AGH DRILLING, OIL, GAS • Vol. 34 • No. 2 • 2017
http://dx.doi.org/10.7494/drill.2017.34.2.439

Tomasz Cieślik*,, Klaudia Metelska****

MODELING OF GAS CONSUMPTION IN THE CITY

1. INTRODUCTION

In recent years rapid growth in demand for natural gas has been observed. The current global natural gas consumption is 3370 bn m^3 a year. It is estimated that natural gas may surpass coal in global consumption and become the second most consumed fuel in the world. At the present time, gas covers 24% of the global energy demand [4, 5]. A gradual, slow increase at the rate of about 3–4% annually is estimated for Poland. The slow growth in the domestic gas consumption can be explained by the price increase for imported gas [6].

In the European Union natural gas has been included as a part of the strategy of building balanced energy systems. Moreover, it is proposed that it is natural gas in particular whose share of the total fuel consumption should be increased at the cost of the fossil fuels more harmful to the environment. The European importers pay more for the natural gas than the importers in the USA.

From the moment of extraction until the reception by the customer, natural gas is transported through a gas pipeline. Providing reliable gas supplies of the right quality requires, among others: providing for the development of the gas infrastructure based on a long term outlook for the gas demand. The gas market is strongly related to weather. It should be mentioned that natural gas can be stored (underground gas storage facilities – USG); however, this generates additional costs. Forecasting the performance of a gas network and gas consumption in relation to a season of the year is exceedingly significant

* Institute of Nuclear Physics, Polish Academy of Sciences, Krakow, Poland
** AGH University of Science and Technology, Faculty of Drilling, Oil and Gas, Krakow, Poland

in the exploitation process. Thanks to the forecast it is possible to predict an increase or decrease in gas consumption, as well as the occurrences of pressure drop in gas pipelines. This information allows for the planning of the renovations of the gas networks and gas facilities, so that the supply fluency remains undisturbed. One other very important benefit of forecasting is an opportunity to project gas consumption in relation to change in temperature, wind speed and wind direction [1, 20]. It paves the way for a rational policy of underground gas storage. Forecasting is defined as predicting the development of an analysed phenomenon in the future and, just as every judgment about the future, it bears a degree of uncertainty. The statistical methods used for the forecasting assessment can be divided into relative and absolute. They are used as a basis for decision-making processes. It is most difficult to forecast the processes which can be characterized by high variability in time and are dependent on many factors of random nature. Forecasting gas consumption in the cities has been carried out in order to provide a long-term and effective projection of the demand for the services of gas fuel transfers. The described process allows for prompt reaction to dynamic changes and for data comparison along the time axis.

Forecasting the gas consumption can be conducted through a variety of methods. Among the ones used for the gas demand forecast the following ones should be mentioned: regression methods, time series methods and genetic algorithm methods. The choice of a given method depends on the aim of the forecast, the area and the horizon of the forecast, as well as on the complexity of the analysed process. The more factors involved, the more complex the forecasting model should be, which makes conducting the calculations significantly more difficult and contributes to an increase in the length of time required for the calculations. The periodic variability of the gas network's workload is mostly dependent on the air temperature. High temperatures in the summer months lead to lower gas network's workload during these months, conversely, low temperatures cause an increase in demand for gas which, in this area, is used to heat buildings' interiors. Another factor which influences the variability of the gas network's workload is the differentiation of gas consumption by the users at the successive hours of day and night. A lower demand is observed at night, whereas a greater one during the day. The periods of increased and subsequently decreased net workloads in the particular days and seasons of the year follow the same patterns in the subsequent years, therefore, they can be subjected to forecasting attempts. Artificial neural network (ANN) is a great tool to generate forecasts and can be a valuable supplement to other methods of mathematical modeling of processes. The neuron structure and the neural network operation are based on the operation of biological nerve cells. The process of forecasting with the use of neural network enables a specific problem to be solved. Before it happens however, decisions have to be made with regard to particular stages of the network's development and operation. First, the network architecture has to be specified, then the number of hidden layers and finally the number of neurons in particular layers. The second stage is about the choice of the learning method, training the network and testing it. The learning itself cannot take too much time, otherwise it might cause the network to memorize the default data. The last stage is applying the network to resolving the presented assignment and assessing its quality. The quality assessment includes, among other things: specifying errors of the particular neurons, determining mean errors for the whole testing set or validation set.

The obtained results are most commonly the basis for determining either the end of the learning process or its continuation [2]. The energy market relies heavily on the course of the weather throughout the heating season. The more difficult fuel storage is, the more precise forecast of demand for the fuel should be. The results of the analysis of the correlation between the electrical energy consumption or natural gas consumption and the number of degree days involving heating gives the answer about the prospect of efficient forecasting of the future consumption. The second method is a degree days method. Degree days are a quantitative indicator which determines the energy demand for heating houses and public facilities. They are calculated on the basis of the observed air temperatures during the day. The method requires the adoption of a base temperature which is the mean temperature inside a building decreased by the energy gains from internal sources. Frequently, a temperature of 18°C (65°F = 18,3°C in the USA) [3] is set as a base temperature.

