Proceedings of the Seventh IASTED International Conference
POWER AND ENERGY SYSTEMS
August 29–31, 2007, Palma de Mallorca, Spain
ISBN Hardcopy: 978-0-88986-689-8 / CD: 978-0-88986-690-4

FORECAST OF THE HEAT DEMAND OF A DISTRICT HEATING SYSTEM

Prof. Dr. Wolfgang Schellong (1), Dipl.-Ing. Francois Hentges (2)
(1) Cologne University of Applied Sciences, Betzdorfer Str. 2, D50679 Cologne, Germany
wolfgang.schellong@fh-koeln.de
(2) Hentges GmbH, Am Meilenstein 4, D53909 Zülpich, Germany
francois.hentges@hentges.biz

ABSTRACT
The paper describes different mathematical modeling methods for the heat demand forecast of a district heating system. Mainly the regression analysis and the design of neural networks are tested on the basis of real consumption data of the heating system. The forecast tools are necessary to control and optimize the operating schedule of a cogeneration plant in combination with the district heating system. The heat demand forecast implemented in an energy management system helps to increase the energy efficiency and supports the sustainable energy development. An analysis of the consumption data and of the main influence factors on the heat demand is necessary in order to obtain suitable forecast models. The paper describes the data management as well as the process of the mathematical modeling. The design of clusters depending on seasonal impacts and the influence of climate factors are investigated. Linear multiple regression models are compared with individually designed neural networks. The experiences of the application of both methods to real data sets are presented.

KEY WORDS
Energy forecast, energy efficiency, modeling, regression analysis, and neural networks

1. Introduction

Much of the energy generated today is produced by large-scale, centralized power plants using fossil fuels (coal, oil and gas), hydropower or nuclear power, with energy being transmitted and distributed over long distances to consumers. The quality of the generation process can be evaluated by the energy efficiency calculating the ratio between energy outputs (services such as electricity and heat) and inputs (primary energy). In accordance with the goals of the World Energy Council [1] energy efficiency programs are necessary for sustainable energy development. Improvements in energy efficiency should be seen as one element of the bigger global energy system challenge to reduce greenhouse gas emissions.

The efficiency of conventional centralized power systems is generally low in comparison with combined heat and power (CHP) technologies (cogeneration) which produce electricity or mechanical power and recover waste heat for process use. CHP systems can deliver energy with efficiencies exceeding 90%, while significantly reducing the emissions of greenhouse gases and other pollutants. Selecting a CHP technology for a specific application depends on many factors, including the amount of power and thermal needs as well as an existing heating network. District heating is a system for distributing heat generated in a centralized location for residential and commercial heating requirements. District heating systems improve the energy efficiency if the heat is obtained by a cogeneration process.

Professional forecast tools of the energy demand are necessary to control and optimize the operating schedule of a cogeneration plant. These tools are based on mathematical models describing the relations between the main influence factors and the power and heat demand. The following experiences present the applications of different mathematical models of the forecast of the heat demand in a district heating system.

2. Mathematical Modeling

2.1 Forecast Model

The energy consumption of the delivery district of a power plant depends on many different influence factors. The heat demand of a district heating system depends strongly on climate factors as outside temperature, wind, global radiation, and humidity. On the other side seasonal factors influence the energy consumption. Usually the power and heat demand is higher on working days than at the weekend. Furthermore vacation and holidays have a significant impact on the energy consumption. Last but not least the heat and power demand in the delivery district is influenced by the operational parameters of enterprises with large energy demand and by the consumer's individual behaviour.

Because of the large number of influence factors and their uncertainty it is impossible to build up an 'exact' physical model for the energy demand. Therefore the energy demand is calculated on the basis of mathematical models describing the influence of climate factors and operating conditions on the energy consumption. The objectives of the mathematical modeling process can be summarized by two main targets: The model should be as simple as possible and on the other side as exact as necessary.

In the special case of the heat demand forecast the heat consumption data of the district heating system are divided into three groups depending on the season: winter, summer, and the transitional period containing spring and autumn. In each cluster the consumption data are separately modeled for working and for holidays. The hourly heat consumption can be calculated by typical time dependent demand profiles which allow splitting up the daily forecast into each hour of the day. Alternatively the hourly consumption can be directly modeled as described in chapter 4.3.

