LOAD FORECASTING APPLICATIONS FOR THE ENERGY SECTOR

BORIS BIZJAK
Univerza v Mariboru, FERI, Smetanova 17, 2000 Maribor

E-mail: boris.bizjak@um.si tel. +386 41 327 348

Abstract – Load forecasting is vitally important for the electricity industry with in a deregulated economy. It has many applications including energy purchasing and generation, load switching, contract evaluation, and infrastructure development [1]. Time series methods are based on the assumption that the data have an internal structure, such as autocorrelation, trend, or seasonal variation. Time series forecasting methods detect and explore such a structure. Time series have been used for decades in such fields as economics, as well as electric load forecasting [2]. We present a short-term 5 days load forecasting applications for industrial plant with an electric arc furnace [3] in the City of Ravne, Slovenia. We present five different load-forecasting techniques: linear regression (off line) [4], ARIMA (off line), Winter’s multiplicative (off line) and real time Data Mining [7] [8] [9]. At short-term load forecasting linear regression, for which we use two time series: energy and production at the electric arc furnace. Next, we divide the forecasting model in two parts: ARIMA with predictor “loads” at the electric arc furnace and Winter's multiplicative without any predictor; in this case the electric arc furnace is off. IBM SPSS software was the tools of selection.

Data Mining provides the means for making sense of tremendous volumes of data by automating the processes of categorising and clustering common elements, identifying trends and anomalies in the data, and predicting what would happen given those factors [7]. In this paper we discuss Data Mining at ARMA (Autoregressive and Moving Average Models) [9] and Data Mining at ART (Autoregressive Tree Models) [8]. For the development and improvements of forecasting we used the Microsoft technology: SQL server, Analysis Server and the WEB server. We briefly discuss and compare predictions with and without any predictors. This means, the model uses only a one time series, energy measurements at the plant:

In addition, we show Long-term load forecasting samples [5] annual predictions for two points at the electrical transmission network in Slovenia. The software tools were IBM SPSS. We present two different load-forecasting techniques: ARIMA and seasonal models. Outliers can occur by forecasting, so we discuss some worst-case scenarios. We show forecasting quality factors for each forecasting model: CL (Confidence Intervals), MAPE (Mean Absolute Percentage Error) and time series value at forecasting time. In conclusion, we compare the results.

Keywords: identify dynamic system, data mining, load forecasting, power systems, market forecasting, outliers, worst-case scenario

1 Load Forecasting using Linear Regression

Many problems in engineering and science involve exploring the relationships between two or more variables. Regression analysis is a statistical technique that is very useful for these types of problems. For example, a load forecasting for an industrial plant with an electric arc furnace in the City of Ravne, Slovenia, supposing that the number of loads at the electric arc furnace is related to the energy consumption at the industrial plant (Figure 1, Figure 2). Regression analysis can be used to build a model for predicting electrical load as a given the number of loads at the electric arc furnace. The resulting model can also be used for business processing, on the electric market, such as buying right quantity of electrical energy for a another day.

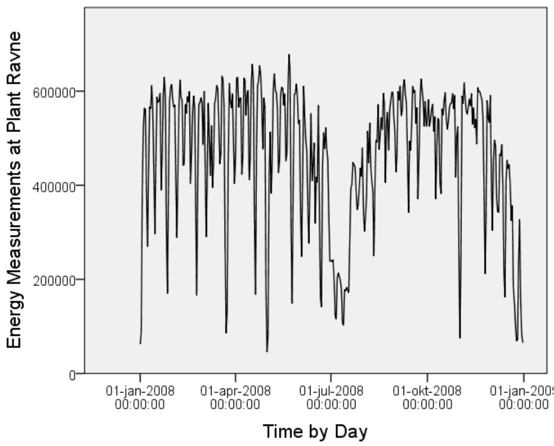


Figure 1: Load curves at plant Ravne for one year

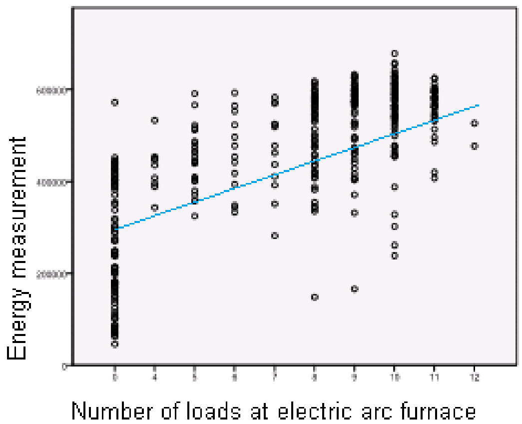


Figure 2: Correlation between number of loads at electric arc furnace and load at plant.

