



PABLO BENALCAZAR*, MAŁGORZATA KRAWCZYK**, JACEK KAMIŃSKI***

Forecasting global coal consumption: An artificial neural network approach

Introduction

The industrialization of services, the increasing population size, and the economic development in Asia have led for the past fifty years to a significant growth in world energy consumption. The U.S. Energy Information Administration (EIA) projects that between 2012 and 2040, the world energy consumption will increase by 48%. By the year 2040, energy consumption in developing non-OECD countries (China, India, and South Africa) is expected to grow drastically by 71%, while in OECD countries consumption will only increase by 18% (EIA 2016). Figure 1 presents the total world energy consumption by energy source (1965–2015) and Figure 2 shows the world energy consumption in 2016.

In the 21st century, energy has become an integral part of our society and of global economic development. Although the world has experienced tremendous technological advancements, fossil fuels (including coal, natural gas, and oil) continue to be the world's primary energy source. It has been estimated that by 2040, coal, natural gas and liquid fuels will account for 78% of total world energy consumption (EIA 2016). Since the beginning of the industrial revolution, coal has played a vital role in the global energy mix, mainly due to its abundance, affordability, and high reliability. At the current production level,

* M.Sc. Eng., *** D.Sc. Eng., Associate Professor of MEERI PAS, Mineral and Energy Economy Research Institute, Polish Academy of Sciences, Krakow, Poland;
e-mail: benalcazar@min-pan.krakow.pl; kaminski@min-pan.krakow.pl

** M.Sc. Eng., Independent researcher

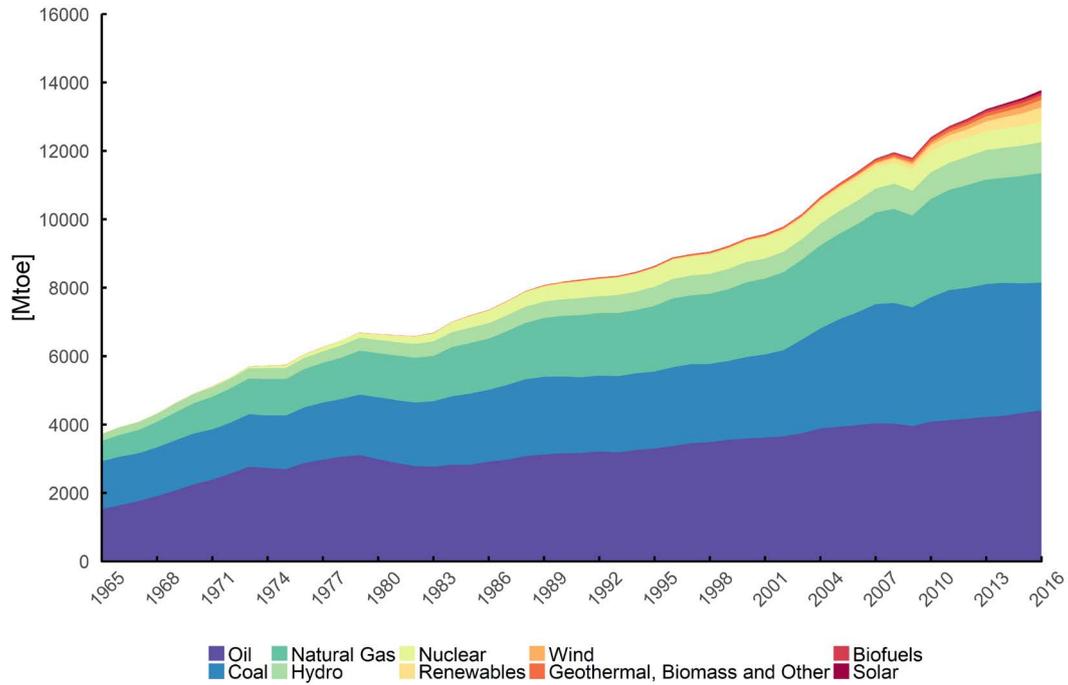


Fig. 1. World energy consumption by energy source, 1965–2015

Rys. 1. Światowe zużycie energii z podziałem na źródła energii w latach 1965–2015

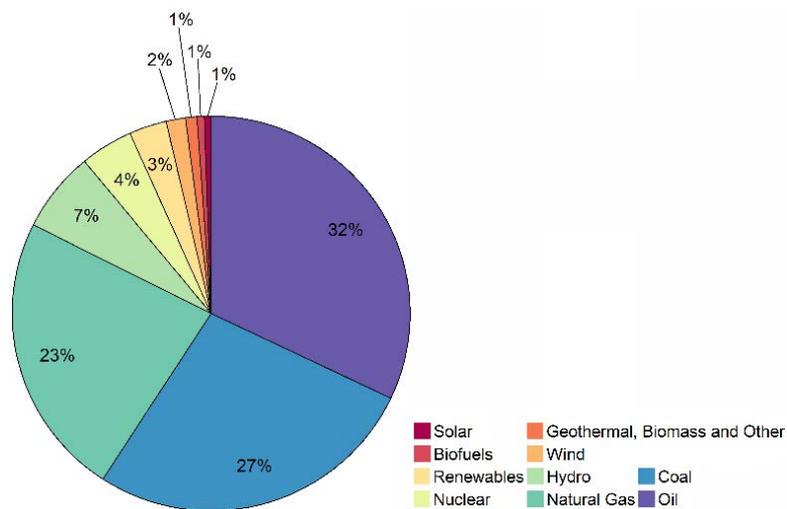


Fig. 2. World energy consumption by energy source, 2016

Rys. 2. Światowe zużycie energii z podziałem na źródła energii, 2016

the IEA estimates that coal reserves (economically recoverable) would last approximately 127.7 years (with the biggest reserves found in the USA, Russia, China, and India), which is much longer than the estimated 53 years for natural gas or the 51 years for oil (BP 2017; IEA 2016). As a result of its wide availability, ease of storage and transportation, coal has become an essential energy source for China, India, the United States, Germany, South Africa, Poland and developing non-OECD countries; particularly countries located in Asia, where coal consumption is expected to rise by approximately 90% in 2040 (EIA 2016). The world coal consumption by region (1965–2015) is shown in Figure 3. Furthermore, the rapidly fluctuating oil and natural gas prices and the unstable socio-political market have allowed coal to position itself as a fundamental and indispensable fuel source in the power, iron, steel, and cement industries. Coal fuels over 40% of the world's electricity generation and nearly 68% of the world's steel production (National Coal Council 2015).