As described above the heat demand mainly depends on climate factors. As a result of a preliminary analysis the outdoor temperature gets the strongest impact among these factors. Additionally the temperature difference of two sequential days represents a significant influence factor modeling heat storage effects of buildings and heating systems. Thus these influence factors represent the basis of the modeling process using the statistical method of the regression analysis described in chapter 2.2 as well as using the algorithm of neural networks which is detailed explained in the chapters 3 and 4.

The data basis for the test of the different forecast models is given by the heat consumption data of the district heating system of the city of Offenbach on Main. The district heating system belongs to a power plant consisting of two cogeneration units and two additional steam blocs. The annual heat consumption amounts to about 460.000 MWh [2]. About 3.000 customers from industry, office buildings, and residential areas are delivered by the system. Thus the consumption behaviour is characterized by a mixed structure. But the main part of the heat consumption is used for room heating purposes.

2.2 Regression model

Following the modeling strategy of chapter 2.1 the heat demand Q_{th} of a district heating system can be simply described by a linear multiple regression model (RM):

$$Q_{th} = a_0 + a_1 t_{out} + a_2 \Delta t_{out} \qquad (1)$$

where t_{out} represents the daily average outside temperature and Δt_{out} describes the temperature difference of two sequential consumption days.

The model (1) can be extended by additional climate factors as wind, solar radiation and others. But in order to get a model based on a simple mathematical structure and because of the dominating impact of the outdoor temperature among the climate factors only the two regression variables are used in (1). The results of the regression analysis for each seasonal and weekday dependent cluster are checked by the correlation coefficients and a residual analysis. The quality of the regression models of the heat consumption strongly depends on seasonal effects (see table 1 in chapter 4.5).

3. Design of neural networks

3.1 Building blocks of neural networks

The basic elements of neural networks (NN) are the neurons, which are simple processing units linked to each other with directed and weighted connections. Depending on their algebraic sign and value the connections weights are inhibiting or enhancing the signal that is to be transferred.

Depending on their function in the net, three types of neurons can be distinguished: The units which receive information from outside the net are called input neurons. The units which communicate information to the outside of the net are called output neurons. The remaining units are called hidden neurons because they only send and receive information from other neurons and thus are not visible from the outside. Accordingly the neurons are grouped in layers. Generally a neural net has one input and one output layer but can have several hidden layers.

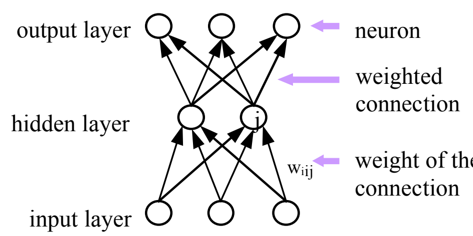
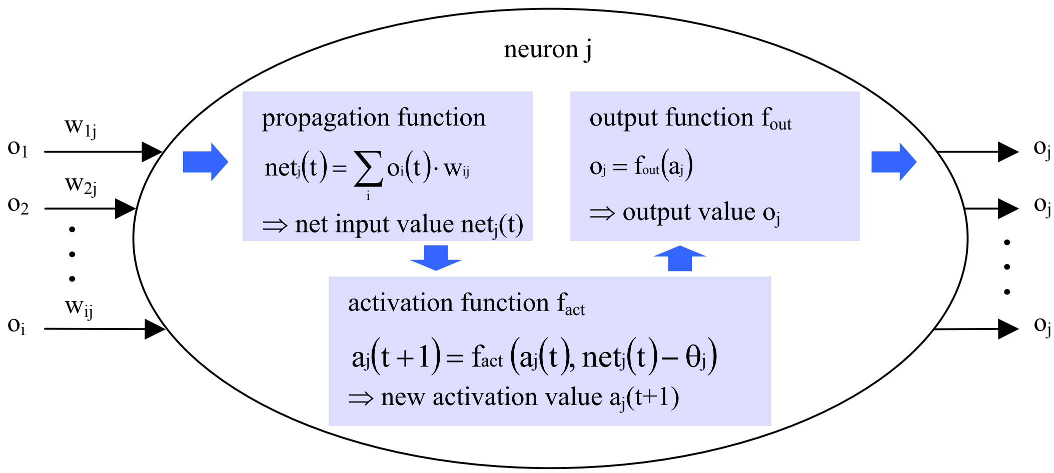