Inspections of Figure 2 indicated, that although no simple curve will pass exactly through all the points that represent relationship. There is a strong indication that the points lie randomly around a straight line. It is reasonable that the mean of random variable Y is related to x by a straight-line relationship:

$$Y = +\beta_1 x + \varepsilon \qquad (1)$$

β_0…….regression coefficient intercept,
β_1 ……regression coefficient slope,
x……...regressor or predictor variable,
Y...........criterion variable,
ε ……..random error term $N(0, \delta^2)$, with mean zero and variance δ^2.

The estimates β_0 and β_1 should result in a line that is a “best fit” to the data. The German scientist Karl Gauss proposed estimating the parameters β_0 and β_1 in order to minimise the sum of the squares of the vertical deviations. We call this criterion for estimating the regression coefficients the method of least squares**.** Advanced reader please refer to [4] to study properties of the least square estimators. Next we explain only the major steps to estimating β_0 and β_1, such as definitions:

$$\widehat{\beta_1} = \frac{S_{xy}}{S_{xx}} \qquad (2)$$

$$\widehat{\beta_0} = \bar{y} - \widehat{\beta_1}.\bar{x} \qquad (3)$$

$\widehat{\beta_0}$…..observed intercept, is an unbiased estimator of the true intercept β_0.,
$\widehat{\beta_1}$….. is an unbiased estimator of the true slope β_1.

$$SS_{xx} = \sum_{i=1}^{n}(x_i - \bar{x})^2 \qquad (4)$$

$$SS_{xy} = \sum_{i=1}^{n} y_i (x_i - \bar{x})^2 \qquad (5)$$

n…..the number of observation in a sample.

Error ε is a random variable normally distributed with a mean of 0 and variance δ^2:

$$e_i = y_i - \hat{y}_i \qquad (6)$$

e_i………...is called the residual.

$$SS_E = \sum_{i=1}^{n} e_i^2 \qquad (7)$$

SS_E……….error sum of squares.

There is actually another unknown parameter in our regression model, δ^2 (the variance of the error term ε):

$$\widehat{\delta^2} = \frac{SS_E}{n-2} \qquad (8)$$

$\widehat{\delta^2}$…..calculated value of ε with properties $N(0, \delta^2)$.

In simple linear regression the estimated standard error of the slope and the estimated standard error of the intercept are:

$$SE(\widehat{\beta_1}) = \sqrt{\frac{\widehat{\delta^2}}{S_{XX}}} \qquad (9)$$

$$SE(\widehat{\beta_0}) = \sqrt{\widehat{\delta^2}\left[\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right]} \qquad (10)$$

respectively, where $\widehat{\delta^2}$ is computed from Equation 8 respectively.

An important part of assessing the adequacy of a linear regression model is testing statistical hypotheses about the model parameters and constructing certain confidence intervals. Hypothesis testing in simple linear regression is discussed in the next section, and presents methods for constructing confidence intervals. To test hypotheses about the slope and intercept of the regression model, we must make the additional assumption that the error component in the model, ε , is normally distributed. Thus, the complete assumptions are that the errors are normally and independently distributed with mean zero and variance δ^2, abbreviated NID(0, δ^2).

$$H_0\colon \beta_1 = \beta_{1,0} \qquad (11)$$
$$H_1\colon \beta_1 \neq \beta_{1,0} \qquad (12)$$

H_0….null hypothesis,
H_1…..alternative hypothesis.

With which probability can we trust the hypothesis, that a calculated slope is equal to a real slope β_1 ($H_0\colon \beta_1 = \beta_{1,0}$)? We use the t-test to the confirm hypothesis:

$$T_o = \frac{\widehat{\beta_1} - \beta_{1,0}}{SE(\widehat{\beta_1})} \qquad (13)$$

We would reject $H_0\colon \beta_1 = \beta_{1,0}$ if:

$$|t_0| > t_{\frac{\alpha}{2}, n-2} \qquad (14)$$

α……………………….... Significance (Sig). Standard values are 0,05 or 0,01.

(1-α)100[%]………Confidence Interval (CI, CL). Standard values are 95 % or 99 %.

A similar procedure can be used for testing hypotheses about the intercept β_0 :

$H_0: \beta_0 = \beta_{0,0}$ (15)
$H_1: \beta_0 \neq \beta_{0,0}$ (16)

Test statistic fort the intercept β_0:

$$T_o = \frac{\widehat{\beta_1} - \beta_{0,0}}{SE(\widehat{\beta_0})} \quad (17)$$

A method called the analysis of variance (ANOVA) can be used to test for significance of regression. The procedure partitions the total variability in the response variable into meaningful components as the basis for the test. At an F-test test statistic have a $F_{1,n-2}$ distribution. The analysis of variance identity is as follows:

$$F_0 = \frac{SS_R}{\frac{SS_E}{n-2}} = \frac{MS_R}{MS_E} \quad (18)$$

$$SS_R = \sum_{i=1}^{n} (\widehat{y_i} - \bar{y})^2 \quad (19)$$

Where SS_R is the sum of the squared errors. We follows the $F_{1,n-2}$ distribution, and we would reject H_0 if:

$$f_0 > f_{\alpha,1,n-2} \quad (20)$$

Fitting a regression model requires several assumptions. Estimation of the model parameters requires the assumption that the errors are uncorrelated random variables with mean zero and constant variance. Tests of hypotheses and interval estimation require that the errors be normally distributed. In addition, we assume that the order of the model is correct; that is, if we fit a simple linear regression model, we are assuming that the phenomenon actually behaves in a linear or first-order manner:

$e_i = y_i - \hat{y}_i$ i=1,2…,n (21)

e_i……residual

$d_i = \frac{e_i}{\sqrt{\hat{\delta}^2}}$ i=1,2…,n (22)

d_i …..standardised residual.