The intricate relationship between economic growth, demographics and energy consumption (particularly in countries with coal intensive industries and heavy reliance on fossil fuels), along with the elevated amounts of greenhouse gases in the atmosphere (mainly

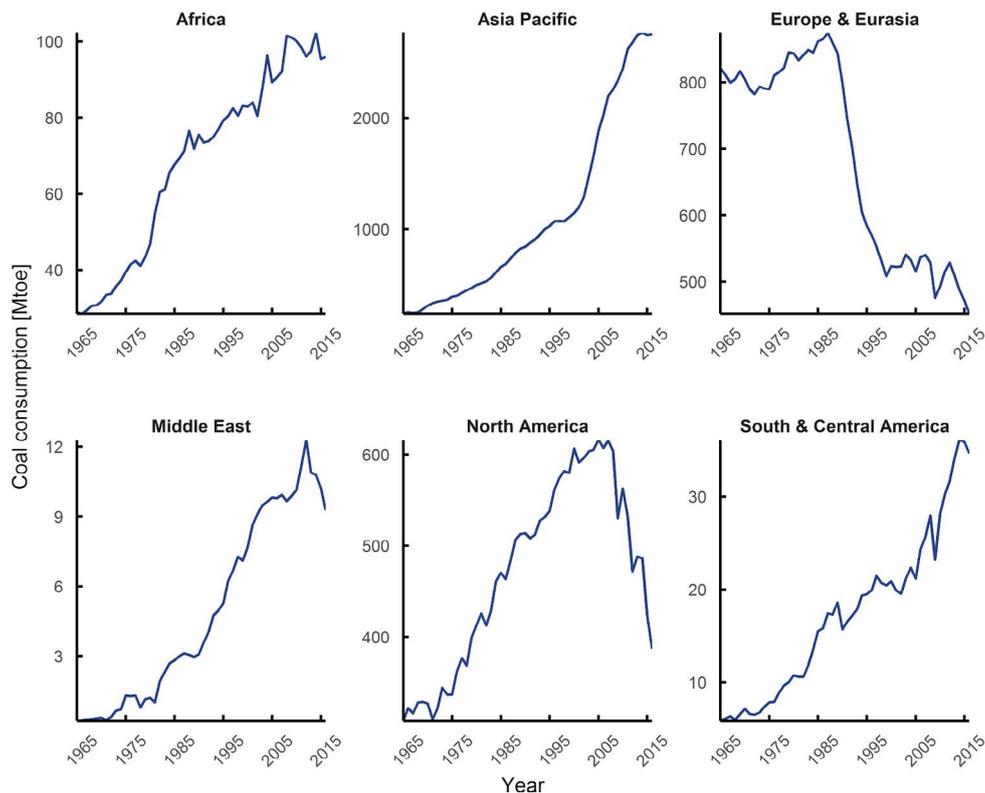


Fig. 3. World coal consumption by region, 1965–2015

Rys. 3. Światowe zużycie węgla według regionów, 1965–2015

attributed to the combustion coal, oil and natural gas), have raised serious concerns within the scientific community about the future of coal. Figure 4 shows the world GDP by region (1965–2015). Thus, various studies have focused on the development and application of forecasting methods to predict the economic prospects of coal, future levels of reserves, production, consumption, and its environmental impact. In the coal sector, long-term forecast models are of prime importance because they offer information on investment strategies, market imbalances, future trade, and allow decision makers to anticipate and assess the potential impact of government policy changes.

Among the various forecasting methods in the energy sector, Computational Intelligence (CI) methods have attracted much attention in the past ten years due to their advantages over linear and nonlinear programming models. CI techniques, which include support vector machines (SVM), fuzzy logic, artificial neural networks (ANN), genetic algorithms and particle swarm optimization, have demonstrated in a wide range of scientific fields a superior capability to handle nonlinear relationships between numerical data when compared to statistical methods (Dreyfus 2005; Rutkowski 2008).

