CENTRAL ASIAN JOURNAL OF SUSTAINABILITY AND CLIMATE RESEARCH (2024) 3(2):1-23

https://doi.org/10.29258/CAJSCR/2024-R1.v3-2/1-23.eng

CAJSCR

© The Author(s) 2024

CO2 emission estimation in Central Asian countries by use of artificial intelligent methods

Balzhan Azibek [a], Andrii Biloshchytskyi [a], Mohammad Alhuyi Nazari [b], Nurkhat Zhakiyev [a]

[a] Astana IT University, Mangilik El avenue, 55/11, Business center EXPO, block C1, Astana, 010000, Kazakhstan
[b] University of Tehran, Enghelab Sq., 16th Azar St., Tehran, 1417935840, Iran

ABSTRACT

Energy consumption and shares of different types of energy sources in the total energy supply play a critical role in CO2 emissions in various countries. Aside from the energy-related factors, economic indicators such as Gross Domestic Product (GDP) can also influence emissions. In the present study, shares of different energy sources in total energy supply and GDP were utilized as inputs to propose a model for estimating CO2 emissions in three Central Asian states, namely Kazakhstan, Uzbekistan, and Turkmenistan. In addition to the modelling outputs, the article likewise describes important characteristics of energy systems in target countries. Based on the comparison of CO2 emissions per GDP unit, the study allowed identifying that this index in the target countries exceeds the global average, which necessitates urgent actions to reduce emissions. The input data for the research were obtained from the International Energy Agency (IEA) and World Bank. The study applied the Group Method of Data Handling (GMDH) and the Multilayer Perceptron (MLP) method. Both models showed significant performance in emissions estimation; however, according to the calculated values for the applied criteria, it was concluded that the GMDH led to better exactness compared to the MLP. The mean absolute relative deviations of the GMDH and MLP modes were approximately 3.69% and 4.28%, respectively. The R^2 value of the mentioned models were 0.9936 and 0.9929, respectively. In the majority of the cases for both models, relative deviations between the predicted and actual CO2 emission were in the range of ±5%.

ARTICLE HISTORY

Received:
July 2, 2024
Accepted:
September 22, 2024
Published:
September 30, 2024

KEYWORDS

Artificial neural network, CO2 emissions, Central Asian countries, GMDH, intelligent methods

CONTACT Nurkhat Zhakiyev ✉ nurkhat.zhakiyev@astanait.edu.kz, Astana IT University, Mangilik El avenue, 55/11, Business center EXPO, block C1, Astana, 010000, Kazakhstan

Published by Kazakh German University, Almaty/Kazakhstan

1. Introduction

There is a growing trend in the emission of greenhouse gases, mainly CO2, due to various factors such as population growth, industrialization, increment in use of fossil fuels etc. Energy plays a critical role in the emission of greenhouse gases and must be considered for making policies for the mitigation of these gases. According to the report of the International Energy Agency (IEA) (CO2 Emissions in 2023, A New Record High, but Is There Light at the End of the Tunnel, 2024), energy-related CO2 emissions in the world increased by 1.1% in 2023. Over ten years ending with 2023, emissions of CO2 in the world have elevated by slightly higher than 0.5% per year. Several approaches and pathways have been proposed for the reduction of energy-related CO2 emission, including the deployment of renewable energies for power generation, transport, and heat production (Renewables, n.d.), development of cleaner energy technologies such as heat pumps for buildings heating and cooling (Heat Pumps, n.d.) and Electric Vehicles (EVs) for transport sector (Electric Vehicles, 2023), development of carbon capture systems (Carbon Capture, Utilisation and Storage, n.d.) etc. The development of models for the prediction of CO2 emission and identification of the critical factors affecting the amount of greenhouse gas emissions help make proper policies towards the reduction of their emission. These policies must be based on complex dynamics between energy consumption, economic growth, and population dynamics. In a study (Turmunkh, 2021), the development of a sustainable energy strategy in Central Asia countries was presented. The author provides statistical analysis and recommendations for policymakers aiming to balance economic development with environmental sustainability in the region. Similar research on panel data econometrics, considering the time-series and cross-sectional dimensions is performed in (Yuldoshboy et al., 2022), the conclusion is made that economic activities also contribute to higher CO2 emissions, reflecting the region's reliance on fossil fuels. The authors recommend investments in energy efficiency and renewable energy sources can be crucial to mitigate CO2 emissions.

Central Asia's economy is deeply intertwined with fossil fuel production and consumption. The region is rich in fossil fuel resources, which plays a critical role in its economic development. Countries like Kazakhstan and Uzbekistan are major exporters of oil and natural gas, and their economies are significantly influenced by fluctuations in global energy markets. Studies have shown that the energy sector in Central Asia is heavily reliant on fossil fuels (Kuziboev et al., 2024; Radovanović et al., 2021), contributing to high levels of CO2 emissions. In another study (Xiu, 2022), a significant positive correlation between CO2 emissions and economic growth in Central Asian countries was reported. The elasticity of CO2 emissions concerning GDP is estimated at 0.82, indicating that a 1% increase in GDP is associated with a 0.82% increase in CO2 emissions.

Projections of CO2 emissions in Central Asia indicate a continuing rise in emissions if current fossil fuel dependency persists. According to (Zhakiyev et al., 2023), the carbon dioxide emission is projected to 30% increase if no energy policy is utilized, therefore policy reforms and the adoption of renewable energy sources are critical to mitigating future emissions. Some positive forecasts are provided (Filipović et al., 2024), by estimating a 30% CO2 reduction in three decades in Central Asia countries considering carbon pricing policy.

According to the literature review on CO2 estimation in the Central Asia area, it can be stated that utilizing a panel data approach for CO2 emission estimation comes with its own set of challenges. This approach can sometimes limit the complexity of the model, as it may struggle to fully capture the nuanced impacts of various socio-economic and technological factors over time and across different regions. The panel approach, while helpful for identifying overarching trends and relationships, might not adequately account for the specific local contexts or temporal dynamics that influence emissions. This can lead to potential biases or oversimplifications in the estimation process. Among different techniques and approaches applicable to the modelling of engineering and environmental problems, intelligent methods have some advantages such as the ability to model with relatively high accuracy, lower computation cost compared with the numerical simulation and fast performance. AI and other advanced technologies can play a crucial role in accurately estimating emissions and guiding effective policy decisions. The continued focus on regional cooperation and policy innovation will be essential in addressing the environmental challenges posed by fossil fuel dependency in Central Asia.

Intelligent methods such as Artificial Neural Network (ANN) are applicable in different energy-related fields of science. In greenhouse emissions estimation, ANN and deep learning neural networks are compared (Altikat, 2021). The study demonstrates that deep learning models are more effective than traditional neural networks in predicting CO2 emissions from greenhouses. However, the techniques have been applied for a variety of systems in ANN-based approaches, such as performance prediction of renewable energy systems (Alhuyi Nazari et al., 2023), forecast of weather data affecting the outputs of energy systems such as wind turbines (Filik & Filik, 2017), energy consumption in different sectors (Luo et al., 2020) and estimation of energy price (Pindoriya et al., 2008). Good evidence is provided in (M. K. & V., 2020) with a multilayered ANN model that includes multiple hidden layers between input and output layers used to capture complex patterns in the data related to CO2 emissions. The ANN model in (Rezaei et al., 2018a) demonstrated a high level of accuracy in predicting CO2 emissions, showing a strong correlation between energy use and emissions. The study found that increases in energy consumption, particularly from non-renewable sources, lead to higher CO2 emissions. Conversely, economic growth driven by renewable energy sources was associated with lower emissions levels. In 2018, Rezaei et al. (Rezaei et al., 2018b) employed the Group Method of Data Handling (GMDH) to forecast the emission of CO2 in four Nordic countries. Their model showed significant accuracy with an R^2 of 0.998. In 2019, Ahmadi et al. (Ahmadi, Jashnani, et al., 2019) developed a model by using shares of different energy sources utilized as primary energy supply in addition to Gross Domestic Product (GDP) as the inputs and GMDH as the intelligent technique for CO2 emission estimation in five Middle Eastern countries. The Average Absolute Relative Deviation (AARD) of the predictive model proposed in their study was 2.3%. In 2020, Ghalandari et al. (Ghalandari et al., 2020) used data-driven techniques for CO2 emission forecasting in four European countries. The use of various energy resources in addition to the GDP were inputs for the ANN-based models. Their models could estimate the CO2 emission of the countries with an R^2 of 0.9999. Birjandi et al. (Komeili Birjandi et al., 2022a) used the Multilayer Perceptron (MLP) method with two different transfer functions namely radial basis and tansig for the estimation of CO2 emission in countries in Southeast Asia. They reported that the use of the radial basis function in optimal structure leads to higher accuracy compared with the optimal model with the tansig function.