2. AIMS OF THE WORK

The aim of the present work is to forecast the consumption of gas in the city with the population of 27,000 inhabitants. The city has 920 gas recipients. The data used for the forecast includes: gas consumption throughout the course of 2 years, temperature and wind speed. In order to improve the quality of the models, artificial data has been used: days, months and holidays. Linear regression and MLP (Multilayer Perceptron) artificial networks have been applied to the forecast of the gas consumption. The calculations have been conducted with the use of Statistica and Gretl software. The quality of the model has been estimated on the basis of mean average percentage error (MAPE).

2.1. Gas consumption in the city

It can be noticed that the gas consumption in the city is very much dependent on the temperature. With a temperature drop, gas consumption increases. Conversely, with a temperature rise, gas consumption decreases. It is worth pointing out that the lowest temperatures in the city have been noticed during the night time, whereas the highest temperatures have been noticed during the day time. The greatest difference between a night temperature and a day temperature was 12°C (Fig. 1).

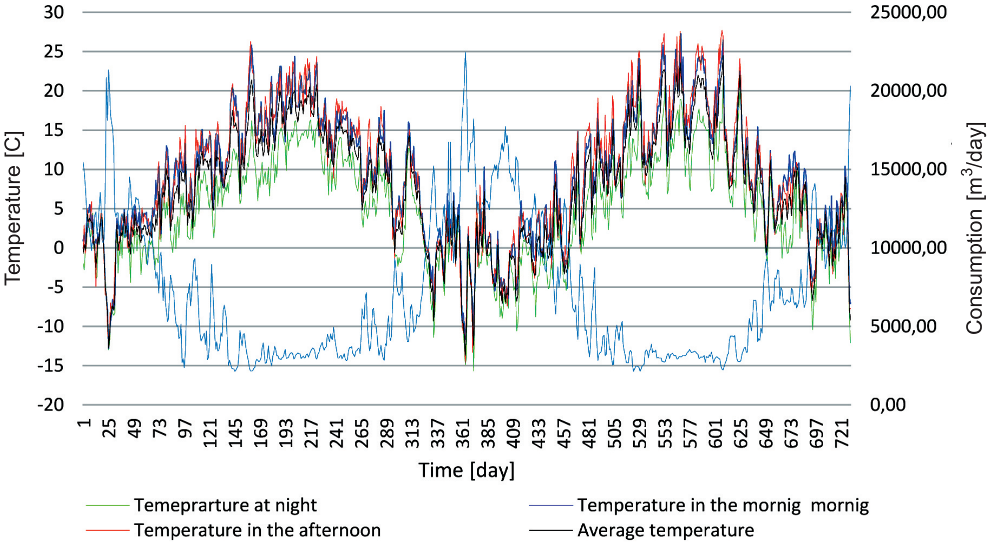

Fig. 1. The influence of temperature on the gas consumption

In 2014 the greatest gas consumption was observed in December and January, amounting to 16% of the annual consumption. These months can be characterized by a mean sub-zero temperature. The lowest gas consumption was observed from June to August, amounting to 4% of the annual consumption. When taking different week days into account, the consumption rises on Friday and Saturday to 15% of the average annual consumption and oscillates around 14% on other days of the week.

In 2015 the mean sub-zero temperature could only be observed in January and February, amounting to 16% of the annual consumption. March and April were exceptionally cold which caused an increase in the gas consumption in comparison to the previous year. Weeks of the year 2015 can be characterized by the greatest gas consumption on Thursday, Saturday and Sunday. The highest consumption on Thursday can be explained by the day's lowest mean temperature (Figs 2 and 3).

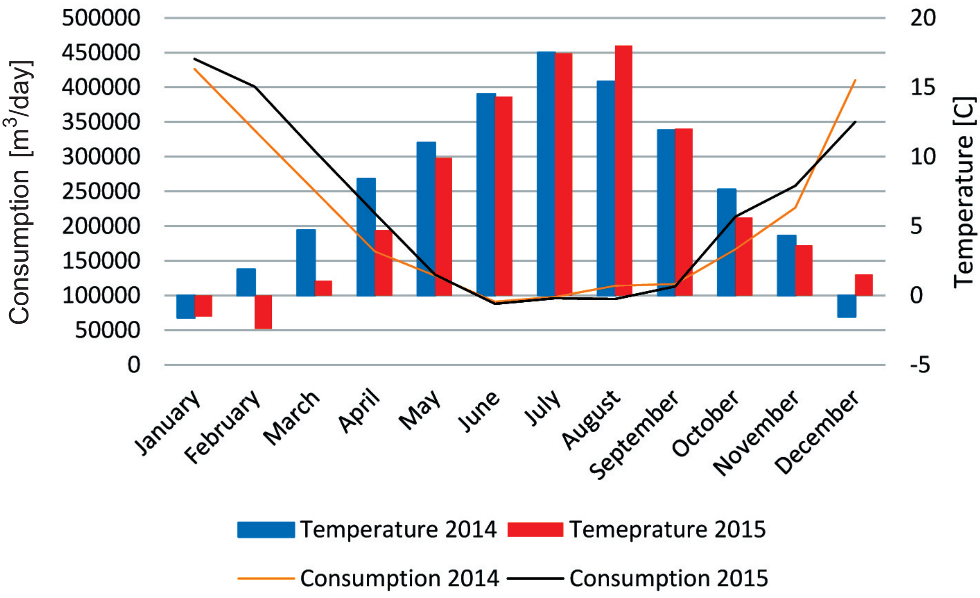

Fig. 2. Analysis of gas consumption during the month in relation to temperature (author's own work)