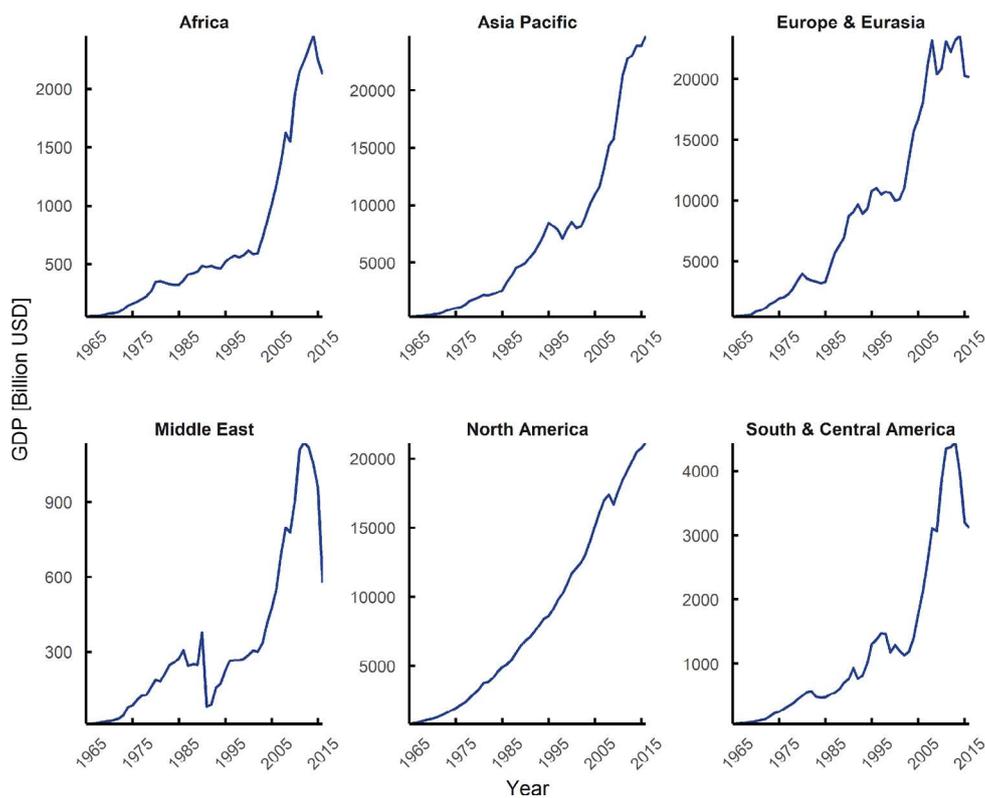


Fig. 4. World GDP by region, 1965–2015

Rys. 4. PKB według regionu, 1965–2015

Neural networks have been widely used in time series prediction and pattern recognition. For instance (Szoplik 2015) used a multilayer perceptron model to forecast the natural gas demand in Szczecin (Poland). (Günay 2016) applied a MLP ANN to predict the annual gross electricity demand of Turkey for the years 2014–2028. (Ermis et al. 2007) used a feed-forward back-propagation neural network to forecast world green energy consumption to the year 2050. (Jebaraj et al. 2011) compared the results of various forecasting methods (i.e., time series regression, double moving average, exponential smoothing, auto regressive integrated moving average (ARIMA), artificial neural networks models) to estimate the annual coal consumption in India for the years 2020–2030 and showed that an ANN model presents the best predictive performance for long-term coal consumption forecasts.

In this context, the aim of this paper is to contribute to the scarce literature on global coal consumption forecasting with the aid of artificial neural network methods. This paper proposes a Multilayer Perceptron neural network (MLP) for the prediction of global coal consumption for the years 2020–2030. The MLP-based model is trained with historical data sets gathered from financial institutions, global energy authorities, and energy statistic agencies, covering the years 1970 through 2016. With this scope in mind, section 2 presents the characteristics of artificial neural networks and section 3 describes the computational details of the MLP model, including data and feature selection. The paper ends with concluding remarks.

1. Methods

1.1. Artificial Neural Networks

Artificial neural networks (ANN) have received a great deal of attention due to their superior performance in nonlinear data modeling and their wide range of potential applications in optimization, forecasting and decision making (Hill et al. 1994; Prieto et al. 2016). ANN are mathematical models derived from the representation of the natural structure and biological function of a neuron (with its three principal components: dendrites, cell body, and axon) (Hagan et al. 1995). Although artificial neural networks are a simplification of the complexity of biological neurons, they present some similarities to their biological counterparts. ANN are highly interconnected networks, generally consisting of an input layer, one or more hidden layers, and an output layer. Each layer comprises small processing units called artificial neurons connected by synaptic weights (Raza and Khosravi 2015). A multiple-input neuron is shown in Figure 5 and its mathematical expression is described in Equation 1.

$$a = f(b + \sum w_i x_i) \quad (1)$$

where a is the neuron output, b is the bias or ‘offset’, w_i the weighted inputs, x_i the individual inputs, n the summer output, and f the activation function or ‘transfer function’. There is

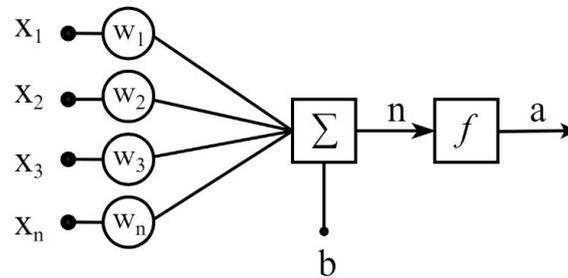


Fig. 5. Multiple-input neuron
Source: adapted from (Hagan et al. 1995)

Rys. 5. Model sztucznego neuronu

a variety of activation functions (linear and nonlinear functions) in use with artificial neural networks. A commonly used activation function is the hyperbolic tangent transfer function, which is defined as:

$$\text{tahn}(n) = \frac{e^n - e^{-n}}{e^n + e^{-n}} \quad (2)$$

where $\text{tahn}(n)$ is the activation function with scaled outputs between -1 and 1 .

With the continuous development and evolution of artificial neural networks and the outstanding growth in computational power, multiple neural network architectures with numerous topologies have been proposed, including multilayer perceptron (MLP), bridged multilayer perceptron (BMLP), among others (Wilamowski 2009). A multilayer perceptron neural network is shown in Figure 6.

