Pakistan Journal of Life and Social Sciences
www.pjlss.edu.pk

https://doi.org/10.57239/PJLSS-2024-22.1.00171

RESEARCH ARTICLE

Estimation of Carbon Dioxide Concentration for Baghdad City Based on Artificial Neural Networks

Saif Hassan[1*], Ali M. Al-Salihi Hasoon[2*], Jasim M. Rajab[3]

[1,2,3] Department of Atmospheric Sciences, College of Science, Mustansiriyah University, Baghdad, Iraq

ARTICLE INFO

Received: May 28, 2024

Accepted: Jun 16, 2024

Keywords

Carbon dioxide

Artificial neural networks

Global warming

Greenhouse gas

Baghdad city

***Corresponding Author:**

saifalshaban202285@gmail.com

ABSTRACT

The study developed a model for predicting CO2 concentration in the center of Iraq's city of Baghdad using artificial neural networks that show the quantities of such emissions that are mainly global warming and can accurately predict greenhouse gas emissions for up to 20 years, supporting environmental management strategies and policies. And using three functions and neural networks. An artificial neural network (ANN) model was used to estimate the ultimate CO2 concentration. Depending on the values of the error and the coefficient of correlation, the ANN model indicated that the suggested model was better appropriate to characterize CO2 in any part of the world. It was shown that the Scaled Conjugate Gradient method (SCG) had the lowest height value of (R) and mean-squared error (MSE). The ideal neuronal count for the SCG hidden layers was 10, with an MSE of 17.24 and R = 0.931. Consequently, the ANN demonstrated outstanding performance in the CO2 prediction value. The results of the study emphasize how important it is to use CO2 prediction as the primary parameter for assessing the impacts of global warming.

INTRODUCTION

The impact of CO_2 concentration on global warming is becoming an important issue, as it is acknowledged that the most urgent problems facing the globe now are related to global warming and climate change in general (Desjardins et al., 2007). Get Accurate data on man-induced CO_2 concentration fluxes into (and out of) the atmosphere is necessary to understand the human role in global climate change (Winiwarter and Rypdal, 2001). The first global attempt to lower CO_2 concentration was the Kyoto Protocol established in 1992 by the United Nations Framework Convention on Climate Change (UNFCCC). Upon the Kyoto Protocol's implementation in 2005, one of its primary goals was to reduce CO_2 concentration by 5.2% relative to 1990 levels. The goal is to reduce CO_2 concentrations by 80% between 2008 and 2012 for sustainable development. However, effective reductions involving all countries will be challenging without the development of appropriate methods and models for simulation. Reducing future CO_2 concentrations is a challenging job. Emerging economies are seeing significant increases in CO_2 concentration due to their rapid economic expansion. Efficiently forecasting greenhouse gas emissions and implementing targeted mitigation actions is a recurring problem. However, ANNs (Artificial Neural Networks) have been employed by several studies and have demonstrated tremendous promise for performing prediction (Nuroğlu, E. 2014; Saleh, C.; Dzakiyullah, N. R.; and Nugroho, J. B. 2016; Mason, K.; Duggan, J.; and

Howley, E. 2018; Kanval et al., 2024). A non-linear computer system termed an artificial neural network (ANN) is composed of several interconnected processing units, or neurons, that simulate the learning processes of the human brain (Balas et al., 2010; Jam et al., 1018). Synapses relate neurons, which are layered structures. The structure of the neural network, including the number of layers, neurons per layer, learning strategy, and other details, is specified by the architecture. The neural network is a widely distributed therapist. that has an innate preference for keeping and retrieving experience data, according to Haykin (1996). Zhang et al. (1998) identified differentiating characteristics of ANN that make them useful for predicting. Because ANNs are so simple to use for modeling, simulation, and prediction, they are widely used in engineering and many other fields. Like this study, the majority of research depended on two metrics to assess the precision of the findings: the (RMSE), which are optimal when it approaches zero. The R, which comes in second, is regarded as reliable when its value falls within the range of 0.75 and 1.