We may also standardise the residuals by computing. If the errors are normally distributed, approximately 95 % of the standardised residuals should fall within the interval (-2, +2). Residuals that are far outside this interval may indicate the presence of an outlier, that is, an observation that is not typical of the rest of the data. We test residual with a scatter chart, we put on x series predicted values and on y series standardised residual. In an ideal situation all scatter plot points are equally distributed within the chart area.

A widely used measure for regression model is following the ratio of the sum of squares R^2 (coefficient of determination):

$R^2 = \frac{SS_R}{SS_T} = 1 - \frac{SS_E}{SS_T}$ (23)
$SS_T = SS_R + SS_E$ (24)

SS_T ……total corrected sum of squares,
R^2……. coefficient of determination.

The values of coefficient of determination R^2 should be at interval $0 \leq R^2 \leq 1$.

We would like to predict complete the load at the industry plant in the City of Ravne, Slovenia for a 5 days in advance. An independent predictor is the number of loads at electric arc furnace. Working time at arc furnace are depend from energy price tariff policy over a week and a business production contracts at the arc furnace. As a result, at a Sunday 88 % of load at plant are depend from arc furnace production. Other production facility at a plant working for 5 days and over weekend is no production. At a five working days we have additive loads form arc furnace and other production facility.

Coefficient of determination R^2 is relatively low 0,6. We explain 60 % of energy consumption at plant in a year is dependent from the number of loads at the electric arc furnace. Therefore 40 % of loads cannot be explained with a linear regression model. State of the art of load forecasting is 90 %, so the linear regression model is robust, and we looking next for a better solution.

Table 1: Model summary

Model Summary [b]				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,771a	,595	,594	94380,865
a. Predictors: (Constant), Number of Loads at Electric Arc Furnace				
b. Dependent Variable: Energy measurements at plant				

Independent variable explain 60% per cent of variance (R Square) in load, which is highly significant, as indicated by F-value of 534,6 in the table below, that F-test say we can trust β_0 and β_1 > 99,9 %.

Table 2: ANOVA test

ANOVA[b]						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	4,762E12	1	4,762E12	534,560	,000[a]
	Residual	3,242E12	364	8,908E9		
	Total	8,004E12	365			
a. Predictors: (Constant), Number of Loads at Electric Arc Furnace						
b. Dependent Variable: Energy measurements at plant						

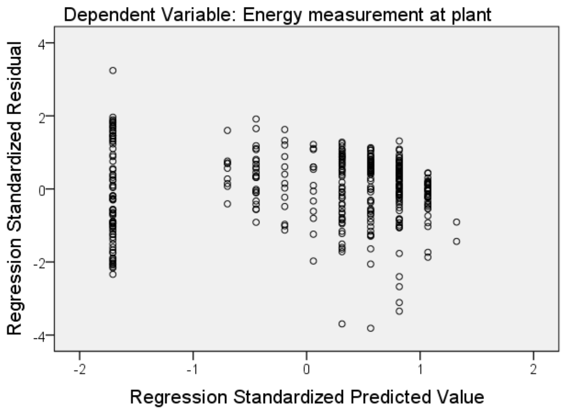
An examination of the t-test indicates that number of loads at electric arc furnace contribute to electric load, t-test say we can trust β_0 > 99,9 % and β_1 > 99,9 %.

Table 3: Coefficients

Coefficients[a]						
		Unstandardized Coefficients		Standardized Coefficients		
Model		B	Std. Error	Beta	t	Sig.
1	(Constant)	266.469,968	9.770,475		27,273	,000
	Number of Loads at Arc Furnace	28.822,853	1.246,633	,771	23,121	,000
a. Dependent Variable: Energy measurements at plant						

Equation of linear regression model is followed:

[Load Forecasting] = 266.470 + 28.823[Number of Loads at Electric Arc Furnace] $+\varepsilon$ ……….. (25)



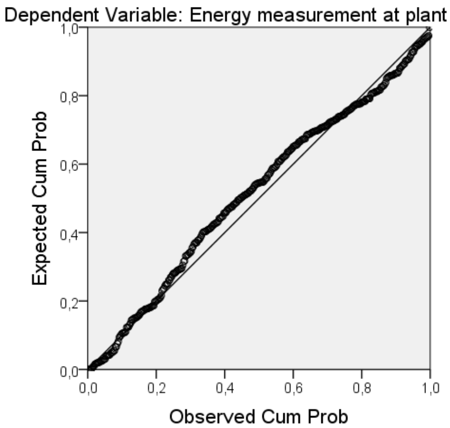


Figure 3: Scatterplot of residual against predicted values.

