# Forecast of heat demand according the Box-Jenkins methodology for specific locality

BRONISLAV CHRAMCOV
Faculty of applied informatics
Tomas Bata University in Zlín
Nad Stráněmi 4511, 760 05 Zlín
CZECH REPUBLIC
chramcov@fai.utb.cz http://web.fai.utb.cz

*Abstract:* - In order to improve the control level of district-heating systems, it is necessary for the energy companies to have reliable optimization routines, implemented in their organizations. However, before a plan of heat production, a prediction of the heat demand first needs to be determined. Forecast of this heat demand course is significant for short-term and long-term planning of heat production. This forecast is most important for technical and economic consideration. In this paper we propose the forecast model of heat demand based on the Box-Jenkins methodology. The model is based on the assumption that the course of DDHD can be described sufficiently well as a function of the outdoor temperature and the weather independent component (social components). Time of the day affects the social components. The time dependence of the load reflects the existence of a daily heat demand pattern, which may vary for different week days and seasons. Forecast of social component is realized by means of Box-Jenkins methodology. This model is used for prediction of heat demand in different locality. The results of heat demand prediction in specific locality and conclusions are presented.

*Key-Words:* - Prediction, District Heating Control, Box-Jenkins, Control algorithms, Time series analysis

## 1 Introduction

The paper deals with the utilization of time series prediction for control of technological process in real time. An improvement of technological process control level can be achieved by time series analysis in order to prediction of their future behaviour. We can find an application of this prediction also by the control in the Centralized Heat Supply System (CHSS), especially for the control of hot water piping heat output [2].

Due to the large operational costs involved, efficient operation control of the production sources and production units in a district heating system is desirable. Knowledge of heat demand is the base for input data for operation preparation of CHSS. Term "heat demand" is instantaneous required heat output or instantaneous consumed heat output by consumers. Term "heat demand" relates to term "heat consumption". It express heat energy, which is the customer supplied in a specific time interval (generally day or year).

The course of heat demand and heat consumption can be demonstrated by means of heat demand diagrams. The most important one is the Daily Diagram of Heat Demand (DDHD) which demonstrates the course of requisite heat output during the day (See Fig. 1).

These diagrams are most important for technical and economic consideration. Therefore forecast of these diagrams course is significant for short-term and long-term planning of heat production. It is possible to judge the question of peak sources and namely the question of optimal distribution loading between cooperative production sources and production units inside these sources according to time course of heat demand. The forecast DDHD is used in this case [2].

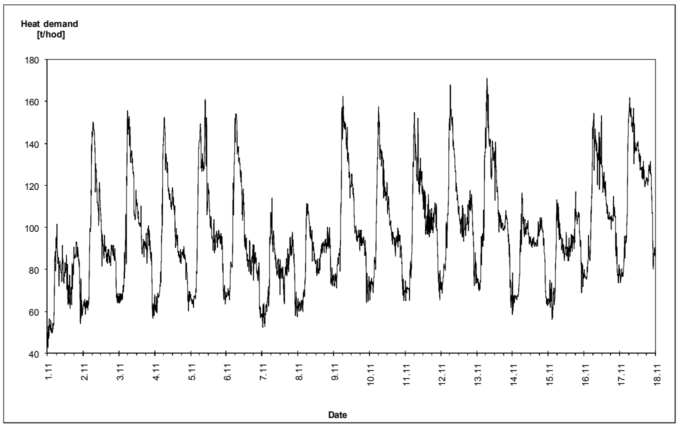


Fig. 1: DDHD for the concrete locality

## 2 Problem Formulation

Most forecasting models and methods for load prediction have already been suggested and implemented with varying degrees of success. They may be classified into two broad categories: classical (or statistical) approaches and artificial intelligence based techniques.

The statistical methods forecast the current value of a variable by using a mathematical combination of the previous values of that variable and previous or current value of exogenous factors, specially weather and social variables. These include linear models, solving by means of non-linear models, spectral analysis method, ARMA models, Box-Jenkins methodology etc. The methods based on artificial intelligence techniques process mass data. These include expert systems, neural networks, fuzzy neural models etc.