There are a lot of studies that use neural networks to predict emissions. Soc (2015), S. A. Abdul-Wahab, S. M. Al-Alawi, Environ (2002), G. Corani, Ecol (2005), U. Brunelli, V. Piazza, L. Pignato, F. Sorbello, S. Vitabile, (2008), S. P. Shukla, Sci (2012) and A. Russo, F. Raischel, P. G. Lind, Atmos (2013). The aim of the study to produce a model to predicting CO2 above Baghdad city for the period 2003 to 2023, using artificial neural networks with widely available indicators to predict more accurately. Modelling was done using three different types of functions; a neurological network (LM), a network (BR), and (SCG).

1. METHODOLOGY

1.1. The AAN and Types

To predict CO_2 concentration, the data was gathered from (AIRS), including the NASA Aqua Satellite for the whole period (2003-2022). from Iraq (Baghdad) provides an artificial web-based model for estimating CO_2 concentration in Iraqi cities. It is a data-processing system consisting of a variety of very basic and overlapping therapeutic parts that have been inspired and established to simulate the possibility of monotherapy for the human brain (Radojević, D.M., Pocajt). To develop an appropriate forecasting model, the variables of the selected inputs need to cover all CO_2 concentrations. It was therefore used as input (carbon dioxide, carbon monoxide, and methane, plus temperature). CO_2 concentration inputs and data can be obtained from the NASA satellite for five Iraqi cities in 2003 to 2023. The General Regression Neural Network (Levenberg-Marquardt, Bayesian Regularization, and Scaled Conjugate Gradient) architecture is the ANN used in this study. This is the location of Levenberg-Marquardt. Usually; this method takes less time but more memory. When a rise in the mean square error of the validation samples indicates that generalization is no longer improving, training automatically ends. However, Bayesian Regularization Although this technique usually takes longer, it can produce high generalization for challenging, tiny, or noisy datasets. Training halts by regularization or adaptive weight reduction. The Scaled Conjugate Gradient is the final. Less memory is needed for this approach. When generalization no longer improves, as seen by a rise in the validation samples' mean square error, training automatically ends. However, several methods, including MAPE, RMSE, and MAE, may be applied as cost functions to assess the accuracy of the forecast. However, the study employs MSE to perform this function. The MSE as the cost function of the study depends on its analytical tractability. The number of epochs shows when the difference between the actual and goal outputs would become insignificant, whilst the regression value gives the variance % of the explanatory variables to the dependent variables in the model.

1.2. Study Area

Iraq, the research region, is situated in southwestern Asia between latitudes 29–38 N and longitudes 39–49 E (Figure 1). The area is 437,072 km2 in it. The landforms may be divided into four main regions: Aljazeera, undulating upland in the north between the Tigris and upper Euphrates rivers;

alluvial plains in the middle and southeast region; desert areas in the west and southwest region; and highlands in the north and northeast, which are typically composed of mountains. (Malinowski, 2002). Iraq's climate is classified as continental and subtropical, with four distinct seasons: a Mediterranean climate in the country's northeast and central regions, and a subtropical summer climate that is more common in the country's center and south, with mean air temperatures in most areas exceeding 48 C. The winter months are pleasant to chilly, with daytime highs of around 16 degrees Celsius and nighttime lows of 2 degrees Celsius (Metz 1993). Compared to the center and southern regions of Iraq, the northern section saw higher winter rainfall levels (Shubbar et al. 2017; Abdulfattah et al. 2020; Rashid et al., 2023). The study period for this research was 20 years, spanning from January 2003 to December 2022. To determine the CO2 distributions and variances for Baghdad city.

