Modelling carbon emission intensity: Application of artificial neural network

Acheampong, Alex O.; Boateng, Emmanuel B.

Published in:
Journal of Cleaner Production

DOI:
10.1016/j.jclepro.2019.03.352

Modelling Energy Demand: Application of Artificial Neural Networks

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

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

Abstract

This paper utilises an artificial neural network to develop models for forecasting energy demand for Australia, China, France, India, and USA. The study used quarterly data that span over the period of 1980Q1 to 2015Q4 to develop and validate the models. Financial development, foreign direct investment, economic growth, industrialisation, population, trade openness, urbanisation, and energy price were used as the inputs for modelling the demand for energy. To ensure stable forecasts, a repeated evaluation approach was used. After several iterations, the optimal models for each country were selected based on predefined criteria. An 8-5-1 multi-layer perceptron with back-propagation algorithm was used for building the models that have been calibrated and validated. The results suggest that the validated models have developed high generalising capabilities with insignificant forecasting deviations. The model for Australia, China, France, India, and USA attained high coefficients of determination (R^2) of 0.9544, 0.9202, 0.8946, 0.9532, and 0.9267 respectively. The results from the Partial Rank Correlation Coefficient (PRCC) further reveal that economic growth has the highest sensitivity weight on energy demand in Australia, France, and USA while industrialisation has the highest sensitivity weight on energy demand in China. Trade openness has the highest sensitivity weight on energy demand in India. After development and validation, the models were then deployed in an operational state for hands-on forecasting of energy demand using a closed-form solution. The models developed in this study could serve as a tool for energy planners and environmental policymakers for forecasting future energy demand as well as designing environmental and energy conservation policies.

Keywords: Energy demand; Neural networks; Forecasting; Deployment; Sensitivity

[*] Corresponding Author: The University of Newcastle, School of Health Sciences, University Drive, Callaghan NSW 2308, Australia or via email at emmanuel.boateng@uon.edu.au

1. Introduction

This study applies the Artificial Neural Network (ANN) to develop models for forecasting energy demand for Australia, China, France, India, and USA. The role of energy in economic development cannot be underestimated since it is a key input in the production function [1, 2]. Optimal forecasting of energy demand is the basis for energy planning and investment for governments and private investors. Underestimation of actual demand for energy could results in energy insecurity, power outages, and higher operating costs for energy suppliers while overestimation of energy demand would result in superfluous idle capacity resulting in wasted financial resources [3-5]. Therefore, accurate prediction of energy demand is crucial to avoid costly mistakes [3, 6]. Thus, developing a reliable model for predicting energy demand could serve as the tool for or energy planners and environmental policymakers for forecasting future energy demand as well as designing environmental and energy conservation policies.

Over the decades, some researchers have used classical statistical and econometric approaches to model energy demand [7]. Regression analysis which is the most popular estimation technique has been used to study the causal relationship between energy demand and other independent variables such as economic growth, population globalisation and among others [5, 8-12]. However, the effectiveness of regression depends on the reliability and availability of independent variables [13]. Additionally, given that variables for modelling energy demand are chaotic, non-stationary, and non-linear, the classical statistical and econometric approaches are not suitable for modelling such a complex behaviour [7, 14-16].

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 energy demand. For accurate forecasting using ARIMA and ARMA models, a large number of historical observation for the variable of interest is required [13, 17]. Other researchers have also employed Grey Model

(GM) prediction especially, GM (1, 1) to forecast energy demand [see 18, 19-22]. Generally, GM performs best with limited data [23]. However, the forecasting accuracy of the GM (1, 1) has been questioned [see 13]. Additionally, comparative studies have shown that ANN produces superior forecasting results relative to ARIMA, ARMA, classical statistical and regression approaches [24-28].

The forecasting ability of ANN has made it received a widespread application in the field of engineering [26, 29, 30], agriculture [15, 22, 31-33], energy [16, 34, 35], and finance [36, 37]. 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 [14, 31, 33, 38]. Also, ANN can handle noisy data, accommodating multiple variables with non-linear, linear, and unknown interactions and make a good generalisation [31, 39, 40]. Despite the forecasting ability of ANN, its application for forecasting energy demand is relatively limited. Therefore, this study utilised ANN to develop models for forecasting energy demand for Australia, China, India, France, and USA. These countries are studied because they are major contributors to global energy[1].

Although some scholars have used ANN to model the demand for electricity, electricity price, load forecasting and power correction techniques [33], however, this study is unique and contributes to the literature in five (5) folds: First, unlike the previous studies which have forecasted energy demand using only economic growth, population and energy price as input variables, this study incorporates other variables such as financial development, FDI, trade openness, industrialisation, and urbanisation, which are found to have important effect on energy demand in our model to prevent underestimation of the actual energy demand. For instance, scholars such as [4, 10, 11, 41, 42] argue that the failure to incorporate the

[1] https://www.iea.org/weo2018/

aforementioned variables in energy demand forecasting models could result in underestimation of actual energy demand. Second, given that previous forecasting studies on energy demand have not conducted sensitivity analysis, this study utilises the Partial Rank Correlation Coefficient (PRCC) to conduct sensitivity analysis to determine the input variable which is most influential in contributing to energy demand for the respective countries. Khoshroo, Emrouznejad [15], Marino, Hogue [43] and Saltelli and Marivoet [44] argue that PRCC is the most reliable and efficient method for sensitivity analysis. Third, to provide an accurate forecasting model, this study used high-frequency data for the modelling. Fourth, unlike previous studies, this study also developed a closed-form solution for practical predictions of energy demand. Finally, given that this study focuses on the high-energy consumption countries, the models that will be developed from this study would help energy and environmental planners in climate change policy decision-making.

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

2. Methodology

This study aims to develop models for forecasting energy demand for high-energy consumption countries such as Australia, China, France, India, and USA. ANN is employed and integrated into the proposed theoretical framework of the model as shown in Fig. 1. Fig. 1 shows the possible relationship between the selected eight (8) determinants of energy demand and energy demand (output).

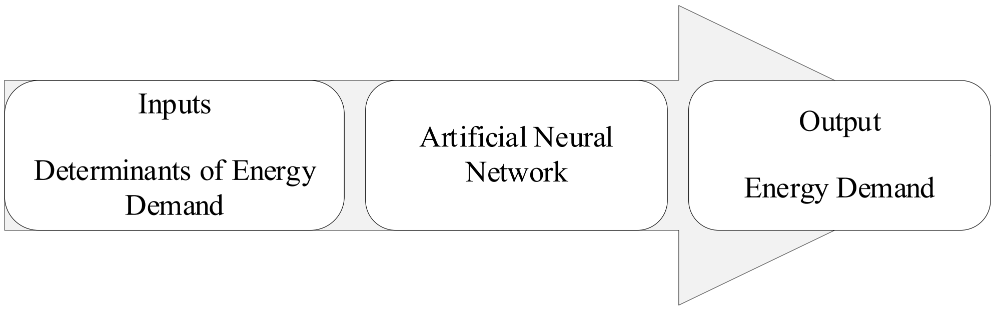

Fig. 1. A conceptual model for predicting energy demand.

2.1. Artificial Neural Network

Artificial neural networks (ANNs) are data processing systems that mimic the way data is processed in the human brain [45]. An ANN consists of numerous processing components called neurons which are likened to the biological neurons in the brain [45]. The neurons are connected by corresponding links between layers with numeric weights on each link. The layers in an ANN consist of input, hidden, and output layers [45]. The weights are the fundamental means of long-term memory in the neural network. A paramount feature of the ANN is the ability to “learn” from patterns in each data point enabling a spontaneous adjustment of the weights. This feature distinguishes the ANN from statistical tools, as it is able to accrue knowledge about the data during training of the network. Giving the experience acquired through the process of training, it is able to respond to new events in the most suitable way [46]. By this means, the ability of the network to generalise is tested with evaluation metrics. In order to achieve high-performance accuracies with negligible or minimal errors during training of the neural network, different architectures with varying tunings to various parameters are experimented. The approach adopted in ANN does not require prior expertise in computer programming to develop and compute solutions as required in other numerical solutions [47].

The advantages of ANNs over traditional statistical methods such as regression have been well substantiated in the literature. ANNs are known to produce better forecasting results than those obtained from statistical regression models [48]. A challenging issue in the development of statistical models is multicollinearity, which is easily dealt with in ANNs [49]. More so, ANNs are specifically fit to find solutions for problems that have fuzzy data and are extremely complicated where individuals often decide using intuition [47]. Also, statistical methods cannot deal effectively with nonlinearity while ANNs are naturally nonlinear nonparametric models that can handle indefinite nonlinearity in a straightforward manner [49]. Besides, unlike most statistical methods, ANNs do not require predefined mathematical expressions of the correlation between the model inputs and corresponding outputs [50]. These merits enable ANNs to overcome the limitations of existing statistical modelling approaches. In spite of these differences between ANNs and statistical approaches, both techniques can be combined into a solid and powerful methodological platform [51]. This is because ANN is like a ‘black box’ and hence lacks self-explanation. As a result, statistical approaches such as descriptive statistics are usually used to generate explanatory outcomes that can be easily be deduced and comprehended.

Neural networks have been applied to various problems such as classification, optimization, pattern recognition, forecasting, clustering, and function approximation. In the field of energy, to mention just a few, ANNs have been successfully used to, forecast; economic dependence between town development policy and increasing energy effectiveness in Poland [52], available ramp up and down capacity of a virtual power plant [53], annual transport energy demand in Iran [54], long-term energy consumption in Greece [46], thermohydraulics of advanced nuclear heat exchangers in USA [55], and future annual electricity demand in Turkey [56]. Achievements in these and other areas suggest that ANN models could serve as a valuable add-on to the toolkits of economists and econometricians [57].

2.2. Choice of optimal model

"Premature optimisation is the root of all evil" [Donald 58 pp. 268].

Premature optimisation is an expression used to explain a state where a programmer allows performance deliberations to influence the design of a specific code. For example, if a programmer intends to increase the performance of a simple neural network by employing a complex neural network to achieve an incremental performance of say 0.2% then it is unnecessary. This is because computational resources, time, money, and skill are valuable, and hence there should be a balance between design and performance. In other words, performing the minimum amount of work to attain high performance with negligible errors is ideal.

In this study, an optimal ANN model is described as a network with high-performance accuracy, reasonably simple architecture (preferably one hidden layer with neurons ≤ to the sum of the average of the input and output parameters) with insignificant or no forecasting errors generating stable results each time the model is executed. Much emphasis is placed on the number of hidden layers because most prominent machine learning (ML) researchers suggest that the use of multiple hidden layers can result in models that are too parsimonious [59-61]. Moreover, the use of even two hidden layers is only justified for the most esoteric applications [59, 62].

Further, consideration should be given to the selection of the hidden layer neurons since it influences the architecture of the model as well as the performance. If the network architecture is too complex, overfitting may occur, and if the architecture is too simple, underfitting may occur [63]. Overfitting causes the model to yield accurate results on the dataset used in training the model but performs poorly on the validation dataset. On the other hand, underfitting is when the model attains high accuracies on the validation dataset and poor results on the training dataset.

Neural network algorithms are stochastic in nature, which is good for the training process to efficiently approximate the function being learned. However, this means that the same network trained on the same data can yield different results due to sources of randomness such as randomness during dropout regularisation, stochastic optimisation, initialisation of weights, and word embedding of layers. Unfortunately, a neural network yielding different results is unstable and hence not reliable. In a quest to stabilise such networks, some researchers and practitioners employ fixed seeds and random number generators. However, it is strongly recommended to conduct repeated evaluation experiments [45], which may take a long time but rather yield consistent and reliable results. Therefore, in developing the robust methodology for selecting the optimal ANN models, the repeated evaluation experiments approach is incorporated in the process.

Fig. 2. shows the robust process employed to select the optimal ANN models for the five countries. Explanatory descriptions of the process are defined in the subsequent sections.

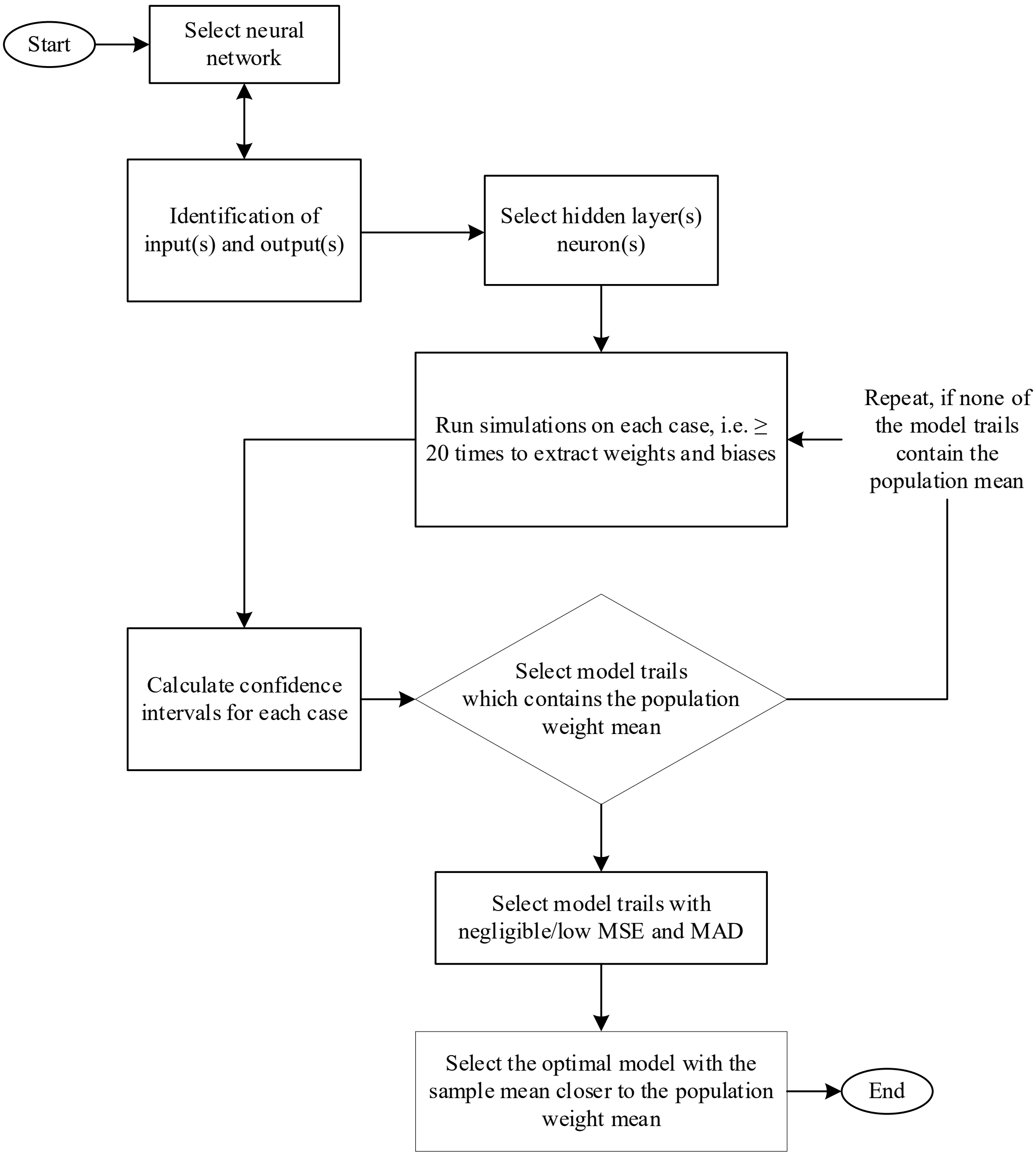

Fig. 2. Flowchart of selecting an optimal model.