Figure 4: Normal P-P plot of regression standardized residual.

From the scatterplot of residuals against predicted values (Figure 3), you can see that there is no clear relationship between the residuals and predicted values, consistent with the assumption of linearity. The normal plot of regression standardised residuals (Figure 4) for dependent variable also indicates a relatively normal distribution.

2 Short term load forecasting with seasonal ARMA

We understand from Section 1 that only 60 % annual energy consumption is directly dependent on number of loads at electric arc furnace. In the next section we show a way of predicting remaining 40 % of energy consumption. The statistic software tool IBM SPSS - Forecast module help as by research. Visual inspection of Figure 1 show as production timetables over a year. First, two quartiles of the year the arc furnace was in normally in operation, next at July the arc furnace get for 3 weeks at the outage, and finally the last 6 months back to the grid with a full production. This gave us the possibility of researching separately two operational cases at the energy complex: production with arc furnace and production without arc furnace. ARMA (Autoregressive Moving Average Models) is the appropriate prediction model for first case, and Winter's multiplicative is the appropriate prediction model for second case.

2.1 Autoregressive Models

A stochastic model [2] that can be extremely useful in the representation of certain practically occurring series is the autoregressive model. In this model, the current value of the process is expressed as a finite, linear aggregate of previous values of the process and a random shock a_t. Let us denote the values of a process at equally spaced times t, t − 1, t − 2, . . . by z_t, z_{t-1}, z_{t-2}, z_{t-3} , ... Also let $\tilde{z}_t = z_t - \mu$ be the series of deviations from μ. Then

$$\tilde{z}_t = \varphi_1 \tilde{z}_{t-1} + \varphi_2 \tilde{z}_{t-2} + \cdot \cdot \cdot + \varphi_p \tilde{z}_{t-p} + a_t \quad (26)$$

is called an autoregressive (AR) process of order p. The reason for this name is that a linear model

$$\tilde{z}_t = \varphi_1 \tilde{x}_1 + \varphi_2 \tilde{x}_2 + \cdot \cdot \cdot + \varphi_p \tilde{x}_p + a_t$$

Relating a "dependent" variable z to a set of "independent" variables x_1, x_2, . . . , x_p, plus a random error term a, is referred to as a regression model, and z is said to be "regressed" on x_1, x_2, . . . , x_p. In (29) the variable z is regressed on previous values of itself; hence the model is autoregressive

The model contains p + 2 unknown parameters μ, φ_1, φ_2 , . . ., φ_p , δ_a^2. Which in practice have to be estimated from the data. The additional parameter is the variance of the white noise process a_t .

2.2 Moving Average Models

The autoregressive model (26) expresses the deviation $\tilde{z}_t$ of the process as a finite weighted sum of p previous deviations $\tilde{z}_{t-1}$, $\tilde{z}_{t-2}$, …,$\tilde{z}_{t-p}$ $+a_t$, of the process, plus a random shock a_t. Equivalently, as we have just seen, it expresses $\tilde{z}_t$ as an infinite weighted sum of the a's. Another kind of model, of great practical importance in the representation of observed time series, is the finite moving average process. Here we take $\tilde{z}_t$ linearly dependent on a finite number q of previous a's. Thus,

$$\tilde{z}_t = = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \cdot \quad \cdot \quad \cdot - \theta_q a_{t-q} \qquad (27)$$

is called a moving average (MA) process of order q. The name "moving average" is somewhat misleading because the weights 1, θ_1 $-\theta_2$ $- \cdot \quad \cdot \quad \cdot -\theta_q$, which multiply the a's, need not have total unity nor need they be positive.

It contains q + 2 unknown parameters μ, $\theta_1, \theta_2, \ldots, \theta_q$, δ_a^2 , which in practice have to be estimated from the data.

2.3 Short-term Load Forecasting Application

Forecasting model design is the process where we start with a coarse model, and through more iteration and reflections get closer to the best fit model. The learning time interval from observation time series, which the model parameter raises ($\mu, \varphi_1, \varphi_2, \ldots, \varphi_p, \delta_a^2$ and $\mu, \theta_1, \theta_2, \ldots, \theta_q, \delta_a^2$), should not be too long or too short. Forecasting models are alive, because they learn permanently parameters from the measurement data.

ARIMA (0,0,1)(0,1,1) model was made from observed data from 1.1.2008 to 7.4.2008. As an independent predictor we used the number of loads at the electric arc furnace.