According to the literature review, it is concluded that intelligent methods, particularly ANN-based ones, are efficient and accurate tools for the prediction of CO2 emission. Their accuracy is dependent on the architecture and applied methods and these factors should be considered to develop the precise model. This article focuses on the CO2 emission modelling of three countries namely Kazakhstan, Uzbekistan and Turkmenistan in central Asia by use of ANN-based methods. The main novelty of the present work is the simultaneous consideration of these three countries for the development of a comprehensive model. Furthermore, some helpful information is provided on the key characteristics of their energy systems and policies considered for the future of their energy systems. In the following section, an overview of the energy systems and CO2 emission status of these countries are provided. Afterwards, the applied methods are described. Subsequently, results and discussion are provided and finally, the conclusion is represented.

2. Overview of Energy Systems in Cases of the Study

Three countries, namely Kazakhstan, Uzbekistan, and Turkmenistan, located in central Asia are considered in the present study for modelling of CO2 emission using ANN-based models. In this section, key features and policies related to the energy systems and environment of these countries are presented.

2.1. Kazakhstan

The Republic of Kazakhstan has a land area of 2,717,300 km^2 (Executive Summary, n.d.) and its population was around 19,621,000 in 2022 according to the World Bank data (Population, Total - Kazakhstan, n.d.). Gas and oil industries and associated sectors were responsible for 17% of GDP in 2020, oil provides the majority of the earnings of the country's exports and is the government revenue chief source (Executive Summary, n.d.). In 2021, oil products had the highest share in Total Final Consumption (TFC) that was followed by coal. High reliance on fossil fuels causes high CO2 emissions per unit of GDP compared with the world and developed countries. In Figure 1, the emission of CO2 between 1990 and 2021 is illustrated (Energy Statistics Data Browser, n.d.-a). It can be seen that although there was a sharp reduction in emissions between 1990 and 2001, there has been an increasing trend in recent years, particularly since 2015. Coal has the highest share in electricity generation in this country and it is followed by natural gas. The industry had the highest share in TFC of Kazakhstan between 2000 and 2019; however, from 2020 the share of the residential sector became higher than the industry sector (Energy Statistics Data Browser, n.d.-a).

Some policies have been made for the energy systems in Kazakhstan in recent decades. For instance, according to the law about the support of the utilization of renewable energy systems, which was adopted in 2009 and the latest amendment in 2021, feed-in tariff (FIT), grants and tax relief were introduced and rules were outlined for land allocation for renewable energy systems utilities (The Law About Support the Use of Renewable Energy Sources (Amended), 2022). Another policy, the law on Energy Conversation and Energy Efficiency, a strategic policy is designed to achieve a remarkable reduction in the energy consumption of industrial and municipal sectors. This law focuses on the enhancement of energy conservation, development of efficient infrastructures, and shifting towards green growth of the national economy (KAZAKHSTAN: Law No. 541-IV of 2012 on Energy Saving and Energy Efficiency (2019 Ed.), 2019). Another policy for the energy sector is entitled to energy efficiency classes of buildings. This policy focuses on the development of rules for the definition and revision of buildings' energy efficiency classes (Energy Efficiency Classes of Buildings, 2022). This policy can lead to improvement in the energy-saving in the buildings and reduction in the energy consumption and consequently decrease in emission of this sector. Furthermore, the United Nations Development Programme in Kazakhstan in partnership with this country's government and with financial help from the Global Environment Facility initiated programmes for supporting the medium- and small-sized enterprises carrying out the energy efficiency and renewable energy systems. The aim of this initiative is promotion of renewable energy projects on a small scale and decrement the financial burden of clean energy systems (Transition to Renewable

Energy Sources: Economic Benefits for Entrepreneurs in Kazakhstan, 2024). Between 2018 and 2020, the U.S. Agency for International Development (USAID) helped Kazakhstan leverage approximately 2 billion USD in private investment, leading to around 60 new projects on renewable energy which would decrease the emissions of CO2 in this country by 11 Mt (Betting Big on Renewables, n.d.). Aside from the decarbonization, there are some investments and plans for the improvement of energy security in Kazakhstan. For instance, Central Asia Regional Economic Cooperation (CAREC) has invested around 8.9 billion USD in different projects that are mainly aimed at bilateral electricity trade expansion and enhancement of power networks in this region (Energy, 2023). Implementation of this plan would lead to the enhancement of energy security in the considered countries, e.g. Kazakhstan.

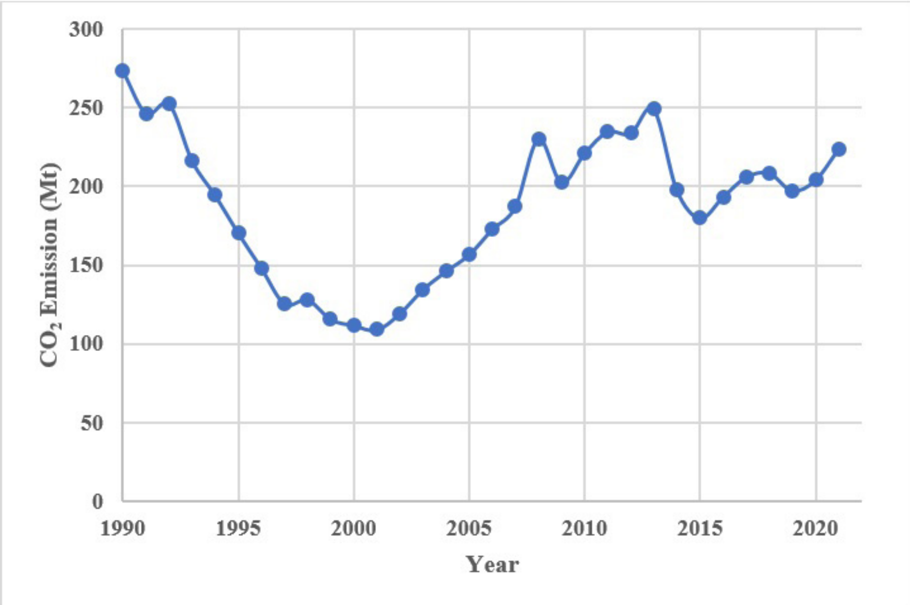

Figure 1. Total CO2 emission in Kazakhstan between 1990 and 2021 (Energy Statistics Data Browser, n.d.-a).

2.2. Uzbekistan

The land area of Uzbekistan is around 447,400 km^2 (Uzbekistan (12/08), n.d.) and its population was approximately 35,648,000 in 2022 based on the data of the World Bank (Population, Total - Uzbekistan, n.d.). Despite being energy self-sufficient owing to its gas sector, the ageing infrastructure of this country struggles to face increasing domestic demand. Overuse, losses and financing remain problems in the energy sector of this country (Uzbekistan, n.d.). In this country, natural gas had the highest share in TFC and was followed by electricity in 2022 (Energy Statistics Data Browser, n.d.-b). CO2 emission per unit of GDP in this country is high in comparison with the world which can be attributed to the low efficiency of the systems and high reliance on fossil fuels. In Figure 2, Uzbekistan's CO2 emission between 1990 and 2021 is represented. It can be seen that since 2015, there has been an increasing trend in the emission of CO2. Natural gas has the highest share in electricity generation in Uzbekistan, and it was followed by hydropower in 2021. This year, the residential sector had the highest share in the TFC of Uzbekistan which was followed by the industry sector (Energy Statistics Data Browser, n.d.-b).