2.3. Description of Dataset

The study used time series data which spans between 1980-2015. However, to develop an accurate model, the study follows Shahbaz, Hoang [64] 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 the selecting the input variables used for the modelling. Except for financial development, all the remaining variables were sourced from World Bank. The financial development index was obtained from the

International Monetary Fund (IMF)[2]. Table 2 also presents the descriptive statistics for variables.

In selecting the input variables, the study follows the literature on estimating energy demand to select the fundamental variables that influence energy demand.

Table 1. Variables for the study.

Variable	Code	Proxies	Reference
Energy consumption	*ENER*	Energy use (kg of oil equivalent per capita)	
Financial development	*FD*	The financial development index is a broad-based measure which comprises bank-based and market-based indicators of financial development.	[4, 5, 9, 65]
Foreign direct investment	*FDI*	Foreign direct investment, net inflows (% of GDP)	[66]
Economic growth	*GDP*	GDP per capita (constant 2010 US$)	[67, 68]
Industrialisation	*INDUS*	Industry, value added (% of GDP)	[9, 42]
Energy price	*PRI*	Dividing West Texas Intermediate crude oil prices by each country's consumer price index	[1]
Population	*POP*	Population, total	[69]
Trade Openness	*TRAD*	Trade (% of GDP)	[11, 70]
Urbanisation	*URB*	Urban population (% of total)	[10, 42]

Table 2: Descriptive statistics.

	Count	Mean	Sd	Min	Max
AUSTRALIA					
FD	144	0.6808346	0.2167618	0.2730694	0.9657203
FDI	144	2.462447	1.749735	-4.378572	7.85027
GDP	144	41787.91	8398.968	29725.5	55179.36
INDUS	144	25.58156	1.129529	22.35634	29.36063
POP	144	1.88e+07	2612337	1.46e+07	2.40e+07
TRAD	144	37.38663	4.790691	28.10575	46.25588
URB	144	1.64e+07	2537482	1.25e+07	2.15e+07
PRI	144	17.57268	18.56677	-4.038635	87.17818
ENER	144	5268.832	423.9819	4533.569	5971.229
CHINA					
FD	144	0.4056464	0.1301184	-0.0558633	0.6457832
FDI	144	2.923733	1.633651	0.2045862	6.701586
GDP	144	2161.015	1822.399	345.7698	6642.671
INDUS	144	44.94903	1.904348	39.84043	49.0031
POP	144	1.21e+09	1.20e+08	9.77e+08	1.37e+09
TRAD	144	36.85259	14.26362	12.00598	64.6305
URB	144	4.37e+08	1.74e+08	1.86e+08	7.70e+08
PRI	144	13.9191	31.12822	-95.32078	127.7941
ENER	144	1103.416	518.0744	596.2954	2238.488
FRANCE					
FD	144	0.587869	0.1710899	0.314396	0.8451784
FDI	144	1.557915	1.073046	0.0826509	4.118247
GDP	144	35446.1	5116.086	26905.43	41759.33
INDUS	144	21.95334	3.073835	17.67301	28.05097

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

POP	144	6.07e+07	3398370	5.53e+07	6.67e+07
TRAD	144	49.48591	6.657419	39.76955	62.35113
URB	144	4.61e+07	3788891	4.05e+07	5.31e+07
PRI	144	108.5753	428.3781	-169.4387	3613.361
ENER	144	3890.573	289.5859	3265.009	4305.708
INDIA					
FD	144	0.3229858	.0974152	0.1856536	0.4696478
FDI	144	0.8443818	.905103	-0.0012212	3.772586
GDP	144	818.1663	392.8301	382.6152	1805.414
INDUS	144	28.71699	1.400506	25.94348	31.7937
POP	144	1.01e+09	1.86e+08	6.91e+08	1.31e+09
TRAD	144	28.94815	14.84953	12.29911	56.58838
URB	144	2.82e+08	7.96e+07	1.59e+08	4.32e+08
PRI	144	6.031658	4.340546	.8578234	19.36692
ENER	144	414.3776	95.80655	283.4874	654.1635
USA					
FD	144	0.7201048	0.1950741	0.2862345	0.8938406
FDI	144	1.295188	0.7861581	0.2494226	3.721406
GDP	144	41135.79	7427.445	28281.24	52366.99
INDUS	144	20.98293	0.7908985	18.93935	23.41365
POP	144	2.74e+08	2.94e+07	2.26e+08	3.22e+08
TRAD	144	22.92577	4.374599	16.49049	31.27016
URB	144	2.14e+08	3.00e+07	1.67e+08	2.63e+08
PRI	144	30.59444	136.7145	-188.2027	1159.07
ENER	144	7562.857	351.4622	6692.11	8074.154

3. Models development and validation

In this study, the ANN approach is used to model the parameters contributing to energy demand for five countries that is Australia, China, France, India, and USA. To perform the necessary network computations, Spyder 3.2.6 software is used to write and execute computer codes scripted in Python programming language. Libraries such as keras, pandas, numpy, and matplotlib are imported to effectively perform the computations, build, and visualise the models.

3.1. Data pre-processing

Data pre-processing is crucial to the performance of the model because the variables are of different units. Feature scaling is applied to features of the dataset to ease the severe computations on the high dimensional data when training the network [45]. Standardisation (x_{stand}) and normalisation (y_{norm}) are performed on the inputs and output data respectively. Standardisation inclines the focus of the input values near zero. Standardising the input data into a smaller collection of variability would possibly aid the operational learning of the neural learning network while refining the numerical state of the optimisation problem [71]. Standardisation is expressed as:

$$x_{stand} = \frac{x-\mu}{\sigma} \tag{1}$$

Where x is the observation of the input parameter, μ is the mean of the observations for each input parameter and σ is the standard deviation of the observations for each input parameter. In normalisation, the features of the observations are scaled between zero (0) and one (1). Therefore, datasets of the output parameter (energy demand) for each country are normalised to confine the dataset variance between 0 and 1 to increase the calibration rate of the network. Normalisation is expressed as:

$$y_{norm} = \frac{y-\min(y)}{\max(y)-\min(y)} \tag{2}$$

Where y is the observation of the output parameter, $\min(y)$ and $\max(y)$ are the minimum and maximum observations of the output parameter respectively. Most importantly, feature scaling eliminates any instances of one variable dominating the other [45]. The 144 observations for individual countries (Australia, China, France, India, and USA) were randomly split into training and validation data sets. Random data division is the most commonly used data division method for ANNs [47]. Moreover, this approach helps to ensure reliability and robustness of the model to suitably predict outcomes over changing times. 80% (115 quarters) of the data were used for training the network, while the remaining 20% (29 quarters) were used to validate the model. Similar data ration has been used in other studies [see 72, 73].

3.2. Configuration of models

The choice of network architecture is the key to building an optimal model. However, this often poses a crucial challenge. The architecture of a neural network could primarily be feedforward, recurrent, or a hybrid. In this study, the feed-forward neural network (FFNN) is employed. In an FFNN, the neurons process the data and feed forward to the subsequent layer [45]. The most commonly used FFNN is the multi-layer perceptron (MLP) structure [74]. Therefore, this study employs the MLP. The sigmoid (logsig) and rectifier (ReLU) functions were selected as the activation functions for the input and output layers respectively. At present, the ReLU is the most commonly used activation function for deep neural networks [75] and the sigmoid function is the most popular activation function for ANNs [76]. The sigmoid and rectifier functions can be expressed 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.

In building the optimal models, the choice of hidden layer neurons is essential to the efficiency of the model [45]. The ideal number of hidden layer neurons generally is mostly found through a trial and error approach [61]. Nevertheless, some general guidelines may be followed. Hecht-Nielsen [77] suggests the following upper limit for the number of hidden layer neurons to ensure that the network can approximate any continuous function:

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

Where N^h is the number of hidden layer neurons and N^i is the number of input parameters. For this study, with eight (8) input parameters, the upper limit of the hidden layer neurons is expressed as:

$$N^h \leq 2(8) + 1$$

$N^h \leq 17$ Hidden layer neurons

This suggests that the number of neurons in the hidden layer should not be more than seventeen (17). Nonetheless, to eliminate instances of overfitting, the relationship between the number of training samples and network size requires concern [61]. Rogers and Dowla [78] propose the following upper limit for the number of hidden layer neurons to satisfy the said criteria:

$$N^h \leq \frac{N^{tr}}{N^i+1} \tag{6}$$

Where N^{tr} is the number of training samples. Therefore, the upper limit for the number of hidden layer neurons may be obtained as the smaller value for N^h from Eqs. (5) and (6). With 115 training samples, the upper limit is taken as:

$$N^h \leq \frac{115}{8+1}$$

$N^h \leq 12.7 \sim 13$ Hidden layer neurons

This suggests that the number of hidden layer neurons should not be more than 13. From experimentations and practice, the problem could be argued as trivial; hence, the hidden layer neurons required in the MLP with a single hidden layer could be determined using a simplified expression:

$$N^h = \frac{N^i+N^o}{2} \tag{7}$$

$$N^h = \frac{8+1}{2}$$

$N^h = 4.5 \cong 5$ Hidden layer neurons

Where N^o is the number of output parameters, in this case, one (1). Accordingly, five (5) neurons in the hidden layer were deemed appropriate in configuring the network to achieve an optimal model. Therefore, an 8-5-1 MLP was sufficient to perform the necessary computations for each country. Fig. 3 depicts the configuration of the three-layer FFNN.

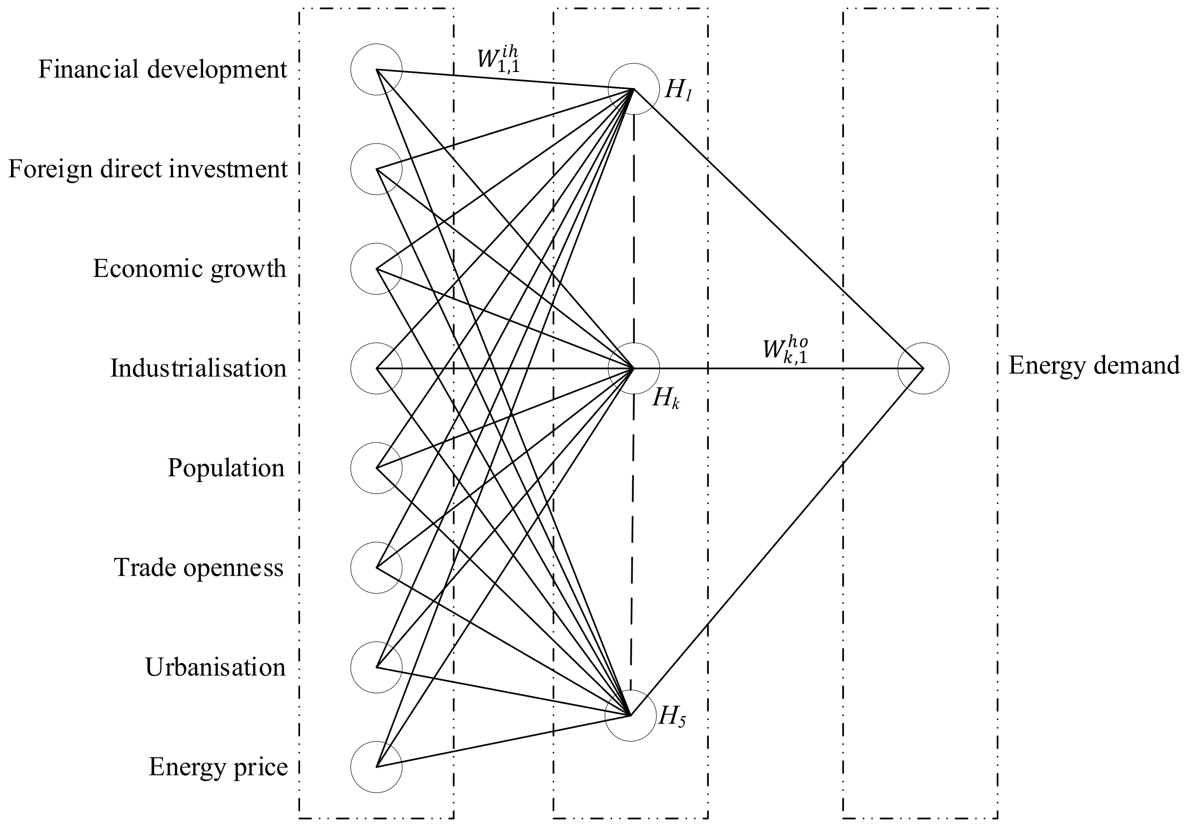

Fig. 3. The configuration of the developed FFNN (8-5-1 MLP).

3.3.Networks calibration

The calibration of the network is to find the ideal set of parameters to depict the input-output data correlation, which is performed by optimisation of learning algorithms [47]. Back-propagation (BP) is selected as the learning algorithm to train the FFNNs. The BP algorithm is mostly used to iteratively minimise the cost function regarding the interconnection weight and neurons thresholds [79]. Thus, the MLP with a BP algorithm can approximate any continuous function to meet the ideal accuracy [80]. By this means, the deviation of the predicted energy demand from the actual energy demand coefficients is minimised. A stochastic gradient descent batching is applied to the 8-5-1 MLP to update the weights after every 5 observations. The stochastic gradient descent is initialised to improve the accuracy and minimise the loss over the epochs [45]. As discussed earlier, during calibration or training of a neural network, there is some randomness involved at each initialisation resulting in unstable results. To ensure reproducibility and expressive results during each initialisation, a robust method is employed by conducting repeated evaluation experiments [45]. In this method, each case is run at least 20 times with different random weights and then the mean is extracted to calculate confidence intervals (CIs). Subsequently, the simulated weight trials for each model for individual countries are computed and presented in Fig. 4a to 4e. For Australia, apart from the 11th trial, the remaining trials contained the population weights mean (see Fig. 4a). In the case of China, except the 4th, 5th, and 15th trials, the rest contained the population weights mean (see Fig. 4b). For France, apart from the 11th, 14th, and 15th trials, the remaining trials contained the population weights mean (see Fig. 4c). In the case of India, except the 1st and 4th trials, the rest contained the population weights mean (see Fig. 4d). In the case of USA, except the 19th trial, the remaining contained the population weights mean (see Fig. 4e). By removing the trials that could not contain the population weights mean, random effects are avoided to ensure stable outcomes.

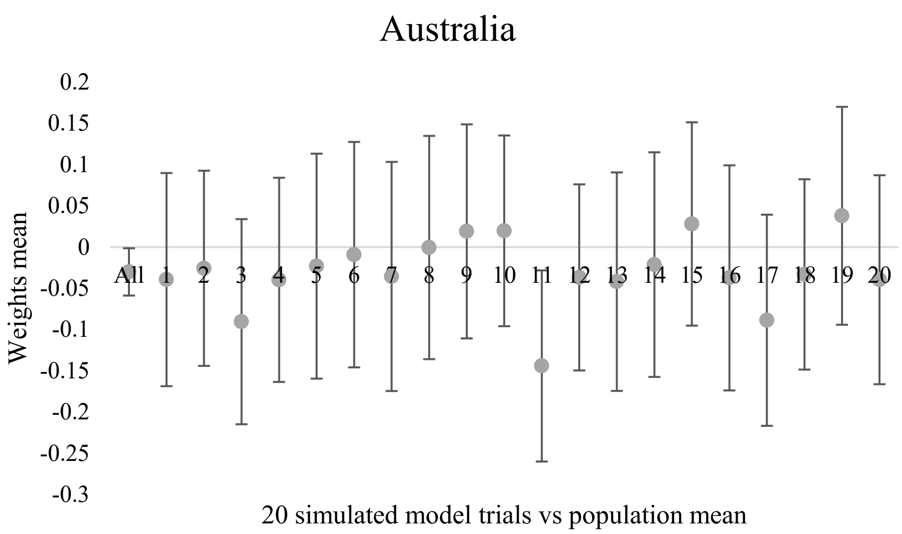

Fig. 4a. Point and interval estimates.