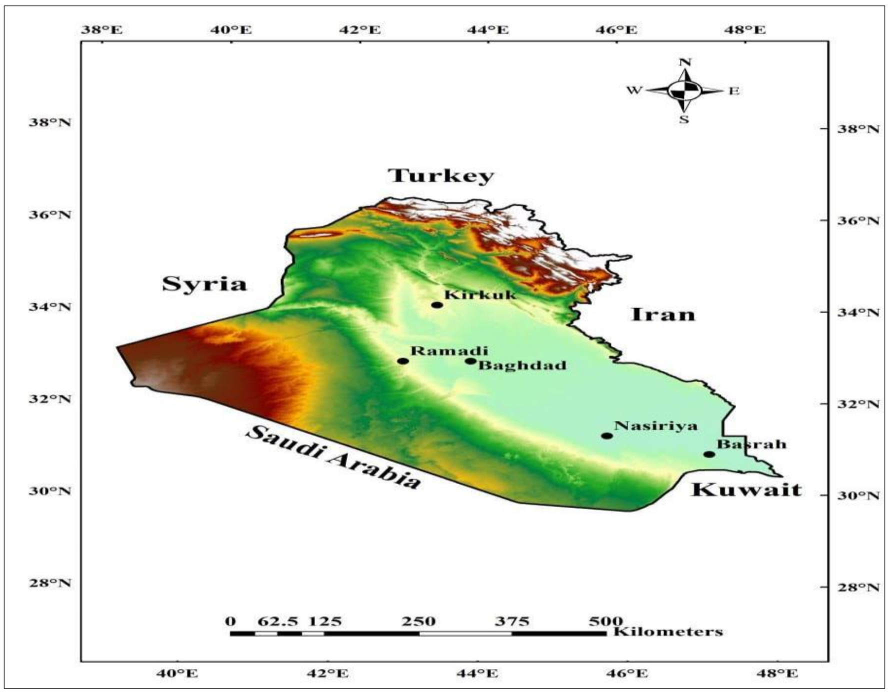

Figure 1: Geographical map of Iraq.

1.3. The Predicting Process

The first step in the process is to sum and choose the variables for the writers' data set. After that, the data is normalized or modified to match its natural logarithms. To construct a model with acceptable bias and variance is the aim of model selection. Sort the data into three groups: test sets, validation sets, and training sets. The study trained the network using an early stopping technique, as shown in Figure 2, to improve generalization.

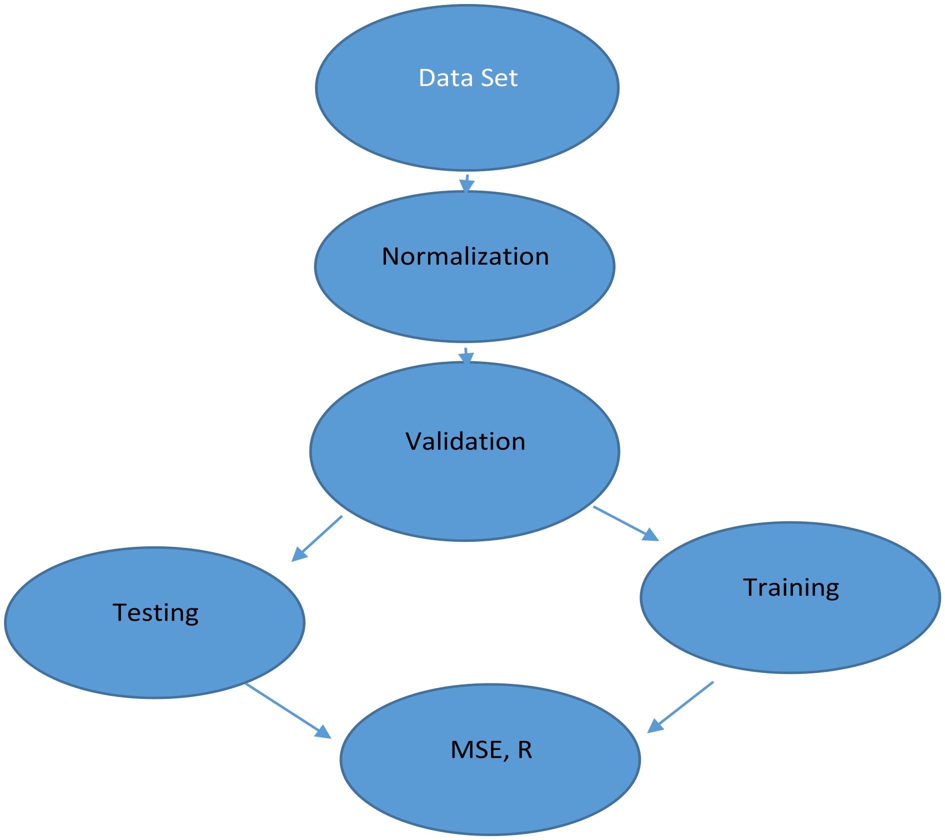

Figure 2: Procedure used for prediction based on ANN.