The predictive model presented in this paper is a multilayer perceptron (MLP) neural network, generally referred to as a multilayer feed-forward network. With the aim of finding the model with the best generalization performance, the MLP-based model developed for this study is computed under various configurations of hidden layers and the number of neurons. The predictive performance of the neural network is measured by statistical methods often used in forecasting studies, such as the Mean Absolute Percentage Error (MAPE):

$$\text{MAPE} = \left(\frac{1}{2} \sum_{t=1}^N \left(\frac{|\hat{Q}_t - Q_t|}{\hat{Q}_t} \right) \right) \cdot 100\% \quad (3)$$

Where N is the number of predictions. \hat{Q}_t and Q_t are the actual and predicted coal consumption values at time t , respectively. Moreover, the Root Mean Square Error (RMSE),

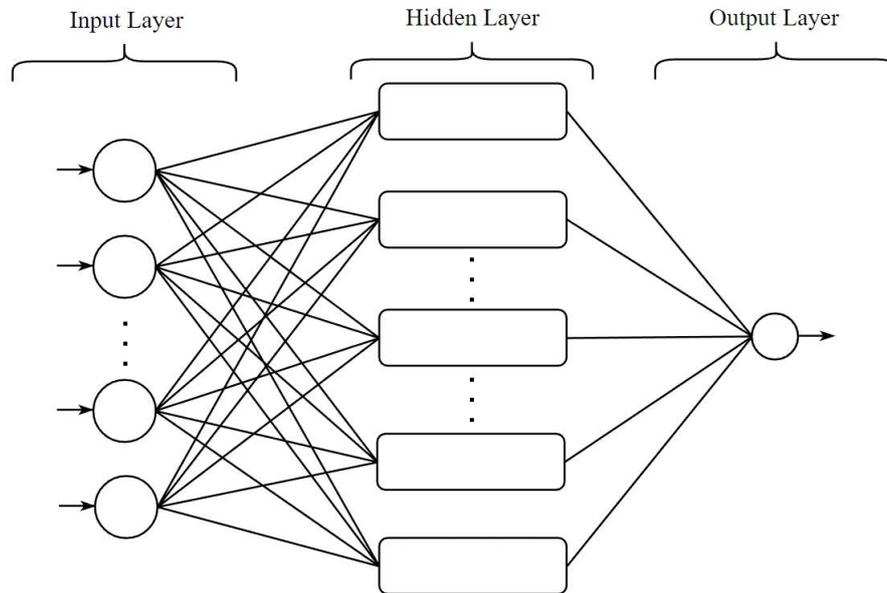


Fig. 6. Multilayer perceptron neural network

Rys. 6. Wielowarstwowa sieć neuronowa perceptronu

which can be interpreted as a measure of the deviation between the predicted and observed values, is used as an additional performance indicator. RMSE is defined as follows:

$$RMSE = \sqrt{\frac{1}{N} \sum_{t=1}^N (\hat{Q}_t - Q_t)^2} \quad (4)$$

Two additional performance metrics are used, Theil's inequality coefficient (U) and the Coefficient of determination R^2 . U yields values between 0 (denoting a perfect fit) and 1 (denoting a poor fit).

$$U = \frac{\sqrt{\frac{1}{N} \sum_{t=1}^N (\hat{Q}_t - Q_t)^2}}{\sqrt{\frac{1}{N} \sum_{t=1}^N (\hat{Q}_t)^2} + \sqrt{\frac{1}{N} \sum_{t=1}^N (Q_t)^2}} \quad (5)$$

2. Computational details

2.1. Data

The historical data sets (1970–2016) used in this study were gathered from various information sources. World population estimates and economic indicators such as Gross Domestic Product (GDP) (current USD), inter alia, and GDP growth (annual percent) were obtained from the World's Bank World Development Indicators ([The World Bank 2017](#)). The world's GDP long-term projections were taken from long-term baseline projections of the Organization for Economic Cooperation and Development (OECD) and the World Economic Outlook database of the International Monetary Fund (IMF). Primary energy consumption and coal consumption data for the years 1970–2016 were derived from BP's Statistical Review of World Energy (2017). World coal consumption forecasts for the years 2020–2040 were obtained from BP's Energy Outlook (2035), the International Energy Outlook 2016 (U.S. Energy Information Administration), and the World Energy Outlook 2016 (International Energy Agency). Historical coal prices (1970–2016) and price projections were taken from the IMF's primary commodity price database and the World Bank.

2.2. Feature selection and model description

In the process of building and evaluating predictive models, the selection of relevant variables or 'inputs' is of paramount importance, particularly for economic and financial processes ([Dreyfus 2005](#)). The input data sets must be identified carefully since it has a significant impact on the sensitivity (throughout the training, validation and testing process) and output of the model. In computational intelligence (CI) literature, it is commonly found that the input selection process is carried out through trial and error. However, in this study, the Boruta algorithm, a feature selection algorithm based on a random forest classification algorithm was used in order to reduce the number of input variables and better understand the relationship between the relevant input features ([Kursa and Rudnicki 2010](#)). The variables subjected to analysis with the Boruta algorithm were energy production (Mtoe), total primary energy supply (Mtoe), TPES/population (toe/capita), TPES/GDP (toe/thousand 2010 USD), Electricity consumption (TWh), CO₂ emissions (Mt of CO₂), to mention a few. From the algorithm's execution, four relevant variables with values from 1970 to 2016, corresponding to 46 data points per variable, were selected as inputs for the MLP model (out of a group of 20 variables): Year, world gross domestic product (current USD), world population, Australian thermal coal price. Moreover, the data was normalized between $[-1, 1]$ and split into two subsets: training and validation subsets. 40 observations were used for the training of the MLP and the remaining points were used for the validation of the model. The historical changes of the abovementioned variables (coal

consumption, world gross domestic product, world population, Australian thermal coal prices) are shown in Figure 7.

