201664 |
| 4 | 3.233742636 | 3.289377947 | 0.017204619 | -0.055635311 | 1.475355053 | 1.444093351 | 0.021189274 | 0.031261702 |
| 5 | 3.039918372 | 3.026970948 | 0.004259135 | 0.012947424 | 1.541428543 | 1.563372641 | 0.014236209 | -0.021944098 |
| 6 | 3.078738732 | 3.018433468 | 0.019587652 | 0.060305264 | 1.652063616 | 1.664560296 | 0.007564285 | -0.01249668 |
| 7 | 3.057637514 | 3.067132259 | 0.003105255 | -0.009494745 | 1.564609994 | 1.562360404 | 0.001437796 | 0.002249591 |
| 8 | 3.328124784 | 3.26535332 | 0.01886091 | 0.062771464 | 1.404028125 | 1.459460063 | 0.039480646 | -0.055431938 |
| 9 | 3.073089603 | 3.016134279 | 0.018533571 | 0.056955324 | 1.662672303 | 1.665176424 | 0.001506082 | -0.002504121 |
| 10 | 3.251017716 | 3.249962095 | 0.000324705 | 0.001055621 | 1.407687907 | 1.455193057 | 0.033746934 | -0.047505151 |
| 11 | 3.052164158 | 3.02073217 | 0.010298263 | 0.031431988 | 1.633045241 | 1.672776636 | 0.024329635 | -0.039731395 |
| 12 | 3.073111027 | 3.051729348 | 0.006957666 | 0.021381679 | 1.559834321 | 1.591888242 | 0.020549568 | -0.032053922 |
| 13 | 3.1051127 | 3.009504513 | 0.030790569 | 0.095608187 | 1.478753755 | 1.486136695 | 0.004992677 | -0.00738294 |
| 14 | 3.067642054 | 3.068035964 | 0.000128408 | -0.00039391 | 1.534474657 | 1.522429385 | 0.007849769 | 0.012045272 |
| 15 | 3.046993661 | 3.035909155 | 0.00363785 | 0.011084506 | 1.562172938 | 1.554131514 | 0.005147589 | 0.008041424 |
| 16 | 3.362218222 | 3.290196929 | 0.021420767 | 0.072021293 | 1.468488625 | 1.489780906 | 0.014499453 | -0.021292282 |
| 17 | 3.249380725 | 3.287585144 | 0.011757446 | -0.038204419 | 1.452999157 | 1.443124254 | 0.006796221 | 0.009874903 |
| 18 | 3.218645018 | 3.221053761 | 0.000748372 | -0.002408744 | 1.568904846 | 1.556687011 | 0.007787492 | 0.012217834 |
| 19 | 3.053339072 | 3.018070233 | 0.011550908 | 0.035268839 | 1.643779428 | 1.674355585 | 0.018601131 | -0.030576157 |
| 20 | 3.083274299 | 3.094359779 | 0.00359536 | -0.01108548 | 1.534667613 | 1.535503758 | 0.000544838 | -0.000836145 |
| 21 | 3.0758051 | 3.076858286 | 0.00034241 | -0.001053186 | 1.768920753 | 1.724573304 | 0.025070342 | 0.044347448 |
| 22 | 3.097988232 | 3.074213878 | 0.007674127 | 0.023774355 | 1.525899465 | 1.568506383 | 0.027922494 | -0.042606918 |
| 23 | 3.328440301 | 3.207311214 | 0.036392147 | 0.121129088 | 1.469795395 | 1.445266386 | 0.016688724 | 0.02452901 |
| 24 | 3.033172721 | 3.071346465 | 0.012585417 | -0.038173744 | 1.60941944 | 1.605598744 | 0.002373959 | 0.003820696 |
| 25 | 3.017621789 | 3.041308013 | 0.007849302 | -0.023686224 | 1.502641929 | 1.522103477 | 0.012951554 | -0.019461548 |
| 26 | 3.046647468 | 3.0608779 | 0.004670849 | -0.014230432 | 1.559945702 | 1.579782966 | 0.012716637 | -0.019837264 |
| 27 | 3.0417623 | 3.07815298 | 0.011963683 | -0.03639068 | 1.627733054 | 1.652092246 | 0.014965103 | -0.024359192 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 28 | 3.112029743 | 2.996052163 | 0.037267504 | 0.11597758 | 1.587177672 | 1.571643103 | 0.009787542 | 0.015534569 |
| 29 | 3.083874659 | 3.102107491 | 0.005912313 | -0.018232832 | 1.617687941 | 1.589754347 | 0.017267604 | 0.027933594 |
| | | Mean | 0.011931939 | 0.017781567 | | | 0.013451706 | -0.005974264 |

Note: $y_a$ = actual $CO_2$ intensity and $y_p$ = predicted $CO_2$ intensity, AD = absolute deviation, $E_y = y_a - y_p$

Appendix IIb. Results of evaluating predictions for $CO_2$ emission intensities on selected optimal models.