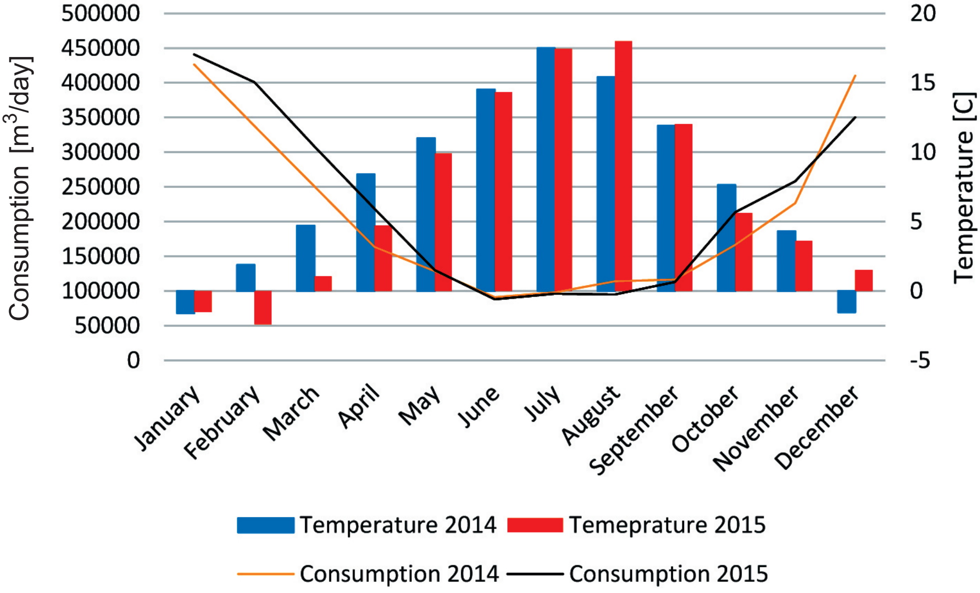

Fig. 3. Analysis of gas consumption during the week in relation to temperature (author's own work)

2.2. Forecasting gas consumption using regression

Regression is a statistical method which examines the relationship between data and uses them to predict unknown values on the basis of the other values [7].

Assumptions [7–10, 19]:

- The model has a linear form – $Y = \beta x + \varepsilon$,
  where:
  - β – parameter,
  - ε – error variable.
- The matrix x is known and is not random
  The rank $x = k$ $\quad\quad T > k + 1$,
  where:
  - k – number of parameters,
  - T – number of observations.
- The expected value of the error variable is zero.
- The variance of the error variable is constant and equals σ^2.

Calculations

The linear model assumes the following general formula [7–10, 19]:
Formula:

$$Y = \beta_0 + \beta_1 \cdot X_1 + \beta_2 \cdot X_2 + \ldots + \beta_k \cdot X_k + \varepsilon \tag{1}$$

where:

- Y – dependent variable (whose value is explained by the model, i.e. it is endogenous),
- $X_1, ..., X_k$ – variables by which we test Y,
- $\beta_0, \beta_1, \beta_2, ..., \beta_k$ – the parameters of the model (numerical values),
- ε – error variable.

Temperature and wind speed during the night, in the morning, afternoon and mean temperature were used for the analysis. In addition, dummy variables such as day, month and holidays were used (Figs 4–7).

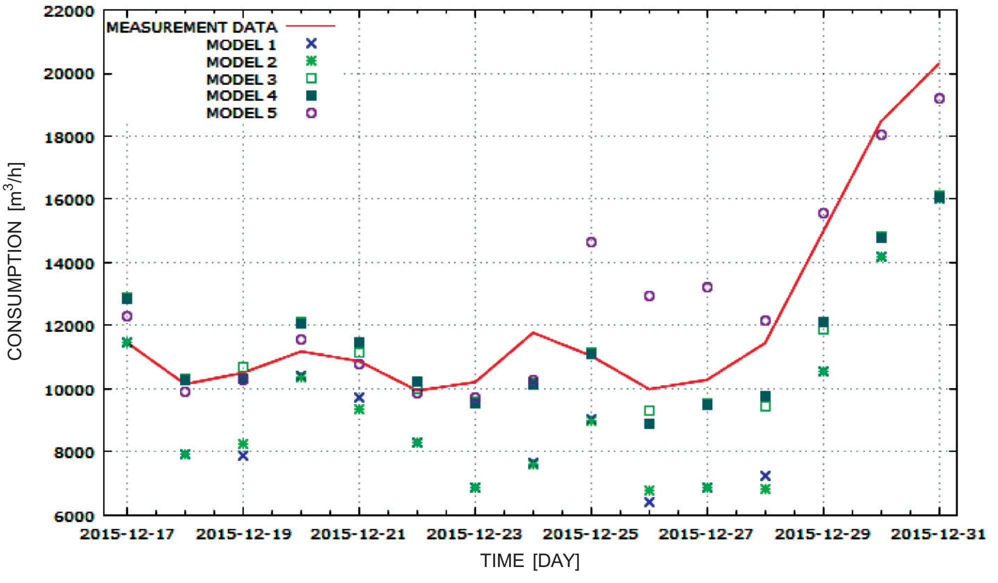

Fig. 4. Forecasting gas consumption with regression using temperature and wind speed during the night. Model 1 – without dummy variables, Model 2 – with days, Model 3 – with days and months, Model 4 – with months, Model 5 – with days, months and holidays