An artificial neural network (ANN) is conceptualized as a network of neurons arranged in layers. As seen in Figure 3, the layers are composed of the (inputs layer), hidden layers, and (outputs layer).

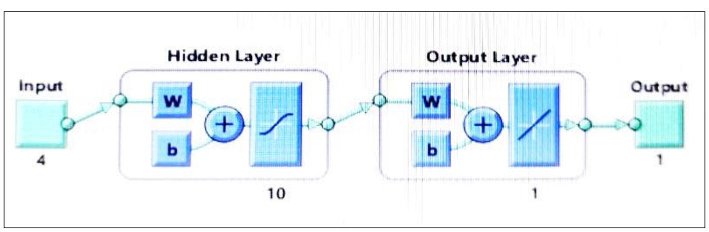

Figure 3: Diagram of the multilayer network.

RESULT

In the present peppers several ANN models including LM, BR, and SCG) in addition to (the MLR) equation for Baghdad city were developed to estimate the monthly mean of CO_2 concentrations employing input data for the study area. This data includes a monthly mean concentration of (CH_4, CO, OLR, and Temperature). For the period from 2003 to 2023. The summary statistical analysis of the input data is tabulated in Table (1).

Table 1: The summary statistical analysis of the input data.

Input parameter	Min	Max	Mean	SD
CO (ppmv)	95.16	148.38	121.90	20.63
CH4 (ppmv)_	1862.58	1888.35	1874.35	7.59
OLR (w/m^2)	275.07	364.65	316.82	33.02
T (K^o)	7.89	31.38	19.60	8.6

Employ the MATLAB 2010 program to develop ANN network input including (CH_4, CO, OLR, and T), while the CO_2 was considered as network output, the data were divided into three sets (Taring, Validation, Testing) with ratios of (80%,10%,10%) respectively. The network was trained using the training set, and at regular intervals during training, its performance was assessed using the validation set. Training ends when the validation set's errors drop to a minimum. Lastly, a test dataset that was not used for training is used to assess the network's performance. Ten (10) neurons were used as the hidden layer in each of the methods under consideration. In addition to (MLR), three algorithms were taken into consideration while selecting the best one. SCG with a least (MSE) is the better algorithm among the proposed algorithms in the present study as tabulated in Table (2).

Table 2: The better algorithm among the proposed algorithms in the present study.

Algorithms	MSE	MAE	RMSE	R	Epoch
LM	24.39	3.632	4.713	0.9431	21
BR	22.18	5.062	6.393	0.9150	4
SCG	17.24	3.928	4.987	0.9311	4
MLR	----	6.69	8.51	0.637	----

The study developed a network by the hidden layer's 10 neurons with chachi algorithms to generate the lowest error. Numerous ANN tests using various algorithms were carried out., and each perfumed neural network was trained, tested, and validated. Table (3) contains a tabulation of the statistical assessment criteria for the models that performed the best.

Table 3: contains a tabulation of the statistical assessment criteria for the models that performed the best

Algorithm	MAE			RMSE			R		
	Training	Validation	Testing	Training	Validation	Testing	Training	Validation	Testing
LM	3.63	0.30	0.26	4.71	0.96	0.93	0.949	0.933	0.926
BR	5.06	0.30	0.51	6.39	1.04	1.79	0.914	0.942	0.891
SCG	3.92	0.31	0.37	4.98	1.01	1.20	0.935	0.944	0.901
MLR	6.69	----	----	8.51	----	----	0.637	----	----

The findings derived from Table (3) indicate that the ANN with (4-10-1). The SCG training algorithm is used to identify the algorithms that perform the best based on statistical assessment criteria and are most suitable for the values of the CO2 gas concentrations under study. For the training, validation, and testing stages, the R was (0.935, 0.944, and 0.901), respectively. It was discovered that the networks performed well in identifying the CO2 concentration that they trained to detect. The ANN demonstrated effectiveness in generalized training by predicting CO2 values for unknown cases, with high R and low MSE. The network had an MSE of 17.24 for the total dataset, with performance diagram 4 showing its three phases of performance and the best epoch.