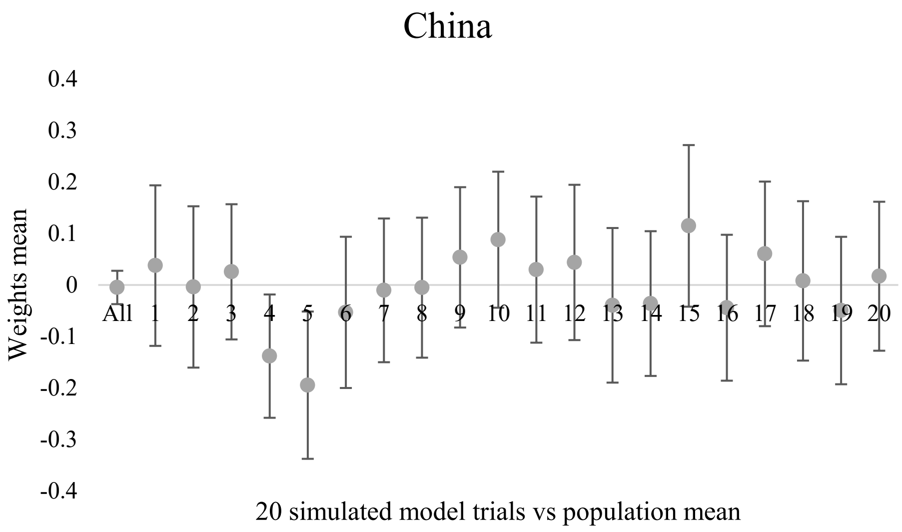

Fig. 4b. Point and interval estimates.

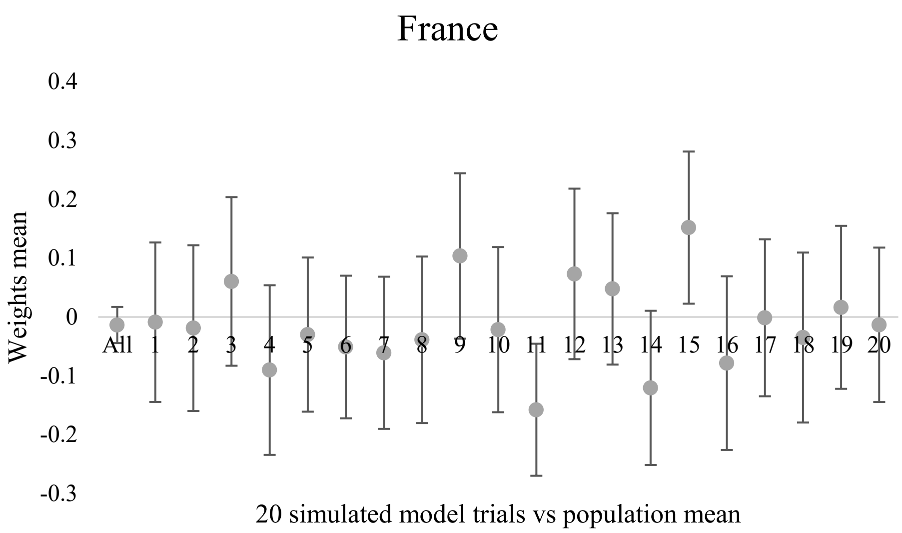

Fig. 4c. Point and interval estimates.

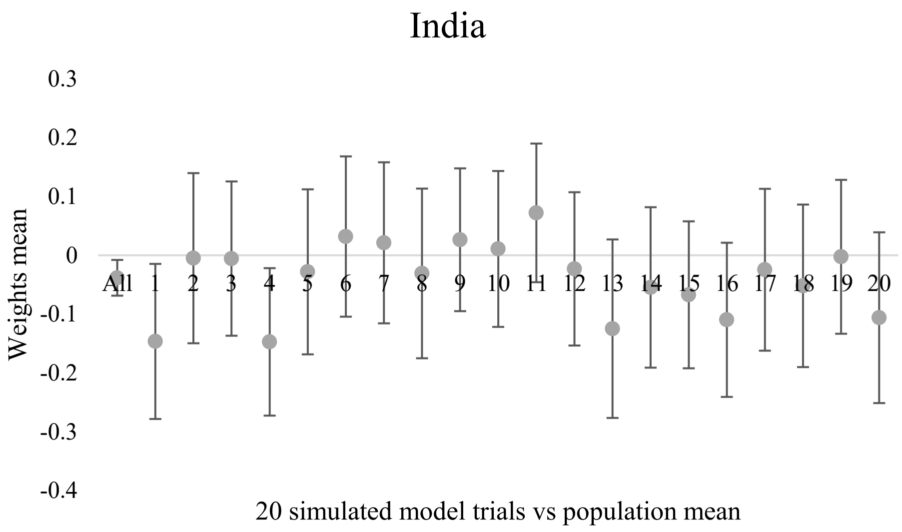

Fig. 4d. Point and interval estimates.

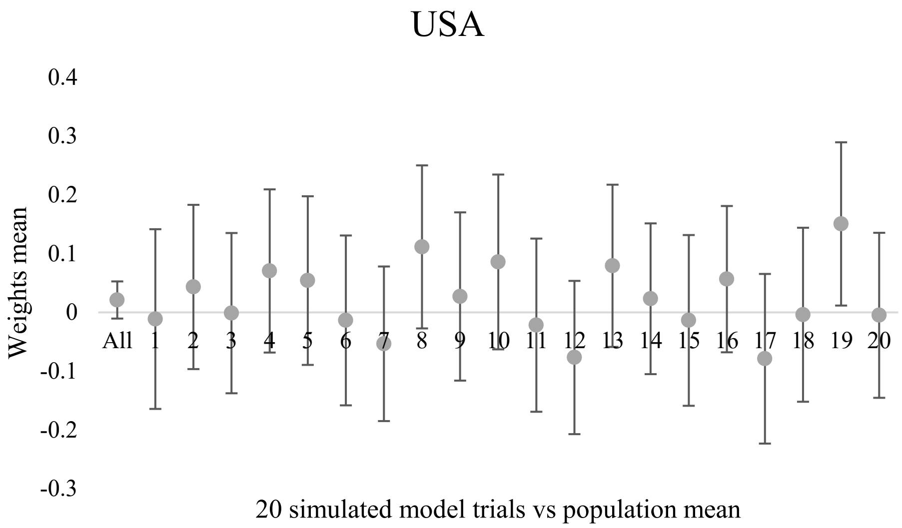

Fig. 4e. Point and interval estimates.

In addition, other metrics such as accuracies, means, standard deviations (SDs), standard errors (SEs) and intervals are assessed to estimate the skill of the stochastic model at a 95% confidence interval while spontaneously evaluating the mean squared errors (MSEs) on both the training and validation data sets. The MSEs are computed using the expression:

$$MSE = \sum_{i=1}^{n} (y_a - y_p)^2 \qquad (8)$$

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 energy demands respectively. The calculated MSEs are shown in Appendix Table Ia to Ie. Coefficient of determination (R^2) was also calculated for each run on the actual and predicted energy demands for the individual countries. The R^2 denotes the fraction of the variation in the dependent variable that is estimated from the input variables. The R^2 value ranges from 0 to 1 and a value closer to 1 indicates a tremendous performance. The R^2 is determined with the expression:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_a - y_p)^2}{\sum_{i=1}^{n}(y_a - \bar{y})^2} \quad (9)$$

Where $\bar{y}$ is the mean of the observations in the output parameter (energy demand). R^2 for each simulated run is shown in Appendix Table Ia to Ie. The energy demands yielded from the iterations for each country is denormalised to obtain the actual and predicted energy demands. Denormalisation is computed using the expression:

$$y = y_{norm}(\max(y) - \min(y)) + \min(y) \quad (10)$$

Mean absolute deviations (MADs) were computed after the output parameter were denormalised for each country. The MADs were calculated using the expression:

$$MAD = \frac{1}{n}\sum_{i=1}^{n}\left|y_a - y_p\right| \quad (11)$$

The MADs are shown in Appendix Table Ia to Ie. The MADs for Australia ranges from 0.015661504 to 0.021326936, 0.0594981239 to 0.0928153048 for China; 0.017056879 to 0.026064956 for France; 0.019431892 to 0.066439005 for India; and 0.009772336 to 0.027289701 for USA. Model trials with insignificant/lower MSEs and MADs were selected for further assessment. Finally, the optimal model for each country was selected based on the trial with the sample weights mean nearer to the population weights mean with insignificant MAD and MSE. Adopting the criteria, the 6th trial for Australia, the 20th trial for China, the 1st

trial for France, the 3rd trial for India, and the 7th trial for USA (see Appendix Table Ia to Ie for bolds) were selected. The selected models and their confidence limits are shown in Table 3. This suggests how precise and near the weight points tend to approach the actual population weights means with given data [45].

Table 3. Selected model trails and their confidence intervals

Country	Model	Lower limit*	Upper limit*	Mean	P_{mean}
Australia	6	-0.145913515	0.127560631	-0.009176442	-0.030105043
China	20	-0.127715424	0.161710193	0.016997384	-0.004821043
France	1	-0.144365111	0.126825358	-0.008769876	-0.013552569
India	3	-0.136732669	0.125549336	-0.005591667	-0.038251271
USA	7	-0.18488519	0.078114453	-0.053385369	0.02112324

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

The MSE coefficients on both the training and validation sets are shown in Fig. 5. The MSEs on the 8-5-1 MLP for each country is approximately zero (0). This indicates that the developed models are sufficient to timely predict the energy demands with insignificant errors.

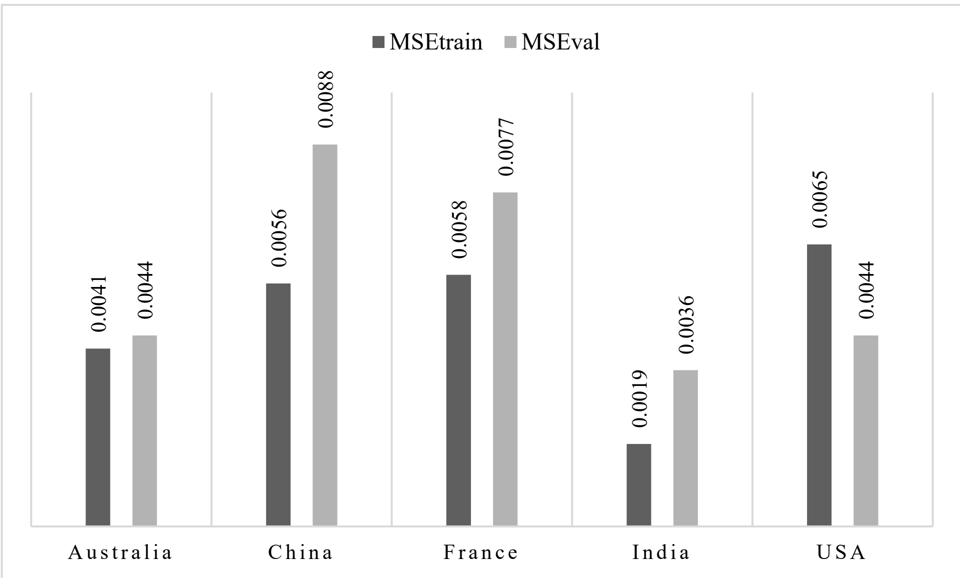

Fig. 5. Regression metric scores.

The R^2 for each 8-5-1 MLP for each country is shown in Fig. 6a to 6e. Australia (0.9544), China (0.9202), France (0.8946), India (0.9532), and USA (0.9267) attained high R^2 coefficients indicating how well the 8-5-1 MLPs fits the data. This suggests that there are strong relationships between the developed FFNNs and the output parameter for each country.

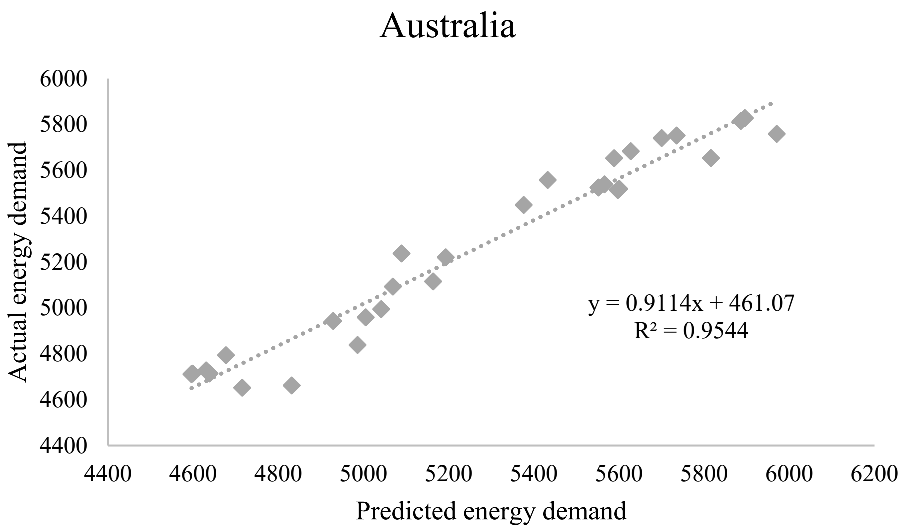

Fig. 6a. Scatter chart of actual and predicted energy demand.

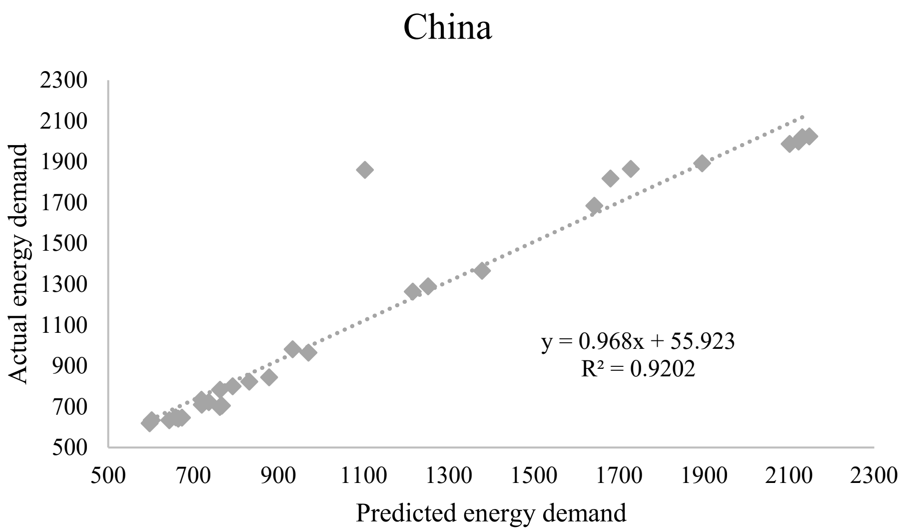

Fig. 6b. Scatter chart of actual and predicted energy demand.