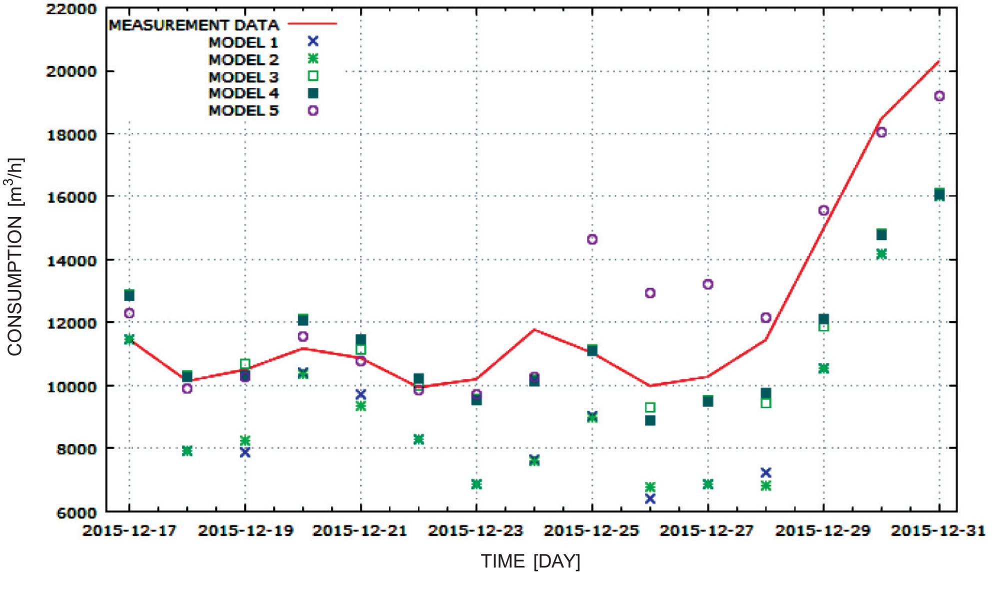

Fig. 5. Forecasting gas consumption with regression using temperature and wind speed during the day. Model 1 – without dummy variables, Model 2 – with days, Model 3 – with days and months, Model 4 – with months, Model 5 – with days, months and holidays

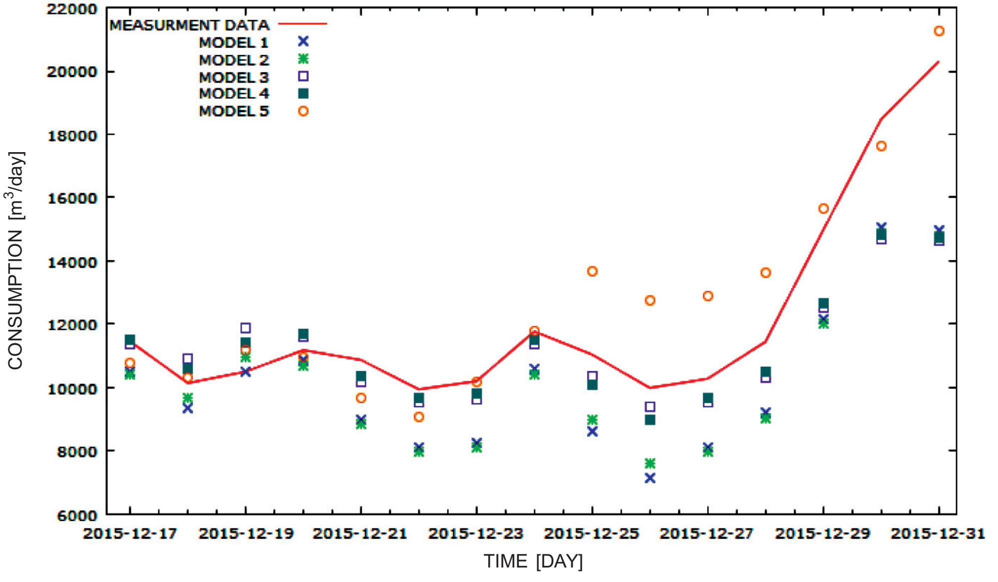

Fig. 6. Forecasting gas consumption using regression using temperature and wind speed in the afternoon. Model 1 – without dummy variables, Model 2 – with days, Model 3 – with days and months, Model 4 – with months, Model 5 – with days, months and holidays