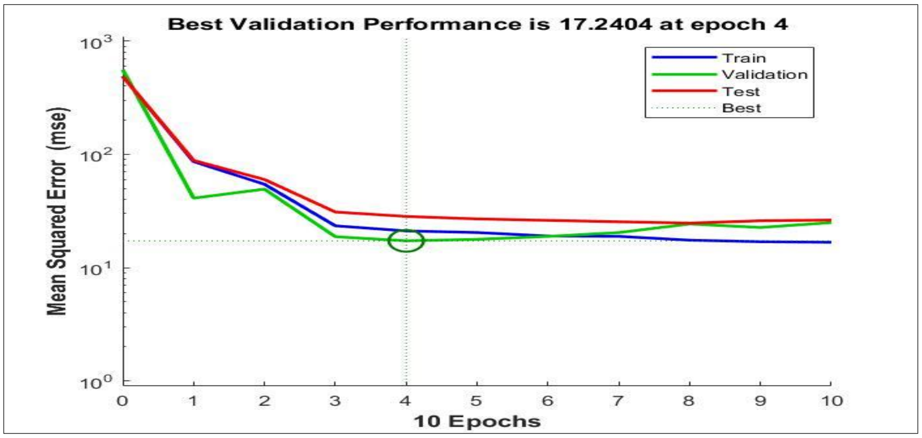

Figure 4: the three-stage Network performance (4-10-1) and the best epoch.

The network's performance is evaluated through metrics that show its ability to recognize and forecast events not covered during the training stage. Validation is an important phase in which the trained ANN is evaluated using unseen data. During training, several examples are provided, and the ANN's internal parameters are adjusted depending on these instances. Validation approaches are used to show the model's performance on previously unknown data. This study indicates if the model needs additional intelligence by emphasizing patterns or whether it has picked up redundant information from the training set, reducing its ability to anticipate new data. ANNs are primarily trained on training data to learn fundamental correlations and patterns for new data. Validation is a crucial stage in creating and assessing ANN models, providing information on their dependability and efficacy. It helps in performance evaluation, generalized estimates, hyperparameter tuning, model selection, and overfitting control. Training with validation ensures the development of reliable, accurate, and well-generalized models that can predict new data. To compare networks, the error function is assessed using separate data from the training set, aiming to identify the best network for fresh data. Minimizing the error function defined by the training data set is essential for training different networks. To compare the performance of each network, the error function is assessed using a different validation set. When the error on the validation set rises, early halting is applied to the network with the lowest error. The candidate models are iterative variants of the same network that are developed over time. Figure 5 depicts the MSE fluctuation by epoch for training, validation, and testing. The best validation performance is 17.24 across four epochs. The validation error is at its lowest in the fourth epoch, thus training was ended at that time.

Figure 5 shows high correlation coefficients between network values and measured CO2 values, indicating the model's ability to simulate the studied phenomenon. Validated findings reveal that ANN output is stable across all climatic data factors, and output CO2 quickly converges to expected values, confirming the usefulness of the proposed ANN modules. The validation step saw a substantial increase, with the network predicting CO2 concentration for previously encountered situations with an accuracy of 17.24.