Fig. 1: Graphical representation of a neural network

The pattern of connection between the neurons is called the network topology. In the most common topology each neuron of a hidden layer is connected to all neurons of the preceding and the following layer. Additionally in so-called feedforward networks the signal is allowed to travel only in one direction from input to output. To calculate its new output depending on the input coming from the preceding units (or from outside) a neuron uses three functions, which are characterizing its behaviour. First the inputs to the neuron j from the preceding units combined with the connection weights are accumulated to yield the net input. This value is subsequently transformed by the activation function f_{act}, which also takes into account the previous activation value and the threshold θ_j (bias) of the neuron to yield the new activation value of the neuron. The final output o_j can be expressed as a function of the new activation value of the neuron. In most of the cases this function f_{out} is not used so that the output of the neurons is identical to their activation values (fig. 2).

Fig. 2: Structure of a neuron



3.2 Learning process in neural nets

A neural network has to be configured in such a way that the application of a set of input values produces the desired output values. There exist various methods to set the weights of each connection in order to reduce the error between the desired output and the actual output. The supervised learning methods incorporate an external teacher, who trains the network by providing it with input and matching output patterns. The most widely used algorithm for supervised learning is the backpropagation rule. It was presented by Rumelhart, Hinton and Williams in 1986 [3].

Backpropagation trains the weights and thresholds of a neural net and can be applied to feedforward networks with monotonic and everywhere differentiable activation functions. Prior to start learning it is necessary to initialize the connection weights by random values. The learning process involves three phases: During the first phase the input values are presented and propagated forward through the network to compute the output values. In the second phase the output values are compared with the desired values to calculate the error. During the third phase the algorithm alters the weights and thresholds in the network in an appropriate way in order to reduce the error (fig. 3).

Typically, the error E of the network is calculated by the sum of the squared individual errors for each pattern of the training set. This error is depending on the connection weights; different sets of weights produce different errors. We can write $E(W) = E(w_{11}, w_{12}, \ldots, w_{nn}) = \sum_p E_p$ with

$$E_p = \frac{1}{2} \sum_j (t_{pj} - o_{pj})^2$$

where E_p is the error for one pattern p, t_{pj} is the desired output from the output neuron j and o_{pj} is the real output from this neuron.

The backpropagation rule is a gradient descent method. With N as the number of weights, the mathematical error function can be plotted in N+1 dimensional space. The backpropagation algorithm tries to reach the global minimum of the error surface by adding to the network weights a fraction of the negative gradient.

A general problem of the training is the overlearning (or overfitting). A network that has been adjusted too many times to the patterns of the training set learns these specific inputs so well that it won't tolerate any slight deviation from them. Thus if too many cycles of the backpropagation algorithm are applied, the network looses his ability to give appropriate outputs to new (not already trained) input patterns. This is the ability of generalization. To avoid overlearning a set of pattern not used in training is presented to the network in regular intervals during the learning process. If the error for this set is not falling anymore but start to rise, overlearning begins and the training has to be stopped.

The gradient descent method has different drawbacks, which results from the fact that the algorithm doesn't know the totality of the error surface. The method aims to find a global minimum with only information about a very limited part of the error surface. Backpropagation can for example be stuck on plateaus where the slope is extremely slight or in deep gaps by oscillation from one side to the other. To allow a faster and more effective learning different extends to the backpropagation method have been published. The most common are the momentum term [5] and the flat spot elimination term [6].

To simulate neural networks specialized software is needed. The results exposed in the next chapter have been obtained by using the free available software JavaNNS [7].