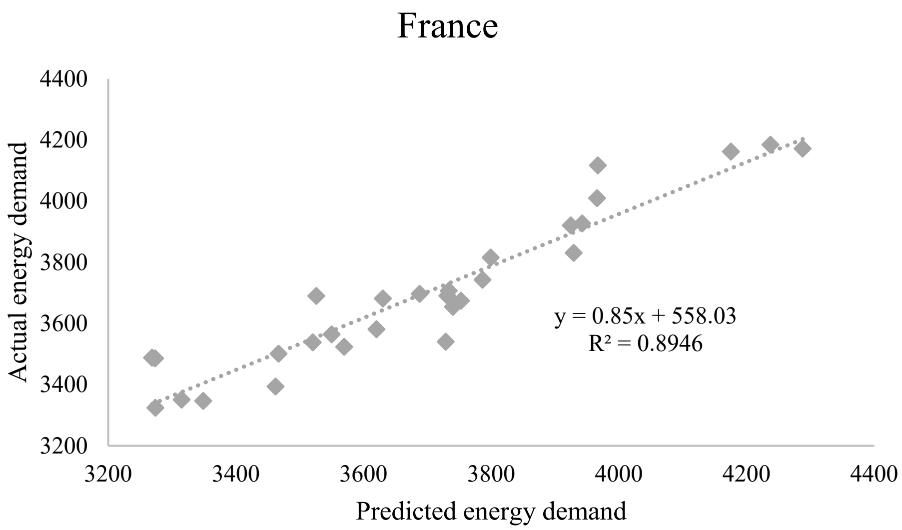

Fig. 6c. Scatter chart of actual and predicted energy demand.

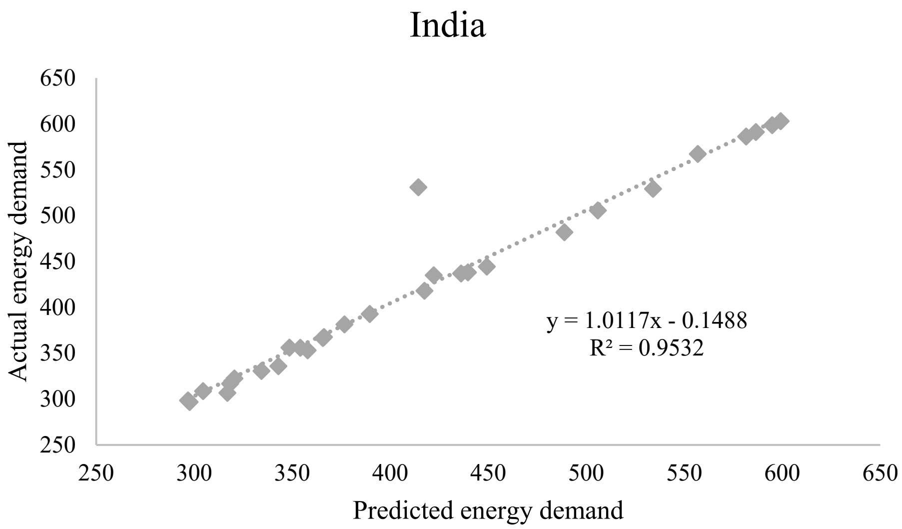

Fig. 6d. Scatter chart of actual and predicted energy demand.

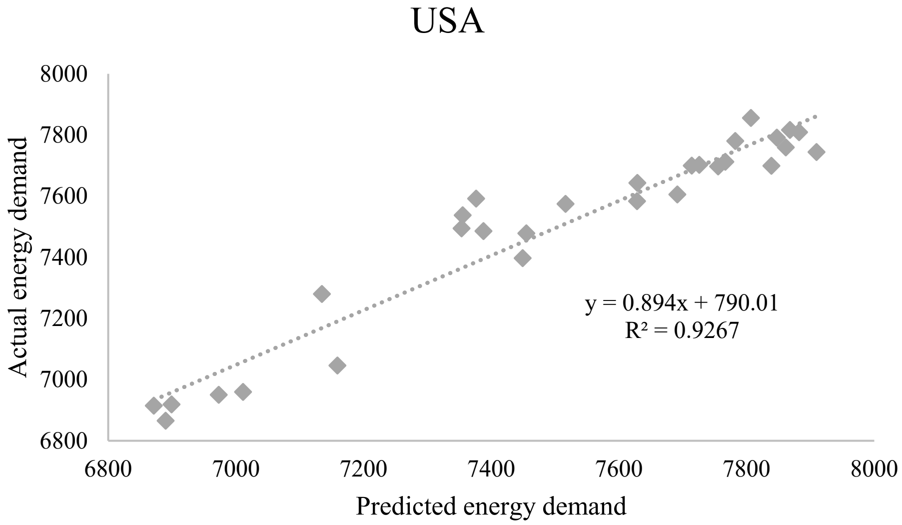

Fig. 6e. Scatter chart of actual and predicted energy demand.

3.4. Validation of the ANN models

The 29 observations in the validation data set representing 29 quarters were used to validate the model. The 8-5-1 MLPs were tasked to forecast the energy demands from the eight input parameters. Absolute percentage deviations (APDs) were computed to validate the models. The APD is the positive percentage ratio of the variance between the actual and predicted observations to the actual observation. The mean absolute percentage deviation (MAPD) for Australia is 1.5661504%, 5.9498124% for China, 1.8411551% for France, 1.9431892% for India, and 0.9772336% for USA (see Appendix Table IIa to IIc). The MAPDs indicate that the developed FFNNs are capable of forecasting energy demand with minimal error. Fig. 7a to 7e presents the actual against predicted energy demands for each country. After validating the models, the optimal models are deployed for the end-user.

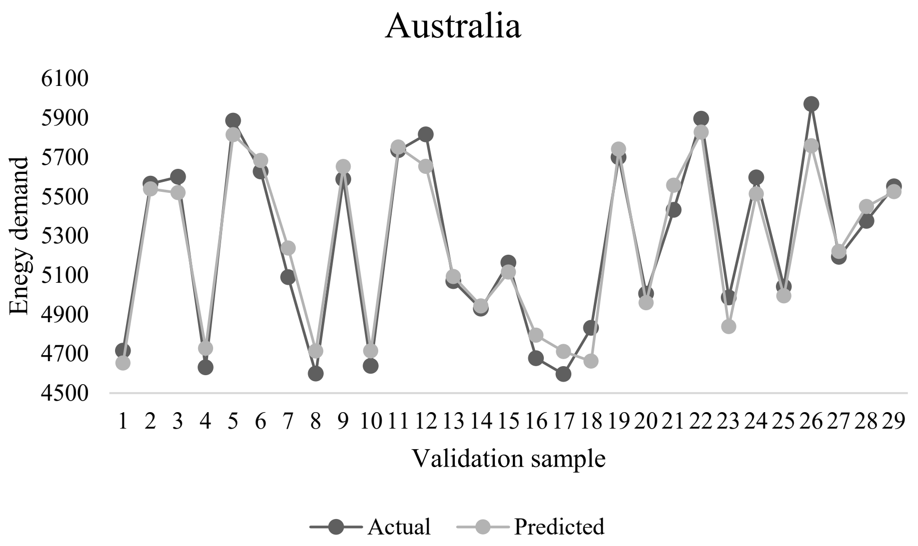

Fig. 7a. Predicted versus actual energy demands.

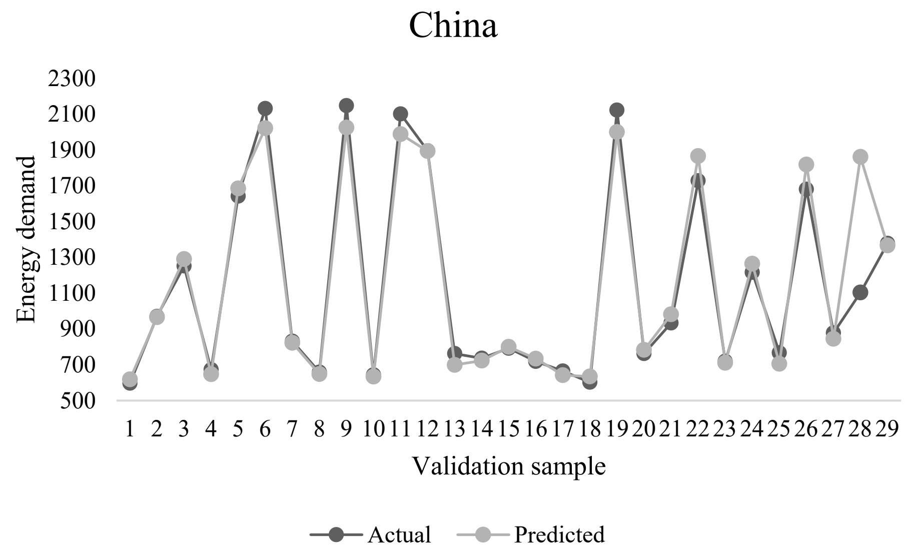

Fig. 7b. Predicted versus actual energy demands.

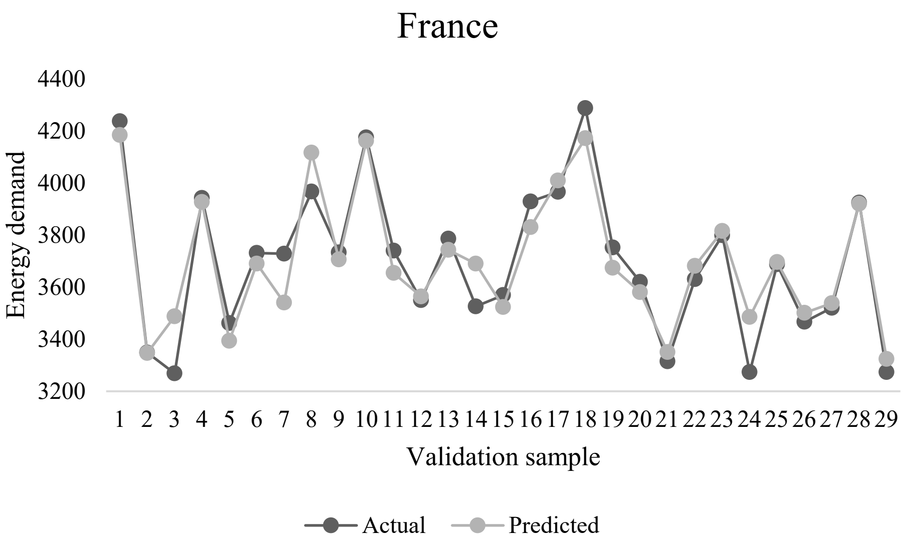

Fig. 7c. Predicted versus actual energy demands.

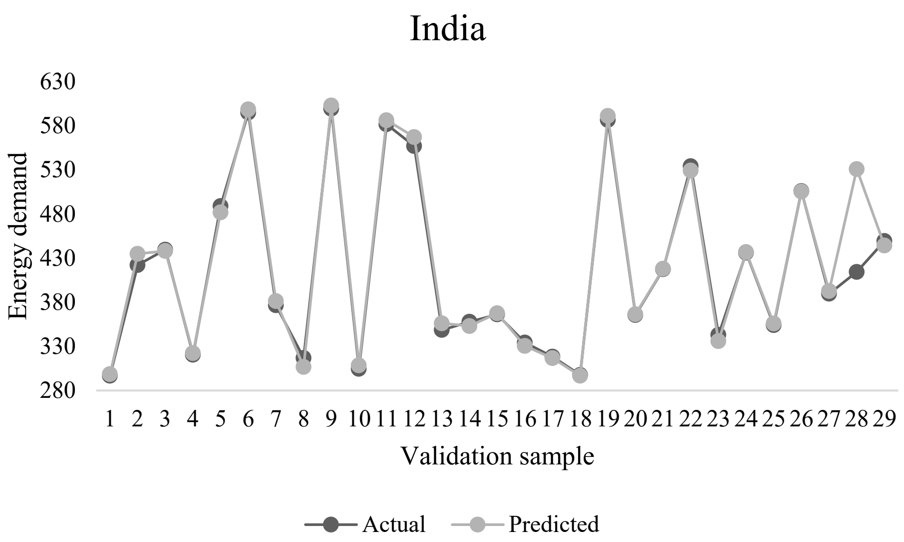

Fig. 7d. Predicted versus actual energy demands.

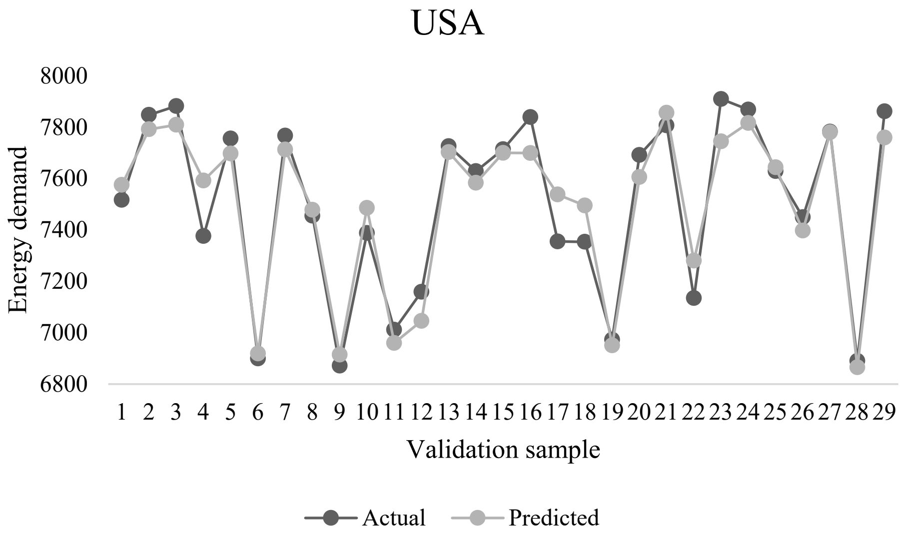

Fig. 7e. Predicted versus actual energy demands.

4. Deployment of the validated models

Over the years, a number of studies conducted have conducted significant investigations into modelling energy-related outcomes using machine learning (ML) approaches such ANNs. The amount of work and successes in these studies reflect the current state of ANN in modelling environmental concerns. However much of this research (outside of the ‘tech giants’) remains just that enquiry, and not part of a production utilisation [81]. In spite of the constant influx of new papers, it remains extremely challenging to truly employ any of these models in production (and that is even if the papers provide code) [81]. Considering the novelty in the contributory factors of energy demand used in this study, previous works that investigated energy demand have not yet addressed the issue of how to successfully deploy the validated model in an operational state for hands-on forecasting of energy demand. This suggests that in the context of these papers, after developing, validating and testing the model by making some ‘black-box’ predictions, the life cycle of the model ends. However, a validated model should be packaged with the end-user in mind. Deploying a model is one of the pressing challenges of gaining value from a validated neural network. Model deployment thereby fills the gap between the development and production of a model for facilitating the decision-making process of a business or entity.