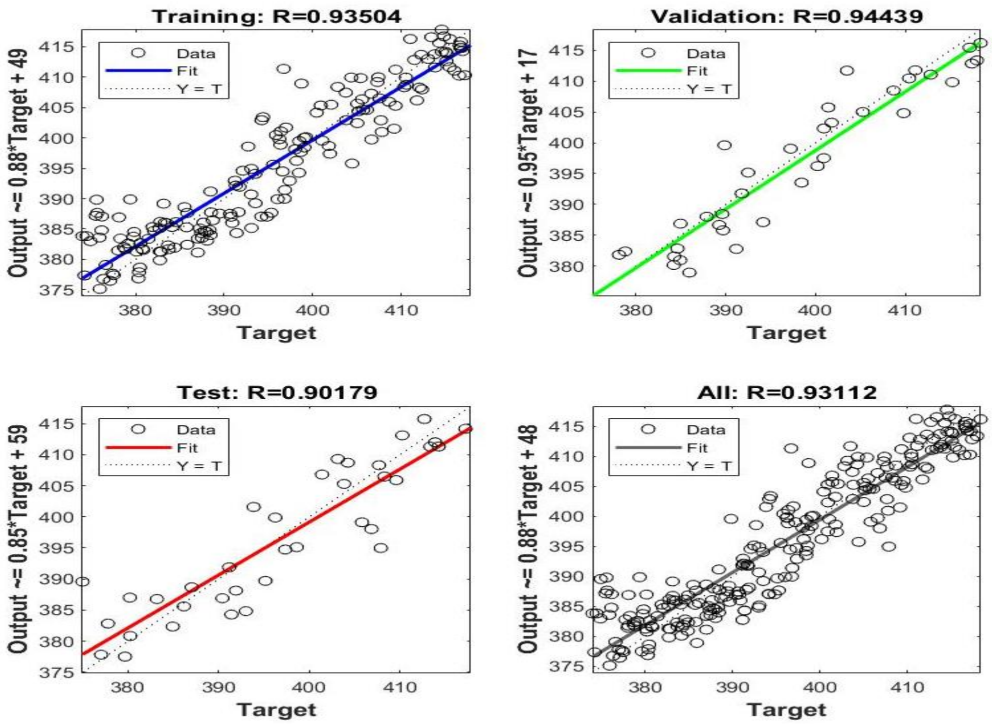

Figure 5: Relationship graphic for the SCG function between CO2 levels and ANN-generated values.

1.4. Compared with measuring CO_2

As seen in Figure 6, the prediction of CO_2 obtained from the ANN model is compared to the measured results of CO_2.

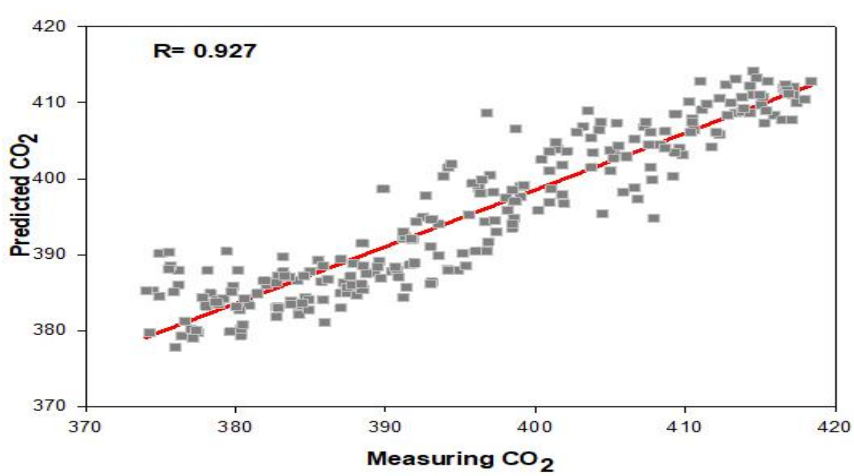

Figure 6: study of regression between values that are measured and predict coefficient of correlation R = 0. 927.

CONCLUSIONS

Depending on temperature, CH4, CO, OLR, and other factors, the ANN with the structure (4-10-1) demonstrated its capacity to compute the CO_2 content of the city of Baghdad, with an MSE of 17.24 and a R of 0.931 for the validation set. The predicted values and the measurements were extremely close. Artificial neural networks (ANN) developed to fill in the gaps in long-term CO_2 concentration data can be used in the study. The best function of CO2 forecasting for Baghdad City is SGC because it has less MSE than the rest of the functions. R=0.927 values between measured and predicted values and that indicates the closeness and strength of the relationship.

REFERENCES

A. Russo, F. Raischel, and P. G. Lind, "Air quality prediction using optimal neural networks with stochastic variables," Atmospheric Environment, vol. 79, pp. 822-830, 2013.

Abdulfattah IS, Rajab JM, Al-Salihi AM, Suliman A, Lim HS (2020) Observed vertical distribution of tropospheric carbon monoxide during 2012 over Iraq. Sci. Rev. Eng. Env. Sci. 29(2):184.

C. Saleh, N. R. Dzakiyullah, and J. B. Nugroho, "Carbon dioxide emission prediction using support vector machine," in IOP conference series: materials science and engineering, 2016, p. 012148.

D. Antanasijević, V. Pocajt, I. Popović, N. Redžić, and M. Ristić, "The forecasting of municipal waste generation using artificial neural networks and sustainability indicators," Sustainability Science, vol. 8, pp. 37-46, 2013.