But most applications in the subject consider the prediction of electrical-power loads. Nevertheless was created several works, which solve the prediction of DDHD and its use for control of District Heating System (DHS). A number of these works are based on mass data processing [7], [6]. But these methods have a big disadvantage. It consists in out of date of real data. From this point of view is available to use the forecast methods according to statistical method. The basic idea of this approach is to decompose the load into two components, whether dependent and whether independent. The weather dependent component is typically modeled as a polynomial function of temperature and other weather factors. The weather independent component is often described by a Fourier series, ARMA model, Box-Jenkins methodology or explicit time function. Previous works on heat load forecasting [1], [5], show that the outdoor temperature, together with the social behaviour of the consumers, has the greatest influence on DDHD (with respect to meteorological influences). Other weather conditions like wind, sunshine etc. have less effect and they are parts of stochastic component.

In this paper we propose the forecast model of DDHD based on the previous approach. The model is based on the assumption that the course of DDHD can be described sufficiently well as a function of the outdoor temperature and the weather independent component (social components). Time of the day affects the social components. The time dependence of the load reflects the existence of a daily heat demand pattern, which may vary for different week days and seasons. Forecast of social component is realized by means of Box-Jenkins methodology. This is described in Section 3. The complete forecast algorithm with inclusion of outdoor temperature is stated in Section 3.1. Section 4 presents some computational results executed in specific locality. Finally, in Section 5, some conclusions are given.

# 3 Forecast Model of Heat Demand

As mentioned above, a number of works are based on mass data processing. But these methods have a big disadvantage. It consists in out of date of real data. From this point of view is available to use the forecast methods according to statistical method in our case the methodology of Box –Jenkins [3]. This method works with fixed number of values, which are update for each sampling period.

This methodology is based on the correlation analysis of time series and it works with stochastic models, which enable to give a true picture of trend component and also of periodic components. Because this method achieves very good results in practice, it was chosen for prediction of social component of HSDD.

The course of time series of HSDD contains two periodic components (daily and weekly period). But general model according to Box-Jenkins enables to describe only one periodic component. We can propose two eventual approaches to calculation of forecast to describe both periodic components [4].

- The method, that uses the model with double filtration
- The method – superposition of models

## 3.1 Forecast algorithm for inclusion of outdoor temperature

Above mentioned methods do not describe sudden fluctuation of meteorological influences. In this case we have to include these influences in calculation of prediction. For inclusion of outdoor temperature influence in calculation of prediction of HSDD was proposed this general plan:

a) The influence of outdoor temperature filter off from time series of HSDD by means of heating characteristic (function that describes the temperature-dependent part of heat consumption)
b) Prediction of HSDD by means of Box-Jenkins method for this filtered time series
c) Filtration of predicted values for the reason of inclusion of outdoor temperature influence (on the base of weather forecast)

From the previous plan is evident that the principal aim is to derive an explicit expression for the temperature-dependent part of the heat load. It is obvious that the temperature dependence is non-linear. For relatively high outdoor temperatures, the temperature has less influence. For example, the load will almost be the same for 25 °C and 27 °C. A corresponding conclusion is also true for relatively low temperatures, e.g. whether the outdoor temperature is -28 °C or -30 °C does not matter, the production units will produce at their maximum rate anyway.

Regarding to previous consideration we can used the temperature-dependent part of heat consumption in the form (1).

$$z_t^{kor} = x_1 \cdot T_t^3 - x_2 \cdot T_t \qquad (1)$$

where: $z_t^{kor}$ is correction value of heat consumption in time $t$ including outdoor temperature influence, $T_t$ is real value of outdoor temperature in time $t$ , $x_1$, $x_2$ are constants

The course of heating characteristic for constants $x_1 = 0.002$, $x_2 = 3.5$ is shown in the figure 2.

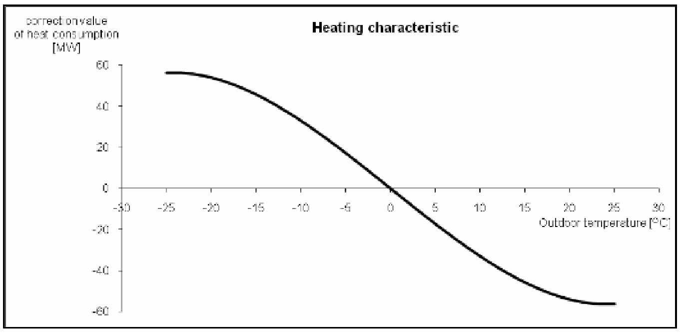


Fig. 2: The sample of heating characteristic (cubic function)