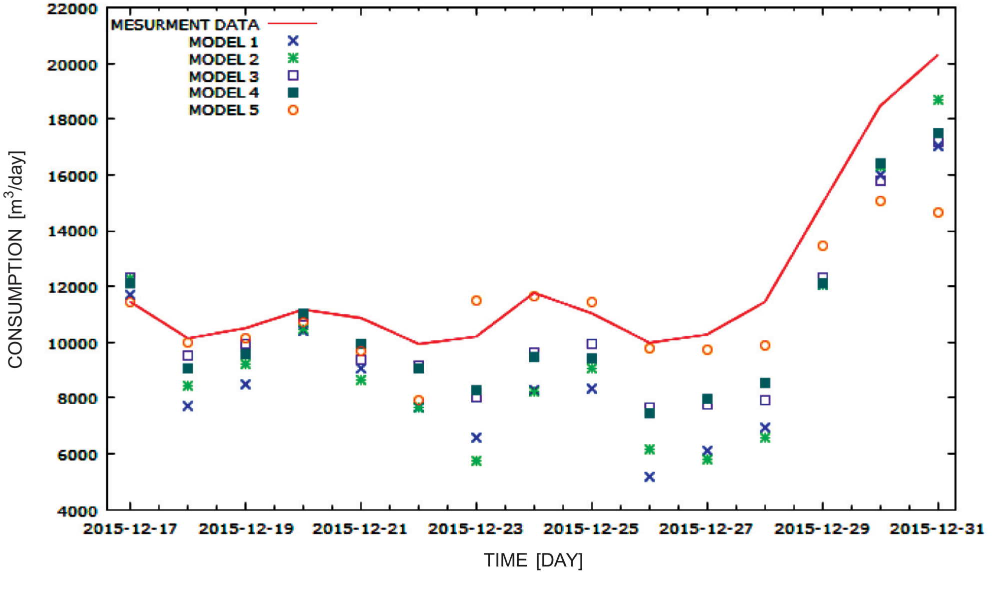

Fig. 7. Forecasting gas consumption with regression using mean temperature and wind speed. Model 1 – without dummy variables, Model 2 – with days, Model 3 – with days and months, Model 4 – with months, Model 5 – with days, months and holidays

2.3. Artificial neural networks (ANNs)

An artificial neural network is composed of artificial neurons, the basic constituents that together form an artificial neural network. It is a «transducer» which receives input signals, multiplies them by appropriate values of weights depending on the significance of the input signal and generates an output value. After multiplication by weights, the signals are totalised and adjusted in the summation layer. This is used to determine the excitation neuron. Signals which cross the activation threshold are routed to the non-linear activation function in order to generate an appropriate output signal [11].

Neural network learning is used when we do not have information on the associations between all inputs and outputs. It is performed by determining weights which give the best solution. There are two types of ANN learning: supervised and unsupervised. In supervised learning, training examples are supplied at the input and supervisory signals are supplied at the output. The network automatically selects the weights in order to learn the function describing the association between the input signals and the output signal. In unsupervised learning, output is generated on the basis of the received input signals. This process is carried out without example weights [11–18].

Calculations

The impact of the number of hidden layer neurons on the quality of the calculations was analysed. It may be noted that a 3-10-1 Multilayer Perceptron (MLP) network which has 3 input neurons and 10 neurons in the hidden layer generates results that are comparable in terms of the MAPE error with a neural network with more neurons per hidden layer (Fig. 8).

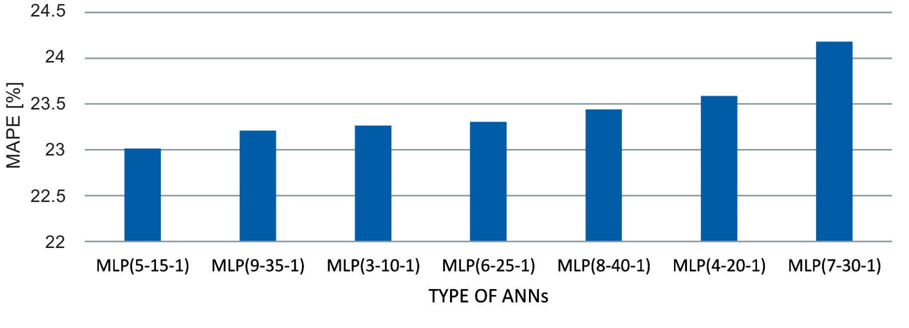

Fig. 8. MAPE errors for different numbers of hidden neurons in the ANN

During simulation 70% of the data set was used to adjust the weights on the neural network, while the test set comprised 20% and the validation set 10% of the data.

Simulation parameters:

- Number of hidden layers – 1.
- Error function – the sum of squares.
- Activation functions of hidden neurons – exponential and logistic.
- Activation functions of output neurons – exponential and logistic.
- Reduction of hidden layer weights – min. –0.0001, max. –0.001.
- Reduction of output layer weights – min. –0.0001, max. –0.001.

Temperature and wind speed during the night, in the morning, afternoon and mean temperature were used for the analysis. In addition, dummy variables such as day, month and holidays were used (Figs 9–12).