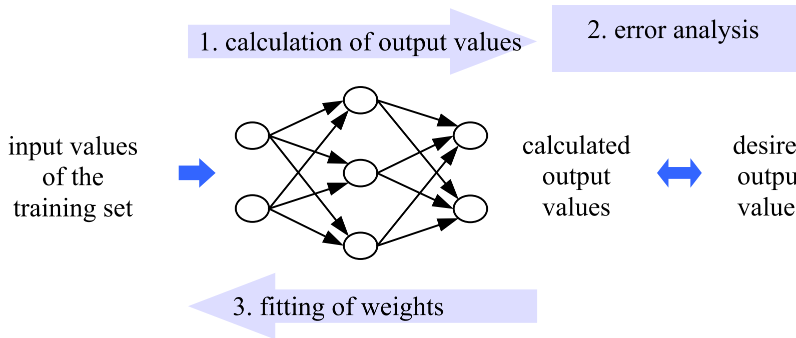


Fig. 3: Backpropagation learning rule [4]

4. Forecast of the heat demand

4.1 Design of NN

The following structure and rules are used for the network design:

- feedforward net with one layer of hidden neurons connected to all neurons of the input and output layer
- backpropagation learning rule with momentum term and flat spot elimination
- topological update mode to calculate new activation values of the neurons

In order to design the complete network the following other attributes have to be defined: The number of neurons in the hidden layer, the activation function and the way to present the input values to the net. To reach the best generalization and to evaluate the quality of a trained network the dataset is split into three same sized parts: The training set is used to train the net, the validation set is used to avoid overlearning and the test set is used to check the performance of the trained neural network. In order to find the most accurate net for the forecast of the heat demand several type of networks are trained and their prediction errors for the test set are compared. Each net is trained three times up to the beginning overlearning phase and then the net with the best forecast is retained. Networks with 3 to 8 hidden neurons (fig.4) are used with three sigmoid (S-shaped) activation functions: The logistic, hyperbolic tangent and limited sine function.

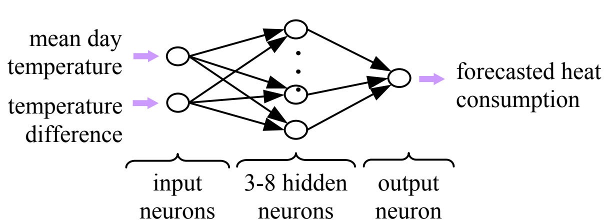


Fig. 4: Network for the daily forecast

The formulas of the activation functions are:

$$f_{\log}(x) = \frac{1}{1+e^{-x}}, \qquad f_{\tanh}(x) = \tanh(x) = \frac{e^{x}-e^{-x}}{e^{x}+e^{-x}},$$

$$f_{\sin}(x) = \begin{cases} 1 & \text{for } x > \pi/2 \\ \sin(x) & \text{for } -\pi/2 \le x \le \pi/2 \\ -1 & \text{for } x < \pi/2 \end{cases}$$

Additionally for the hourly heat forecast two ways of presentation of the hour are used.

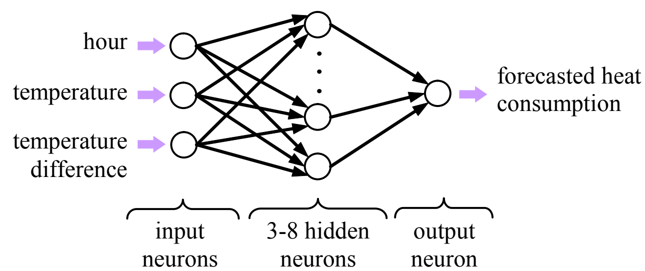


Fig. 5: Hourly forecast with coding by standardization

The coding by standardization (fig. 5) implies that a dedicated value is assigned to every hour. This value is presented to one input neuron. The 1 out of n coding (fig. 6) requires that a neuron is assigned to each hour. To enter a defined hour the neuron that is allocated to this hour is set to 1 while the other neurons for the hour input are set to 0.

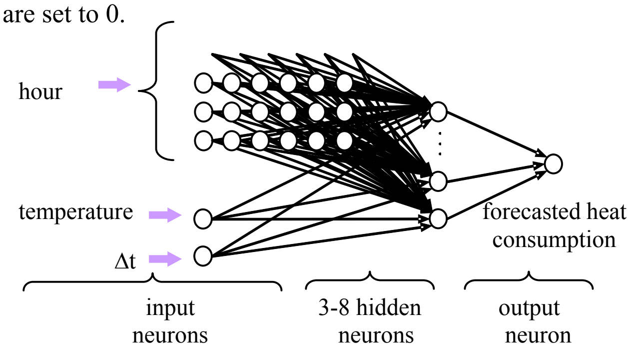


Fig. 6: Hourly forecast with 1 out of n coding

Before starting the training of the networks, the optimal parameter values of the backpropagation with momentum term and flat spot elimination are defined by testing different values and retaining the values which require the lowest number of training cycles to the beginning overlearning.