A three-layer MLP with a 4-5-1 architecture (five hidden neurons) was trained with the selected input data and the Bayesian regularization backpropagation algorithm was implemented as the training algorithm. This function is frequently used in ANN modeling since it updates the weight and bias values according to Levenberg-Marquardt optimization, preventing possible overfitting and poor generalization (Beale et al. 2017; MacKay 1992). Furthermore, various transfer functions were tested (Log-Sigmoid, Positive linear, Hyperbolic Tangent Sigmoid) and it was found that the Hyperbolic Tangent Sigmoid performs better than other transfer functions. The selection of the optimal number of hidden neurons was done using an exhaustive search method and a comparative analysis of their RMSEs. The neural network structure developed for this study is summarized in Table 1 and the effect of the number of hidden neurons on the predictive performance of the MLP is presented in Figure 8.

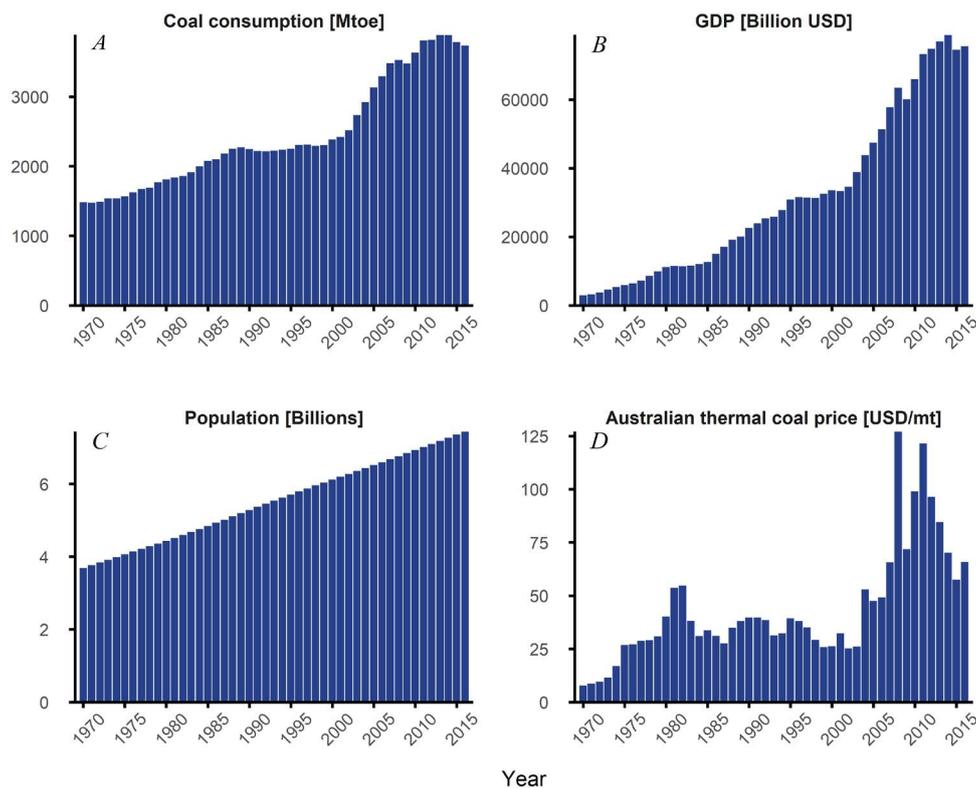


Fig. 7. Data sets: World coal consumption (a), world gross domestic product (b), world population (c), Australian thermal coal prices (d)

Rys. 7. Zestaw danych: Światowe zużycie węgla (a), produkt światowy brutto (b), ludność świata (c), ceny węgla australijskiego (d)

Table 1. MLP structure

Tabela 1. Struktura MLP

	Parameters	Components
Structure	Input layer	<ul style="list-style-type: none"> ◆ Year ◆ GDP (Billion USD) ◆ World population (Millions) ◆ Australian thermal coal price (USD/mt)
	Hidden layer	Neurons: 5
	Output layer	Neurons: 1
Activation function	Hidden layer	Hyperbolic Tangent Sigmoid
	Output layer	Linear
Training algorithm	Bayesian regularization backpropagation	

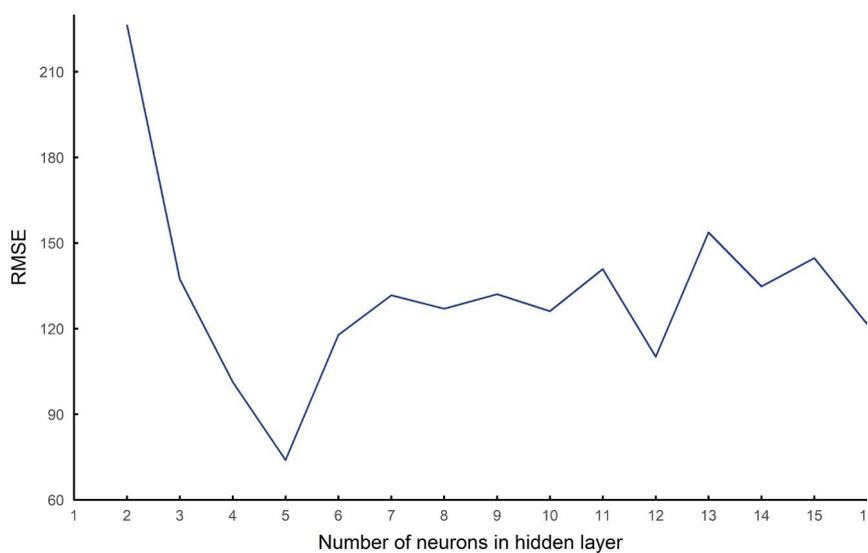


Fig. 8. RMSE vs number of hidden neurons

Rys. 8. Pierwiastek błędu średniokwadratowego vs. ukryte neurony

3. Computational results

The predictive performance of the MLP is reported in Table 2. Forty years of global coal consumption data was used to train the neural network, while consumption data for the years 2011–2016 were used for testing and measuring the accuracy of the model. Results from the MLP-based model for global coal consumption for the years (1970–2015) are presented in Table 3. The actual and estimated values presented in Table 3 demonstrates the acceptable level of accuracy of the trained neural network.