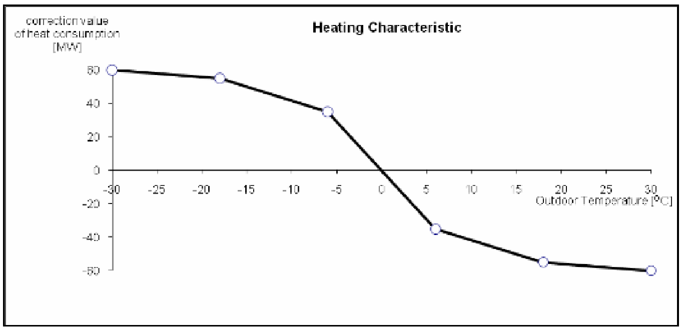


Fig. 3: The sample of heating characteristic (piecewise linear function)

The temperature dependent part can assumed to vary as a piecewise linear function [5], see the illustrating example in Fig. 3. Here a function with five segments is used, but the number of segments can of course be chosen arbitrarily.

Given the number of segments as a $N_s$. and the temperature levels as $\tau_i$, $i = 1,2..., N_s+1$. The parameters of heating characteristic are changed in these temperature levels. Now we can consider the temperature-dependent part of heat consumption in the form(2).

$$\begin{aligned} &z_t^{kor} = \alpha_i \cdot T_t + \beta_i, \\ &\tau_i < T_t < \tau_{i+1},\ i = 1,...,N_s \end{aligned} \qquad (2)$$

where: $z_t^{kor}$ is correction value of heat consumption in time $t$ including outdoor temperature influence, $T_t$ is real value of outdoor temperature in time $t$ ,$\alpha_i$ is the slope of $i$-th segment, $\beta_i$ is absolute equation term of $i$-th segment

Constants ($x_1$, $x_2$ and ,$\alpha_i$, $\beta_i$) have to be determined for concrete locality empirically.

Filtration time series of HSDD that input in prediction model is defined in the form (3).

$$z_t^{filtr} = z_t - z_t^{kor} \qquad (3)$$

where: $z_t^{filtr}$ is heat consumption in time $t$ with filtering off the influence of outdoor temperature, $z_t^{kor}$ is correction value of heat consumption in time $t$ including outdoor temperature influence, $z_t$ is real value of heat consumption in time $t$

After prediction calculation of filtering off time series is necessary to filtrate the predicted values for the reason of inclusion of outdoor temperature influence (on the base of weather forecast). We can define this operation in the form (4).

$$z_t^+ = z_t^{filtr+} + z_t^{kor} \qquad (4)$$

where: $z_t^{filtr+}$ is predicted value of filter off time series of heat consumption in time $t$, $z_t^{kor}$ is correction value of heat consumption in time $t$ including outdoor temperature influence, $z_t^+$ is predicted value of heat consumption in time $t$.

The value $z_t^{filtr+}$ is obtained by application of the equation (1) or (2) for this operation. We use weather forecast (temperature forecast).

# 4 Calculation of forecast for specific locality

Pursuant to the mentioned theory and literature a program was created in Matlab, which enables to choose available mathematical statistical model for calculation of prediction of HSDD course. All testing is based on lot of real data. These data were obtained in specific locality and they are processed for next using in text file form. The program is drawn in user's menu and by help of that it is possible to choose many parameters of forecast calculation.

Selection of calculation method of prediction of HSDD course is a other possibility of submitted program. We can realize the calculation of prediction by means of the method that uses model with double filtration and the method – superposition of models.

After choosing one of the methods the calculation of prediction is started. At first in the course of calculation it is searched for the most suitable model, it is for optimum number of autoregression parameters and optimum number of parameters of moving average process. After following calculation of prediction, resulting graphic window is displayed (see Fig. 4.). In this window there is drawn course of HSDD, course of predicted data and probability limit. The result can be represented in concrete value form. These values are followed by calculation and they can be displayed in resulting window, which is shown on Fig. 6. In this window it is possible to find also optimum number of autoregression parameters and optimum number of parameters of moving average process.

## 4.1 Results of heat demand forecast in concrete locatity

It is necessary to stress that the real data are used for all experiments and tests of proposed forecast model. The real data were obtained due to close cooperation of our research workplace with energy plant operations. In our case it is close cooperation with company MST a.s. – Power and Heating plant Olomouc, Power and Heating Plant Otrokovice, a.s. and company United Energy a.s. - Power and Heating plant Most-Komořany.

The accuracy of proposed forecast model is presented on data from the locality Most-Komořany. The model was tested on data from two following weeks, i.e. from the January 13 to January 26, 2009. 24 hours-ahead and 12 hours-ahead forecast were made twice a day at 6.00 AM and 6.00 PM. Accuracy of the forecast is analyzed and summarized by means of Mean Absolute Percent Error (MAPE) and Root Mean Squared Error (RMSE) (see Table 1). The samples of results of heat demand prediction are shown in the Fig. 4 and Fig. 5.