Deploying machine learning (ML) models such as ANNs for production could be simple or complex depending on the requirements. However, the process is typically inefficient and often involves rewriting the model to run on the target production platform that is costly and requires considerable time and effort [82]. Nevertheless, some common approaches such as Kubernetes, custom REST-API with Flask/Django, AWS Lambda, Apache Beam, and Spark/Flink are used [81]. These approaches have demonstrated promising capabilities but requiring at least technical expertise, team coordination, time, and computational resources that may not be readily and timely available to the environmental policymaker, consultant, or reader.

To ensure scalability of the validated models, a closed-form formula is used to deploy the validated 8-5-1 MLPs suitable for spreadsheet or hand calculations. For the purpose of environmental policymakers and consultants, these closed-form expressions Eqs. (12) and (13) can be used for forecasting energy demand. The simplified closed-form equations require the values of the inputs, weights of the connections between the nodes in different layers, and the biases of the output and input nodes [80, 83].

$$O_1 = \frac{1}{1+e^{-\left(bias_o+\sum_{k=1}^{r}\frac{w_{k,1}^{ho}}{1+e^{-H_k}}\right)}} \quad (12)$$

$$H_k = \sum_{j=1}^{q} w_{j,k}^{ih} \times I_j + bias_k \quad (13)$$

Where q and r are the numbers of input parameters and the number of hidden neurons/nodes respectively; $bias_k$ and $bias_o$ are the biases of the kth hidden H_k and the bias of the output neuron respectively; $w_{j,k}^{ih}$ and $w_{k,1}^{ho}$ are the weights on the connection between I_j and H_k, and the weight of the connection between H_k and O_1, respectively. For each country, the weights and biases are shown in Tables 4a to 4e. The closed-form formula can be used to deploy the validated models to practically forecast energy demand based on the previous input values. An example is illustrated in the next section.

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

Connection	Weight/bias	Number of hidden layer neuron (k)				
		1	2	3	4	5
Input to hidden layer	$w_{1,k}^{ih}$	0.509292	0.421938	0.288192	0.64516	-0.559387
	$w_{2,k}^{ih}$	-0.232718	0.280403	0.340961	-0.0326693	-0.125697
	$w_{3,k}^{ih}$	-0.378145	0.556411	-0.852792	0.423308	-0.516781
	$w_{4,k}^{ih}$	0.483088	-0.046592	-0.16578	0.398452	0.227115
	$w_{5,k}^{ih}$	0.0993522	0.146396	-0.823605	-0.693957	-0.780821
	$w_{6,k}^{ih}$	0.706024	-0.373422	0.221026	0.745271	0.0683481
	$w_{7,k}^{ih}$	0.208317	-0.632872	0.353302	-0.2279	0.0433821
	$w_{8,k}^{ih}$	-0.640674	0.189851	-0.479431	-0.347889	-0.107604
	$bias_k$	0.286806	0.294798	0.049036	0.320431	0.0740741
Hidden layer to output	$w_{k,1}^{ho}$	0.116912	0.646581	-0.657789	0.739453	-0.59495
	$bias_o$	0.139423				

Table 4b. Weights and biases for a neural network (China).

Connection	Weight/bias	Number of hidden layer neuron (k)				
		1	2	3	4	5
Input to hidden layer	$w_{1,k}^{ih}$	-0.112064	0.270548	0.59587	-0.158364	-0.155859
	$w_{2,k}^{ih}$	-0.0321051	-0.425258	-0.195451	-0.488375	0.850921
	$w_{3,k}^{ih}$	0.102293	-0.824211	0.304097	0.89765	-0.0158014
	$w_{4,k}^{ih}$	-0.165776	0.325526	0.377118	0.397825	-0.40444
	$w_{5,k}^{ih}$	0.59321	-0.662105	0.715025	0.388705	0.0512346
	$w_{6,k}^{ih}$	0.451143	-0.513727	-0.171652	0.277802	-1.31755
	$w_{7,k}^{ih}$	0.193602	-0.0525023	-0.207932	0.197155	0.315539
	$w_{8,k}^{ih}$	-0.338704	0.127376	-0.157266	0.0339765	0.182348
	$bias_k$	0.22206	0.440728	-0.278885	-0.101303	0.224534
Hidden layer to output	$w_{k,1}^{ho}$	-0.488387	-0.903122	0.895781	0.766772	-0.755983
	$bias_o$	-0.464927				

Table 4c. Weights and biases for a neural network (France).

Connection	Weight/bias	Number of hidden layer neuron (k)				
		1	2	3	4	5
Input to hidden layer	$w_{1,k}^{ih}$	-0.690149	-0.0137212	0.11637	-0.171358	0.49787
	$w_{2,k}^{ih}$	-0.0553914	-0.0493834	0.407739	0.261679	-0.619998
	$w_{3,k}^{ih}$	-0.56049	0.053813	0.213573	0.822512	0.369216

Connection	Weight/bias	1	2	3	4	5
	$w_{4,k}^{ih}$	-0.33915	-0.488688	-0.598979	-0.737855	0.0723782
	$w_{5,k}^{ih}$	-0.725156	-0.193512	-0.00148144	-0.0498245	-0.21723
	$w_{6,k}^{ih}$	0.478151	0.549664	-0.330196	0.403128	0.670449
	$w_{7,k}^{ih}$	-0.468772	-0.122841	-0.678324	0.137127	0.302162
	$w_{8,k}^{ih}$	0.324845	0.125505	-0.51119	0.0788173	0.168152
	$bias_k$	-0.375433	-0.218708	0.590543	0.408459	-0.169195
Hidden layer to output	$w_{k,1}^{ho}$	0.976791	0.747297	-0.59376	-0.652652	0.698219
	$bias_o$	-0.408505				

Table 4d. Weights and biases for a neural network (India).

Connection	Weight/bias	Number of hidden layer neuron (k)				
		1	2	3	4	5
Input to hidden layer	$w_{1,k}^{ih}$	0.280724	0.011403	-0.0126624	-0.268785	-0.248447
	$w_{2,k}^{ih}$	-0.0537839	0.0750604	0.221635	0.658448	-0.128913
	$w_{3,k}^{ih}$	0.510083	-0.0289934	-0.0671551	0.282494	0.275239
	$w_{4,k}^{ih}$	0.363306	-0.114916	-0.14992	-0.625166	-0.242974
	$w_{5,k}^{ih}$	-0.688233	-0.434811	0.314119	-0.443372	0.762582
	$w_{6,k}^{ih}$	-0.480949	-0.443136	-0.0525896	-0.133071	0.914542
	$w_{7,k}^{ih}$	0.344458	-0.57423	0.15572	-0.791084	-0.120384
	$w_{8,k}^{ih}$	-0.085283	-0.104931	0.330133	0.605117	-0.459608
	$bias_k$	-0.207296	0.29265	-0.29862	-0.194299	-0.17153
Hidden layer to output	$w_{k,1}^{ho}$	0.601151	-0.717325	0.352868	-0.828952	0.988967
	$bias_o$	-0.370174				

Table 4e. Weights and biases for a neural network (USA).

Connection	Weight/bias	Number of hidden layer neuron (k)				
		1	2	3	4	5
Input to hidden layer	$w_{1,k}^{ih}$	-0.138569	-0.421648	0.378325	0.471562	-0.405861
	$w_{2,k}^{ih}$	-0.342688	0.203618	-0.0519162	0.433887	-0.377822
	$w_{3,k}^{ih}$	-0.358456	-0.0324796	-0.340861	-0.105785	-0.177031
	$w_{4,k}^{ih}$	-0.742548	-0.523445	-0.313529	0.0350867	0.11611
	$w_{5,k}^{ih}$	0.537183	0.188922	0.305726	0.480243	-0.0556034
	$w_{6,k}^{ih}$	0.0403225	0.374124	-0.819562	-0.465874	0.409478
	$w_{7,k}^{ih}$	-0.333649	0.64791	-0.329841	-0.236031	-0.0349339
	$w_{8,k}^{ih}$	0.217216	0.263172	-0.105597	0.24736	0.0997923
	$bias_k$	-0.207296	0.29265	-0.29862	-0.194299	-0.17153
Hidden layer to output	$w_{k,1}^{ho}$	-0.190911	-0.203757	-0.31329	0.836326	0.261818
	$bias_o$	0.50382				

4.1. Illustrative example

For Australia, consider the eight input parameters (I_1 to I_8) for determining the energy demand for 2015Q1 (Table 5). The energy demand for the subsequent quarter (2015Q2), O_1, may be obtained by the following procedures:

Table 5. Input values and the output value for 2015Q1, Australia

Period	2015Q1(Un-normalised)	2015Q1(normalised)
Financial development (I_1)	0.870131541	0.861996
Foreign direct investment (I_2)	2.921859621	0.596985
Economic growth (I_3)	54850.98896	0.9871
Industrialisation (I_4)	24.60117326	0.320494
Population (I_5)	23722602.84	0.97272
Trade openness (I_6)	42.29998679	0.782046
Urbanisation (I_7)	21201575	0.971714
Energy price (I_8)	47.07179259	0.560318
Energy demand (O)	5376.163424	0.586087
Energy demand for 2015Q2	y_p	O_1?

First procedure: Insert the normalised values of the input parameters (Table 5), and weights and biases of input to the hidden layer (Table 4a) into Eq. (13) to compute H_1 to H_5.

H_1 = (0.509292I_1 - 0.232718I_2 - 0.378145I_3 + 0.483088I_4 + 0.0993522I_5 + 0.706024I_6 + 0.208317I_7 - 0.640674I_8) + 0.286806

H_2 = (0.421938I_1 + 0.280403I_2 + 0.556411I_3 - 0.046592I_4 + 0.146396I_5 - 0.373422I_6 - 0.632872I_7 + 0.189851I_8) + 0.294798

H_3 = (0.288192I_1 + 0.340961I_2 - 0.852792I_3 - 0.16578I_4 - 0.823605I_5 + 0.221026I_6 + 0.353302I_7 - 0.479431I_8) + 0.049036

H_4 = (0.64516I_1 - 0.0326693I_2 + 0.423308I_3 + 0.398452I_4 - 0.693957I_5 + 0.745271I_6 - 0.2279I_7 - 0.347889I_8) + 0.320431

H_5 = (-0.559387I_1 - 0.125697I_2 - 0.516781I_3 + 0.227115I_4 - 0.780821I_5 + 0.0683481I_6 + 0.0433821I_7 - 0.107604I_8) + 0.0740741

The values of H_1 to H_5 are obtained as 0.8607, 0.7020, -0.9475, 0.894, and -1.6447 respectively.

Second procedure: Insert the values of H_1 to H_5 and the weights and biases of the hidden to output layer (Table 4a) into Eq. (12). The value of the predicted output O_1 is 0.7107.

$$O_1 = \frac{1}{1+e^{-\left(0.139423+\frac{0.116912}{1+e^{-H_1}}+\frac{0.646581}{1+e^{-H_2}}-\frac{0.657789}{1+e^{-H_3}}+\frac{0.739453}{1+e^{-H_4}}-\frac{0.59495}{1+e^{-H_5}}\right)}}$$

Third procedure: The value obtained from the sigmoid function is the normalised value. Using Eq. (10) to denormalise O_1.

$$y_p = 0.7107(5971.2290 - 4533.5692) + 4533.5692$$

The actual energy demand (y_a) for 2015Q2 is 5593.1945 and the predicted energy demand (y_p) is 5555.2775. The absolute percentage deviation (APD) for this forecast is 0.6779%.

5. Sensitivity analysis

Sensitivity analysis is conducted to identify the extent to which each input variable contributes to energy demand in Australia, China, France, India, and USA. To conduct the sensitivity analysis, the Partial Rank Correlation Coefficient (PRCC) between energy demand and each input variable is calculated for each country. Fig. 7a-7e depicts the normalised sensitivity weight of each input variable for each country. Fig. 7a shows that in Australia, economic growth has the highest sensitivity weight, followed by trade openness, population, energy price, and industrialisation. As depicted in Fig 7a, the PRCC results show that economic growth (0.6166), trade openness (0.3951), population (0.3535), energy price (0.3038) and industrialisation (0.1032) increase energy demand while financial development (-0.0392), foreign direct investment (-0.0534) and urbanisation (-0.435) reduce energy demand in Australia.

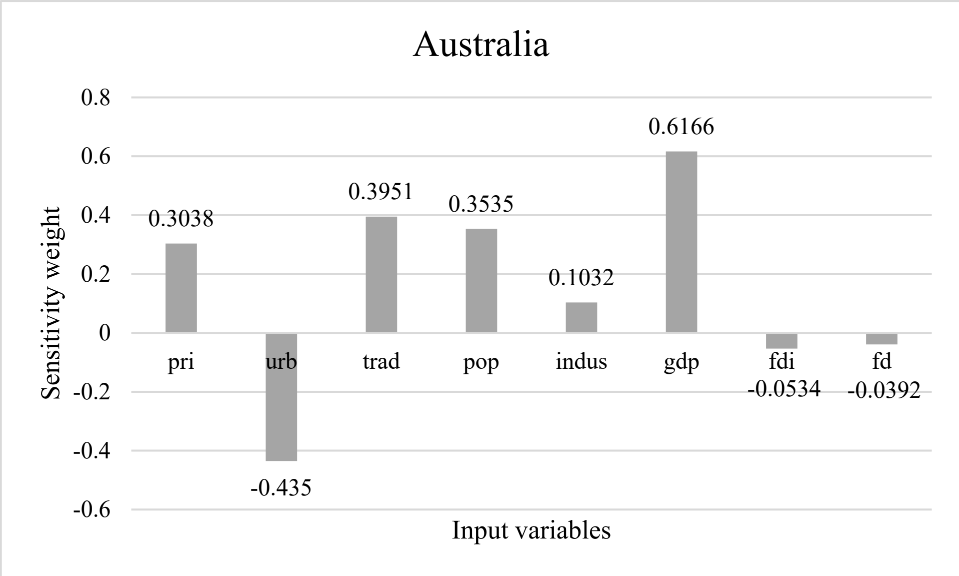

Fig. 7a. Sensitivity analysis of energy demand determinants.

In China, as depicted Fig. 7b, industrialisation has the highest sensitivity weight followed by trade openness, economic growth, and population. The PRCC results as shown in Fig. 7b reveals that industrialisation (0.4263), trade openness (0.3733), economic growth (0.3682), and population (0.2374) increase energy demand in China while financial development (-0.1212), urbanisation (-0.2604), energy price (-0.2726) and foreign direct investment (-0.3468) reduce the demand for energy.

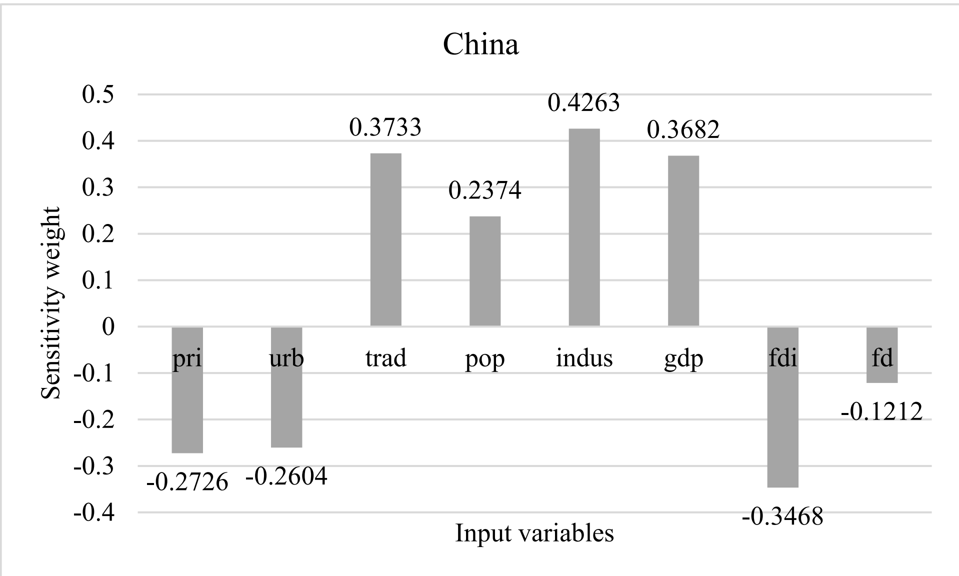

Fig. 7b. Sensitivity analysis of energy demand determinants.