Figure 9 presents a comparison between the actual and estimated coal consumption and Figure 10 shows the percent error for the years 1970–2015. It can be observed from Figure 8 that the maximum error throughout the training and testing procedure occurred in 2004 (2.68%).

Table 2. Performance results for the years 1970–2016

Tabela 2. Wyniki osiągnięte na lata 1970–2016

Training (1970–2010)				Testing (2011–2016)			
RMSE	R2	MAPE	U	RMSE	R2	MAPE	U
21.45	0.99	0.73	0.0046	26.99	0.76	0.52	0.0035

Table 3. Actual and estimated MLP results

Tabela 3. Rzeczywiste i szacowane wyniki MLP

Coal Consumption [Mtoe]										
Year	1970	1975	1980	1985	1990	1995	2000	2005	2010	2015
Actual	1 482.76	1 566.45	1 812.74	2 077.28	2 246.41	2 249.91	2 384.72	3 134.29	3 635.64	3 784.65
Estimated	1 476.46	1 571.80	1 815.53	2 027.50	2 249.99	2 269.55	2 396.97	3 143.53	3 623.03	3 784.97

Table 4. Comparison of projections for the years 2020–2030

Tabela 4. Porównanie prognoz na lata 2020–2030

Coal Consumption [Mtoe]				
Year	ANN model	BP (2017)	IEA (2016)	EIA (2016)
2020	3 932.00	3 896.84	3 906.00	4 249.03
2025	4 069.41	4 066.99	3 955.00	4 363.31
2030	4 182.02	4 072.32	4 039.70	4 395.99

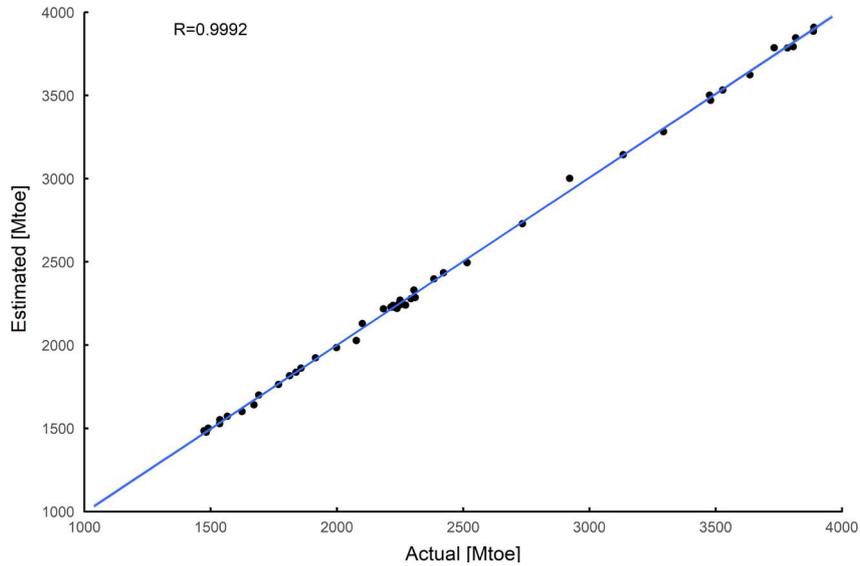


Fig. 9. Comparison between actual and estimated coal consumption

Rys. 9. Porównanie rzeczywistego i szacowanego zużycia węgla

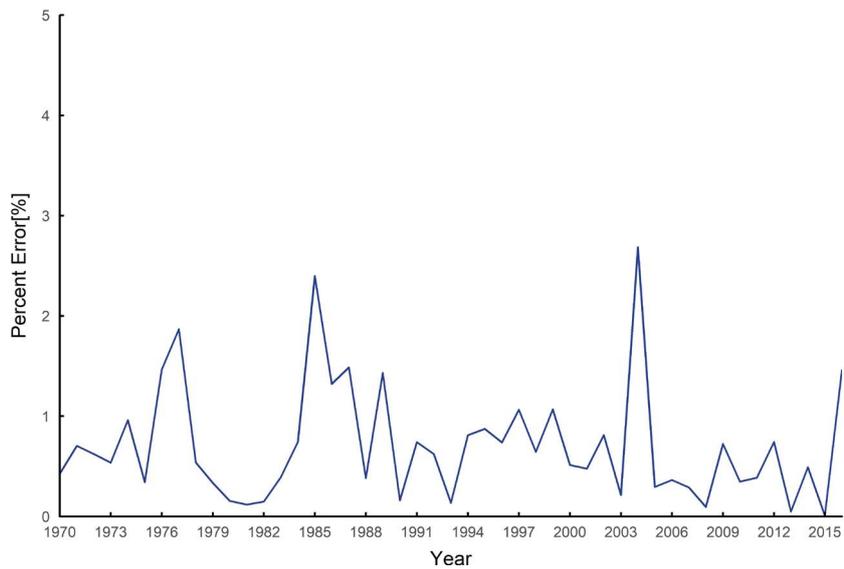


Fig. 10. Error percent (training and testing)

Rys. 10. Błąd procentu (szkolenie i testowanie)

A comparison between coal consumption forecasts published by various agencies (EIA, IEA, BP) and the MLP forecast is presented in Table 4. Figure 11 shows the forecasts of coal consumption obtained by the ANN model and other models.