Table 1: Accuracy of the forecast model for 24, 12 hours ahead forecasts

| Date, Time | 24 hours-ahead forecast | | 12 hours-ahead forecast | |
|---|---|---|---|---|
| | MAPE [%] | RMSE [MW] | MAPE [%] | RMSE [MW] |
| 13.1.2009, 6:00 AM | 1,75 | 2,97 | 1,78 | 3,21 |
| 13.1.2009, 6:00 PM | 2,95 | 5,24 | 1,74 | 2,73 |
| 14.1.2009, 6:00 AM | 3,99 | 6,07 | 4,17 | 6,88 |
| 14.1.2009, 6:00 PM | 4,35 | 6,26 | 4,84 | 6,63 |
| 15.1.2009, 6:00 AM | 3,33 | 4,89 | 3,36 | 5,27 |
| 15.1.2009, 6:00 PM | 3,62 | 4,90 | 3,08 | 4,20 |
| 16.1.2009, 6:00 AM | 4,72 | 5,53 | 3,86 | 5,14 |
| 16.1.2009, 6:00 PM | 6,90 | 9,38 | 5,63 | 6,37 |
| 17.1.2009, 6:00 AM | 5,09 | 7,60 | 7,41 | 10,25 |
| 17.1.2009, 6:00 PM | 4,35 | 6,36 | 2,01 | 2,27 |
| 18.1.2009, 6:00 AM | 5,65 | 7,08 | 6,78 | 8,82 |
| 18.1.2009, 6:00 PM | 5,79 | 7,01 | 4,39 | 4,56 |
| 19.1.2009, 6:00 AM | 4,95 | 6,03 | 5,98 | 7,14 |
| 19.1.2009, 6:00 PM | 3,33 | 4,18 | 3,02 | 3,36 |
| 20.1.2009, 6:00 AM | 2,99 | 3,64 | 3,40 | 4,45 |
| 20.1.2009, 6:00 PM | 2,70 | 3,28 | 2,63 | 2,62 |
| 21.1.2009, 6:00 AM | 3,48 | 4,13 | 2,76 | 3,85 |
| 21.1.2009, 6:00 PM | 6,55 | 7,37 | 4,36 | 4,54 |
| 22.1.2009, 6:00 AM | 5,58 | 6,82 | 8,02 | 9,07 |
| 22.1.2009, 6:00 PM | 7,51 | 9,78 | 3,56 | 3,68 |
| 23.1.2009, 6:00 AM | 9,47 | 10,75 | 10,59 | 12,54 |
| 23.1.2009, 6:00 PM | 8,26 | 8,29 | 8,75 | 8,61 |
| 24.1.2009, 6:00 AM | 6,59 | 7,38 | 7,72 | 7,93 |
| 24.1.2009, 6:00 PM | 4,15 | 5,18 | 4,75 | 6,01 |
| 25.1.2009, 6:00 AM | 5,40 | 7,07 | 3,23 | 3,79 |
| 25.1.2009, 6:00 PM | 9,57 | 11,16 | 7,93 | 9,52 |
| 26.1.2009, 6:00 AM | 7,34 | 8,63 | 9,83 | 10,88 |
| 26.1.2009, 6:00 PM | 5,39 | 6,29 | 4,99 | 5,73 |
| Average value | 5,21 | 6,55 | 5,02 | 6,07 |