Fig. 7c indicates that in France, economic growth has the highest the sensitivity weight followed by foreign direct investment, population while trade openness has the lowest sensitivity weight. The PRCC results show that in France (see Fig.7c), economic growth (0.5275), foreign direct investment (0.253) and population (0.0274) increase energy demand while financial development (-0.0799), energy price (-0.1174), urbanisation (-0.1452), industrialisation (-0.2469) and trade openness (-0.4063) reduce the demand for energy.

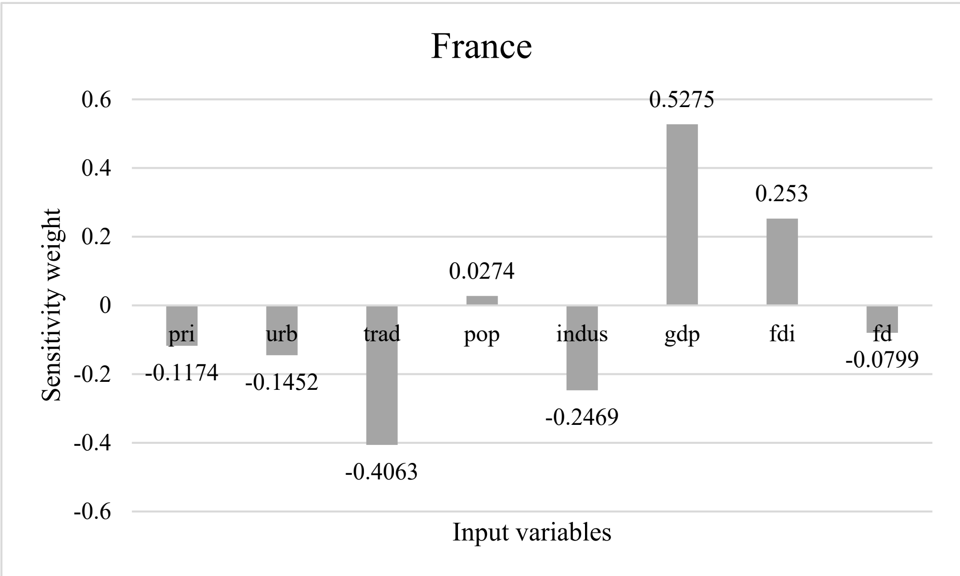

Fig. 7c. Sensitivity analysis of energy demand determinants.

For India, Fig. 7d shows that trade openness has the highest sensitivity weight followed by economic growth, financial development and population. As depicted in Fig. 7d, the factors responsible for the increase in energy demand in India include trade openness (0.5783), economic growth (0.0647), financial development (0.0538) and population (0.036) while urbanisation (-0.0205), industrialisation (-0.0793), foreign direct investment (-0.304) and energy price (-0.573) contribute to the reduction in energy demand.

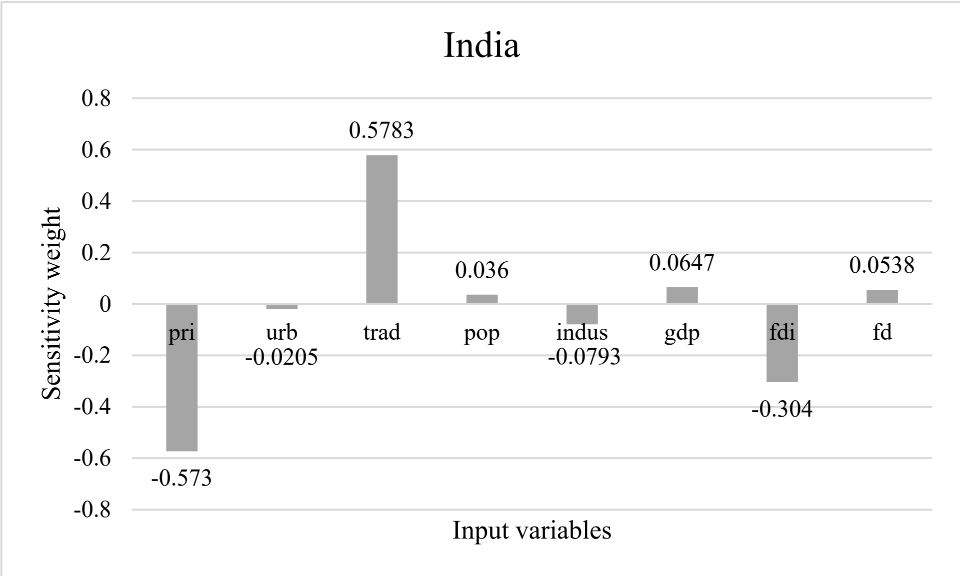

Fig. 7d. Sensitivity analysis of energy demand determinants.

For USA, economic growth has the highest sensitivity weight, followed by financial development, trade openness, urbanisation, industrialisation and energy price (see Fig. 7e). Fig. 7e shows that the factors contributing to the increase in energy demand in USA include economic growth (0.7134), financial development (0.5063), trade openness (0.3868), urbanisation (0.3335), industrialisation (0.0135) and energy price (0.0011) while foreign direct investment (-0.2978) and population (-0.4663) reduce energy demand.

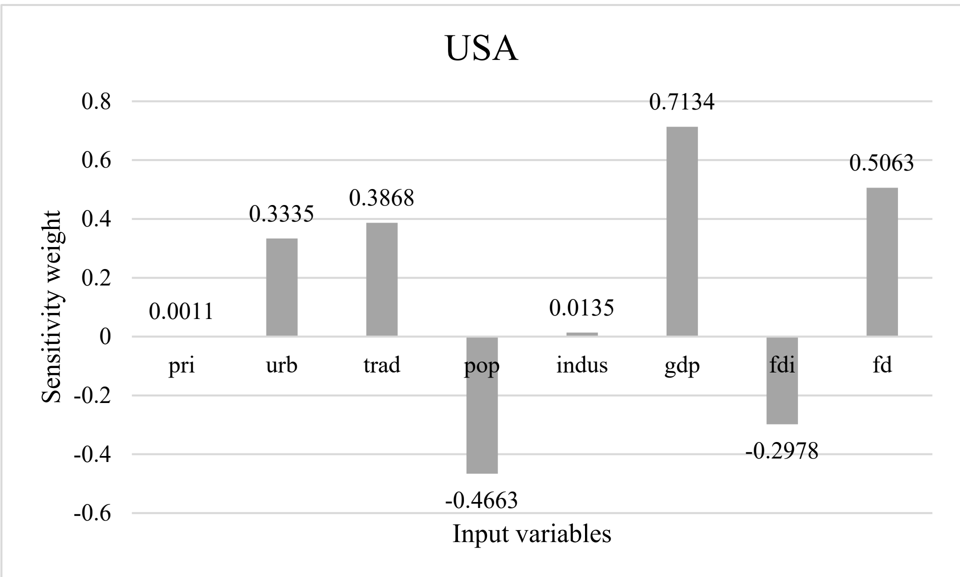

Fig. 7e. Sensitivity analysis of energy demand determinants.

6. Conclusion and policy implications

In this paper, an artificial neural network was employed to develop optimal models for forecasting energy demand for Australia, China, France, India, and USA. The feedforward multi-layer perceptron (FFMLP) was selected as the neural network type for each country. An 8-5-1 FFMLP with back-propagation algorithm was sufficient in building the models. Each developed model requires financial development, foreign direct investment, economic growth, industrialisation, population, trade openness, urbanisation, and energy price as the inputs. Five ANN models were developed and trained on 114 quarters of energy demand data. Twenty-nine quarters of energy demand data were used to validate the performance of the developed models.

The results suggest that the validated models have developed high generalising abilities with negligible errors. The model for Australia (0.9544), China (0.9202), France (0.8946), India (0.9532), and USA (0.9267) attained high coefficients of determination (R^2). The models developed from this study have been validated and reliable to forecast energy demand with an insignificant error. Another significant contribution in this paper is that it bridges the gap between model development and production by effectively developing a closed-form solution for practical forecasting by the end user. Considering the random effects inherent in neural networks, this study further proposes a robust methodology in selecting an optimal model for stable results. The results from the Partial Rank Correlation Coefficient (PRCC) further reveal that economic growth has the highest sensitivity weight on energy demand in Australia, France, and USA while industrialisation has the highest sensitivity weight on energy demand in China. Trade openness has the highest sensitivity weight on energy demand in India.

Environmental policymakers, Energy Ministries, and consultants could employ the validated models to forecast the energy demand in Australia, China, France, India, and USA. This would help in planning and framing policy decisions for sustainable energy management. There exist several extensions of this study. Further research could consider developing a clustering system to map the indicators of energy demand in different regions/countries. Such a system would certainly assist in an easier depiction and understanding of the similarities and disparities of the determinants of energy demand in those regions/countries. Other high-energy consuming countries could adopt the methodology of this study to develop and deploy robust neural networks for planning their future energy demand requirements as a dais to identify proactive conservation measures. Data engineers could also employ the closed-form expression, the extracted weights and biases, and model architecture to develop an energy demand predictive application on accessible platforms using any readily available programming language.

Appendix

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

Model	SD	SE	MSE_{train}	MSE_{val}	MAD	ME	R^2
1	0.442493656	0.06596306	0.0058	0.0078	0.020400881	-6.656196215	0.9175
2	0.405226868	0.060407655	0.0055	0.0065	0.018832997	10.07192712	0.9327
3	0.425850812	0.063482091	0.0062	0.0066	0.01830603	4.223413127	0.9298
4	0.42391547	0.063193587	0.0061	0.0075	0.020195406	-9.78708706	0.9195
5	0.467066075	0.0696261	0.0058	0.0073	0.020032702	-21.76673913	0.9253
6	**0.467989883**	**0.069763813**	**0.0041**	**0.0044**	**0.015661504**	**4.961973657**	**0.9544**
7	0.475668976	0.070908544	0.0064	0.0066	0.019267501	-16.6456562	0.9304
8	0.463448398	0.069086808	0.0048	0.0055	0.016798934	-13.00365747	0.9425
9	0.444332946	0.066237245	0.0059	0.006	0.018526535	-2.664647865	0.9359
10	0.395922441	0.059020633	0.0063	0.0068	0.018434702	1.927014378	0.9281
11	0.396710574	0.059138121	0.0077	0.0071	0.019623807	6.172924402	0.9254
12	0.386333527	0.057591202	0.0062	0.0063	0.018858701	-1.761336482	0.933
13	0.453635128	0.067623932	0.005	0.0055	0.01700007	-4.926389098	0.9411
14	0.466305023	0.069512649	0.0062	0.0078	0.020693466	-1.432325564	0.917
15	0.422128908	0.062927262	0.0074	0.0078	0.019474174	-14.76762636	0.9179
16	0.467284158	0.069658609	0.0056	0.0068	0.019395606	-22.12711582	0.9292
17	0.438294968	0.065337156	0.0063	0.0065	0.018080621	14.05653009	0.9327
18	0.395114384	0.058900175	0.0057	0.0065	0.018364674	0.690056863	0.9311
19	0.452345757	0.067431724	0.0059	0.0083	0.021326936	-22.78996621	0.9144
20	0.433790157	0.064665619	0.0064	0.0066	0.017863191	1.922126335	0.9301

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 the selected optimal model and other ANN models (China).

Model	SD	SE	MSE_{train}	MSE_{val}	MAD	ME	R^2
1	0.533766346	0.079569189	0.0057	0.0092	0.077014644	-21.55867819	0.9187
2	0.536316665	0.079949368	0.0042	0.0079	0.076446285	-20.34933927	0.9287
3	0.449476714	0.067004032	0.005	0.0093	0.08417018	-23.92914354	0.9199
4	0.409986452	0.061117172	0.0163	0.018	0.092815305	34.00348738	0.8395
5	0.48962278	0.072988655	0.0107	0.0122	0.087875853	35.78730491	0.8948
6	0.502733436	0.074943076	0.0093	0.0109	0.064652615	-14.11231055	0.8981
7	0.477816155	0.071228627	0.0078	0.0105	0.087967063	-9.8244591	0.9007
8	0.4655164	0.069395088	0.0077	0.0096	0.065707507	-25.83002675	0.9126
9	0.466571148	0.06955232	0.0047	0.01	0.086948348	-27.79364737	0.9201
10	0.452206189	0.067410919	0.0107	0.0126	0.078098446	-15.73404764	0.8814
11	0.4856737	0.07239996	0.0042	0.0077	0.067100291	-16.07629756	0.9312
12	0.516292	0.076964267	0.0034	0.0089	0.074567325	-6.455608623	0.9197
13	0.513971542	0.076618354	0.0058	0.0093	0.080030605	-25.4420276	0.918
14	0.481154202	0.071726234	0.0064	0.0095	0.084367159	-21.8451955	0.9136
15	0.537249644	0.080088448	0.0026	0.0082	0.066814599	-14.15935484	0.9276
16	0.485083696	0.072312008	0.0029	0.0089	0.07402778	-25.37372938	0.9239
17	0.480707016	0.071659571	0.0048	0.0081	0.071810285	-27.95029556	0.93
18	0.529873786	0.07898892	0.0063	0.0084	0.077106836	-14.56317566	0.9254
19	0.490134063	0.073064872	0.0065	0.0103	0.085732993	-37.5893544	0.9132

20	**0.495287261**	**0.073833066**	**0.0056**	**0.0088**	**0.059498124**	**-19.38679924**	**0.9202**

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 the selected optimal model and other ANN models (France).

Model	SD	SE	MSE_{train}	MSE_{val}	MAD	ME	R^2
1	**0.464081881**	**0.069181242**	**0.0058**	**0.0077**	**0.018411551**	**-3.444530747**	**0.8946**
2	0.482313751	0.071899089	0.0055	0.0077	0.019030718	3.64574167	0.8927
3	0.490511197	0.073121092	0.0053	0.0074	0.017582985	0.700527242	0.8991
4	0.493572647	0.073577466	0.0047	0.0056	0.017056879	-3.26881767	0.9234
5	0.448613155	0.066875301	0.0071	0.0103	0.022182386	-1.35055079	0.8553
6	0.414777568	0.061831389	0.0071	0.01	0.023097238	1.857547827	0.8609
7	0.442663013	0.065988306	0.0065	0.0093	0.020959881	-14.46409551	0.8734
8	0.484555508	0.07223327	0.0062	0.0092	0.020110908	-9.039515885	0.8744
9	0.481193243	0.071732053	0.0126	0.0152	0.025079664	23.48251499	0.8213
10	0.480370199	0.071609361	0.0053	0.0081	0.019385486	0.656128864	0.8881
11	0.384146921	0.057265242	0.0074	0.0111	0.023643238	-2.387420866	0.8442
12	0.496016803	0.073941819	0.0054	0.0073	0.01863976	-3.214884334	0.8991
13	0.440325291	0.065639819	0.0056	0.0074	0.01947255	-8.14659603	0.9014
14	0.448503719	0.066858987	0.0075	0.0129	0.025849505	-5.179315192	0.8195
15	0.442734732	0.065998997	0.0057	0.0076	0.018951601	5.802238935	0.8947
16	0.505264094	0.075320324	0.0062	0.0079	0.019829511	-0.008536963	0.8898
17	0.456683568	0.068078367	0.0065	0.0097	0.0222057	3.211339498	0.8635
18	0.494396153	0.073700227	0.006	0.0101	0.023450177	8.309943448	0.8602
19	0.474069243	0.07067007	0.0101	0.0109	0.020352005	22.19316413	0.8779
20	0.449323095	0.066981132	0.0125	0.016	0.026064956	25.19474139	0.8253

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 the selected optimal model and other ANN models (India).