| Test samples | China | | | | India | | | |
|---|---|---|---|---|---|---|---|---|
| | y_test | $y_p$ | $AD_{9\text{-}5\text{-}1}$ | $E_y$ | y_test | $y_p$ | $AD_{9\text{-}5\text{-}1}$ | $E_y$ |
| 1 | 2.4742588 | 2.476217685 | 0.000791706 | -0.001958885 | 1.616775083 | 1.629716932 | 0.008004731 | -0.012941849 |
| 2 | 3.04116614 | 3.006094331 | 0.011532356 | 0.035071809 | 2.286198631 | 2.325464318 | 0.017175099 | -0.039265687 |
| 3 | 3.177417239 | 3.150691432 | 0.008411173 | 0.026725807 | 2.320039421 | 2.355122234 | 0.015121645 | -0.035082813 |
| 4 | 2.895104235 | 2.760091773 | 0.04663475 | 0.135012462 | 1.796703515 | 1.775702718 | 0.011688515 | 0.021000797 |
| 5 | 3.275505221 | 3.297245832 | 0.006637331 | -0.021740611 | 2.462073205 | 2.453567944 | 0.003454512 | 0.008505261 |
| 6 | 3.455590533 | 3.433745292 | 0.00632171 | 0.02184524 | 2.642951673 | 2.608650372 | 0.012978407 | 0.034301302 |
| 7 | 3.158607293 | 3.14742997 | 0.003538687 | 0.011177323 | 2.186125698 | 2.17358942 | 0.005734473 | 0.012536278 |
| 8 | 2.860617925 | 2.796813336 | 0.022304478 | 0.063804588 | 1.736322211 | 1.737024707 | 0.000404588 | -0.000702496 |
| 9 | 3.448835101 | 3.433540262 | 0.004434784 | 0.015294839 | 2.661675841 | 2.622687942 | 0.014647877 | 0.038987899 |
| 10 | 2.649050948 | 2.698519267 | 0.018673978 | -0.049468319 | 1.647965031 | 1.678007653 | 0.018230133 | -0.030042621 |
| 11 | 3.480602117 | 3.431597758 | 0.014079276 | 0.049004359 | 2.558507729 | 2.576938057 | 0.007203546 | -0.018430329 |
| 12 | 3.343434723 | 3.383338484 | 0.01193496 | -0.039903761 | 2.500872208 | 2.490311314 | 0.004222885 | 0.010560895 |
| 13 | 2.807543608 | 2.940877781 | 0.047491399 | -0.133334173 | 2.007911528 | 2.002885412 | 0.002503156 | 0.005026117 |
| 14 | 3.014717229 | 2.963396239 | 0.017023484 | 0.05132099 | 2.060385124 | 2.06252965 | 0.001040837 | -0.002144526 |
| 15 | 3.103787048 | 3.140877455 | 0.011950049 | -0.037090407 | 2.144148797 | 2.147372376 | 0.001503431 | -0.003223579 |
| 16 | 2.985550082 | 2.983903498 | 0.000551518 | 0.001646584 | 1.87473487 | 1.862534407 | 0.006507834 | 0.012200463 |
| 17 | 2.872467409 | 2.742890364 | 0.045110014 | 0.129577045 | 1.76787526 | 1.756087159 | 0.006667949 | 0.011788102 |
| 18 | 2.546181743 | 2.565442968 | 0.007564748 | -0.019261224 | 1.591922905 | 1.631702439 | 0.024988355 | -0.039779534 |
| 19 | 3.482958909 | 3.433740008 | 0.014131347 | 0.049218901 | 2.583927212 | 2.595336286 | 0.004415401 | -0.011409075 |
| 20 | 3.082007333 | 3.052040145 | 0.00972327 | 0.029967188 | 2.130558157 | 2.132698979 | 0.001004817 | -0.002140822 |
| 21 | 2.950989792 | 3.009949746 | 0.019979722 | -0.058959955 | 2.325705068 | 2.335725201 | 0.004308428 | -0.010020133 |
| 22 | 3.389394276 | 3.355821091 | 0.009905364 | 0.033573186 | 2.639833037 | 2.594654779 | 0.017114059 | 0.045178257 |
| 23 | 2.989548173 | 2.939642315 | 0.016693445 | 0.049905858 | 1.967782845 | 1.960001069 | 0.003954591 | 0.007781777 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 24 | 3.167345389 | 3.135017181 | 0.01020672 | 0.032328208 | 2.318672702 | 2.3458725 | 0.011730762 | -0.027199797 |
| 25 | 2.827133515 | 2.952901646 | 0.044486095 | -0.125768131 | 2.042094307 | 2.037462032 | 0.002268394 | 0.004632275 |
| 26 | 3.420218576 | 3.338701059 | 0.023834008 | 0.081517517 | 2.619170524 | 2.606677457 | 0.004769856 | 0.012493067 |
| 27 | 3.194873351 | 3.177663486 | 0.005386712 | 0.017209864 | 2.213806995 | 2.250382457 | 0.016521522 | -0.036575462 |
| 28 | 3.065350205 | 3.068575734 | 0.001052254 | -0.003225528 | 2.181732939 | 2.243029092 | 0.028095168 | -0.061296153 |
| 29 | 3.239661906 | 3.255905802 | 0.005014072 | -0.016243897 | 2.364280606 | 2.340317968 | 0.010135276 | 0.023962638 |
| | | Mean | 0.0153586 | 0.011284375 | | | 0.009186077 | -0.00280344 |

Note: $y_a$ = actual $CO_2$ intensity and $y_p$ = predicted $CO_2$ intensity, AD = absolute deviation, $E_y = y_a - y_p$

Appendix IIc. Results of evaluating predictions for $CO_2$ emission intensities on selected optimal models.

| Test samples | USA | | | |
|---|---|---|---|---|
| | y_test | $y_p$ | $AD_{9-5-1}$ | $E_y$ |
| 1 | 2.574457916 | 2.58678845 | 0.004789565 | -0.012330535 |
| 2 | 2.500005366 | 2.503391864 | 0.001354597 | -0.003386499 |
| 3 | 2.494462043 | 2.504325183 | 0.003954015 | -0.00986314 |
| 4 | 2.538573152 | 2.532567461 | 0.002365774 | 0.006005691 |
| 5 | 2.476437727 | 2.479761468 | 0.001342146 | -0.003323741 |
| 6 | 2.384014635 | 2.401271511 | 0.007238578 | -0.017256875 |
| 7 | 2.490700071 | 2.497796243 | 0.002849067 | -0.007096172 |
| 8 | 2.530982965 | 2.540565841 | 0.003786227 | -0.009582877 |
| 9 | 2.375633112 | 2.400213481 | 0.010346871 | -0.02458037 |
| 10 | 2.558776552 | 2.559767265 | 0.000387182 | -0.000990712 |
| 11 | 2.411677746 | 2.405923085 | 0.002386165 | 0.005754661 |
| 12 | 2.437965418 | 2.42529669 | 0.005196435 | 0.012668729 |
| 13 | 2.527679338 | 2.528293237 | 0.00024287 | -0.000613898 |
| 14 | 2.498110568 | 2.51145488 | 0.005341762 | -0.013344312 |
| 15 | 2.510632326 | 2.498814312 | 0.004707186 | 0.011818014 |
| 16 | 2.549300685 | 2.548590773 | 0.000278473 | 0.000709912 |
| 17 | 2.536693642 | 2.531737129 | 0.001953926 | 0.004956512 |
| 18 | 2.56326721 | 2.578462067 | 0.005927926 | -0.015194858 |
| 19 | 2.402602563 | 2.403664347 | 0.000441931 | -0.001061784 |

