Short-term heat load forecasting in district heating systems using artificial neural networks

P Benalcazar1,*, J Kamiński1
1 Mineral & Energy Economy Research Institute, Polish Academy of Sciences, Department of Policy and Strategic Research, Division of Energy Economics, Krakow, Poland.

E-mail: benalcazar@min-pan.krakow.pl

Abstract. With the advent of sustainable energy systems based on renewable energy sources (RES) and the development of a new generation of district heating systems (4GDH), it has become imperative for cogeneration and RES plant operators, as well as district heating (DH) operators, to apply new tools that lead to improvements in production planning, energy efficiency, and at the same time, reduce costs of heat generation. In recent years, machine learning (ML) methods used for the estimation and forecasting of energy demand have drawn considerable attention due to their advantage over linear and nonlinear programming models. In this context, the paper presents an artificial neural network (ANN) approach for the prediction of short-term heat load in a district heating system. The ANN model is trained with past heat load data, weather data and social behavior components. The predictive performance of the neural network model is measured by the mean absolute percentage error (MAPE) and the root mean square error (RMSE).

1. Introduction

The global increase in energy demand and the ever-increasing role of fossil fuels in the world economy has motivated the scientific community to investigate and develop intelligent methods that could help optimize the use of natural resources as well as reduce greenhouse gas (GHG) emissions. The development of world energy consumption by source in the period 1985–2015 is presented in Figure 1 and Figure 2 shows the total world energy consumption by source in 2016.

Since the early 1980s, district heating systems (DHS), in particular those that integrate combined heat and power (CHP) units, have been considered effective means of addressing climate change and viable mechanisms for enhancing energy efficiency through the use of waste heat [1]. Research in energy management and advances in scheduling and planning techniques have led to a better utilization of fossil fuels (improving the operability of energy systems) and have provided significant economic benefits to power plant owners and holding companies (due to a reduction in production costs).
In the state-of-the-art area of energy management and operation planning, models based on mathematical programming approaches, i.e., Mixed-Integer Linear Programming (MILP) and Mixed-Integer Nonlinear Linear Programming (MINLP), have been extensively discussed and often presented as adequate solution methods for short-term planning and scheduling of DHS and CHP units. Furthermore, there is a plethora of papers published on the topic of heat load prediction based on statistical models, time series models, as well as hybrid models (mainly implementing autoregressive integrated moving average models (ARIMA)) [3,4,5,6]. However, the aforementioned models exhibit some limitations when it comes to nonlinear systems with seasonal and/or daily variations and weather-related variables. In recent years, several research studies have proposed the use of Machine Learning (ML) methods, i.e., Support Vector Machines (SVMs), Artificial Neural Networks (ANNs), Fuzzy Logic (FL), as effective tools for modeling nonlinear systems [7]. ML techniques, notably ANNs, offer several advantages and performance benefits over other computational methods [8]. Their intrinsic capability of analyzing data and model dependencies between complex nonlinear features have shown a great potential in the area of energy management and operational planning. With the improvement in computational capabilities, ANNs have found their way into applications in production planning, scheduling of operations and costs optimization of CHP units and district heating system networks. For instance, [9] applied a number of supervised machine learning techniques (SVM, ANN, MLR) to forecast heat load in a district heating system located in Sweden. [10] presented a method based on Extreme Learning Machine for heat load prediction in the second biggest district heating system in Serbia. Several case studies [11,12,13,14] have successfully demonstrated the use of various types of neural networks for the prediction of heat load in CHP plants and DHS.

This paper develops an artificial neural network (ANN) model for the prediction of short-term heat...
load in a district heating system. The model is trained with past heat load data, weather data and social behavior components. The predictive performance of the neural network is measured by the mean absolute percentage error (MAPE) and the root mean square error (RMSE).

2. Problem statement
Due to the strong correlation between ‘heat load’ (also referred to as ‘thermal load’) and fuel cost, the prediction of thermal load plays a vital role in the net income and short-term operation planning of CHP plants and DHS [15,16,3]. Short-term forecasting of heat load poses a number of challenges due to the fact that several technical factors and physical phenomena, i.e. outdoor temperature, wind, precipitation, air humidity, solar radiation, social factors, have a significant impact on hourly heating demand [17,4]. For large CHP and DHS operators, the implementation of advanced methods has led to better day-ahead short planning and consequently lower costs of electricity and heat production, hence increasing profits. Yet, for small DHS and independent power producers (cogeneration units), these advanced systems are in many cases considered inaccessible tools due to their elevated costs, special software requirements and long hours of technical training. Thus, the short-term forecasting approach presented in this paper, based on an artificial neural network model, is aimed to fill this gap in the literature and offer a reliable forecasting method for small DHS and independent power producers. Since ANNs are largely driven by data, the recommended short-term forecasting approach can be treated as a grey-box model, allowing operators to make effective operational decisions without the need of understanding the technical relations between descriptive and target features.