The time series model supports both exponential smoothing and ARIMA model. Exponential smoothing model types are listed by their commonly used names such as Holt and Winters' Multiplicative. ARIMA model for our example is using the standard notation ARIMA(0,0,1)(0,1,1), where 0 is the order of autoregression, 0 is the order of differencing (or integration), and 1 is the order of moving-average, and (0,1,1) are their seasonal counterparts.

Forecasting model has determined that short-term load is best described by a seasonal ARIMA model with no order of differencing. The seasonal nature of the model accounts for the seasonal peaks that we saw in the week series plot (Figure 6).

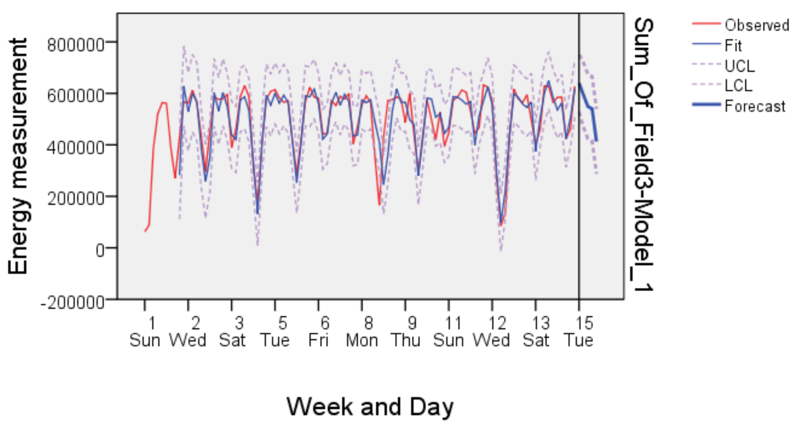


Figure 5: Observed, model assessment and prediction for 5 days at 15th week, Day 2, 2008.

The model statistics Table 5 provides summary information and goodness-of-fit statistics for estimated model. First, notice that the model contains one predictor. Model offers a number of different goodness-of-fit statistics, we opted only for the stationary R^2 value. This statistic provides an estimate of the proportion of the total variation in the series that is explained by the model and is preferable to ordinary R^2 when there is a trend or seasonal pattern, as is the case here. Larger values of stationary R-squared indicate better fit. A value of 0,87 means that the model does an excellent job of explaining the observed variation in the series.

The Ljung-Box statistic, also known as the modified Box-Pierce statistic, provides an indication of whether the model is correctly specified. A significance value less than 0,05 implies that there is structure in the observed series which is not accounted for by the model. The value of 0,77 shown here is not significant, so we can be confident that the model is correctly specified.

Table 5: Model statistic

Model	Number of Predictors	Model Fit statistics	Ljung-Box Q(18)			Number of Outliers
		Stationary R-squared	Statistics	DF	Sig.	
Sum_Of_Field3-Model_1	1	,866	11,670	16	,766	0

Table 6: Prediction statistic for a five days – MAPE is 6%.

Date	Week, Day	Prediction	UCL	LCL	Observed	Prediction Error [%]
10.4.2008	15 Tue	639.944	750359	529529	623.695	2,6
11.4.2008	15 Wed	593.950	719671	468228	548.449	8,3
12.4.2008	15 Thu	549.363	675084	423641	594.051	-7,5
13.4.2008	15 Fri	539.868	665590	414146	601.958	-10,3
14.4.2008	15 Sat	413.232	538954	287510	410.700	0,6

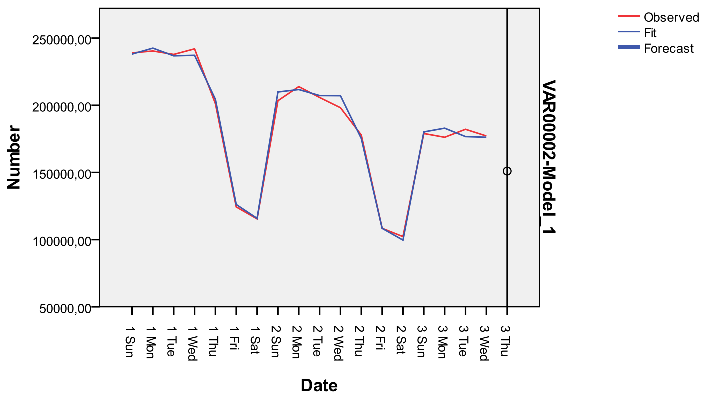


Figure 6: Seasonal week model (7,5) in observed data at July 2008.

3 Short Term Load Forecasting when Data Mining

Section three introduces data mining applications of automatic forecasting, that means we will link knowledge and conclusion from section 1 and section 2 into one new nonlinear prediction model. Introducing of automatic prediction at load prediction called data mining. Linear regression model [4] and seasonal ARMA [2] from the previous section were replaced with a nonlinear prediction model. This nonlinear model was realised according to the methodology of ART [8] (Autoregressive Tree Models), ARMA and sessional ARMA [7] [9]. ART and ARMA methodology can be explained in several ways, including the theory of digital filters [10], so also the spectral method in other words, he does not require additional amplitude predictors to the basic time series.