Model	SD	SE	MSE_{train}	MSE_{val}	MAD	ME	R^2
1	0.451148653	0.06725327	0.0032	0.0045	0.032574327	-0.4811	0.9374
2	0.494901488	0.07377556	0.0045	0.0044	0.035707149	-1.9618	0.9371
3	**0.448837035**	**0.06690867**	**0.0019**	**0.0036**	**0.019431892**	**-4.7231**	**0.9532**
4	0.428972882	0.0639475	0.0088	0.0075	0.045378388	5.84895	0.9052
5	0.480027326	0.07155825	0.0024	0.0034	0.020266066	-3.9675	0.9549
6	0.466357794	0.06952052	0.0033	0.004	0.027225913	-1.1125	0.944
7	0.468443482	0.06983143	0.0042	0.0045	0.031650905	-1.7232	0.9366
8	0.493553294	0.07357458	0.0025	0.0042	0.032673806	-3.4424	0.9466
9	0.415394867	0.06192341	0.0037	0.0045	0.030573815	-3.4038	0.9422
10	0.45344273	0.06759525	0.0049	0.0045	0.035242929	-2.2916	0.9355
11	0.403297343	0.06012002	0.0092	0.0094	0.066439005	-16.249	0.892
12	0.445957861	0.06647947	0.0042	0.0045	0.034417212	-0.703	0.9383
13	0.519235608	0.07740307	0.0023	0.0038	0.023661327	-2.7814	0.9458
14	0.466744936	0.06957823	0.0038	0.0042	0.02940075	-1.0284	0.9414
15	0.427512558	0.06372981	0.0032	0.004	0.027270123	-1.8716	0.9436
16	0.448448582	0.06685077	0.0028	0.0046	0.032679177	-4.2566	0.9357
17	0.470838785	0.0701885	0.0031	0.0035	0.027781034	-1.7141	0.953
18	0.472789628	0.07047932	0.0053	0.0064	0.044944645	-1.6595	0.908

19	0.447675978	0.06673559	0.0039	0.0041	0.029604946	-4.7522	0.9495
20	0.496811803	0.07406033	0.0031	0.004	0.021646826	-2.2155	0.9427

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 the selected optimal model and other ANN models (USA).

Model	SD	SE	MSE_{train}	MSE_{val}	MAD	ME	R^2
1	0.523243385	0.078000518	0.0174	0.0252	0.0242788	-36.813691	0.7133
2	0.478220979	0.071288974	0.0074	0.0079	0.0132648	-8.4677346	0.8667
3	0.466624458	0.069560267	0.017	0.0187	0.0219123	-36.892849	0.8151
4	0.47515084	0.070831305	0.0076	0.0077	0.0130722	-8.2289747	0.8722
5	0.491021815	0.07319721	0.0093	0.0064	0.010567	-12.90867	0.8932
6	0.494563661	0.073725198	0.0067	0.0039	0.0099505	7.65981091	0.9401
7	**0.450065112**	**0.067091746**	**0.0065**	**0.0044**	**0.0097723**	**4.97874127**	**0.9267**
8	0.47472341	0.070767588	0.0102	0.0109	0.0159319	19.0283602	0.8245
9	0.489843117	0.0730215	0.0076	0.0073	0.0128135	7.54629269	0.882
10	0.50900121	0.07587742	0.0056	0.005	0.0108478	1.92418621	0.9154
11	0.504199895	0.075161683	0.0108	0.0112	0.0151899	3.91089799	0.8249
12	0.446280233	0.066527529	0.0067	0.006	0.0116303	12.1143763	0.9007
13	0.472420304	0.070424261	0.0115	0.0084	0.012867	-4.6958035	0.8579
14	0.438974249	0.065438417	0.0118	0.0083	0.0129459	1.02509524	0.8588
15	0.49713429	0.074108404	0.0268	0.0338	0.0272897	-50.766186	0.5812
16	0.42594	0.063495386	0.0116	0.0092	0.0135404	1.60803178	0.8435
17	0.493899273	0.073626157	0.0216	0.0276	0.0252514	-53.179282	0.7606
18	0.506634488	0.07552461	0.005	0.005	0.0106538	5.14644514	0.9244
19	0.475367461	0.070863597	0.0092	0.0087	0.0135756	-2.6359388	0.8575
20	0.480473192	0.071624715	0.0105	0.0073	0.0128772	20.1931371	0.8803

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

Appendix IIa. Results of evaluating predictions for energy demand on selected optimal models.

Validation	Australia				China			
samples	y_a	y_p	$AD_{8\text{-}5\text{-}1}$	E_y	y_a	y_p	$AD_{8\text{-}5\text{-}1}$	E_y
1	4714.853797	4652.119787	0.013305611	62.73400956	597.3418243	618.9532297	0.036179294	-21.61140537
2	5565.88944	5539.008078	0.004829662	26.88136251	970.5330299	965.289509	0.005402723	5.243520893
3	5600.358771	5519.697432	0.014402888	80.66133945	1251.682968	1290.074146	0.030671647	-38.39117803
4	4630.332197	4727.696411	0.021027479	-97.36421451	673.1958358	646.8737879	0.039100135	26.32204791
5	5886.263296	5814.873424	0.012128216	71.38987155	1642.286683	1685.024744	0.026023509	-42.73806177
6	5627.750503	5683.121969	0.009839005	-55.37146631	2130.986771	2021.342499	0.051452347	109.6442717
7	5089.392922	5237.513571	0.029103795	-148.1206492	831.6281709	822.947541	0.010438114	8.680629954
8	4599.035062	4713.787053	0.024951319	-114.7519907	658.8067802	648.3629281	0.015852679	10.44385212
9	5588.666284	5652.745656	0.011465951	-64.07937158	2147.525293	2024.850223	0.057123923	122.6750698
10	4638.206691	4714.426812	0.016433101	-76.22012087	643.6508207	633.9006307	0.01514826	9.750189978
11	5735.36365	5751.797539	0.002865361	-16.43388891	2100.985555	1988.344282	0.05361354	112.6412731
12	5815.786338	5653.624066	0.027883121	162.1622722	1895.571897	1893.200571	0.001250982	2.371326079
13	5069.189489	5093.071893	0.004711287	-23.88240422	762.326002	699.2163771	0.082785612	63.10962489
14	4929.013348	4943.175734	0.00287327	-14.16238646	736.314657	722.8378394	0.018303068	13.47681759
15	5163.387833	5115.648902	0.00924566	47.73893056	792.3321448	800.0849359	0.009784774	-7.752791148
16	4677.052402	4793.815822	0.024965173	-116.7634205	719.1777206	734.3005078	0.021027886	-15.12278721
17	4595.969253	4711.12882	0.025056644	-115.1595672	664.6403409	642.244447	0.03369626	22.3958939
18	4831.614773	4662.118423	0.035080684	169.4963498	602.3558357	634.3487851	0.05311304	-31.99294942
19	5700.092105	5740.895765	0.007158421	-40.80365979	2122.097583	1999.819924	0.057621129	122.2776592
20	5005.321453	4959.135195	0.009227431	46.186258	762.8597146	782.9650783	0.026355257	-20.1053637
21	5432.32941	5557.687591	0.023076322	-125.3581819	933.3570744	981.6391783	0.05172951	-48.28210391
22	5895.738911	5827.523393	0.011570309	68.21551876	1727.803861	1865.363764	0.079615462	-137.5599033
23	4986.120069	4838.079929	0.029690448	148.0401402	719.9782894	710.468188	0.013208872	9.510101423
24	5596.986021	5513.729706	0.014875205	83.25631535	1215.766574	1264.419813	0.04001857	-48.65323943
25	5041.694246	4994.587885	0.009343359	47.10636026	767.101498	705.4305981	0.080394707	61.67089998
26	5970.504408	5758.4194	0.035522126	212.085008	1680.208194	1818.456176	0.082280269	-138.247982
27	5193.213522	5220.543435	0.005262621	-27.32991236	877.9938361	844.658969	0.037967086	33.33486709

28	5376.162917	5448.871123	0.013524182	-72.70820557	1103.415969	1860.877294	0.686469425	-757.4613254
29	5551.531532	5525.078592	0.004764981	26.4529399	1378.378125	1366.224258	0.008817513	12.15386725
		Mean	0.015661504	4.961973657			0.059498124	-19.38679924
		%mean	1.5661504				5.9498124	

Note: y_a = actual energy demand and y_p = predicted energy demand, AD = absolute deviation, $E_y = y_a – y_p$

Appendix IIb. Results of evaluating predictions for energy demand on selected optimal models.

Validation samples	France				India			
	y_a	y_p	$AD_{8\text{-}5\text{-}1}$	E_y	y_a	y_p	$AD_{8\text{-}5\text{-}1}$	E_y
1	4237.300109	4184.312913	0.012504943	52.98719627	296.8237352	298.5147226	0.005696941	-1.690987389
2	3348.871808	3346.546678	0.000694302	2.325129999	422.1817653	434.9081885	0.030144417	-12.72642321
3	3268.575677	3487.996734	0.067130481	-219.4210569	439.6380158	438.0841415	0.00353444	1.553874293
4	3942.422173	3927.119733	0.003881482	15.30244003	320.5146468	322.2960421	0.005557922	-1.781395295
5	3462.218617	3393.681294	0.019795781	68.5373227	488.8474947	481.8194755	0.014376711	7.028019224
6	3731.293459	3690.379413	0.010965111	40.91404565	594.8371416	598.5083179	0.006171733	-3.671176287
7	3728.738542	3540.561407	0.050466702	188.1771355	376.6582897	381.3829275	0.012543565	-4.724637818
8	3967.025341	4117.043162	0.037816199	-150.0178205	316.9405135	306.7595232	0.032122717	10.1809903
9	3733.441462	3706.586221	0.00719316	26.85524109	599.1829484	602.8285481	0.006084285	-3.645599635
10	4175.558552	4162.155387	0.003209909	13.40316411	304.5339467	308.5978544	0.013344679	-4.063907635
11	3739.886511	3654.475293	0.022837917	85.41121841	581.5895478	586.2018707	0.007930547	-4.612322954
12	3549.95892	3564.185277	0.004007471	-14.22635713	556.9054836	567.1858151	0.018459742	-10.2803315
13	3786.425536	3743.017975	0.011463994	43.40756077	348.5766083	356.1005921	0.021584879	-7.523983872
14	3525.974968	3690.191047	0.04657324	-164.2160787	357.8954059	353.0855126	0.013439383	4.809893326
15	3569.506372	3523.755157	0.012817239	45.75121521	366.1796465	367.6037842	0.003889178	-1.424137651
16	3929.20009	3830.780133	0.025048345	98.41995755	334.3293009	330.5899203	0.011184723	3.739380693
17	3965.817089	4009.499395	0.011014705	-43.68230534	318.3186873	316.684191	0.00513478	1.634496349
18	4287.861436	4171.959814	0.02703017	115.9016216	297.6812573	296.7782532	0.00303346	0.903004095
19	3752.449831	3674.510833	0.020770164	77.93899865	586.5491942	591.0317806	0.007642302	-4.482586312
20	3620.158239	3581.004016	0.010815611	39.15422342	365.6058399	366.513255	0.002481949	-0.90741514
21	3314.992576	3350.639851	0.01075335	-35.64727549	417.4237666	417.9768154	0.00132491	-0.55304877
22	3630.293608	3681.748895	0.014173864	-51.45528715	534.0859204	529.1388769	0.009262636	4.94704349

23	3799.162653	3815.748275	0.004365599	-16.58562206	342.9738387	335.8790978	0.020685954	7.094740926
24	3273.437147	3485.62394	0.064820793	-212.186793	436.11511	436.8438592	0.001671002	-0.728749251
25	3687.793276	3696.682928	0.002410561	-8.88965198	354.0996825	355.9126594	0.005119962	-1.8129769
26	3466.756065	3501.091852	0.009904298	-34.33578644	505.9293623	505.4875164	0.000873335	0.441845934
27	3520.609123	3538.349921	0.005039127	-17.74079809	389.5926622	392.707083	0.007994044	-3.114420755
28	3924.776078	3920.513375	0.001086101	4.262703642	414.3775503	530.8973644	0.281192391	-116.5198141
29	3273.887624	3324.123358	0.015344367	-50.23573347	449.2852325	444.3241033	0.011042271	4.961129182
		Mean	0.018411551	-3.444530747			0.019431892	-4.723086091
		%mean	1.8411551				1.9431892	

Note: y_a = actual energy demand and y_p = predicted energy demand, AD = absolute deviation, $E_y = y_a - y_p$

Appendix IIc. Results of evaluating predictions for energy demand on selected optimal models.

Validation samples	USA			
	y_a	y_p	$AD_{8\text{-}5\text{-}1}$	E_y
1	7516.513395	7574.864666	0.007763077	-58.35127088
2	7847.465894	7791.36459	0.007148971	56.10130359
3	7882.137226	7808.953861	0.009284711	73.18336484
4	7376.270502	7591.986807	0.02924463	-215.716305
5	7755.399664	7697.301307	0.007491343	58.09835687
6	6899.251177	6918.801568	0.002833698	-19.55039146
7	7766.678524	7712.398754	0.006988801	54.27976987
8	7454.970985	7478.277772	0.003126342	-23.30678648
9	6871.441692	6914.672022	0.006291304	-43.23032977
10	7387.878287	7485.699347	0.013240751	-97.82105949
11	7011.181522	6959.895259	0.007314924	51.28626302
12	7159.047769	7045.995205	0.015791564	113.0525641
13	7725.870917	7702.450803	0.003031388	23.42011407
14	7628.421627	7583.534227	0.005884232	44.88740027
15	7714.015745	7699.761345	0.001847857	14.25439966
16	7838.84194	7699.037154	0.017834878	139.8047858
17	7355.183277	7537.599237	0.024801008	-182.4159599

18	7353.422554	7494.808397	0.019227216	-141.3858439
19	6973.067518	6950.457281	0.003242509	22.61023641
20	7691.598995	7605.583353	0.011183064	86.01564144
21	7806.902908	7855.718077	0.006252821	-48.81516873
22	7134.820541	7280.207402	0.020377087	-145.3868606
23	7909.505839	7744.336404	0.020882396	165.1694355
24	7867.953311	7816.162602	0.006582488	51.79070901
25	7628.814127	7643.3477	0.001905089	-14.5335725
26	7449.283875	7397.179442	0.006994556	52.10443295
27	7782.33708	7780.157597	0.000280055	2.179483058
28	6889.903033	6865.385576	0.003558462	24.51745684
29	7861.550302	7759.408973	0.012992517	102.1413284
		Mean	0.009772336	4.97874127
		%mean	0.9772336	

Note: y_a = actual energy demand and y_p = predicted energy demand, AD = absolute deviation, $E_y = y_a - y_p$

References

[1] Asafu-Adjaye, J., The relationship between energy consumption, energy prices and economic growth: time series evidence from Asian developing countries. Energy Economics, 2000. 22(6): p. 615-625.