3. Method
In 1943, Warren McCulloch and Walter Pitts developed a computational model that could emulate the characteristics of biological neurons [18]. Since then, and with the outstanding growth in computational power, artificial neural networks have received great attention in the literature. ANN present a superiority over statistical methods due to its capability to model nonlinear relationships between numerical data (inputs and outputs) without the need of developing complex mathematical descriptions [19]. The architecture of a neural network consists of a set of highly interconnected processing elements (neurons) arranged in several layers (input layer, hidden layer, output layer). The basic elements of a multiple-input neuron are depicted in Figure 3.

The network architecture employed for this study is a Multilayer Perceptron (MLP) with a single hidden layer. Figure 4 shows a generic representation of a MLP artificial neural network.

![Figure 3. Basic elements of a multiple-input neuron.](image)

![Figure 4. MLP with one hidden layer.](image)

The predictive performance of the neural network model is measured by statistical methods often used in forecasting studies, such as the Mean Absolute Percentage Error (MAPE):
\[ MAPE = \left( \frac{1}{N} \sum_{t=1}^{N} \left( \frac{\hat{P}_t - P_t}{P_t} \right) \right) \times 100\% \] (1)

Where \( N \) is the number of predictions. \( \hat{P}_t \) and \( P_t \) are the actual and predicted load values at time \( t \), respectively. Moreover, the Root Mean Square Error (RMSE), which can be interpreted as a measure of the standard deviation of the residuals, is used as an additional performance indicator. RMSE is defined as follows:

\[ RMSE = \sqrt{\frac{1}{N} \sum_{t=1}^{N} (\hat{P}_t - P_t)^2} \] (2)

4. Input selection and neural network description

In order to generate a suitable input data set for the ANN-based predictive model, several candidate input features were investigated. The analyzed hourly data consisted of historical heat load data from a DHS and weather data obtained from MERRA2, a numerical weather reanalysis [20]. The data set was normalized to \([0,1]\), correlation coefficients were calculated to determine the dependency between input features, and a principal component analysis (PCA) approach was taken to reduce the dimensionality of the data, helping identify uncorrelated input components [21]. It is worth noting that feature selection is an essential step in the process of data mining and the formulation of a neural-based forecasting model since it plays an essential role in the generalization capabilities of the network. The descriptive features analyzed by PCA are presented in Table 1.

| Table 1. Features analyzed by PCA. |
|-----------------------------------|
| Load for previous day              | Variable – Month |
| Outdoor temperature               | Variable – Hour of day |
| Outdoor temperature for previous day | Variable – Day of week |
| Dew point temperature             | Variable – Day of month |
| Wet bulb temperature              | Variable – Day of year |
| Specific humidity                 | Binary variable – Holidays |
| Solar irradiance                  | Binary variable – Working day |

Based on the amount of variation explained by each component, PCA helps remove redundant variables and identify input features. From the use of the PCA approach, the number of candidate features were reduced to nine (from the original set of 14 features).

| Table 2. Model inputs. |
|------------------------|
| Heat load for previous day | Variable – Day of week |
| Outdoor temperature      | Variable – Day of year |
| Outdoor temperature for previous day | Binary variable – Holidays |
| Variable – Month         | Binary variable – Working day |
| Variable – Hour of day   |                         |

The number of neurons in the input layer of the MLP was set to be equal to the number of independent input features. The total number of neurons in the hidden layer was determined through a trial and error approach (varying the number of neurons from 3 to 30). Various activation functions, including logistic sigmoid, hyperbolic tangent and exponential were tested in the MLP. Although a number of independent ANN models were developed, one single ANN model that consisted of nine inputs, one hidden layer...
with 19 neurons and one output layer was selected as the most optimal model. The MLP architecture used for this study is shown in Figure 5 and its structure is summarized in Table 3 along with its corresponding coefficients of determination and calculated MSEs (calculated from the normalized data set).