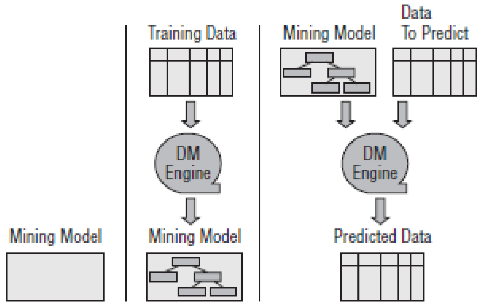


Figure 7: Data mining process

It took entrusted machine learning to determine the optimally parameters for the selected forecast method. For all data stored in a database use machine learning and statistics, so we have to build a structure called data mining [7].

Data mining is an interdisciplinary field of computer science, seeking patterns in large amounts of data through the use of artificial intelligence, machine learning, and statistics and data warehouse. The main goal of data mining is to extract information from data in an understandable form for later use. The basic idea of data mining is that the algorithm automatically extracts the characteristic pattern of sample data, the pattern is then used to calculate for example forecasts. The architecture of data mining is compiled from data warehouses and associated servers for analyses. The data visualising is performed with a WEB server. Analysis Server communicates with the data server via XMLA data formats (XML for Analysis). XMLA is an industry standard for data transmission in a systems for the analysis, which is format independent of the source and recipient data.

3.1 Real-Time Data Mining on the WEB

Linear regression forecasting method and ARMA forecasting, for the described application were first time published directories at the conferences ERK 2009 [6] and CIRED 2011 [5], then we have a project carried out with the software IBM SPSS, which means manually preparing data and forecast for each steps. This method is does not do well when applied for the daily work in a modern enterprise energy-consuming and for the modern IT applications is somewhat out-dated. So we made a technological leap to the latest IT technology and use of the data mining technology. The applied IT system has the following structure:

- Module to load observed data to the data warehouse.
- Time-sharing controller and communication controller.
- SQL data server.
- The server for analysis with data mining structure.
- WEB server.

We are learning the forecast model in sliding mode with 160 historical data. Load forecasting statistic is a real time WEB server applications (Table 7, Figure 8 and Figure 9), that show quality forecasting criteria: MAE - an absolute error of prediction, MAPE - relative error of prediction, RMSE - standard deviation of forecast errors.

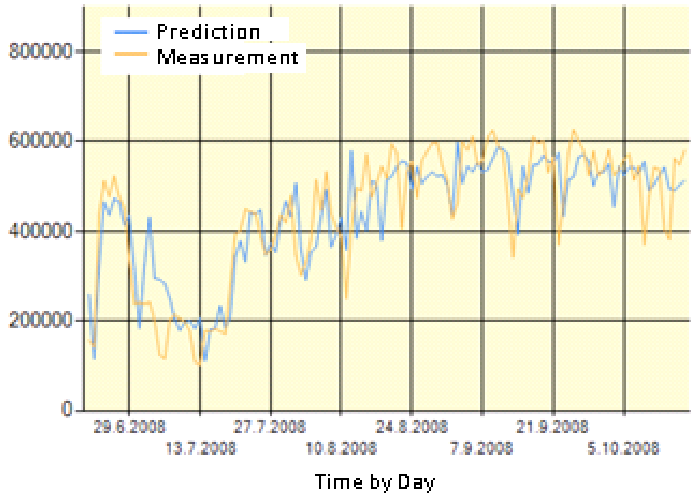


Figure 8: Time series of prediction and observed data 1 day ahead.

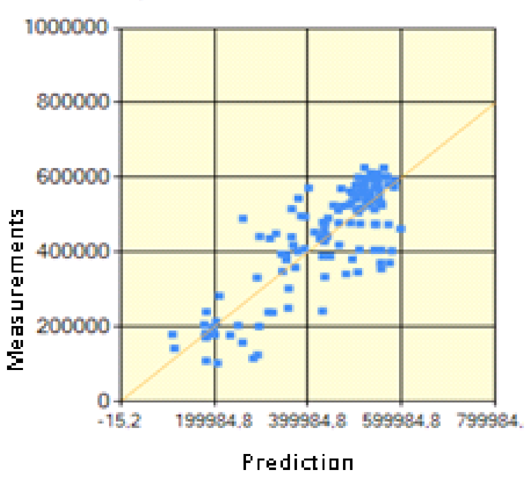


Figure 9: The correlation between the observed and prediction 1 day ahead.

.

One problem at work may be missing measurement data. Missing measurement data can be replaced with the previous measurement or mean value of time series. A substitution works well to some extent. Machine learning algorithms themselves react to the data holes in the measurement data. For time series with missing data forecast models automatically become simpler. Autoregressive tree (ART) become a lower grades, which returns lower predictive significance.