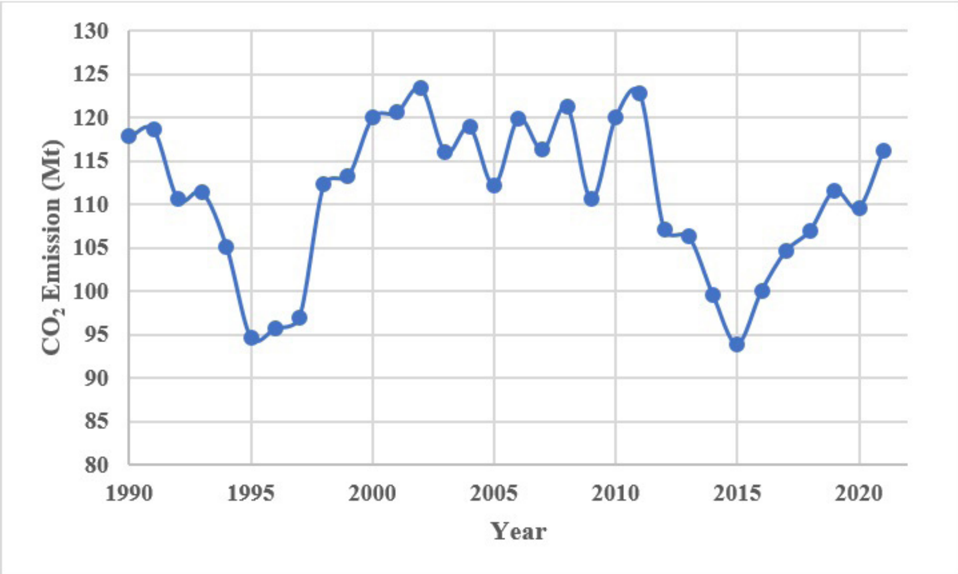

Figure 2. Total CO2 emission in Uzbekistan between 1990 and 2021 (Energy Statistics Data Browser, n.d.-b).

Some policies and plans have been considered in the energy sector of Uzbekistan in recent years. The general items focused on by the government are designed in a way to achieve fuel independence by incrementing the natural gas and petroleum condensate output, creating raw materials base for the energy sector, maximising public access to electricity, liquefied and natural gas and the latest kinds of fuels, promotion of financial stability and attraction of more investment in this sector, utilization of energy resources in a more efficient way, development and legal framework for the energy sector and enhance the related financial-taxation system and promotion of competition in the energy sector (The Outlook for the Development of Renewable Energy in Uzbekistan, 2014). In addition to the general policies, some of the plans and policies are more specific. For instance, according to the Decree of the President of the Republic of Uzbekistan, primary directions for more development of the fuel and energy industry of the country are defined. There are different aims for it including the introduction of advanced resources and technologies applicable for energy saving in the economy and household sectors in addition to the development of new energy sources (Decree of the President of the Republic of Uzbekistan “On Measures to Radically Improve the Management System of the Fuel and Energy Industry of the Republic of Uzbekistan” Dated 01.02.2019, №UP-5646, 2022). According to another policy, a law is presented that aims to strengthen the energy security of Uzbekistan, fuel diversification, and balance of energy in terms of electricity generation, heat and biogas utilizing renewable energies (Law of the Republic of Uzbekistan “On the Use of Renewable Energy Sources” Dated May 21, 2019, No. ZRU-539, 2022). Furthermore, this country set a 25% target for renewable energy generation by 2030 and carbon neutrality by 2050. USAID is cooperating with the stakeholders of the energy sector and the Ministry of Energy to initiate a Green Hydrogen Hub to support the efforts of the government to reach the mentioned goals (USAID Energizes Uzbekistan’s First Green Hydrogen Hub, n.d.).

2.3. Turkmenistan

Another case study of the present study is Turkmenistan located in Central Asia. The land area of this country is 469,930 km^2 and the population of this country was around 6,431,000 in 2022 (Population, Total - Turkmenistan, n.d.). The government of Turkmenistan was continuously investing in gas and oil to expand and modernize the heat and electricity sector by 2020. In addition, the energy sector is approximately fully subsidized (Turkmenistan, n.d.). The highest share of TFC belonged to natural gas in 2021 which was followed by the oil products (Energy Statistics Data Browser, n.d.-c). CO2 emission per unit of GDP is high compared with the world and is higher than the two other countries considered in the present study. In Figure 3, the emission of CO2 between 1990 and 2021 in Turkmenistan is illustrated. Between 1998 and 2011, there was a significant increase in emissions; however, its increasing trend has reduced in recent years. Among different sectors of this country, commercial and public services had the highest share of TFC in 2021 (Energy Statistics Data Browser, n.d.-c).

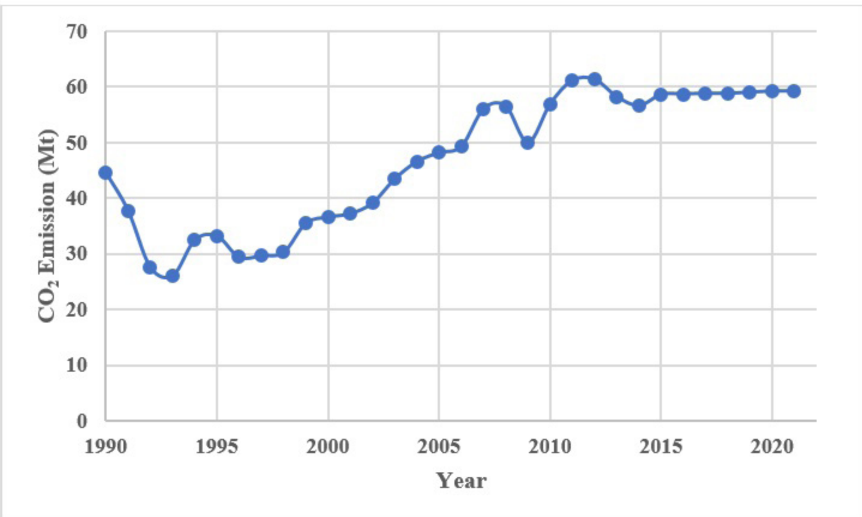

Figure 3. Total CO2 emission in Turkmenistan between 1990 and 2021 (Energy Statistics Data Browser, n.d.-c).

There are some policies and plans for environmental protection and modification of energy systems in Turkmenistan. For instance, according to one of the policies on the protection of atmospheric air, a law is adopted to provide the set standards for air quality (On Protection of the Atmospheric Air, 2022). According to another policy, Turkmenistan's Nationally Determined Contribution (NDC) reflects decrement of 20% in the emission of greenhouse gases in 2030 in comparison with 2010 emissions under the business-as-usual scenario (Nationally Determined Contribution (NDC) to the Paris Agreement (2022 Update): Turkmenistan, n.d.). A project entitled "Sustainable Cities in Turkmenistan: Integrated Green Urban Development in Ashgabat and Avaza", which is funded by the UNDP and Global Environmental Fund aims to promote sustainable cities development and decrease unfavourable effects of urban growth in Turkmenistan (UNDP Continues to Support Turkmenistan in Improving Energy Efficiency and Developing Renewable Energy Sources, n.d.). As another international cooperation and support, USAID aids in the identification of opportunities related to renewable energy systems, helps in the development of large-scale and comprehensive low-carbon strategies in the sector of energy and takes part in the decrement of emissions of methane from the gas and oil sector (USAID Power Central Asia, n.d.).

3. Methods

There are various kinds and structures based on ANN for the development of predictive models. MLP is one of the most conventional ANN-based models that have been applied for the modelling of multiple systems and problems in recent decades. The general structure of an MLP ANN with one hidden layer is shown in Figure 4. A description of this model is presented here according to the study by Komeili et al. (Komeili Birjandi et al., 2022a). In various layers of this structure, neurons are located. Input information is received by the input layer and transferred to the next layer through the neurons. The communication ability of a neuron in comparison with the other neurons represents its weight. The quantity of neurons in each layer is associated with the weight value and value of the network's former layer. The principal processor component of this network is a neuron.