| | | | | |
|---|---|---|---|---|
| 20 | 2.495939797 | 2.499950654 | 0.001606952 | -0.004010857 |
| 21 | 2.508659751 | 2.495357814 | 0.005302408 | 0.013301937 |
| 22 | 2.438174074 | 2.423021859 | 0.006214575 | 0.015152215 |
| 23 | 2.546923136 | 2.545363046 | 0.000612539 | 0.00156009 |
| 24 | 2.497267631 | 2.50378933 | 0.002611534 | -0.006521699 |
| 25 | 2.509936628 | 2.515705771 | 0.002298522 | -0.005769143 |
| 26 | 2.463329963 | 2.44271247 | 0.008369765 | 0.020617493 |
| 27 | 2.481426423 | 2.494025421 | 0.005077321 | -0.012598998 |
| 28 | 2.497681992 | 2.40902637 | 0.03549516 | 0.088655622 |
| 29 | 2.498347921 | 2.497707739 | 0.000256242 | 0.000640182 |
| | | Mean | 0.004577094 | 0.001183262 |

Note: $y_a$ = actual $CO_2$ intensity and $y_p$ = predicted $CO_2$ intensity, AD = absolute deviation, $E_y = y_a - y_p$

**Acknowledgements**

The authors wish to deeply thank the Editorial Board of the journal for their time and support. We are also thankful to the two anonymous reviewers for their valuable comments that help us to improve the quality of this paper. Nevertheless, the authors are only responsible for all remaining errors or imperfections.

**References**

Abidoye, R. B., & Chan, A. P. (2018). Improving property valuation accuracy: a comparison of hedonic pricing model and artificial neural network. *Pacific Rim Property Research Journal, 24*(1), 71-83.

Acheampong, A. O. (2018). Economic growth, $CO_2$ emissions and energy consumption: What causes what and where? *Energy Economics, 74*, 677-692. doi:https://doi.org/10.1016/j.eneco.2018.07.022

Adnan Aziz, M., & Dar, H. A. (2006). Predicting corporate bankruptcy: where we stand? *Corporate Governance: The international journal of business in society, 6*(1), 18-33.

Ahmad, N., Du, L., Lu, J., Wang, J., Li, H.-Z., & Hashmi, M. Z. (2017). Modelling the CO2 emissions and economic growth in Croatia: Is there any environmental Kuznets curve? *Energy, 123*, 164-172. doi:10.1016/j.energy.2016.12.106

Ahmadi, M. A. (2011). Prediction of asphaltene precipitation using artificial neural network optimized by imperialist competitive algorithm. *Journal of Petroleum Exploration and Production Technology, 1*(2), 99-106. doi:10.1007/s13202-011-0013-7

Ahmadi, M. A., Soleimani, R., Lee, M., Kashiwao, T., & Bahadori, A. (2015). Determination of oil well production performance using artificial neural network (ANN) linked to the particle swarm optimization (PSO) tool. *Petroleum, 1*(2), 118-132. doi:https://doi.org/10.1016/j.petlm.2015.06.004

Ahmed, K., Rehman, M. U., & Ozturk, I. (2017). What drives carbon dioxide emissions in the long-run? Evidence from selected South Asian Countries. *Renewable and Sustainable Energy Reviews, 70*, 1142-1153. doi:https://doi.org/10.1016/j.rser.2016.12.018

Alaka, H. A., Oyedele, L. O., Owolabi, H. A., Kumar, V., Ajayi, S. O., Akinade, O. O., & Bilal, M. (2018). Systematic review of bankruptcy prediction models: Towards a framework for tool selection. *Expert Systems with Applications, 94*, 164-184. doi:10.1016/j.eswa.2017.10.040

Almeida, T. A. d. N., Cruz, L., Barata, E., & García-Sánchez, I.-M. (2017). Economic growth and environmental impacts: An analysis based on a composite index of environmental damage. *Ecological Indicators, 76*, 119-130. doi:https://doi.org/10.1016/j.ecolind.2016.12.028

Al-Mulali, U., & Ozturk, I. (2015). The effect of energy consumption, urbanization, trade openness, industrial output, and the political stability on the environmental degradation in the MENA (Middle East and North African) region. *Energy, 84*, 382-389. doi:https://doi.org/10.1016/j.energy.2015.03.004

Al-Mulali, U., Ozturk, I., & Lean, H. H. (2015). The influence of economic growth, urbanization, trade openness, financial development, and renewable energy on pollution in Europe. *Natural Hazards, 79*(1), 621-644. doi:10.1007/s11069-015-1865-9

Al-Mulali, U., Ozturk, I., & Solarin, S. A. (2016). Investigating the environmental Kuznets curve hypothesis in seven regions: The role of renewable energy. *Ecological Indicators, 67*, 267-282. doi:https://doi.org/10.1016/j.ecolind.2016.02.059

Al-Mulali, U., Tang, C. F., & Ozturk, I. (2015). Does financial development reduce environmental degradation? Evidence from a panel study of 129 countries. *Environ Sci Pollut Res Int, 22*(19), 14891-14900. doi:10.1007/s11356-015-4726-x

Alvanitopoulos, P. F., Andreadis, I., & Elenas, A. (2010). Neuro-fuzzy techniques for the classification of earthquake damages in buildings. *Measurement, 43*(6), 797-809. doi:10.1016/j.measurement.2010.02.011

Antweiler, W., Copeland, B. R., & Taylor, M. S. (2001). Is Free Trade Good for the Environment? *The American economic review, 91*(4), 877-908.