In large production systems with arc furnaces it is difficult to measure production data, arc furnace technology process is too complex. It is interesting to note, that the load forecasting by consideration predictor "number of loads at electric arc furnace" and without this predictor differ only by a few %.

Table 7: Prediction statistic for 130 samples.

Prediction for a Day	R^2	AVG measurements	STDEV measurements	AVG prediction	STDEV prediction	MAE	MAPE	RMSE
+5	0,36	441.054	141.209	431.911	111.789	92.377	27,5	113.107
+4	0,28	440.322	142.076	432.756	109.034	97.250	29,6	120.284
+3	0,37	440.265	143.129	433.972	109.567	90.999	27,9	113.231
+2	0,43	439.630	143,994	439.262	113.404	85.064	27,1	108.680
+1	0,72	438.210	144.951	437.343	123.829	59.186	17,6	77.032

Machine learning is carried out automatically, hourly for short-term forecast, once a day for a load forecasting average or daily extremes. Basic parameter for ARMA and ART are determined automatically by machine learning. After prediction we calculate the MAE, MAPE and RMSE of prediction. We developed a small simulation program for advanced optimisation. The simulation shows us how to allocate the weights to smooth the combination of seasonal ARMA and ART. Simulation algorithm is carried out step by step, first to move to the database to query learning time series, next learn a model, predict and save data back to database. If the simulation of model ideas are good, we accomplished switch to real-time forecasting and only the data source will be changed to real-time measurement. The simulation program for daily hourly extremes for 6 months be carried out for 1 hour with standard hard drives at server.

Finally the result was MAPE 17 % for all production conditions in one year. Please make a focus, that a observed time series is non-stationary and hides three different models inside. Figure 9 show us how data mining predicts cover switch from full arc furnace production, to three weeks stop. It is one outlier in prediction time series, after that prediction becomes stationary with low error ahead.

4 Long-Term Load Forecasting of Seasonal Models

4.1 Problem of Power Flow

The problem of power flow is the calculation of the voltage and phase angles of the voltage at each node in the fixed symmetrical three-phase system. Consequently, from this calculated flow of active and reactive power in transmission lines are transformers. We can also count losses. Starting point, input computer analysis of power flows, the single-pole circuit diagram consisting of: data transmission lines, data hubs and data transformers. As shown in Figure 10, each node belongs to one of 4 variables V_k,δ_k,P_k,Q_k. After two variables representing the input data for the power flow computation programs, two are calculated output variables. Due to the transparency, transported to the power node, divided into the generator and load:

$$Q_k = Q_{Gk} - Q_{Lk} \qquad (28)$$
$$P_k = P_{Gk} - P_{Lk} \qquad (29)$$

Each node is categorised in one of three types of nodes:

- Swing, slack node - normally is only one "swing" node, which for convenience is marked with 1. Swing node is a reference node to which the $V_1 = 1$ and phase angle $\delta_k = 0^0$. The programs for power flow account P_l and Q_l.
- Load node - input data are P_k and Q_k. The program for power flow account V_k and δ_k. Most of the nodes are a type of "load."
- Voltage controlled node- input data are V_k and P_k. The classic program for power flow account $\delta_k$and Q_k. An example of such a node is connected to a generator or static compensator.

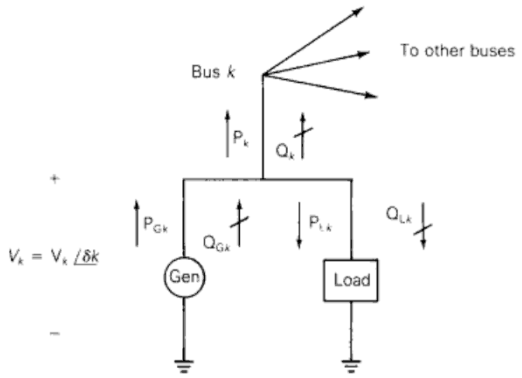


Figure 10: Node variables V_k,δ_k,P_k,Q_k

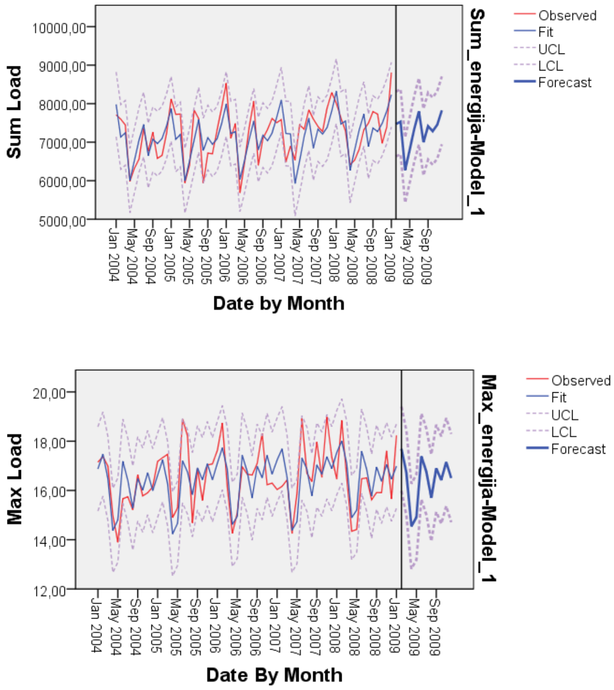
We are limited to the basic theoretical methods of implementation, to the extent necessary for an understanding of the article.