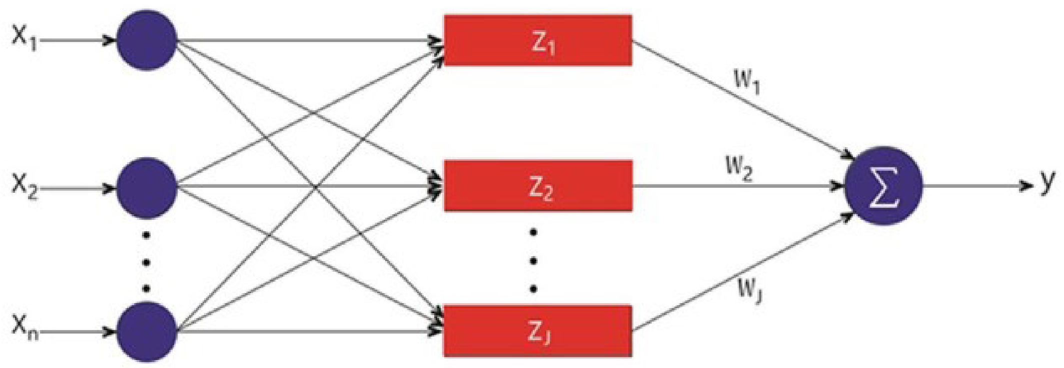

Figure 4. Structure of an MLP ANN with a single hidden layer (Komeili Birjandi et al., 2022a).

Considering X vector (X_1, X2, …, X_n) as the network input, the weight of the neurons which stands for the jth neuron synapse power as ω_j, u_j as the linear combination of output and y_j as the network output, the following equations are used for determination of the output.

$$u_j = \sum_{i=1}^{n} \omega_{ji} X_i \quad (1)$$
$$y_j = \varphi(u_j + b_j) \quad (2)$$
$$v_j = u_j + b_j \quad (3)$$

In these equations, *φ(0)* and b_j are activation function and bias, respectively. A variety of activation functions are applicable in this network. tansig as one of the most conventional activation functions, as presented in Equation (4) (Komeili Birjandi et al., 2022a), is applied in the present study.

$$tansig\ (x) = \frac{2}{1+\exp\ (-2x)} - 1 \quad (4)$$

In the applied model, Levenberg-Marquardt is used as the training function for optimization of the network parameters. In addition, one hidden layer with different numbers of neurons is considered to obtain the optimal structure. The other hyperparameter values such number of epochs are selected based on the default value of MATLAB software.

Another technique applied in the present work for modelling is GMDH. This approach is described here according to the study by Kim et al. (Kim & Okuyucu, 2022). This method maps an input vector, X=(x_1, x_2,…,x_n) to the forecasted value denoted by $\hat{y}_i$ as illustrated in Figure 5. It is anticipated to have as near as possible the value of output to its actual value. M results determined for pairs of data in a network with single output and multiple inputs are observed as follows (Oh & Pedrycz, 2002).

$$y_i = f(x_{i1}, x_{i2}, \cdots, x_{in}) \quad i = 1, 2, \ldots, M$$

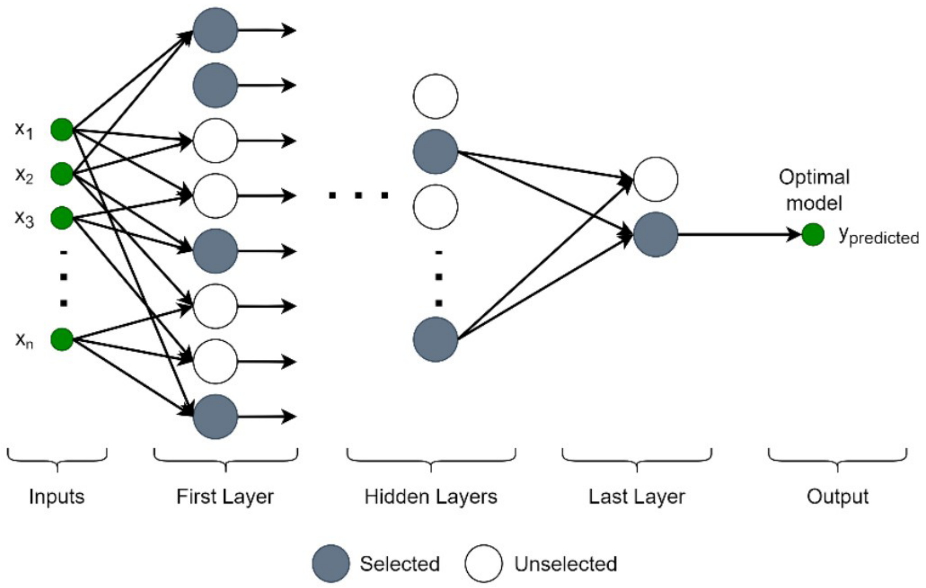

Figure 5. Schematic of GMDH ANN (Kim & Okuyucu, 2022)

The predicted value of the output, $\bar{y}_i$, from the inputs is represented as follows:

$$\bar{y}_i = \bar{f}\ (x_{i1}, x_{i2}, \ldots, x_{in}) \quad i = 1,2, \ldots, M \tag{5}$$

The least-squares approach is utilized between the predicted and actual outputs to obtain the model. The following equation is used for minimization of the model error.

$$\sum\nolimits_{i=1}^{M} (\bar{f}(x_{i1}, x_{i2}, \ldots, x_{in}) - y_i)^2 \rightarrow minimum \tag{6}$$

This type of ANN is identified based on input and output factors emphasized in the form of the Kolmogorov-Gabor polynomial function, which is as Equation (7) (Farlow, 2020):

$$\bar{y} = a_0 + \sum\nolimits_{i=0}^{n} a_i x_i + \sum\nolimits_{i=1}^{n} \sum\nolimits_{j=1}^{n} a_{ij} x_i x_j + \sum\nolimits_{i=1}^{n} \sum\nolimits_{j=1}^{n} \sum\nolimits_{k=1}^{n} a_{ijk} x_i x_j x_k + \cdots \tag{7}$$

Equation (7), a refers to the coefficients of a polynomial and $(i,j,k) \in 1,2,\ldots,n)$. In general, this polynomial can be written in quadratic polynomial form with just two variables as follows (Elbaz et al., 2021):

$$\bar{y} = G(x_i, x_j) = a_0 + a_1 x_i + a_2 x_j + a_3 x_i x_j + a_4 x_i^2 + a_5 x_j^2 \quad (8)$$

This model forecasts the output for each input parameter set and is applied for the prediction of the coefficient and decreased value of Root Mean Square Error (RMSE) between the forecasted and actual outputs. Minimization of RMSE is as Equation (9):

$$RMSE = \frac{\sum_{i=1}^{M} (\bar{y}_i - y_i)^2}{M} \rightarrow Minimum \quad (9)$$

In the principle form of this algorithm, all binary probabilities of the independent variables provide regression architecture utilizing the form of polynomials represented in Equation (8). The count of cells in the networks' hidden layer is calculated by (n/2). Afterwards, the creation of M data triples is possible to form the actual output as $(y_i, x_{ip}, q_{iq})(p,q) \in (1,2,\ldots n)$. The obtained matrix is as follows:

$$\begin{bmatrix} x_{1p} & x_{1q} & \vdots\, y_1 \\ x_{2p} & x_{2q} & \vdots\, y_2 \\ \cdots & \cdots & \cdots\cdots \\ x_{Mp} & x_{Mq} & \vdots\, y_M \end{bmatrix} \quad (10)$$

GMDH algorithm essential form is written in matrix form and Equation (7) could be applied to rewrite it as follows:

$$Y = Aa \quad (11)$$

where $Y=\{y_1, y_2, \ldots, y_M\}^T$ denotes the actual value of the output vector and $a=\{a_1, a_2, \ldots, a_5\}$ is the vector of the unknown coefficient of the polynomial. The forecasted matrix for various p and q are as Equation (12).