Apergis, N., & Ozturk, I. (2015). Testing Environmental Kuznets Curve hypothesis in Asian countries. *Ecological Indicators, 52*(Supplement C), 16-22. doi:https://doi.org/10.1016/j.ecolind.2014.11.026

Awaworyi Churchill, S., Inekwe, J., Smyth, R., & Zhang, X. (2019). R&D intensity and carbon emissions in the G7: 1870–2014. *Energy Economics, 80*, 30-37. doi:https://doi.org/10.1016/j.eneco.2018.12.020

Begum, R. A., Sohag, K., Abdullah, S. M. S., & Jaafar, M. (2015). CO2 emissions, energy consumption, economic and population growth in Malaysia. *Renewable and Sustainable Energy Reviews, 41*, 594-601. doi:https://doi.org/10.1016/j.rser.2014.07.205

Bekhet, H. A., & Othman, N. S. (2017). Impact of urbanization growth on Malaysia CO2 emissions: Evidence from the dynamic relationship. *Journal of Cleaner Production, 154*, 374-388. doi:https://doi.org/10.1016/j.jclepro.2017.03.174

Ben Jebli, M., Ben Youssef, S., & Ozturk, I. (2016). Testing environmental Kuznets curve hypothesis: The role of renewable and non-renewable energy consumption and trade in OECD countries. Ecological Indicators, 60, 824-831. doi:https://doi.org/10.1016/j.ecolind.2015.08.031

Birdsall, N. (1992). *Another look at population and global warming* (Vol. 1020): World Bank Publications.

Boateng, E. B., Pillay, M., & Davis, P. (2019). Predicting the Level of Safety Performance Using an Artificial Neural Network. In T. Ahram, W. Karwowski, & R. Taiar (Eds.), *Human Systems Engineering and Design* (pp. 705-710). Cham: Springer International Publishing.

Boussabaine, A. H. (1996). The use of artificial neural networks in construction management: A review. *Construction Management and Economics, 14*(5), 427-436. doi:10.1080/014461996373296

Boutabba, M. A. (2014). The impact of financial development, income, energy and trade on carbon emissions: Evidence from the Indian economy. *Economic Modelling, 40*, 33-41. doi:10.1016/j.econmod.2014.03.005

Breheny, M. (2001). Densities and sustainable cities: the UK experience. *Cities for the new millennium*, 39-51.

Burton, E. (2000). The Compact City: Just or Just Compact? A Preliminary Analysis. *Urban Studies, 37*(11), 1969-2006. doi:10.1080/00420980050162184

Capello, R., & Camagni, R. (2000). Beyond Optimal City Size: An Evaluation of Alternative Urban Growth Patterns. *Urban Studies, 37*(9), 1479-1496. doi:10.1080/00420980020080221

Colwell, J. D. (1994). *Estimating fertilizer requirements: A Quantitative approach*: CAB International, Oxon (EUA).

Deb, C., Zhang, F., Yang, J., Lee, S. E., & Shah, K. W. (2017). A review on time series forecasting techniques for building energy consumption. *Renewable and Sustainable Energy Reviews, 74*, 902-924. doi:https://doi.org/10.1016/j.rser.2017.02.085

Debnath, K. B., & Mourshed, M. (2018). Forecasting methods in energy planning models. *Renewable and Sustainable Energy Reviews, 88,* 297-325. doi:https://doi.org/10.1016/j.rser.2018.02.002

Destek, M. A., & Sarkodie, S. A. (2019). Investigation of environmental Kuznets curve for ecological footprint: The role of energy and financial development. Science of The Total Environment, 650, 2483-2489. doi:https://doi.org/10.1016/j.scitotenv.2018.10.017

Detienne, K. B., Detienne, D. H., & Joshi, S. A. (2003). Neural networks as statistical tools for business researchers. *Organizational Research Methods, 6*(2), 236-265.

Dinda, S. (2004). Environmental Kuznets Curve Hypothesis: A Survey. *Ecological Economics, 49*(4), 431-455. doi:https://doi.org/10.1016/j.ecolecon.2004.02.011

Ding, S., Dang, Y.-G., Li, X.-M., Wang, J.-J., & Zhao, K. (2017). Forecasting Chinese $CO_2$ emissions from fuel combustion using a novel grey multivariable model. *Journal of Cleaner Production, 162,* 1527-1538. doi:https://doi.org/10.1016/j.jclepro.2017.06.167

Dogan, E., & Turkekul, B. (2016). CO2 emissions, real output, energy consumption, trade, urbanization and financial development: testing the EKC hypothesis for the USA. *Environ Sci Pollut Res Int, 23*(2), 1203-1213. doi:10.1007/s11356-015-5323-8

Dong, K., Hochman, G., Zhang, Y., Sun, R., Li, H., & Liao, H. (2018). CO2 emissions, economic and population growth, and renewable energy: Empirical evidence across regions. *Energy Economics, 75,* 180-192. doi:https://doi.org/10.1016/j.eneco.2018.08.017

El Kadi, H. (2006). Modeling the mechanical behavior of fiber-reinforced polymeric composite materials using artificial neural networks - A review. *Composite Structures, 73*(1), 1-23. doi:10.1016/j.compstruct.2005.01.020

Falat, L., & Pancikova, L. (2015). Quantitative Modelling in Economics with Advanced Artificial Neural Networks. *Procedia Economics and Finance, 34,* 194-201. doi:https://doi.org/10.1016/S2212-5671(15)01619-6

Fernández Fernández, Y., Fernández López, M. A., & Olmedillas Blanco, B. (2018). Innovation for sustainability: The impact of R&D spending on CO2 emissions. *Journal of Cleaner Production, 172,* 3459-3467. doi:https://doi.org/10.1016/j.jclepro.2017.11.001

Gallo, C., Contò, F., & Fiore, M. (2014). A neural network model for forecasting $CO_2$ Emission. *Agris on-line Papers in Economics and Informatics, 6*(2), 31.