[2] Acheampong, A.O., Economic growth, CO_2 emissions and energy consumption: What causes what and where? Energy Economics, 2018. 74: p. 677-692.

[3] Kaytez, F., et al., Forecasting electricity consumption: A comparison of regression analysis, neural networks and least squares support vector machines. International Journal of Electrical Power & Energy Systems, 2015. 67: p. 431-438.

[4] Sadorsky, P., Financial development and energy consumption in Central and Eastern European frontier economies. Energy Policy, 2011. 39(2): p. 999-1006.

[5] Sadorsky, P., The impact of financial development on energy consumption in emerging economies. Energy Policy, 2010. 38(5): p. 2528-2535.

[6] Lee, Y.-S. and L.-I. Tong, Forecasting energy consumption using a grey model improved by incorporating genetic programming. Energy Conversion and Management, 2011. 52(1): p. 147-152.

[7] Suganthi, L. and A.A. Samuel, Energy models for demand forecasting—A review. Renewable and Sustainable Energy Reviews, 2012. 16(2): p. 1223-1240.

[8] Stern, D.I., Economic growth and energy, in Encyclopedia of Energy, C.J. Cleveland, Editor. 2004, Elsevier: Amsterdam. p. 35-51.

[9] Shahbaz, M. and H.H. Lean, Does financial development increase energy consumption? The role of industrialization and urbanization in Tunisia. Energy Policy, 2012. 40: p. 473-479.

[10] Salim, R.A. and S. Shafiei, Urbanization and renewable and non-renewable energy consumption in OECD countries: An empirical analysis. Economic Modelling, 2014. 38: p. 581-591.

[11] Sadorsky, P., Energy consumption, output and trade in South America. Energy Economics, 2012. 34(2): p. 476-488.

[12] Mielnik, O. and J. Goldemberg, Foreign direct investment and decoupling between energy and gross domestic product in developing countries. Energy Policy, 2002. 30(2): p. 87-89.

[13] Zhou, P., B.W. Ang, and K.L. Poh, A trigonometric grey prediction approach to forecasting electricity demand. Energy, 2006. 31(14): p. 2839-2847.

[14] Sözen, A., Future projection of the energy dependency of Turkey using artificial neural network. Energy Policy, 2009. 37(11): p. 4827-4833.

[15] Khoshroo, A., et al., Sensitivity analysis of energy inputs in crop production using artificial neural networks. Journal of Cleaner Production, 2018. 197: p. 992-998.

[16] Debnath, K.B. and M. Mourshed, Forecasting methods in energy planning models. Renewable and Sustainable Energy Reviews, 2018. 88: p. 297-325.

[17] Pao, H.-T., H.-C. Fu, and C.-L. Tseng, Forecasting of CO2 emissions, energy consumption and economic growth in China using an improved grey model. Energy, 2012. 40(1): p. 400-409.

[18] Pao, H.-T. and C.-M. Tsai, Modeling and forecasting the CO_2 emissions, energy consumption, and economic growth in Brazil. Energy, 2011. 36(5): p. 2450-2458.

[19] Lin, C.-S., F.-M. Liou, and C.-P. Huang, Grey forecasting model for CO2 emissions: A Taiwan study. Applied Energy, 2011. 88(11): p. 3816-3820.

[20] Wu, L., et al., Modelling and forecasting CO2 emissions in the BRICS (Brazil, Russia, India, China, and South Africa) countries using a novel multi-variable grey model. Energy, 2015. 79: p. 489-495.

[21] Ding, S., et al., Forecasting Chinese CO2 emissions from fuel combustion using a novel grey multivariable model. Journal of Cleaner Production, 2017. 162: p. 1527-1538.

[22] Safa, M., et al., Predicting CO_2 emissions from farm inputs in wheat production using artificial neural networks and linear regression models-Case study in Canterbury, New Zealand. 2016.

[23] Yin, M.-S. and H.-W.V. Tang, On the fit and forecasting performance of grey prediction models for China's labor formation. Mathematical and Computer Modelling, 2013. 57(3): p. 357-365.

[24] Falat, L. and L. Pancikova, Quantitative Modelling in Economics with Advanced Artificial Neural Networks. Procedia Economics and Finance, 2015. 34: p. 194-201.

[25] Prybutok, V.R., J. Yi, and D. Mitchell, Comparison of neural network models with ARIMA and regression models for prediction of Houston's daily maximum ozone concentrations. European Journal of Operational Research, 2000. 122(1): p. 31-40.
[26] Valipour, M., M.E. Banihabib, and S.M.R. Behbahani, Comparison of the ARMA, ARIMA, and the autoregressive artificial neural network models in forecasting the monthly inflow of Dez dam reservoir. Journal of Hydrology, 2013. 476: p. 433-441.
[27] Kohzadi, N., et al., A comparison of artificial neural network and time series models for forecasting commodity prices. Neurocomputing, 1996. 10(2): p. 169-181.
[28] Stamenković, L.J., et al., Modeling of methane emissions using artificial neural network approach. Journal of the Serbian Chemical Society, 2015. 80(3): p. 421-433.
[29] Ahmadi, M.A., Prediction of asphaltene precipitation using artificial neural network optimized by imperialist competitive algorithm. Journal of Petroleum Exploration and Production Technology, 2011. 1(2): p. 99-106.
[30] Ahmadi, M.A., et al., Determination of oil well production performance using artificial neural network (ANN) linked to the particle swarm optimization (PSO) tool. Petroleum, 2015. 1(2): p. 118-132.
[31] Safa, M. and S. Samarasinghe, Determination and modelling of energy consumption in wheat production using neural networks: "A case study in Canterbury province, New Zealand". Energy, 2011. 36(8): p. 5140-5147.
[32] Soltanali, H., A. Nikkhah, and A. Rohani, Energy audit of Iranian kiwifruit production using intelligent systems. Energy, 2017. 139: p. 646-654.
[33] Azadeh, A., S.F. Ghaderi, and S. Sohrabkhani, Annual electricity consumption forecasting by neural network in high energy consuming industrial sectors. Energy Conversion and Management, 2008. 49(8): p. 2272-2278.
[34] Jebaraj, S. and S. Iniyan, A review of energy models. Renewable and Sustainable Energy Reviews, 2006. 10(4): p. 281-311.
[35] Deb, C., et al., A review on time series forecasting techniques for building energy consumption. Renewable and Sustainable Energy Reviews, 2017. 74: p. 902-924.
[36] Moghaddam, A.H., M.H. Moghaddam, and M. Esfandyari, Stock market index prediction using artificial neural network. Journal of Economics, Finance and Administrative Science, 2016. 21(41): p. 89-93.
[37] Kara, Y., M. Acar Boyacioglu, and Ö.K. Baykan, 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, 2011. 38(5): p. 5311-5319.
[38] Wong, S.L., K.K.W. Wan, and T.N.T. Lam, Artificial neural networks for energy analysis of office buildings with daylighting. Applied Energy, 2010. 87(2): p. 551-557.
[39] Hagan, M.T., et al., Neural network design. Vol. 20. 1996: Pws Pub. Boston.
[40] Colwell, J.D., Estimating fertilizer requirements: A Quantitative Approach. 1994: CAB International, Oxon (EUA).
[41] Omri, A. and B. Kahouli, Causal relationships between energy consumption, foreign direct investment and economic growth: Fresh evidence from dynamic simultaneous-equations models. Energy Policy, 2014. 67: p. 913-922.
[42] Sadorsky, P., Do urbanization and industrialization affect energy intensity in developing countries? Energy Economics, 2013. 37: p. 52-59.
[43] Marino, S., et al., A methodology for performing global uncertainty and sensitivity analysis in systems biology. Journal of Theoretical Biology, 2008. 254(1): p. 178-196.
[44] Saltelli, A. and J. Marivoet, Non-parametric statistics in sensitivity analysis for model output: A comparison of selected techniques. Reliability Engineering & System Safety, 1990. 28(2): p. 229-253.
[45] Boateng, E.B., M. Pillay, and P. Davis, Predicting the Level of Safety Performance Using an Artificial Neural Network, in Human Systems Engineering and Design, T. Ahram, W. Karwowski, and R. Taiar, Editors. 2019, Springer International Publishing: Cham. p. 705-710.
[46] Ekonomou, L., Greek long-term energy consumption prediction using artificial neural networks. Energy, 2010. 35(2): p. 512-517.

[47] Ghritlahre, H.K. and R.K. Prasad, Application of ANN technique to predict the performance of solar collector systems - A review. Renewable and Sustainable Energy Reviews, 2018. 84: p. 75-88.
[48] Rumelhart, D.E., B. Widrow, and M.A. Lehr, The basic ideas in neural networks. Communications of the ACM, 1994. 37(3): p. 87-93.
[49] Detienne, K.B., D.H. Detienne, and S.A. Joshi, Neural networks as statistical tools for business researchers. Organizational Research Methods, 2003. 6(2): p. 236-265.
[50] Shahin, M. and M. Elchalakani, Neural networks for modelling ultimate pure bending of steel circular tubes. Journal of Constructional Steel Research, 2008. 64(6): p. 624-633.
[51] Karlaftis, M.G. and E.I. Vlahogianni, Statistical methods versus neural networks in transportation research: Differences, similarities and some insights. Transportation Research Part C: Emerging Technologies, 2011. 19(3): p. 387-399.
[52] Skiba, M., M. Mrówczyńska, and A. Bazan-Krzywoszańska, Modeling the economic dependence between town development policy and increasing energy effectiveness with neural networks. Case study: The town of Zielona Góra. Applied Energy, 2017. 188: p. 356-366.
[53] Macdougall, P., et al. Value assessment of aggregated energy flexibility when traded on multiple markets. 2017. IEEE Computer Society.
[54] Kazemi, A., et al., A hierarchical artificial neural network for transport energy demand forecast: Iran case study. Neural Network World, 2010. 20(6): p. 761-772.
[55] Ridluan, A., M. Manic, and A. Tokuhiro, EBaLM-THP - A neural network thermohydraulic prediction model of advanced nuclear system components. Nuclear Engineering and Design, 2009. 239(2): p. 308-319.
[56] Günay, M.E., Forecasting annual gross electricity demand by artificial neural networks using predicted values of socio-economic indicators and climatic conditions: Case of Turkey. Energy Policy, 2016. 90: p. 92-101.
[57] Kuan, C.-M. and H. White, Artificial neural networks: An econometric perspective. Econometric reviews, 1994. 13(1): p. 1-91.
[58] Knuth, D.E., Structured Programming with go to Statements. ACM Computing Surveys (CSUR), 1974. 6(4): p. 261-301.
[59] Sarle, W.S., Neural networks and statistical models. 1994.
[60] Tamura, S.i. and M. Tateishi, Capabilities of a four-layered feedforward neural network: four layers versus three. IEEE Transactions on Neural Networks, 1997. 8(2): p. 251-255.
[61] Maier, H.R. and G.C. Dandy, Neural network based modelling of environmental variables: a systematic approach. Mathematical and Computer Modelling, 2001. 33(6-7): p. 669-682.
[62] Masters, T., Practical neural network recipes in C++. Academic Press, San Diego, CA. Practical neural network recipes in C++. Academic Press, San Diego, CA., 1993: p. -.
[63] Hippert, H.S., C.E. Pedreira, and R.C. Souza, Neural networks for short-term load forecasting: A review and evaluation. IEEE Transactions on power systems, 2001. 16(1): p. 44-55.
[64] Shahbaz, M., et al., Energy consumption, financial development and economic growth in India: New evidence from a nonlinear and asymmetric analysis. Energy Economics, 2017. 63: p. 199-212.
[65] Karanfil, F., How many times again will we examine the energy-income nexus using a limited range of traditional econometric tools? Energy Policy, 2009. 37(4): p. 1191-1194.
[66] Sbia, R., M. Shahbaz, and H. Hamdi, A contribution of foreign direct investment, clean energy, trade openness, carbon emissions and economic growth to energy demand in UAE. Economic Modelling, 2014. 36: p. 191-197.
[67] Akarca, A.T. and T.V. Long, On the relationship between energy and GNP: A reexamination. Journal of Energy Development 1980. 5(2): p. 326–331.
[68] Huang, B.-N., M.J. Hwang, and C.W. Yang, Causal relationship between energy consumption and GDP growth revisited: A dynamic panel data approach. Ecological Economics, 2008. 67(1): p. 41-54.
[69] Lozano, S. and E. Gutiérrez, Non-parametric frontier approach to modelling the relationships among population, GDP, energy consumption and CO2 emissions. Ecological Economics, 2008. 66(4): p. 687-699.

[70] Sadorsky, P., Trade and energy consumption in the Middle East. Energy Economics, 2011. 33(5): p. 739-749.
[71] StatSoft Inc., A Short Course in Data Mining. 2008, StatSoft, Inc.
[72] Abidoye, R.B. and A.P. Chan, Improving property valuation accuracy: a comparison of hedonic pricing model and artificial neural network. Pacific Rim Property Research Journal, 2018. 24(1): p. 71-83.
[73] Morano, P., F. Tajani, and C.M. Torre, Artificial intelligence in property valuations: an application of artificial neural networks to housing appraisal. Advances in Environmental Science and Energy, Planning, 2015: p. 23-29.
[74] Ghaedi, A.M. and A. Vafaei, Applications of artificial neural networks for adsorption removal of dyes from aqueous solution: A review. Advances in Colloid and Interface Science, 2017. 245: p. 20-39.
[75] LeCun, Y., Y. Bengio, and G. Hinton, Deep learning. nature 521 (7553): 436. Nature, 2015. 521(7553): p. 436-444.
[76] Alvanitopoulos, P.F., I. Andreadis, and A. Elenas, Neuro-fuzzy techniques for the classification of earthquake damages in buildings. Measurement, 2010. 43(6): p. 797-809.
[77] Hecht-Nielsen, R. Kolmogorov's mapping neural network existence theorem. in Proceedings of the international conference on Neural Networks. 1987. IEEE Press.
[78] Rogers, L.L. and F.U. Dowla, Optimization of groundwater remediation using artificial neural networks with parallel solute transport modeling. Water Resources Research, 1994. 30(2): p. 457-481.
[79] El Kadi, H., Modeling the mechanical behavior of fiber-reinforced polymeric composite materials using artificial neural networks - A review. Composite Structures, 2006. 73(1): p. 1-23.
[80] Patel, D.A. and K.N. Jha, Neural Network Model for the Prediction of Safe Work Behavior in Construction Projects. Journal of Construction Engineering and Management, 2015. 141(1).
[81] Godfried, I. Deploying deep learning models: Part 1 an overview. 2018 20 June 2018 [cited 2018 4 December]; Available from: https://towardsdatascience.com/deploying-deep-learning-models-part-1-an-overview-77b4d01dd6f7.
[82] Barga, R., et al., Predictive analytics with Microsoft Azure machine learning. 2015: Springer.
[83] Tadesse, Z., et al., Neural networks for prediction of deflection in composite bridges. Journal of Constructional Steel Research, 2012. 68(1): p. 138-149.