$$A = \begin{bmatrix} 1 & x_{1p} & x_{1q} & x_{1p}x_{1q} & x_{1p}^2 & x_{1q}^2 \\ 1 & x_{2p} & x_{2q} & x_{2p}x_{2q} & x_{2p}^2 & x_{2q}^2 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & x_{Mp} & x_{Mq} & x_{Mp}x_{Mq} & x_{MP}^2 & x_{Mq}^2 \end{bmatrix} \quad (12)$$

The following normal equation must be solved for multi-regression analysis based on the least-squares method. For comparison of the models, Relative Deviation (RD) and Average Absolute Relative Deviation (AARD) are considered that are determined as follows (Komeili Birjandi et al., 2022b):

$$RD = \frac{y_{actual} - y_{modeled}}{y_{actual}} \quad (13)$$

$$AARD = \frac{\sum_{i=1}^{n} \left| \frac{y_{i,actual} - y_{i,modeled}}{y_{i,actual}} \right|}{n} \quad (14)$$

For the abovementioned parameters, it is preferred to have values closer to 0 as much as possible. Lower values of AARD and RD refers to higher accuracy of the proposed model and lower deviations between the predicted and actual outputs. In addition to these criteria, R^2, calculated by use of Equation (15), is used as another parameter for evaluation (Komeili Birjandi et al., 2022b). Closer values to 1 means higher accuracy and for the current models (Dossumbekov et al., 2024), achieving R^2 values of higher than 0.95 would be acceptable.

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_{i,actual} - y_{i,modeled})^2}{\sum_{i=1}^{n} (y_{i,actual} - \bar{y}_{actual})^2} \quad (15)$$

4. Results and Discussion

In the present study, two intelligent methods, namely GMDH and MLP ANN, are applied for the modelling of CO2 emissions. GMDH Shell Software is used for the development of the GMDH model. The inputs of the proposed models are the total energy supply of different sources including 1. coal, 2. natural gas, 3. oil, 4. hydro, 5. biofuels and waste, 6. other renewables (wind, solar etc) and 7. GDP. The supplies of different energy sources are used as the inputs as the important factors of energy systems specifications and GDP is used as an economic index that can influence the emissions of greenhouse gases. The data are gathered for years between 1990 and 2021. The total number of datasets is equal to 96, that is not very big to make difficulties in analysis. Moreover, there is no missing data in the considered period. Furthermore, it is assumed that all of the obtained data are correct. Due to the mentioned reasons, no preprocessing is done for the present analysis and modelling(García et al., 2015, 2016). It should be noted that all of the raw data are used and no normalization is applied. Data related to the supply of different energy sources are obtained from the IEA website (Energy Statistics Data Browser, n.d.-c; Energy Statistics Data Browser, n.d.-a; Energy Statistics Data Browser, n.d.-b) and the GDP data are gathered from the WorldBank website (GDP (Current US$) - Kazakhstan, n.d.; GDP (Current US$) - Turkmenistan, n.d.; GDP (Current US$) - Uzbekistan, n.d.). For training and testing of the proposed models, 70% and 30% of datasets were used, respectively. It should be indicated that validation dataset is not considered in the present study. In Figure 6, a comparison between the modelled and actual values of

CO2 emission is presented. Most of the data are very close to Y=X line which means the prediction is precise and accurate. The R^2 value of this model for overall data is 0.9936. This value of R^2 is very close to 1 and it means that the precision of the model is acceptable. In Figure 7, RD of the different data versus the corresponding actual value of CO2 emission is depicted. It can be observed that for the majority of cases, the value of RD is in the range of ±5%, indicating high precision of the model, and the maximum absolute value is around 20.5%. High values of RD in some years of the models have been reported in other studies on the similar problems. For instance, Ahmadi et al. (Ahmadi, Dehghani Madvar, et al., 2019). reported that for their model for CO2 emission prediction in Latin American countries by use of PSO-LSSVM method, maximum value of RD was around 22%. The obtained values for the developed model in this study indicate that the model is acceptable for the estimation of CO2 emission in the majority of the cases and could be applied for this purpose.

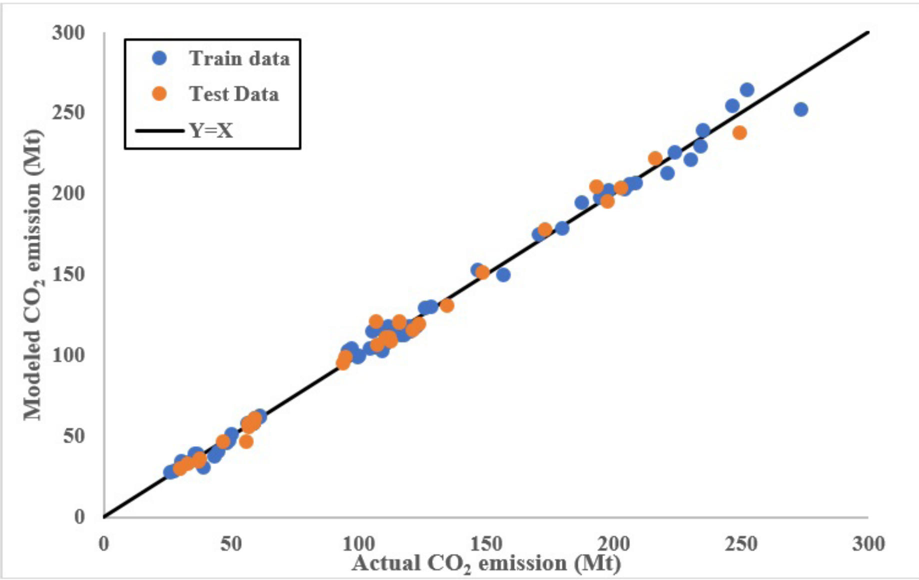

Figure 6. Comparison of modelled and actual values of CO2 emission for the GMDH model.

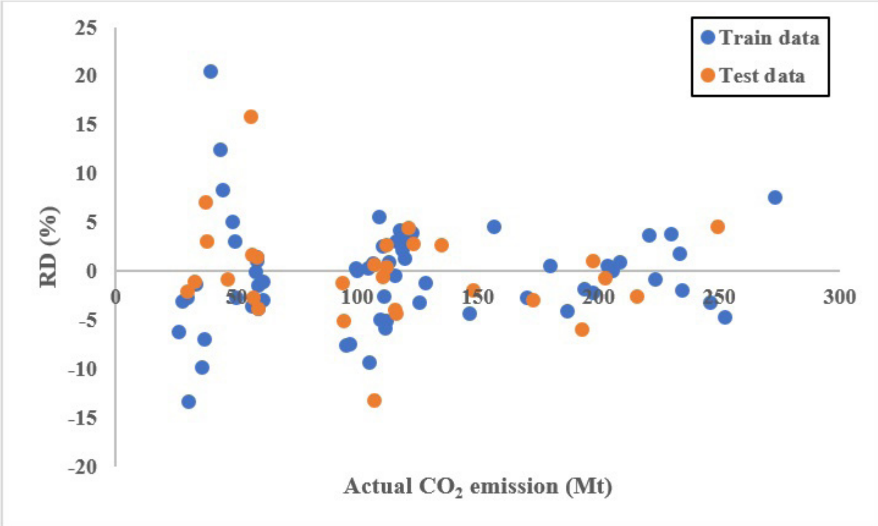

Figure 7. RD value vs actual values of CO2 emission for the GMDH model.

Similar data are applied for the modelling of CO2 emissions in the mentioned countries by use of MLP ANN. In the proposed model, just a single hidden layer is considered and various numbers of neurons, from 4 to 15, were applied to reach the maximum precision. In Table I, MSE values of different model structures are represented. It can be observed that the consideration of 5 neurons in the network provides the model with maximum precision among the tested networks. In this case, the MSE value of the network is 29.55 and R^2 is 0.9929. In Figure 8, the comparison is provided between the actual and modelled values for the MLP model it can be observed that the majority of the predictions are close to Y=X line, meaning the high exactness of the model. Similar to the previous model, values of RD vs the actual value of the emission are depicted in Figure 9. It can be observed that the maximum absolute RD is higher than the GMDH model; however, the majority of the RD values are in the range of ±5%. To have a better understanding of both models, the values of AARD for the models are compared in Figure 10. According to the obtained values of AARD, approximately 4.28% and 3.69% for the MLP and GMDH models, the latter model is a bit more accurate.