Ghaedi, A. M., & Vafaei, A. (2017). Applications of artificial neural networks for adsorption removal of dyes from aqueous solution: A review. *Advances in Colloid and Interface Science, 245,* 20-39. doi:10.1016/j.cis.2017.04.015

Ghani, G. M. (2012). Does trade liberalization effect energy consumption? *Energy Policy, 43,* 285-290. doi:10.1016/j.enpol.2012.01.005

Ghritlahre, H. K., & Prasad, R. K. (2018). Application of ANN technique to predict the performance of solar collector systems - A review. *Renewable and Sustainable Energy Reviews, 84,* 75-88. doi:10.1016/j.rser.2018.01.001

Gouldson, A., & Murphy, J. (1997). Ecological modernization: economic restructuring and the environment. *The Political Quarterly, 68*(5), 74-86.

Grossman, G. M., & Helpman, E. (1991). *Innovation and Growth in the Global Economy.* Cambridge, MA: MIT Press.

Grossman, G. M., & Krueger, A. B. (1995). Economic growth and the environment. *The Quarterly Journal of Economics, 110*(2), 353-377. doi:10.2307/2118443

Grove, R. H. (1992). Origins of Western Environmentalism. *Scientific American, 267*(1), 42-47.

Hagan, M. T., Demuth, H. B., Beale, M. H., & De Jesús, O. (1996). *Neural network design* (Vol. 20): Pws Pub. Boston.
Halicioglu, F. (2009). An econometric study of CO2 emissions, energy consumption, income and foreign trade in Turkey. *Energy Policy, 37*(3), 1156-1164. doi:https://doi.org/10.1016/j.enpol.2008.11.012
Hecht-Nielsen, R. (1987). *Kolmogorov's mapping neural network existence theorem.* Paper presented at the Proceedings of the international conference on Neural Networks.
Hippert, H. S., Pedreira, C. E., & Souza, R. C. (2001). Neural networks for short-term load forecasting: A review and evaluation. *IEEE Transactions on power systems*, *16*(1), 44-55.
Hussain, A., & Reynolds, W. (1975). Measurements in fully developed turbulent channel flow. *Journal of Fluids Engineering*, *97*(4), 568-578.
International Energy Agency. (2018). *Global energy and $CO_2$ emissions status report*. Retrieved from https://www.iea.org/geco/emissions/
Jahangir Alam, M., Ara Begum, I., Buysse, J., & Van Huylenbroeck, G. (2012). Energy consumption, carbon emissions and economic growth nexus in Bangladesh: Cointegration and dynamic causality analysis. *Energy Policy*, *45*, 217-225. doi:10.1016/j.enpol.2012.02.022
Jani, D. B., Mishra, M., & Sahoo, P. K. (2017). Application of artificial neural network for predicting performance of solid desiccant cooling systems – A review. *Renewable and Sustainable Energy Reviews*, *80*, 352-366. doi:10.1016/j.rser.2017.05.169
Jebaraj, S., & Iniyan, S. (2006). A review of energy models. *Renewable and Sustainable Energy Reviews*, *10*(4), 281-311. doi:https://doi.org/10.1016/j.rser.2004.09.004
Jiao, J., Jiang, G., & Yang, R. (2018). Impact of R&D technology spillovers on carbon emissions between China's regions. Structural Change and Economic Dynamics.
Kalogirou, S. A. (2000). Artificial neural networks in renewable energy systems applications: A review. *Renewable and Sustainable Energy Reviews, 5*(4), 373-401. doi:10.1016/S1364-0321(01)00006-5
Kara, Y., Acar Boyacioglu, M., & Baykan, Ö. K. (2011). Predicting direction of stock price index movement using artificial neural networks and support vector machines: The sample of the Istanbul Stock Exchange. *Expert Systems with Applications*, *38*(5), 5311-5319. doi:https://doi.org/10.1016/j.eswa.2010.10.027
Karlaftis, M. G., & Vlahogianni, E. I. (2011). Statistical methods versus neural networks in transportation research: Differences, similarities and some insights. *Transportation Research Part C: Emerging Technologies*, *19*(3), 387-399.
Kartalopoulos, S. V., & Kartakapoulos, S. V. (1997). *Understanding neural networks and fuzzy logic: basic concepts and applications*: Wiley-IEEE Press.
Khoshroo, A., Emrouznejad, A., Ghaffarizadeh, A., Kasraei, M., & Omid, M. (2018). Sensitivity analysis of energy inputs in crop production using artificial neural networks. *Journal of Cleaner Production, 197*, 992-998. doi:https://doi.org/10.1016/j.jclepro.2018.05.249
Kohzadi, N., Boyd, M. S., Kermanshahi, B., & Kaastra, I. (1996). A comparison of artificial neural network and time series models for forecasting commodity prices. *Neurocomputing*, *10*(2), 169-181. doi:https://doi.org/10.1016/0925-2312(95)00020-8
Köne, A. Ç., & Büke, T. (2010). Forecasting of $CO_2$ emissions from fuel combustion using trend analysis. *Renewable and Sustainable Energy Reviews, 14*(9), 2906-2915. doi:https://doi.org/10.1016/j.rser.2010.06.006
Lam, K. C., Yu, C., & Lam, K. (2008). An artificial neural network and entropy model for residential property price forecasting in Hong Kong. *Journal of Property Research, 25*(4), 321-342.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. nature 521 (7553): 436. *Nature, 521*(7553), 436-444.

Lee, J. W. (2013). The contribution of foreign direct investment to clean energy use, carbon emissions and economic growth. *Energy Policy, 55*, 483-489. doi:https://doi.org/10.1016/j.enpol.2012.12.039

Li, K., & Lin, B. (2015). Impacts of urbanization and industrialization on energy consumption/CO2 emissions: Does the level of development matter? *Renewable and Sustainable Energy Reviews, 52*, 1107-1122. doi:https://doi.org/10.1016/j.rser.2015.07.185

Lin, C.-S., Liou, F.-M., & Huang, C.-P. (2011). Grey forecasting model for $CO_2$ emissions: A Taiwan study. *Applied Energy, 88*(11), 3816-3820. doi:https://doi.org/10.1016/j.apenergy.2011.05.013

Liu, X., & Bae, J. (2018). Urbanization and industrialization impact of CO2 emissions in China. *Journal of Cleaner Production, 172*, 178-186. doi:https://doi.org/10.1016/j.jclepro.2017.10.156

Maier, H. R., & Dandy, G. C. (2001). Neural network based modelling of environmental variables: a systematic approach. *Mathematical and Computer Modelling, 33*(6-7), 669-682.