4.2 The daily forecast by NN

The comparison of the mean forecast error for the 6 categories in which the days were divided shows that a network with a logistic activation function delivers the best results. For winter days the maximum difference of the forecast errors between the nets with different activation functions is 1%, for days of the transition period it is generally lower than 5% but for summer days the difference is generally higher than 10%. Concerning the number of neurons in the hidden layer, the comparison shows that for days of winter and the transitional period nets with 6 neurons in the hidden layer deliver the best forecasts. For summer days all nets with logistic activation function delivers the nearly same mean forecast error. For each period a net was trained 3 times to the beginning overlearning, and the net with the best forecast was retained.

4.3 The hourly forecast by NN

The analysis of the mean forecast error shows that the following three network types with 3 to 6 hidden neurons deliver the best predictions: Nets with logistic or hyperbolic tangent activation function with coding by standardization and networks with limited sine activation function with 1 out of n coding. The coding method of standardization implicates significant smaller networks than the 1 out of n coding. Considering that smaller neural networks can be computed much faster than larger nets, only the two networks with coding by standardization and 3 hidden neurons are retained from the nets mentioned above. The analysis of the mean forecast error for the important winter period delivered from the two retained networks shows that a net with 3 hidden neurons and logistic activation function with coding by standardization delivers the best hourly forecasts. Again for each period a net was trained 3 times to the beginning overlearning and the net with the best forecast was retained.

4.4 The daily forecast by RM

Corresponding to the modeling aspects described in chapter 2.2 for each season and each weekday a regression model (see equation 1) is calculated. The models describe the dependence of the daily heat demand on the outdoor temperature and the temperature difference of two sequential days. In order to estimate the regression parameters of the model (1) the database of the reference year is split up into the training set and the test set. The regression parameters are calculated by solving the corresponding least squares optimization on the basis of the training set. The quality of the model is checked by the comparison between the forecasted and the real heat consumption for the test dataset. The correlation coefficients and the mean prediction errors (table 1) are used as quality parameter. For the reference year the correlation coefficients range from 0.81 for the summer time to 0.93 for the winter season. The modeling results show that the quality of the models for the summer and transitional seasons is worse in comparison with the winter time.

4.5 Comparison of the results

The following table shows the mean prediction errors of the tested models for the different seasons and workdays of the reference year. The mean error is calculated for each model by:

$$\varepsilon = \frac{1}{n}\sum_{i=1}^{n}\frac{|Q_{th} - Q_{real}|}{Q_{real}} \cdot 100\% \ , \qquad (3)$$

where n is the number of test data.

The results of table 1 show that the daily heat forecast models by neural networks and by linear regression have similar qualities. Only for the weekend days of the transitional period the daily NN produces better results than the RM. The results of the hourly NN are generally worse than the others, especially for the summer and the transitional periods. All methods obtain the best results in the winter period. Additional tests show that the neural networks have advantages in the case of data sets with large deviations, which is typical for the summer and the transitional periods.

season	summer		transitional period		winter	
day type	workdays	weekend / public holiday	workdays	weekend / public holiday	workdays	weekend / public holiday
Daily NN	16.1	12.0	15.0	15.8	5.6	5.7
Hourly NN	25.5	24.5	20.5	24.2	7.7	7.7
Daily RM	16.0	12.0	12.9	19.8	5.5	5.6