Table I. Values of MSE for different network structures.

Number of neurons	Data set	MSE value
4	Train	736.94
	Test	459.23
	Overall	653.05
5	**Train**	**35.50**
	Test	**22.74**
	Overall	**29.55**
6	Train	59.51
	Test	30.58
	Overall	50.77
7	Train	67.33
	Test	15.18
	Overalls	51.57
8	Train	26.55
	Test	176.54
	Overall	71.86
9	Train	144.71
	Test	121.40
	Overall	137.67
10	Train	152.78
	Test	139.91
	Overall	148.89
11	Train	57.44
	Test	4109
	Overall	52.51
12	Train	71.70
	Test	25.99
	Overall	57.89
13	Train	116.54
	Test	7.32
	Overall	83.55
14	Train	43.67
	Test	8.06
	Overall	32.92
15	Train	696.00
	Test	569.56
	Overall	657.80

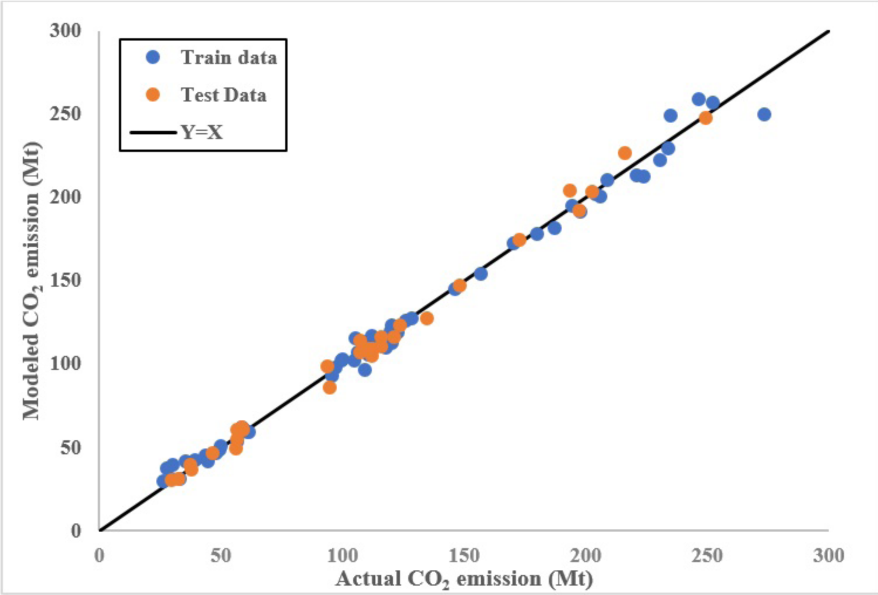

Figure 8. Comparison of modelled and actual values of CO2 emission for the MLP model.

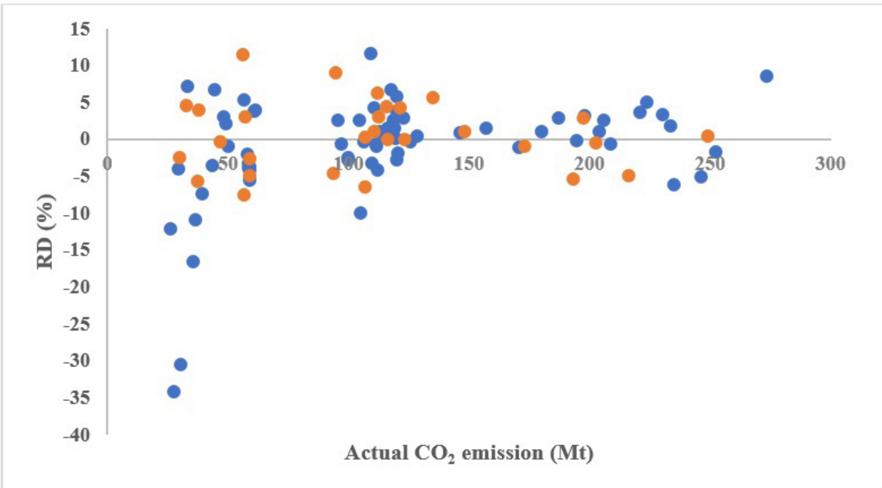

Figure 9. RD value vs actual values of CO2 emission for the MLP model.

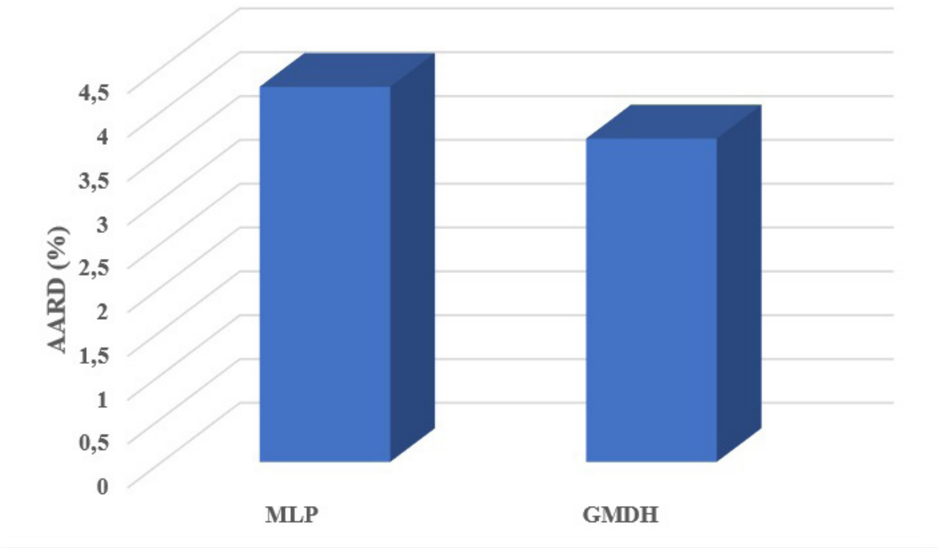

Figure 10. AARD of the proposed models.

According to the obtained values for R^2, AARD and RD, it can be denoted that the proposed models have acceptable performance in the estimation of emissions. It should be noted that in the present model, it is assumed that the emission is a function of supplies of different energy sources and GDP while some other factors such as weather condition and social parameters could be effective on the emission. This assumption is considered due to the availability of the data and simplification of the proposed. The availability of all influential data on the emission of CO2 and quantifying the detail specifications of the energy system specifications are among the most important limitations to propose more accurate and detailed models. The use of these models can be beneficial for policymakers and scholars to estimate emissions in upcoming years and evaluate different scenarios. By use of these models, different scenarios can be evaluated and compared in term of CO2 emissions.

Algorithm synthesis ability via the learning process and obtaining solution of the cases of the nonlinear problems in addition to the models robustness are the principle advantages of ANNs; however, need to training process for each problem, necessity of applying multiple tests for finding the best structure and requirement for datasets with large size for network training are the main disadvantages of these techniques (Alhuyi Nazari et al., 2021; Navarro, 2013). In this regard, some other methods can be considered. In future studies, other intelligent methods including ANFIS, LSSVM etc could be tested and compared with the present model. In some cases, depending on the problems and their characteristics, the employment of other models can result in higher exactness and more accurate predictions. The main advantages of ANFIS are its reliable performance in capturing the nonlinear architecture of a procedure and fast capacity of learning. Furthermore, this technique has both numerical and linguistic knowledge (Şahin & Erol, 2017). Ability of obtaining nonlinear solutions is among the most important advantages of SVM-based approaches; however, knowledge requirement related to the kernel function is the main problem of this method (Alhuyi Nazari et al., 2021). Aside from use of other methods, it would be an applicable recommendation to use different activation functions in the MLP model to investigate if those models could be more precise or not. In addition, it is suggested to consider more countries in the region to improve the comprehensiveness of the models.