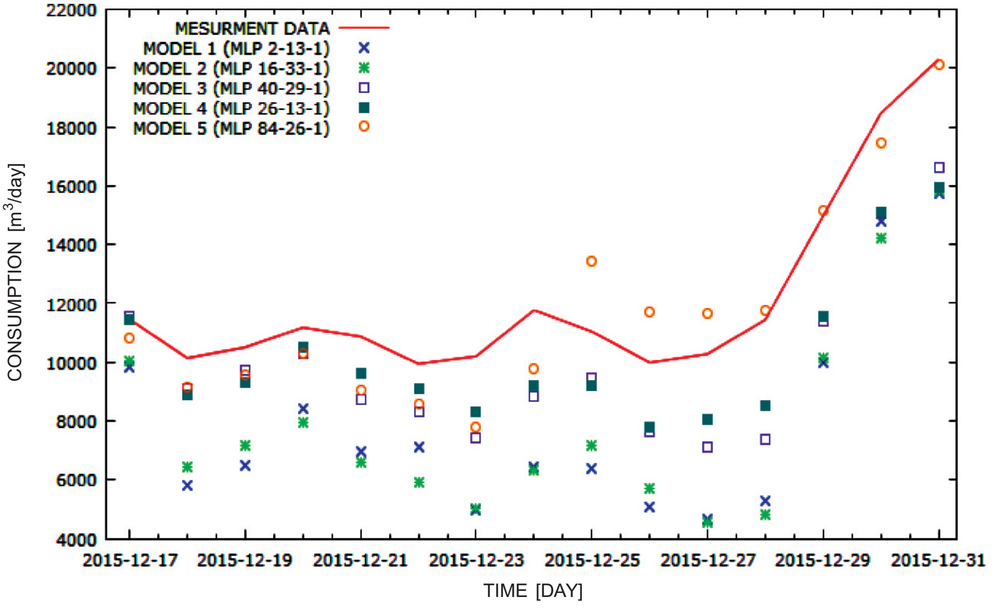

Fig. 9. Forecasting gas consumption with ANNs using temperature and wind speed during the night. Model 1 – without dummy variables, Model 2 – with days, Model 3 – with days and months, Model 4 – with months, Model 5 – with days, months and holidays

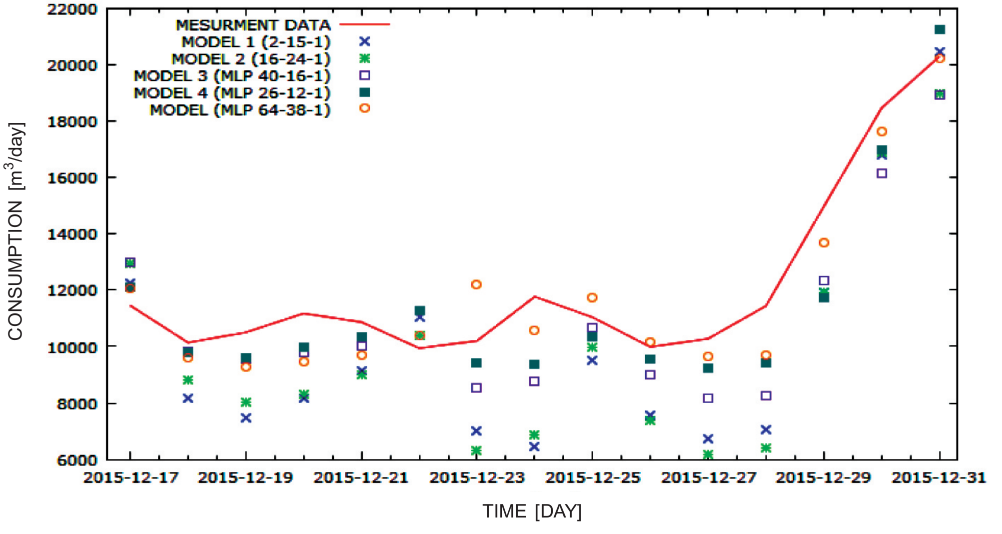

Fig. 10. Forecasting gas consumption with ANNs using temperature and wind speed in the morning. Model 1 – without dummy variables, Model 2 – with days, Model 3 – with days and months, Model 4 – with months, Model 5 – with days, months and holidays

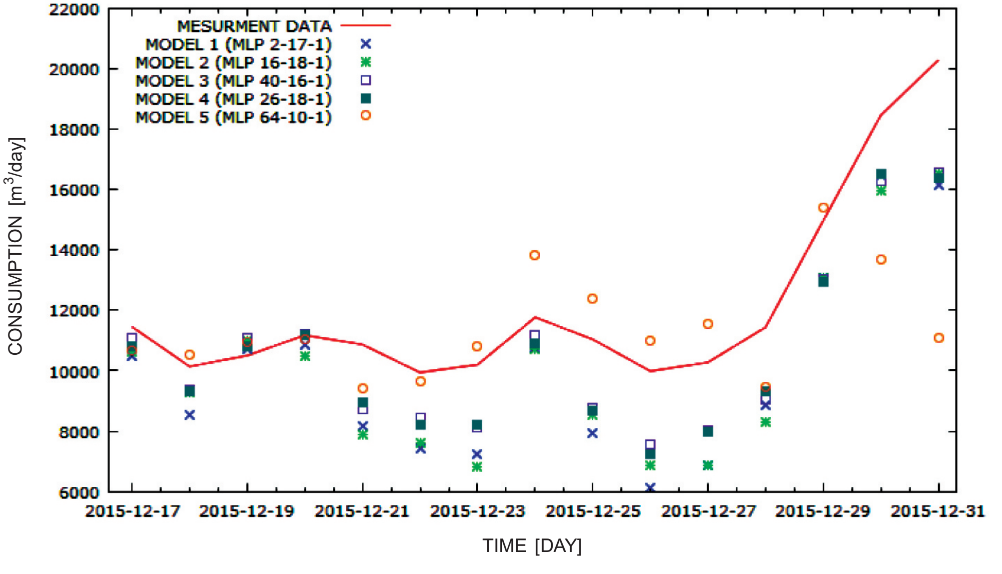

Fig. 11. Forecasting gas consumption with ANNs using temperature and wind speed in the afternoon. Model 1 – without dummy variables, Model 2 – with days, Model 3 – with days and months, Model 4 – with months, Model 5 – with days, months and holidays