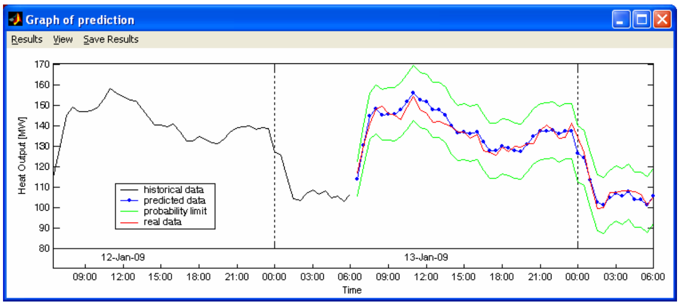


Fig. 4: 24 hours ahead forecast of heat demand on 13.1.2009 6:00 AM

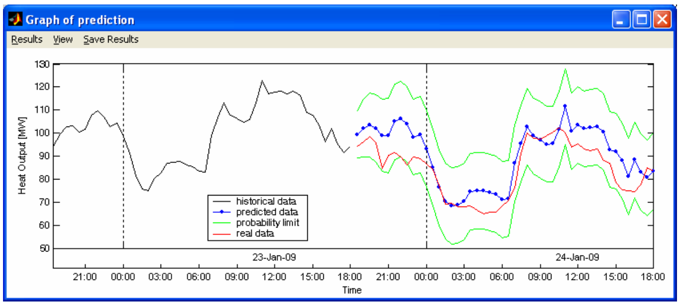


Fig. 5: 24 hours ahead forecast of heat demand on 23.1.2009 6:00 PM

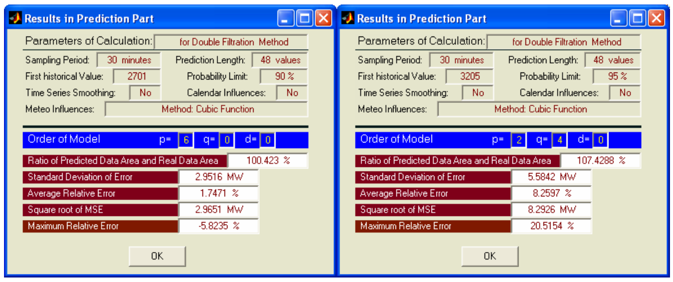


Fig. 6: Result windows for 24 hours ahead forecast of heat demand on 13.1.2009 6:00 AM and 23.1.2009 6:00 PM

From the results, we conclude that the MAPE for the test period is at any time less than 10 percent and RMSE seldom exceed the value of 10MW. Average value of MAPE in the test period is approximately 5% and average value of RMSE is approximately 6 %. Obviously, we also observe that the value of MAPE and

RMSE are lower for a half a day ahead forecast than for a day ahead forecast. A deeper analysis of the results shows that the worse prediction was achieved on the days of weekend.

# 5 Conclusion

This paper presents the Box-Jenkins methodology for building up the forecast model of time series of DDHD and the possibility of improvement of this forecast model with help of inclusion of outdoor temperature influence. The proposed forecast method was successfully applied to real data from concrete locality. The effectiveness of proposed forecast model was demonstrated through a comparison of the real heat demand data with short-term (24, 12 hours) forecasted values. In term of the average MAPE in the test period our approach achieved 5% error.

Heat demand forecast plays an important role in power system operation and planning. Accurate heat demand prediction saves costs by improving economic load dispatching, unit commitment, etc. Model described should prove useful for the control in the Centralized Heat Supply System (CHSS), especially for the qualitative-quantitative control method of hot-water piping heat output – the Balátě System [2].

*Acknowledgement:*
This work was supported in part by the Ministry of Education of the Czech Republic under grant No. MSM7088352102 and National Research Programme II No. 2C06007.

*References:*

[1] ARVASTSON, Lars. *Stochastic modelling and operational optimization in district-heating systems*. Lund, 2001. 240 s. Doctoral thesis. Lund University, Centre for Mathematical Sciences. ISBN 91-628-4855-0.

[2] BALATE, Jaroslav. *Design of Automated Control System of Centralized Heat Supply*. Brno, 1982. Thesis of DrSc (Doctor of Science) Work. TU Brno, Faculty of Mechanical Engineering.

[3] BOX, George E; JENKINS, Gwilym M. *Time series analysis : forecasting and control*. Rev. ed. San Francisco : Holden-Day, 1976. 575 s. ISBN 0-8162-1104-3..

[4] DOSTÁL, Petr. *Machine Processing of Daily Diagram Course Prediction of Loading the Centralized Heat Supply System*, Brno, 1986. Doctoral thesis. TU Brno, Faculty of Mechanical Engineering.

[5] DOTZAUER, Erik. Simple model for prediction of loads in district-heating systems. *Applied Energy*. November-December 2002, Volume 73, Issues 3-4, s. 277-284. ISSN 0306-2619.

[6] HIPPERT, H.S.; PEDREIRA, C.E.; SOUZA, R.C. Neural networks for short-term load forecasting : a review and evaluation. *IEEE Transactions on Power Systems* . February 2001, Volume 16, Issue 1, s. 45-55. ISSN 0885-8950.

[7] LEHTORANTA, O., SEPPÄLÄ, J., KOIVISTO, H., KOIVO, H. Adaptive district heat load forecasting using neural networks. *Proceedings of Third international Symposium on Soft Computing for Industry*, Maui, USA, June, 2000.