Maji, I. K., Habibullah, M. S., & Saari, M. Y. (2017). Financial development and sectoral CO2 emissions in Malaysia. *Environ Sci Pollut Res Int, 24*(8), 7160-7176. doi:10.1007/s11356-016-8326-1

Marino, S., Hogue, I. B., Ray, C. J., & Kirschner, D. E. (2008). A methodology for performing global uncertainty and sensitivity analysis in systems biology. *Journal of Theoretical Biology, 254*(1), 178-196. doi:https://doi.org/10.1016/j.jtbi.2008.04.011

McGranahan, G. (2010). *The citizens at risk: from urban sanitation to sustainable cities*: Earthscan.

Meadows, D., Randers, J., & Meadows, D. (1992). *The limits to growth*. New York: Universe Books.

Mobley, B. A., Leasure, R., & Davidson, L. (1995). Artificial neural network predictions of lengths of stay on a post-coronary care unit. *Heart & Lung: The Journal of Acute and Critical Care, 24*(3), 251-256.

Moghaddam, A. H., Moghaddam, M. H., & Esfandyari, M. (2016). Stock market index prediction using artificial neural network. *Journal of Economics, Finance and Administrative Science, 21*(41), 89-93. doi:https://doi.org/10.1016/j.jefas.2016.07.002

Mol, A. P. J., & Spaargaren, G. (2000). Ecological modernisation theory in debate: A review. *Environmental Politics, 9*(1), 17-49. doi:10.1080/09644010008414511

Morano, P., Tajani, F., & Torre, C. M. (2015). Artificial intelligence in property valuations: an application of artificial neural networks to housing appraisal. *Advances in Environmental Science and Energy, Planning*, 23-29.

Narayan, P. K., & Narayan, S. (2010). Carbon dioxide emissions and economic growth: Panel data evidence from developing countries. *Energy Policy, 38*(1), 661-666. doi:https://doi.org/10.1016/j.enpol.2009.09.005

Omri, A. (2013). CO2 emissions, energy consumption and economic growth nexus in MENA countries: Evidence from simultaneous equations models. *Energy Economics, 40*, 657-664. doi:10.1016/j.eneco.2013.09.003

Omri, A., Daly, S., Rault, C., & Chaibi, A. (2015). Financial development, environmental quality, trade and economic growth: What causes what in MENA countries. *Energy Economics, 48*, 242-252. doi:10.1016/j.eneco.2015.01.008

Pao, H.-T., & Tsai, C.-M. (2011). Modeling and forecasting the $CO_2$ emissions, energy consumption, and economic growth in Brazil. *Energy, 36*(5), 2450-2458. doi:https://doi.org/10.1016/j.energy.2011.01.032

Pao, H.-T., & Tsai, C.-M. (2011). Multivariate Granger causality between CO2 emissions, energy consumption, FDI (foreign direct investment) and GDP (gross domestic product): Evidence from a panel of BRIC (Brazil, Russian Federation, India, and China) countries. *Energy, 36*(1), 685-693. doi:https://doi.org/10.1016/j.energy.2010.09.041

Pao, H.-T., Fu, H.-C., & Tseng, C.-L. (2012). Forecasting of $CO_2$ emissions, energy consumption and economic growth in China using an improved grey model. *Energy, 40*(1), 400-409. doi:https://doi.org/10.1016/j.energy.2012.01.037

Patel, D. A., & Jha, K. N. (2015). Neural network model for the prediction of safe work behavior in construction projects. *Journal of Construction Engineering and Management, 141*(1). doi:10.1061/(ASCE)CO.1943-7862.0000922

Patel, D. A., & Jha, K. N. (2016). Evaluation of construction projects based on the safe work behavior of co-employees through a neural network model. *Safety Science, 89*, 240-248. doi:10.1016/j.ssci.2016.06.020

Pérez-Sánchez, B., Fontenla-Romero, O., & Guijarro-Berdiñas, B. (2016). A review of adaptive online learning for artificial neural networks. *Artificial Intelligence Review, 49*(2), 281-299. doi:10.1007/s10462-016-9526-2

Poumanyvong, P., & Kaneko, S. (2010). Does urbanization lead to less energy use and lower $CO_2$ emissions? A cross-country analysis. Ecological Economics, 70(2), 434-444. doi:10.1016/j.ecolecon.2010.09.029

Prybutok, V. R., Yi, J., & Mitchell, D. (2000). Comparison of neural network models with ARIMA and regression models for prediction of Houston's daily maximum ozone concentrations. *European Journal of Operational Research, 122*(1), 31-40. doi:https://doi.org/10.1016/S0377-2217(99)00069-7

Ren, S., Yuan, B., Ma, X., & Chen, X. (2014). International trade, FDI (foreign direct investment) and embodied $CO_2$ emissions: A case study of Chinas industrial sectors. China Economic Review, 28, 123-134. doi:https://doi.org/10.1016/j.chieco.2014.01.003

Rogers, L. L., & Dowla, F. U. (1994). Optimization of groundwater remediation using artificial neural networks with parallel solute transport modeling. *Water Resources Research, 30*(2), 457-481.

Rudlin, D., & Falk, N. (1999). *Building the 21st century home: the sustainable urban neighbourhood*: Butterworth-Heinemann.