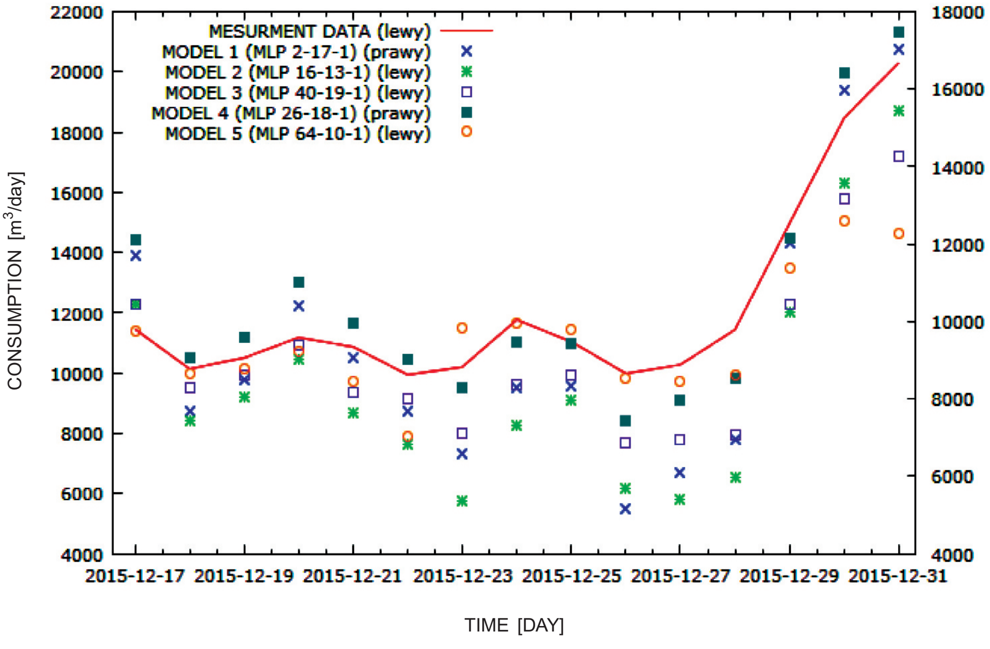

Fig. 12. Forecasting gas consumption with ANNs using mean temperature and wind speed. Model 1 – without dummy variables, Model 2 – with days, Model 3 – with days and months, Model 4 – with months, Model 5 – with days, months and holidays

3. RESULTS

The results of all models were compared with each other using MAPE (Figs 13 and 14) formula (2):

$$\mathrm{MAPE} = \frac{100}{n} \sum_{i=1}^{n} \left| \frac{e_{(i)}}{y_{(i)}} \right| \tag{2}$$

where:

- n – number of observations,
- $e_{(i)}$ – non-standardized residual for subsequent observations,
- $y_{(i)}$ – dependent variable values (y).