5. Conclusion

In conclusion, this study highlights the intricate relationship between CO2 emissions, influenced significantly by energy consumption patterns, energy source compositions, and economic indicators. CO2 emission is dependent on a variety of factors and parameters; however, energy consumption and shares of different sources are very influential. In addition to them, GDP could be an affecting factor that should be considered in the estimation and forecasting of emissions. The main contribution of the present work is applying data-driven methods based on artificial intelligence to develop predictive models for the CO2 emissions. This study considered three countries namely Kazakhstan, Uzbekistan and Turkmenistan as the cases. Sources such as IEA, the WorldBank and USAID were used for data gathering. It also provided valuable insights into the energy systems of these nations, along with their respective policies aimed at enhancing energy efficiency and reducing emissions. Central Asian countries are increasingly recognizing the pivotal role of economic policies and energy strategies in shaping their emission trajectories. Policies that integrate economic growth with sustainable energy practices are crucial. Measures such as incentivizing renewable energy investments, promoting energy efficiency in industries and households, and adopting cleaner technologies can significantly mitigate CO2 emissions while fostering economic development.

The application of the Group Method of Data Handling (GMDH) and Multilayer Perceptron (MLP) models demonstrated strong performance in estimating emissions, with the GMDH showing a marginal advantage in accuracy over MLP. Comparison analysis indicated that both models have acceptable accuracy; however, making use of GMDH is preferred in terms of exactness compared with the MLP. The high coefficients of determination (R^2) and low average absolute relative deviations (AARD) for both models—0.9936 and 3.69% for GMDH, and 0.9929 and 4.28% for MLP—underscore their reliability. The proposed models can serve as valuable tools for predicting future emissions, offering insights that can inform policymakers in formulating effective strategies for sustainable development and environmental stewardship. The policymakers can consider these models to evaluate different scenarios for prediction of emissions in future and compare them with each other. In future studies, other structures of networks and functions can be applied to develop more accurate models. Moreover, some other data-driven techniques such as LSSVM and ANFIS can be considered for modelling the emission of CO2 to achieve more accurate predictions.

Acknowledgements

This research was supported by the Committee of Science of the Ministry of Science and Higher Education of the Republic of Kazakhstan (Grant No. BR21882258).

References

Ahmadi, M. H., Dehghani Madvar, M., Sadeghzadeh, M., Rezaei, M. H., Herrera, M., & Shamshirband, S. (2019). Current Status Investigation and Predicting Carbon Dioxide Emission in Latin American Countries by Connectionist Models. *Energies*, 12(10), 1916. https://doi.org/10.3390/en12101916

Ahmadi, M. H., Jashnani, H., Chau, K. W., Kumar, R., & Rosen, M. A. (2019). Carbon dioxide emissions prediction of five Middle Eastern countries using artificial neural networks. *Energy Sources, Part A: Recovery, Utilization and Environmental Effects*. https://doi.org/10.1080/15567036.2019.1679914

Alhuyi Nazari, M., Mukhtar, A., Yasir, A. S. H. M., Rashidi, M. M., Ahmadi, M. H., Blazek, V., Prokop, L., & Misak, S. (2023). Applications of intelligent methods in solar heaters: an updated review. *In Engineering Applications of Computational Fluid Mechanics* (Vol. 17, Issue 1). Taylor and Francis Ltd. https://doi.org/10.1080/19942060.2023.2229882

Alhuyi Nazari, M., Salem, M., Mahariq, I., Younes, K., & Maqableh, B. B. (2021). Utilization of Data-Driven Methods in Solar Desalination Systems: A Comprehensive Review. *Frontiers in Energy Research*, 0, 541. https://doi.org/10.3389/FENRG.2021.742615

Altikat, S. (2021). Prediction of CO2 emission from greenhouse to atmosphere with artificial neural networks and deep learning neural networks. *International Journal of Environmental Science and Technology*, 18(10), 3169–3178. https://doi.org/10.1007/s13762-020-03079-z

Betting Big on Renewables. (n.d.). USAID. https://www.usaid.gov/stories/betting-big-on-renewables#:~:text=From 2018 to 2020%2C USAID,the road for a year.

Carbon Capture, Utilisation and Storage. (n.d.). IEA. https://www.iea.org/energy-system/carbon-capture-utilisation-and-storage

CO2 Emissions in 2023 A new record high, but is there light at the end of the tunnel? (2024).

Decree of the President of the Republic of Uzbekistan "On measures to radically improve the management system of the fuel and energy industry of the Republic of Uzbekistan" dated 01.02.2019 №UP-5646. (2022). *IEA*. https://www.iea.org/policies/13314-decree-of-the-president-of-the-republic-of-uzbekistan-on-measures-to-radically-improve-the-management-system-of-the-fuel-and-energy-industry-of-the-republic-of-uzbekistan-dated-01022019-up-5646

Dossumbekov, Y. K., Zhakiyev, N., Nazari, M. A., Salem, M., & Abdikadyr, B. (2024). Sensitivity analysis and performance prediction of a micro plate heat exchanger by use of intelligent approaches. *International Journal of Thermofluids*, 22(February), 100601. https://doi.org/10.1016/j.ijft.2024.100601

Elbaz, K., Shen, S., Zhou, A., Yin, Z., & Lyu, H. (2021). Prediction of Disc Cutter Life During Shield Tunneling with AI via the Incorporation of a Genetic Algorithm into a GMDH-Type Neural Network. *Engineering*, 7(2), 238–251. https://doi.org/10.1016/j.eng.2020.02.016

Electric Vehicles. (2023). *IEA*. https://www.iea.org/energy-system/transport/electric-vehicles

Energy. (2023). *CAREC*. https://www.carecprogram.org/?page_id=16#:~:text=Key Projects,support growth in ongoing trade.

Energy efficiency classes of buildings. (2022). *IEA*. https://www.iea.org/policies/7040-energy-efficiency-classes-of-buildings

Energy Statistics Data Browser. (n.d.-a). *IEA*. https://www.iea.org/data-and-statistics/data-tools/energy-statistics-data-browser?country=KAZ

Energy Statistics Data Browser. (n.d.-b). *IEA*. https://www.iea.org/data-and-statistics/data-tools/energy-statistics-data-browser?country=UZB

Energy Statistics Data Browser. (n.d.-c). *IEA*. https://www.iea.org/data-and-statistics/data-tools/energy-statistics-data-browser?country=TKM

Executive summary. (n.d.). *IEA*. https://www.iea.org/reports/kazakhstan-2022/executive-summary

Farlow, S. J. (2020). Self-Organizing Methods in Modeling: GMDH Type Algorithms. CRC Press, Boca Raton.

Filik, Ü. B., & Filik, T. (2017). Wind Speed Prediction Using Artificial Neural Networks Based on Multiple Local Measurements in Eskisehir. *Energy Procedia*, 107, 264–269. https://doi.org/10.1016/J.EGYPRO.2016.12.147

Filipović, S., Orlov, A., & Panić, A. A. (2024). Key forecasts and prospects for green transition in the region of Central Asia beyond 2022. *Energy, Sustainability and Society*, 14(1), 25. https://doi.org/10.1186/s13705-024-00457-0

García, S., Luengo, J., & Herrera, F. (2015). Data Preprocessing in Data Mining (Vol. 72). *Springer International Publishing*. https://doi.org/10.1007/978-3-319-10247-4

García, S., Ramírez-Gallego, S., Luengo, J., Benítez, J. M., & Herrera, F. (2016). Big data preprocessing: methods and prospects. *Big Data Analytics*, 1(1), 9. https://doi.org/10.1186/s41044-016-0014-0

GDP (current US$) - Kazakhstan. (n.d.). *WorldBank*. https://data.worldbank.org/indicator/NY.GDP.MKTP.CD?locations=KZ

GDP (current US$) - Turkmenistan. (n.d.). *WorldBank*. https://data.worldbank.org/indicator/NY.GDP.MKTP.CD?locations=TM

GDP (current US$) - Uzbekistan. (n.d.). *WorldBank*. https://data.worldbank.org/indicator/NY.GDP.MKTP.CD?locations=uz