Saboori, B., Sulaiman, J., & Mohd, S. (2012). Economic growth and $CO_2$ emissions in Malaysia: A cointegration analysis of the Environmental Kuznets Curve. Energy Policy, 51(Supplement C), 184-191. doi:https://doi.org/10.1016/j.enpol.2012.08.065

Sadorsky, P. (2010). The impact of financial development on energy consumption in emerging economies. *Energy Policy, 38*(5), 2528-2535. doi:10.1016/j.enpol.2009.12.048

Sadorsky, P. (2011). Financial development and energy consumption in Central and Eastern European frontier economies. *Energy Policy, 39*(2), 999-1006. doi:10.1016/j.enpol.2010.11.034

Sadorsky, P. (2013). Do urbanization and industrialization affect energy intensity in developing countries? *Energy Economics, 37*, 52-59. doi:https://doi.org/10.1016/j.eneco.2013.01.009

Sadorsky, P. (2014). The effect of urbanization on $CO_2$ emissions in emerging economies. Energy Economics, 41, 147-153. doi:https://doi.org/10.1016/j.eneco.2013.11.007

Safa, M., & Samarasinghe, S. (2011). Determination and modelling of energy consumption in wheat production using neural networks: "A case study in Canterbury province, New Zealand". *Energy, 36*(8), 5140-5147. doi:https://doi.org/10.1016/j.energy.2011.06.016

Safa, M., Nejat, M., Nuthall, P. L., & Greig, B. J. (2016). Predicting $CO_2$ emissions from farm inputs in wheat production using artificial neural networks and linear regression models-Case study in Canterbury, New Zealand.

Salahuddin, M., Alam, K., Ozturk, I., & Sohag, K. (2018). The effects of electricity consumption, economic growth, financial development and foreign direct investment on CO2 emissions in Kuwait. *Renewable and Sustainable Energy Reviews, 81*, 2002-2010. doi:https://doi.org/10.1016/j.rser.2017.06.009

Saltelli, A., & Marivoet, J. (1990). Non-parametric statistics in sensitivity analysis for model output: A comparison of selected techniques. *Reliability Engineering & System Safety, 28*(2), 229-253. doi:https://doi.org/10.1016/0951-8320(90)90065-U

Sarkodie, S. A., & Strezov, V. (2018). Assessment of contribution of Australia's energy production to $CO_2$ emissions and environmental degradation using statistical dynamic approach. Science of The Total Environment, 639, 888-899. doi:https://doi.org/10.1016/j.scitotenv.2018.05.204

Sarkodie, S. A., & Strezov, V. (2019). Effect of foreign direct investments, economic development and energy consumption on greenhouse gas emissions in developing countries. Science of The Total Environment, 646, 862-871. doi:https://doi.org/10.1016/j.scitotenv.2018.07.365

Sbia, R., Shahbaz, M., & Hamdi, H. (2014). A contribution of foreign direct investment, clean energy, trade openness, carbon emissions and economic growth to energy demand in UAE. *Economic Modelling, 36*, 191-197. doi:10.1016/j.econmod.2013.09.047

Sehrawat, M., Giri, A. K., & Mohapatra, G. (2015). The impact of financial development, economic growth and energy consumption on environmental degradation. *Management of Environmental Quality: An International Journal, 26*(5), 666-682. doi:10.1108/meq-05-2014-0063

Shahbaz, M., Hoang, T. H. V., Mahalik, M. K., & Roubaud, D. (2017). Energy consumption, financial development and economic growth in India: New evidence from a nonlinear and asymmetric analysis. *Energy Economics, 63*, 199-212. doi:https://doi.org/10.1016/j.eneco.2017.01.023

Shahbaz, M., Hye, Q. M. A., Tiwari, A. K., & Leitão, N. C. (2013). Economic growth, energy consumption, financial development, international trade and CO2 emissions in Indonesia. *Renewable and Sustainable Energy Reviews, 25*, 109-121. doi:https://doi.org/10.1016/j.rser.2013.04.009

Shahbaz, M., Kumar Tiwari, A., & Nasir, M. (2013). The effects of financial development, economic growth, coal consumption and trade openness on CO2 emissions in South Africa. *Energy Policy, 61*, 1452-1459. doi:10.1016/j.enpol.2013.07.006

Shahbaz, M., Nasir, M. A., & Roubaud, D. (2018). Environmental degradation in France: The effects of FDI, financial development, and energy innovations. Energy Economics, 74, 843-857. doi:https://doi.org/10.1016/j.eneco.2018.07.020

Shahbaz, M., Nasreen, S., Abbas, F., & Anis, O. (2015). Does foreign direct investment impede environmental quality in high-, middle-, and low-income countries? *Energy Economics, 51*, 275-287. doi:https://doi.org/10.1016/j.eneco.2015.06.014

Shahbaz, M., Shahzad, S. J. H., Ahmad, N., & Alam, S. (2016). Financial development and environmental quality: The way forward. *Energy Policy, 98*, 353-364. doi:10.1016/j.enpol.2016.09.002

Shahin, M., & Elchalakani, M. (2008). Neural networks for modelling ultimate pure bending of steel circular tubes. *Journal of Constructional Steel Research, 64*(6), 624-633.

Shi, A. (2003). The impact of population pressure on global carbon dioxide emissions, 1975–1996: evidence from pooled cross-country data. *Ecological Economics, 44*(1), 29-42. doi:https://doi.org/10.1016/S0921-8009(02)00223-9

Soltanali, H., Nikkhah, A., & Rohani, A. (2017). Energy audit of Iranian kiwifruit production using intelligent systems. *Energy, 139*, 646-654. doi:https://doi.org/10.1016/j.energy.2017.08.010

Sözen, A. (2009). Future projection of the energy dependency of Turkey using artificial neural network. *Energy Policy, 37*(11), 4827-4833. doi:https://doi.org/10.1016/j.enpol.2009.06.040

Stamenković, L. J., Antanasijević, D. Z., Ristić, M. Đ., Perić-Grujić, A. A., & Pocajt, V. V. (2015). Modeling of methane emissions using artificial neural network approach. *Journal of the Serbian Chemical Society, 80*(3), 421-433.

Stanley, H. E. (1997). Turbulence: The legacy of an kolmogorov. *Journal of Statistical Physics, 88*(1-2), 521-523.

StatSoft Inc. (2008). A Short Course in Data Mining. In: StatSoft, Inc.