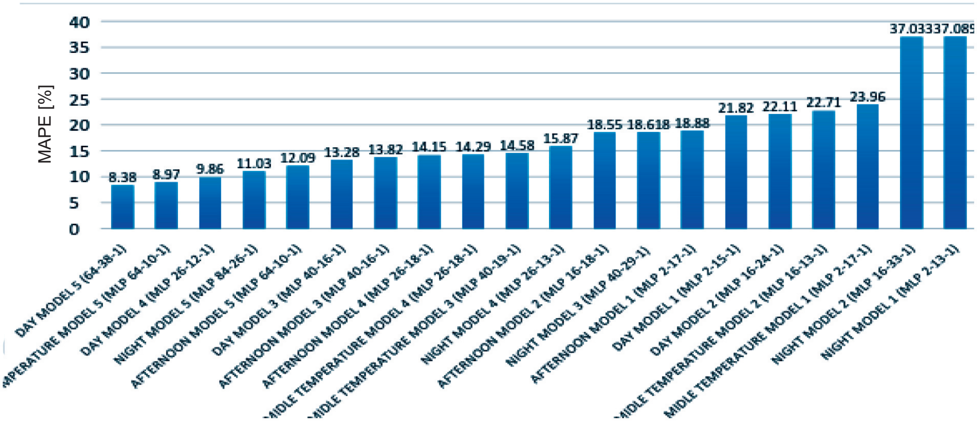

Fig. 13. Estimation of neural network models

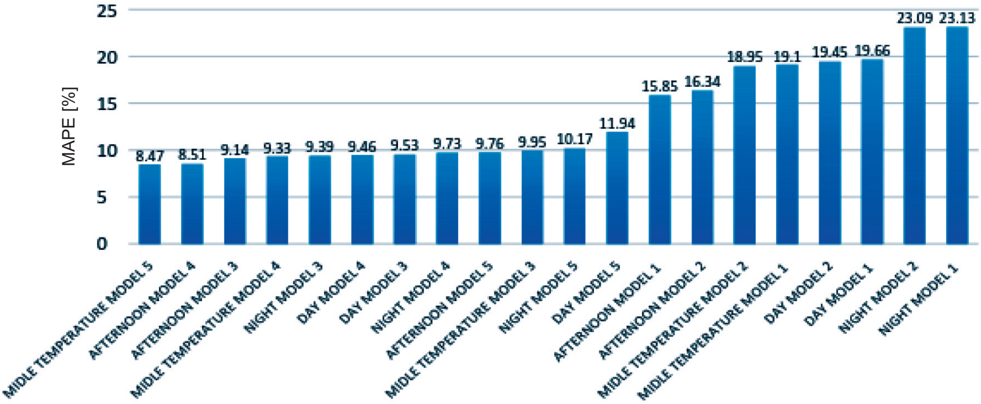

Fig. 14. Estimation of linear regression models

4. CONCLUSIONS

In the course of the analysis it was observed that gas consumption depends not only on the temperature but also on specific days of the week and months (Figs 1–3). In the summer months consumption gradually falls whereas in the winter months consumption increases markedly. It may also be noted that consumption peaks during the weekend and is the lowest at the beginning of the week. One exception from the rule was the mean gas consumption on Thursday in 2015. It was at a record high because the average temperature on that day was the lowest of the whole year.

In addition to temperature and wind speed measurements dummy variables were used in the regression analysis. Significant dummy variables included: change to daylight saving time, All Saints' Day, the winter holiday break, Catholic Church holidays and days preceding Christmas and Easter.

The best model that we achieved was an MLP neural network (MLP 64-38-1) that used mean 24 hours temperature and wind speed. Additionally, dummy variables to represent the days of the week were included in the model. High deviation of certain models can be explained by the fact that the forecast was performed for the last fourteen days of the year (17–31 December) when the gas consumption is not increased due to a change in temperature but due to Christmas and New Year's Eve celebrations.

REFERENCES

[1] Cieślik T., Kogut K.: *Prognozowanie pracy sieci gazowej za pomocą sztucznych sieci neuronowych*. Nafta-Gaz, nr 6, 2016.

[2] http://www.msn.uph.edu.pl/smp/msn/32/ambroch.pdf [19.02.2017].

[3] Dopke J.: *Zależność zużycia gazu ziemnego na ogrzewanie pomieszczeń od liczby stopniodni grzania*. Rynek Energii, nr 5, 2006

[4] Wilczyński M.: *Gaz ziemny wsparciem dla niskoemisyjnej gospodarki*. Fundacja Instytut na Rzecz Ekorozwoju, Warszawa 2014, ISBN: 978-83-89495-97-6.

[5] http://www.bp.com/en/global/corporate/energy-economics/statistical-review-of-world-energy.html [01–19.02.2017].

[6] Torjanowska M.: *Wykorzystanie modeli strukturalnych do prognozowania zapotrzebowania odbiorców wiejskich na gaz ziemny*. Inżynieria Rolnicza, 9(107), 2008.

[7] Koronacki J., Ćwik J.: *Statystyczne systemy uczące się*. Wydawnictwa Naukowo-Techniczne, Warszawa 2005.

[8] Rabiej M.: *Statystyka z programem Statistica*. Wydawnictwo Helion, Gliwice 2012.

[9] Stanisz A.: *Przystępny kurs statystyki zastosowaniem programu STATISTICA PL na przykładach z medycyny*. T. 2. StatSoft Polska, Kraków 2007.

[10] Goryl A.: *Wprowadzenie do ekonometrii w przykładach i zadaniach*. PWN, Warszawa 2000.

[11] Żurada J.: *Sztuczne sieci neuronowe*. PWN, Warszawa 1996.

[12] Tadeusiewicz R.: *Sieci neuronowe*. Akademicka Oficyna Wydawnicza, Warszawa 1993.

[13] Tadeusiewicz R.: *Elementarne wprowadzenie do technik sieci neuronowych z przykładowymi programami*. Warszawa 1998.

[14] http://www.cs.put.poznan.pl/kkrawiec/aied/NNMLW2.pdf [01–19.02.2017].

[15] Wójcik M.: *Model sieci gazowniczej oparty o sztuczne sieci neuronowe*. AGH, Kraków 2005.

[16] Kogut K.: *Analiza możliwości modelowani sieci przesyłowej gazu ziemnego*. AGH, Kraków 2007.

[17] Maciejasz M.: *Zastosowanie sieci neuronowy do analizy pracy sieci przesyłowych*. AGH, Kraków 2006.

[18] Rymarczyk M. (ed.): *Decyzje, symulacje, sieci neuronowe*. Wydawnictwo Wyższej Szkoły Bankowej w Poznaniu 1997.

[19] Metelska K., Cieślik T.: *A review of currently available LNG facilities and means of transport used for supplying terminals in natural gas with a projection of LNG prices in the USA*. AGH Drilling, Oil, Gas, vol. 33, No. 1, 2016, pp. 49–67.

[20] Szoplik J.: *Zastosowanie metody sztucznych sieci neuronowych do prognozowania obciążenia sieci rurociągów do transportu gazu ziemnego*. Inż. Ap. Chem., 52, 6, 2013, pp. 572–574.