Table 1: Mean prediction errors ε for tested models

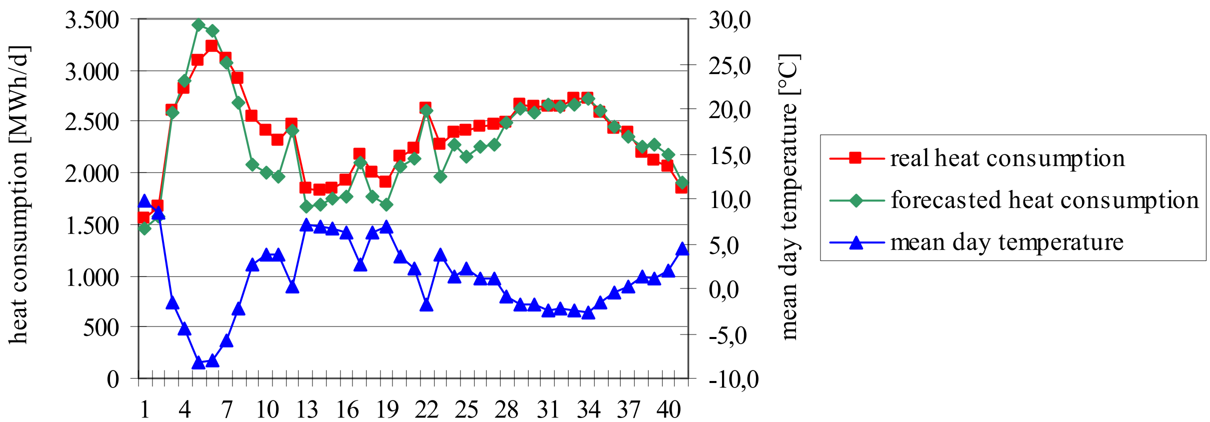


Fig. 7: Real and forecasted heat demands of sequential winter workdays

The diagram in fig. 7 shows the winter workdays of January and February. It allows the comparison between the forecasted and the real heat consumption. The forecasted values tend to overshoot in booth directions. With rapidly falling mean temperatures the neural network predicts higher heat consumption than the real values and vice versa for rapidly increasing temperatures. In periods with only small changes of the mean temperature the predictions of the neural net are extremely close to the real values.

5. Conclusion

The tested methods of the neural networks as well as the classical regression analysis represent useful and applicable tools for the forecast of the heat demand of a district heating system. For both methods it is necessary to divide the annual data set into clusters depending on seasons and weekday types. Regarding only the influence of the outdoor temperature together with the temperature difference of sequential days delivers sufficient modeling results. Generally the models for the winter period are better than those ones for the summer and the transitional period. The reason for this modeling effect lies in the fact that the deviation of the heat consumption increase with higher outdoor temperatures. In this case the individual consumer's behavior dominates the impact of climate factors.

Generally there exists no best modeling method. The regression analysis has the advantage of an easy numerical management whereas the individually designed neural networks are able to model highly deviated data sets. The effort for the optimal design of a neural network is much higher than for the regression analysis. The modeling is to be improved by taking into account more data clusters which are able to fit the consumer's behavior. Additionally it is possible to use more climate factors in the heat demand model assuming that there are detailed climate data available for each cluster. The forecast tools of the heat demand are suitable for the implementation into an energy management system to control and optimize the operating schedule of a cogeneration plant.

References

[1] WEC Statements, Energy efficiencies – pipe-dream or reality, World Energy Council, London, 2006

[2] W. Schellong, K. Steinmetz, Rechnergestützte Kraftwerksberichterstattung und Betriebsdatenanalyse, *Euroheat & Power, 24*, 1995, 542-550

[3] D.E. Rumelhart, G.E. Hinton, R.J. Williams, Learning representations by back-propagation errors, *Nature, 323*, 1998, 533-536

[4] K. Hufendiek, J. Albiger, Künstliche neuronale Netze bei der kurzfristigen Lastprognose für die Kraftwerkseinsatzoptimierung, *VDI Berichte, 1352,* 1997

[5] D.E. Rumelhart, G.E. Hinton, R.J. Williams, *Learning internal representations by error propagation*, in D.E. Rumelhart, J.L. McClelland (Eds.), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, Vol. 1, MIT Press, Cambridge, MA, 1986, 318-362

[6] S.E. Fahlman, An empirical study of learning speed in back-propagation networks, in D. Touretzky, G. Hinton, T. Sejnowski (Eds.): *Proc. of the 1988 Connectionist Models Summer School*, Carnegie Mellon University, Morgan Kaufmann

[7] Java Neural Network Simulator, http://www-ra.informatik.uni-tuebingen.de/SNNS/