Stern, D. I. (2004). The Rise and Fall of the Environmental Kuznets Curve. *World Development, 32*(8), 1419-1439. doi:https://doi.org/10.1016/j.worlddev.2004.03.004

Stern, D. I., & Common, M. S. (2001). Is There an Environmental Kuznets Curve for Sulfur? *Journal of Environmental Economics and Management, 41*(2), 162-178. doi:https://doi.org/10.1006/jeem.2000.1132

Tadesse, Z., Patel, K., Chaudhary, S., & Nagpal, A. (2012). Neural networks for prediction of deflection in composite bridges. *Journal of Constructional Steel Research, 68*(1), 138-149.

Tamazian, A., Chousa, J. P., & Vadlamannati, K. C. (2009). Does higher economic and financial development lead to environmental degradation: Evidence from BRIC countries. Energy Policy, 37(1), 246-253. doi:10.1016/j.enpol.2008.08.025

Tamazian, A., & Bhaskara Rao, B. (2010). Do economic, financial and institutional developments matter for environmental degradation? Evidence from transitional economies. Energy Economics, 32(1), 137-145. doi:10.1016/j.eneco.2009.04.004

Valipour, M., Banihabib, M. E., & Behbahani, S. M. R. (2013). Comparison of the ARMA, ARIMA, and the autoregressive artificial neural network models in forecasting the monthly inflow of Dez dam reservoir. *Journal of Hydrology, 476*, 433-441. doi:https://doi.org/10.1016/j.jhydrol.2012.11.017

Wang, Q., Wu, S.-d., Zeng, Y.-e., & Wu, B.-w. (2016). Exploring the relationship between urbanization, energy consumption, and CO2 emissions in different provinces of China. *Renewable and Sustainable Energy Reviews, 54*, 1563-1579. doi:https://doi.org/10.1016/j.rser.2015.10.090

Wang, Z., Shi, C., Li, Q., & Wang, G. (2011). Impact of Heavy Industrialization on the Carbon Emissions: An Empirical Study of China. Energy Procedia, 5, 2610-2616. doi:https://doi.org/10.1016/j.egypro.2011.03.324

Weber, H., & Sciubba, J. D. (2018). The Effect of Population Growth on the Environment: Evidence from European Regions. *European Journal of Population*. doi:10.1007/s10680-018-9486-0

World Bank. (2016). The world development indicators. Retrieved 15th August, 2017 https://data.worldbank.org/data-catalog/world-development-indicators

Wu, L., Liu, S., Liu, D., Fang, Z., & Xu, H. (2015). Modelling and forecasting $CO_2$ emissions in the BRICS (Brazil, Russia, India, China, and South Africa) countries using a novel multi-variable grey model. *Energy, 79*, 489-495. doi:https://doi.org/10.1016/j.energy.2014.11.052

Wu, Y., Shen, J., Zhang, X., Skitmore, M., & Lu, W. (2016). The impact of urbanization on carbon emissions in developing countries: a Chinese study based on the U-Kaya method. *Journal of Cleaner Production, 135*, 589-603. doi:https://doi.org/10.1016/j.jclepro.2016.06.121

Yadav, A. K., & Chandel, S. S. (2014). Solar radiation prediction using Artificial Neural Network techniques: A review. *Renewable and Sustainable Energy Reviews*, *33*, 772-781. doi:10.1016/j.rser.2013.08.055.

Yin, M.-S., & Tang, H.-W. V. (2013). On the fit and forecasting performance of grey prediction models for China's labor formation. *Mathematical and Computer Modelling, 57*(3), 357-365. doi:https://doi.org/10.1016/j.mcm.2012.06.013

Zhang, C., & Lin, Y. (2012). Panel estimation for urbanization, energy consumption and CO2 emissions: A regional analysis in China. *Energy Policy, 49*, 488-498. doi:https://doi.org/10.1016/j.enpol.2012.06.048

Zhang, C., & Zhou, X. (2016). Does foreign direct investment lead to lower $CO_2$ emissions? Evidence from a regional analysis in China. Renewable and Sustainable Energy Reviews, 58, 943-951. doi:https://doi.org/10.1016/j.rser.2015.12.226

Zhang, X.-P., & Cheng, X.-M. (2009). Energy consumption, carbon emissions, and economic growth in China. *Ecological Economics, 68*(10), 2706-2712. doi:10.1016/j.ecolecon.2009.05.011

Zhang, Y.-J., Peng, Y.-L., Ma, C.-Q., & Shen, B. (2017). Can environmental innovation facilitate carbon emissions reduction? Evidence from China. *Energy Policy, 100*, 18-28. doi:https://doi.org/10.1016/j.enpol.2016.10.005

Zhao, X., & Du, D. (2015). Forecasting carbon dioxide emissions. *Journal of Environmental Management, 160*, 39-44. doi:https://doi.org/10.1016/j.jenvman.2015.06.002

Zhao, X., Han, M., Ding, L., & Calin, A. C. (2018). Forecasting carbon dioxide emissions based on a hybrid of mixed data sampling regression model and back propagation neural network in the USA. Environmental Science and Pollution Research, 25(3), 2899-2910. doi:10.1007/s11356-017-0642-6

Zhou, P., Ang, B. W., & Poh, K. L. (2006). A trigonometric grey prediction approach to forecasting electricity demand. *Energy, 31*(14), 2839-2847. doi:https://doi.org/10.1016/j.energy.2005.12.002

Zhou, X., Zhang, J., & Li, J. (2013). Industrial structural transformation and carbon dioxide emissions in China. *Energy Policy, 57*, 43-51. doi:https://doi.org/10.1016/j.enpol.2012.07.017

Zhu, Q., & Peng, X. (2012). The impacts of population change on carbon emissions in China during 1978–2008. Environmental Impact Assessment Review, 36, 1-8. doi:https://doi.org/10.1016/j.eiar.2012.03.003

**HIGHLIGHTS**

- ANN models were developed to predict carbon emissions for five countries.
- Stochastic gradient descent batching was employed to train the models.
- Predicted versus actual carbon emissions shows approximately zero forecasting errors.
- Sensitivity analysis shows significant contributory variables for each country.
- A simplified closed-form formula for hands-on prediction of carbon emissions.