E. Nuroğlu, "Estimating and forecasting trade flows by panel data analysis and neural networks," İstanbul Üniversitesi İktisat Fakültesi Mecmuası, vol. 64, pp. 85-111, 2014.

G. Corani, "Air quality prediction in Milan: feed-forward neural networks, pruned neural networks and lazy learning," Ecological modeling, vol. 185, pp. 513-529, 2005.

J. J. Vredenburgh, A. Desjardins, J. E. Herndon, J. Marcello, D. A. Reardon, J. A. Quinn, et al., "Bevacizumab plus irinotecan in recurrent glioblastoma multiforme," Journal of clinical oncology, vol. 25, pp. 4722-4729, 2007.

J. Segalàs, D. Ferrer-Balas, and K. F. Mulder, "What do engineering students learn in sustainability courses? The effect of the pedagogical approach," Journal of Cleaner production, vol. 18, pp. 275-284, 2010.

Jam, F. A., Singh, S. K. G., Ng, B., & Aziz, N. (2018). The interactive effect of uncertainty avoidance cultural values and leadership styles on open service innovation: A look at malaysian healthcare sector. *International Journal of Business and Administrative Studies*, *4*(5), 208-223.

K. Mason, M. Duggan, E. Barrett, J. Duggan, and E. Howley, "Predicting host CPU utilization in the cloud using evolutionary neural networks," Future Generation Computer Systems, vol. 86, pp. 162-173, 2018.

K. P. Singh, S. Gupta, A. Kumar, and S. P. Shukla, "Linear and nonlinear modeling approaches for urban air quality prediction," Science of the Total Environment, vol. 426, pp. 244-255, 2012.

Kanval, N., Ihsan, H., Irum, S., & Ambreen, I. (2024). Human Capital Formation, Foreign Direct Investment Inflows, and Economic Growth: A Way Forward to Achieve Sustainable Development. *Journal of Management Practices, Humanities and Social Sciences*, *8*(3), 48-61.

Malinowski JC (2002) Iraq: A geography. United States Military Academy, West Point Department of Geography & Environmental Engineering, Building 600, West Point, New York 10996.

Metz, H. C. (1993). Iraq: a country study, Federal Research Division, Library of Congress.

R. M. Shubbar, H. H. Salman, and D. I. Lee, "Characteristics of climate variation indices in Iraq using a statistical factor analysis," International Journal of Climatology, vol. 37, pp. 918-927, 2017.

Rashid, A., Jehan, Z., & Kanval, N. (2023). External Shocks, Stock Market Volatility, and Macroeconomic Performance: An Empirical Evidence from Pakistan. *Journal of Economic Cooperation & Development*, *44*(2), 1-26.

S. A. Abdul-Wahab and S. M. Al-Alawi, "Assessment and prediction of tropospheric ozone concentration levels using artificial neural networks," Environmental Modelling & Software, vol. 17, pp. 219-228, 2002.

S. A. Abdul-Wahab, C. S. Bakheit, and S. M. Al-Alawi, "Principal component and multiple regression analysis in modeling of ground-level ozone and factors affecting its concentrations," Environmental Modelling & Software, vol. 20, pp. 1263-1271, 2005.

S. Haykin, "Neural networks expand SP's horizons," IEEE Signal Processing Magazine, vol. 13, pp. 24-49, 1996.

T. Zhu, J. Zhang, and S. N. Atluri, "A meshless local boundary integral equation (LBIE) method for solving nonlinear problems," Computational Mechanics, vol. 22, pp. 174-186, 1998.

U. Brunelli, V. Piazza, and L. Pignato, "A neural network model for three-hours-ahead prediction of ozone concentration in the urban area of Palermo," WIT Transactions on Ecology and the Environment, vol. 86, 2006.

U. Brunelli, V. Piazza, L. Pignato, F. Sorbello, and S. Vitabile, "Three hours ahead prevision of SO2 pollutant concentration using an Elman neural based forecaster," Building and Environment, vol. 43, pp. 304-314, 2008.

W. Winiwarter and K. Rypdal, "Assessing the uncertainty associated with national greenhouse gas emission inventories: a case study for Austria," Atmospheric environment, vol. 35, pp. 5425-5440, 2001.

.