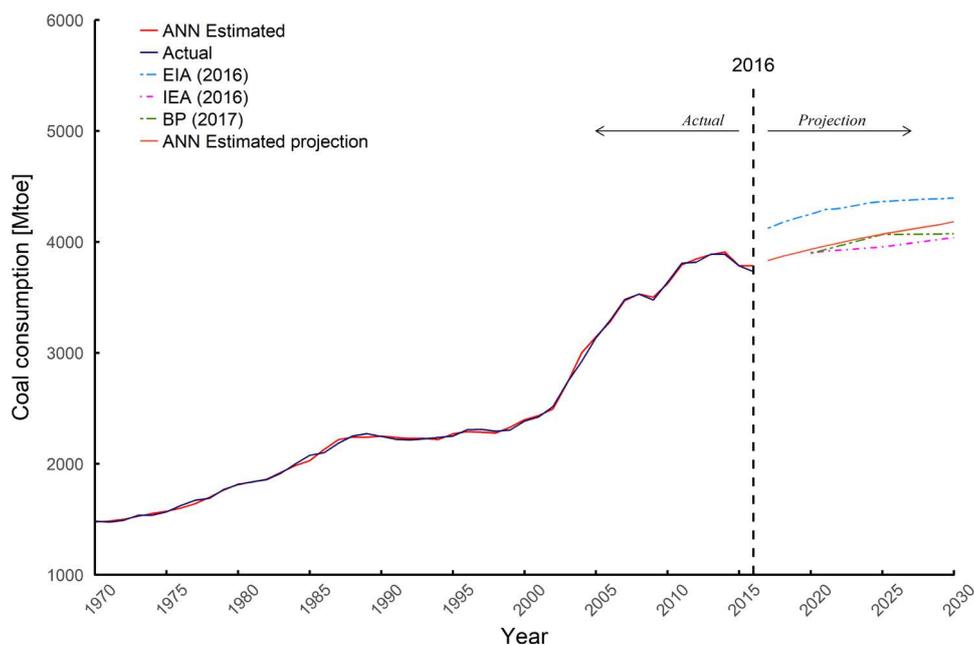


Fig. 11. Comparison of coal consumption forecasts

Rys. 11. Porównanie prognoz konsumpcji węgla

Conclusion and discussion

In this paper, we aim to contribute to the literature on coal consumption forecasting and demonstrate the application of a multilayer perceptron neural network model to estimate global coal consumption for the years 2020 to 2030. The forecast was carried out using historical data sets (1970–2016) gathered from the World's Bank World Development Indicators, inter alia, the World Economic Outlook database of the International Monetary Fund (IMF).

A two-layer MLP with a 4-5-1 architecture (five hidden neurons) was trained with four relevant variables (year, world GDP, world population, Australian thermal coal price). In order to prevent overfitting and poor generalization, the Bayesian regularization backpropagation algorithm was implemented as the training algorithm. The results of this study show

a deceleration in coal consumption for the years 2025 and 2030, from 3 831.33 Mtoe (2017) to 4 069.41 Mtoe (2025) and 4 182.02 Mtoe in 2030. It is worth pointing out that the results obtained in this study are in line with the forecasted consumption values published by the U.S. Energy Information Administration and BP. It can be seen from the proposed coal forecasting method that MLP neural networks offer some advantages over other mathematical programming models due to its ease of modeling linear and nonlinear dependencies between variables solely from observed data. For further research, in order to improve the performance of the forecast, it would be desirable to use a hybrid ANN based technique and include additional environmental and sociological variables.

Acknowledgements

We would like to thank Anna Karnasiewicz for her assistance in collecting the data presented in this study.

The work was carried out as part of the statutory activity of the Mineral and Energy Economy Research Institute Polish Academy of Sciences.

REFERENCES

- Beale et al. 2017 – Beale, M.H., Hagan, M.T. and Demuth, H.B. 2017. Neural Network Toolbox™ Reference. The MathWorks, Inc., Natick.
- BP, 2017. *BP Statistical Review of World Energy 2017*. London.
- Dreyfus, G. 2005. *Neural networks: methodology and applications*, 1st ed. Springer-Verlag Berlin Heidelberg, Berlin, Heidelberg. doi: 10.1007/3-540-28847-3.
- EIA, 2016. *International Energy Outlook 2016*. International Energy Outlook 2016. U.S. Energy Information Administration. [Online] Available at: [www.eia.gov/forecasts/ieo/pdf/0484\(2016\).pdf](http://www.eia.gov/forecasts/ieo/pdf/0484(2016).pdf) [Accessed: 7.06.2017].
- Ermis et al. 2007 – Ermis, K., Midilli, A., Dincer, I. and Rosen, M.A. 2007. Artificial neural network analysis of world green energy use. *Energy Policy* 35, 1731–1743. doi: 10.1016/j.enpol.2006.04.015.
- Günay, M.E. 2016. Forecasting annual gross electricity demand by artificial neural networks using predicted values of socio-economic indicators and climatic conditions: Case of Turkey. *Energy Policy* 90, 92–101. doi: 10.1016/j.enpol.2015.12.019.
- Hagan et al. 1995 – Hagan, M.T., Demuth, H.B. and Beale, M.H. 1995. Neural Network Design. Bost. *Massachusetts PWS* 2, 734. doi: 10.1007/1-84628-303-5.
- Hill et al. 1994 – Hill, T., Marquez, L., O'Connor, M. and Remus, W. 1994. Artificial neural network models for forecasting and decision making. *Int. J. Forecast.* 10, 5–15. doi: 10.1016/0169-2070(94)90045-0.
- IEA, 2016. *Coal Information 2016*. International Energy Agency, Paris.
- Jebaraj et al. 2011 – Jebaraj, S., Iniyan, S. and Goic, R. 2011. Forecasting of Coal Consumption Using an Artificial Neural Network and Comparison with Various Forecasting Techniques. *Energy Sources, Part A Recover. Util. Environ. Eff.* 33, 1305–1316. doi: 10.1080/15567030903397859.
- Kursa, M.B. and Rudnicki, W.R. 2010. Feature Selection with the Boruta Package. *J. Stat. Softw.* 36, Issue 11, 1–13, doi: 10.18637/jss.v036.i11.
- MacKay, D.J.C., 1992. Bayesian Interpolation. *Neural Comput.* 4, 415–447. doi: 10.1162/neco.1992.4.3.415.
- National Coal Council, 2015. *Coal = Reliable Energy*. National Coal Council, Washington, DC.
- Prieto et al. 2016 – Prieto, A., Prieto, B., Ortigosa, E.M., Ros, E., Pelayo, F., Ortega, J. and Rojas, I. 2016. Neural networks: An overview of early research, current frameworks and new challenges. *Neurocomputing* 214, 242–268. doi: 10.1016/j.neucom.2016.06.014.