Ghalandari, M., Forootan Fard, H., Komeili Birjandi, A., & Mahariq, I. (2020). Energy-related carbon dioxide emission forecasting of four European countries by employing data-driven methods. *Journal of Thermal Analysis and Calorimetry*, 1–10. https://doi.org/10.1007/s10973-020-10400-y

Heat Pumps. (n.d.). *IEA*. https://www.iea.org/energy-system/buildings/heat-pumps

KAZAKHSTAN: Law No. 541-IV of 2012 on Energy Saving and Energy Efficiency (2019 Ed.). (2019). *Asia-Pacific Energy*. https://policy.asiapacificenergy.org/node/135

Kim, M., & Okuyucu, O. (2022). Prediction of Undrained Shear Strength by the GMDH-Type Neural Network Using SPT-Value and Soil Physical Properties. *Materials*, 15, 6385. https://doi.org/ https://doi.org/10.3390/ ma15186385

Komeili Birjandi, A., Fahim Alavi, M., Salem, M., Assad, M. E. H., & Prabaharan, N. (2022a). Modeling carbon dioxide emission of countries in southeast of Asia by applying artificial neural network. *International Journal of Low-Carbon Technologies*, 17, 321–326. https://doi.org/10.1093/ijlct/ctac002

Komeili Birjandi, A., Fahim Alavi, M., Salem, M., Assad, M. E. H., & Prabaharan, N. (2022b). Modeling carbon dioxide emission of countries in southeast of Asia by applying artificial neural network. *International Journal of Low-Carbon Technologies*, 17, 321–326. https://doi.org/10.1093/ijlct/ctac002

Kuziboev, B., Saidmamatov, O., Khodjaniyazov, E., Ibragimov, J., Marty, P., Ruzmetov, D., Matyakubov, U., Lyulina, E., & Ibadullaev, D. (2024). CO2 Emissions, Remittances, Energy Intensity and Economic Development: The Evidence from Central Asia. *Economies*, 12(4), 95. https://doi.org/10.3390/economies12040095

Law of the Republic of Uzbekistan "On the use of renewable energy sources" dated May 21, 2019 No. ZRU-539. (2022). *IEA*. https://www.iea.org/policies/13310-law-of-the-republic-of-uzbekistan-on-the-use-of-renewable-energy-sources-dated-may-21-2019-no-zru-539

Luo, X. J., Oyedele, L. O., Ajayi, A. O., Akinade, O. O., Owolabi, H. A., & Ahmed, A. (2020). Feature extraction and genetic algorithm enhanced adaptive deep neural network for energy consumption prediction in buildings. *Renewable and Sustainable Energy Reviews*, 131, 109980. https://doi.org/10.1016/J.RSER.2020.109980

M. K., A. N., & V., M. A. (2020). Role of energy use in the prediction of CO2 emissions and economic growth in India: evidence from artificial neural networks (ANN). *Environmental Science and Pollution Research*, 27(19), 23631–23642. https://doi.org/10.1007/s11356-020-08675-7

Nationally Determined Contribution (NDC) to the Paris Agreement (2022 Update): Turkmenistan. (n.d.). *IEA*. https://www.iea.org/policies/17050-nationally-determined-contribution-ndc-to-the-paris-agreement-2022-update-turkmenistan

Navarro, R. I. (2013). Study of a neural network-based system for stability augmentation of an airplane Annex 1 Introduction to Neural Networks and Adaptive Neuro-Fuzzy Inference Systems (ANFIS).

Oh, S., & Pedrycz, W. (2002). The design of self-organizing Polynomial Neural Networks. 141, 237–258.

On protection of the atmospheric air. (2022). *IEA*. https://www.iea.org/policies/11440-on-protection-of-the-atmospheric-air

Pindoriya, N. M., Singh, S. N., & Singh, S. K. (2008). An adaptive wavelet neural network-based energy price forecasting in electricity markets. *IEEE Transactions on Power Systems*, 23(3), 1423–1432. https://doi.org/10.1109/TPWRS.2008.922251

Population, total - Kazakhstan. (n.d.). *WorldBank*. https://data.worldbank.org/indicator/SP.POP.TOTL?locations=KZ

Population, total - Turkmenistan. (n.d.). *WorldBank*. https://data.worldbank.org/indicator/SP.POP.TOTL?locations=TM

Population, total - Uzbekistan. (n.d.). *WorldBank*. https://data.worldbank.org/indicator/SP.POP.TOTL?locations=UZ

Radovanović, M., Filipović, S., & Andrejević Panić, A. (2021). Sustainable energy transition in Central Asia: status and challenges. *Energy, Sustainability and Society*, 11(1), 49. https://doi.org/10.1186/s13705-021-00324-2

Renewables. (n.d.). *IEA*. https://www.iea.org/energy-system/renewables

Rezaei, M. H., Sadeghzadeh, M., Alhuyi Nazari, M., Ahmadi, M. H., & Astaraei, F. R. (2018a). Applying GMDH artificial neural network in modeling CO2 emissions in four nordic countries. *International Journal of Low-Carbon Technologies*, 13(3), 266–271. https://doi.org/10.1093/ijlct/cty026

Rezaei, M. H., Sadeghzadeh, M., Alhuyi Nazari, M., Ahmadi, M. H., & Astaraei, F. R. (2018b). Applying GMDH artificial neural network in modeling CO2 emissions in four nordic countries. *International Journal of Low-Carbon Technologies*, 13(3), 266–271. https://doi.org/10.1093/ijlct/cty026

Şahin, M., & Erol, R. (2017). A Comparative Study of Neural Networks and ANFIS for Forecasting Attendance Rate of Soccer Games. *Mathematical and Computational Applications* 2017, Vol. 22, Page 43, 22(4), 43. https://doi.org/10.3390/MCA22040043

The Law About Support the Use of Renewable Energy Sources (amended). (2022). *IEA*. https://www.iea.org/policies/5407-the-law-about-support-the-use-of-renewable-energy-sources-amended

The Outlook for the Development of Renewable Energy in Uzbekistan. (2014).

Transition to renewable energy sources: economic benefits for entrepreneurs in Kazakhstan. (2024). *UNDP.* https://www.undp.org/kazakhstan/stories/transition-renewable-energy-sources-economic-benefits-entrepreneurs-kazakhstan

Turkmenistan. (n.d.). *IEA*. https://www.iea.org/countries/turkmenistan

Turmunkh, B.-E. (2021). Renewable and Non-Renewable Energy Consumption, Carbon Dioxide Emissions, and Economic Growth: Empirical Evidence from Central Asian Countries. *Journal of Economics and Development Studies*, 9(1). https://doi.org/10.15640/jeds.v9n1a7

UNDP continues to support Turkmenistan in improving energy efficiency and developing renewable energy sources. (n.d.). *IEA*. https://www.undp.org/turkmenistan/press-releases/undp-continues-support-turkmenistan-improving-energy-efficiency-and-developing-renewable-energy-sources

USAID Energizes Uzbekistan's First Green Hydrogen Hub. (n.d.). *USAID*. https://uz.usembassy.gov/usaid-energizes-uzbekistans-first-green-hydrogen-hub/

USAID Power Central Asia. (n.d.). *IEA*. https://www.usaid.gov/central-asia-regional/fact-sheets/usaid-power-central-asia

Uzbekistan. (n.d.). *IEA*. https://www.iea.org/countries/uzbekistan

Uzbekistan (12/08). (n.d.). *U.S. Department of State*. https://2009-2017.state.gov/outofdate/bgn/uzbekistan/113251.htm

Xiu, Z. Wei. (2022). Environmental implications of economic transition in Central Asia: A study of energy consumption and carbon emissions. *Top Academic Journal of Economics and Statistics*, 7(3), 19–39.

Yuldoshboy, S., Karimov, M., & Kuralbaev, J. (2022). The association between CO2 and economic growth in Central Asian countries: Panel data approach. *Journal of Positive School Psychology*, 5587–5601.

Zhakiyev, N., Khamzina, A., Zhakiyeva, S., De Miglio, R., Bakdolotov, A., & Cosmi, C. (2023). Optimization Modelling of the Decarbonization Scenario of the Total Energy System of Kazakhstan until 2060. *Energies*, 16(13), 5142. https://doi.org/10.3390/en16135142