![Figure 5. ANN architecture used for this study.](image)

Table 3. Predictive model structure.

| Neurons | Inputs: 9 |
|---------|-----------|
|         | Hidden layer: 19 |
|         | Output: 1 |
| Activation function | Hyperbolic Tangent |
| Coefficient of determination $R^2$ | Mean square error (MSE) |
| Training data set | Test data set | Validation data set | Training data set | Test data set | Validation data set |
| 0.965 | 0.960 | 0.962 | 0.023 | 0.011 | 0.009 |

5. Computational results

The predictive performance for 24 hours in autumn and winter (normalized data set) is reported in Table 4. Figure 6–7 present a comparison between the predicted heat load and the actual heat load for the selected hours.

Table 4. Performance results for 24 hours.

| Autumn | Winter |
|--------|--------|
| RMSE   | MAPE   | RMSE   | MAPE   |
| 0.062  | 7.482  | 0.033  | 2.987  |
The results presented in Table 4, as well as Figures 6–7, suggest that the method implemented in this study is a promising approach for heat load forecasting in small cogeneration units and DHS.

6. Conclusion and future work
In recent years, machine learning (ML) methods used for the estimation and forecasting of energy demand have drawn considerable attention due to their advantage over linear and nonlinear programming models. In this paper, we develop a short-term forecasting model based on an artificial neural network approach. The article aims to fill the gap in the literature on heat load forecasting for small cogeneration units and DHS. In order to generate a suitable input data set for the ANN-based predictive model, several candidate input features were investigated. A principal component analysis (PCA) approach was applied to reduce the dimensionality of the data and for the identification of uncorrelated input components. An artificial neural network model with 19 neurons in the hidden layer obtained the best prediction values. The predicted hourly results show that the approach implemented in this study can be employed as a quick and reliable forecasting tool in small cogeneration units and DHS. Future work includes the study of additional meteorological descriptive features and improvements in network complexity.

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

References
[1] Lund H., Werner S., Wiltshire R., Svendsen S., Thorsen J.E., Hvelplund F., Vad Mathiesen B., 2014, *Energy* 68 1-11
[2] BP Statistical Review of World Energy 2017, June 2017.
[3] Short M., Crosbie T., Dawood M., Dawood N., 2017, *Appl. Energy* 186 304-320
[4] Fang T., Lahdelma R., 2016, *Appl. Energy*, 179 544–552
[5] Dotzauer E., 2002, *73* (3-4) 277-284
[6] Grosswindhager S., Voigt A., Kozek M., 2011, *Proc. 31st Int. Symp. Forecast.* 1 1-8
[7] Rahimikhoob A., 2010, *Renew. Energy* 35 (9), 2131–2135
[8] Livingstone D.J., Manallack D.T., Tetko I.V., 1997, *J. Comput. Aided. Mol. Des.* 11 (2), 135-142
[9] Idowu S., Saguna S., Ahlund C., Schelen O., 2016, *Energy Build.* 133 478-488
[10] Sajjadi S., Shamshirband S., Alizamir M., Yee P.L., Mansor Z., Manaf A.A., Altameem T.A., Mostafaepour A., 2016, *Energy Build.* 122 222-227
[11] Wojdyga K., 2014, *Int. J. Energy Power Eng.* 3 (5) 237-244
[12] Dalipi F., Vildirim Yayilgan S., Gebremedhin A., 2016, *Appl. Comput. Intell. Soft Comput.*, 2016 1–11
[13] Sakawa K. I. M., Katagiri H., Matsui T., 2010, *Sci. Math. Jpn. Online* 449–464
[14] Chen Y., Luh P.B., Guan C., Zhao Y., Michel L.D., Coolbeth M.A., Friedland P.B., Rourke S.J., 2010, *IEEE Trans. Power Syst.* **25** (1) 322-330

[15] Ommen T., Markussen W. B., Elmegaard B., 2014, *Energy* **74** (1) 109-118

[16] Lund H., Möller B., Mathiesen B.V., Dyrelund A., 2010, *Energy* **35** (3), 1381-1390

[17] Wojdyga K., 2008, *Energy Build.* **40** (11) 2009-2014

[18] Raza M.Q., Khosravi A., 2015, *Renew. Sustain. Energy Rev.* **50** 1352–1372

[19] Dreyfus G., *Neural networks: methodology and applications*, 1st ed. Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg, 2005.

[20] Goddard Earth Sciences Data and Information Services Center (GES DISC), *Global Modeling and Assimilation Office (GMAO)* (2015), MERRA-2 tavg1_2d_rad_Nx: 2d,1-Hourly,Time-Averaged,Single-Level,Assimilation,Radiation Diagnostics V5.12.4., Greenbelt, MD, USA, 2015.

[21] Abdi H., Williams L.J., 2010, *Wiley Interdiscip. Rev. Comput. Stat.* **2** (4) 433–459