- Raza, M.Q. and Khosravi, A. 2015. A review on artificial intelligence based load demand forecasting techniques for smart grid and buildings. *Renew. Sustain. Energy Rev.* 50, pp. 1352–1372. doi: 10.1016/j.rser.2015.04.065.
- Rutkowski, L. 2008. *Computational Intelligence, Computational Intelligence*. Springer Berlin Heidelberg, Berlin, Heidelberg. doi: 10.1007/978-3-540-76288-1.
- Szoplik, J. 2015. Forecasting of natural gas consumption with artificial neural networks. *Energy* 85, 208–220. doi: 10.1016/j.energy.2015.03.084.
- The World Bank, 2017. World Development Indicators 2017 [WWW Document]. URL. [Online] Available at: <http://data.worldbank.org/data-catalog/world-development-indicators> [Accessed 8.21.2017].
- Wilamowski, B.M. 2009. Neural networks architectures and learning algorithms. *Ieee Ind. Electron. Mag.* 56–63.

**PROGNOZOWANIE ŚWIATOWEGO ZUŻYCIA WĘGLA
Z ZASTOSOWANIEM PODEJŚCIA SZTUCZNYCH SIECI NEURONOWYCH**

Słowa kluczowe

prognozowanie, zużycie węgla, sztuczne sieci neuronowe

Streszczenie

W XXI wieku energia stała się integralną determinantą społeczeństwa i światowego rozwoju gospodarczego. Pomimo tego, że świat doświadczył zdecydowanych postępów technologicznych, paliwa kopalne (tj. węgiel, gaz ziemny i ropa naftowa) nadal są głównym źródłem energii na świecie. Szacuje się, że przy obecnym poziomie produkcji zasoby węgla (ekonomicznie możliwe do pozyskania) powinny wystarczyć na około 130 lat (największe rezerwy węgla znajdują się w Stanach Zjednoczonych, Rosji, Chinach i Indiach). Zawile relacje między wzrostem gospodarczym, demografią, zużyciem energii (szczególnie w krajach o przemysłach silnie uzależnionych od paliw kopalnych) oraz podwyższone ilości gazów cieplarnianych w atmosferze wzbudziły poważne obawy w społeczności naukowej na temat przyszłości węgla. W świetle powyższego, dotychczasowe badania skupiały się na opracowywaniu i stosowaniu metod prognozowania rozwoju rynku węgla, przyszłych poziomów rezerw, produkcji, konsumpcji i wpływu na środowisko. Biorąc pod uwagę powyższe uwarunkowania, celem niniejszego artykułu jest opracowanie metody prognozowania światowego zapotrzebowania na węgiel z wykorzystaniem metody sztucznej sieci neuronowej. W poniższej pracy zastosowano podejście oparte na sieci neuronowej Perceptron Multilayer (MLP) do prognozowania światowego zużycia węgla w latach 2020–2030. Model MLP bazuje na historycznych zestawach danych zebranych od instytucji finansowych, organów i agencji statystycznych pozyskujących dane na temat energii obejmujących lata od 1970–2016. Wyniki niniejszego badania wskazują na spowolnienie światowego zużycia węgla do poziomu 3932 Mtoe w 2020 r., 4069 Mtoe w 2025 r. oraz 4182 Mtoe w 2030 roku.

**FORECASTING GLOBAL COAL CONSUMPTION:
AN ARTIFICIAL NEURAL NETWORK APPROACH****Key words**

forecasting, coal consumption, artificial neural networks

Abstract

In the 21st century, energy has become an integral part of our society and of global economic development. Although the world has experienced tremendous technological advancements, fossil fuels (including coal, natural gas, and oil) continue to be the world's primary energy source. At the current production level, it has been estimated that coal reserves (economically recoverable) would last approximately 130 years (with the biggest reserves found in the USA, Russia, China, and India). The intricate relationship between economic growth, demographics and energy consumption (particularly in countries with coal intensive industries and heavy reliance on fossil fuels), along with the elevated amounts of greenhouse gases in the atmosphere, have raised serious concerns within the scientific community about the future of coal. Thus, various studies have focused on the development and application of forecasting methods to predict the economic prospects of coal, future levels of reserves, production, consumption, and its environmental impact. With this scope in mind, the goal of this article is to contribute to the scarce literature on global coal consumption forecasting with the aid of an artificial neural network method. This paper proposes a Multilayer Perceptron neural network (MLP) for the prediction of global coal consumption for the years 2020–2030. The MLP-based model is trained with historical data sets gathered from financial institutions, global energy authorities, and energy statistic agencies, covering the years 1970 through 2016. The results of this study show a deceleration in global coal consumption for the years 2020 (3 932 Mtoe), 2025 (4 069 Mtoe) and 2030 (4 182 Mtoe).