4.2 Long Term Load Forecasting at Two Nodes

The aim of the project was to carry out forecasting monthly average power flow for a period of one year and to predict the monthly hourly maxima of flows. Geographic distributed power systems are difficult to identify an appropriate predictor, or arrive at an appropriate database for predictors. For long-term forecasting we had available only historical time series of basic phenomenon - the flow of power.

We carried out a long-term forecast for the two energy nodes. Node 1 is called Ajdovščina, node 2 is called Krško. The aim was to carry out prediction of 12 months. Today's state of art prediction is 6 month, but our experiment for twelve months showed good result at MAPE from 5,4 % to 7,1%. In charts (Figure 11, 12, 13, 14) SUM are the average monthly value of P, MAX are monthly hour extreme value of P. Input average time series and maximum time series was calculated with SQL aggregate functions from hourly measurements of P and Q at the interval from 1.1.2004 to 31.12.2008. The model did not use an independent predictor. Visual inspection of the shape at input power flow time series, for different year's, shows that they are other somewhat similar shapes to each other, but certainly we found appropriate patterns of seasonal type, for the node 1 with the optimal model Winter's additive and node 2 with the simple seasonal forecast model. Predicted time interval was from 1.1.2009 to 31.12.2009. The result was calculated with an expert modeller at forecast module of a statistic software IBM SPSS. Figure 11 and Figure 13 show long term load forecasting at 12 months with upper and lower 95% confidence intervals.

We also got real measurement data from 1.1.2009 to 31.12.2009 and they were compared with predicted values. Quality of long term forecasting determines a relative difference between predicted data and measurements data, relative prediction error, for year 2009, follow at Figure 12 and Figure 14.



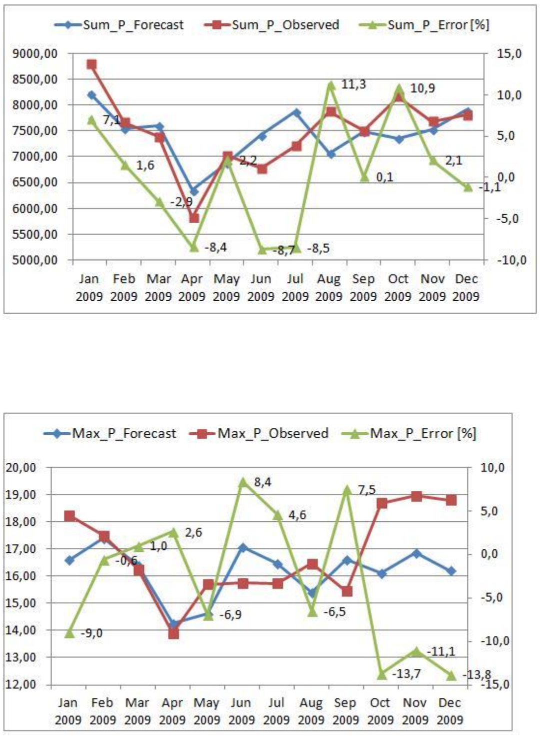
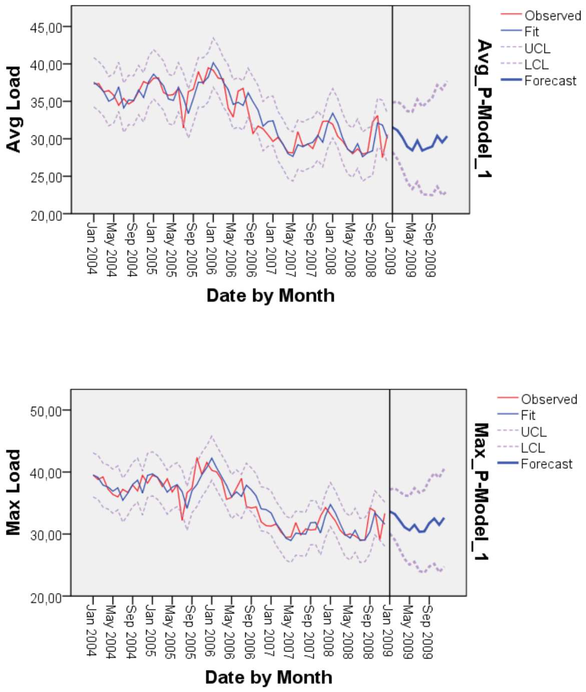


Figure 11: Observed and forecast for twelve months - node 1 Ajdovščina.

Figure 12: Load forecast error at node1.



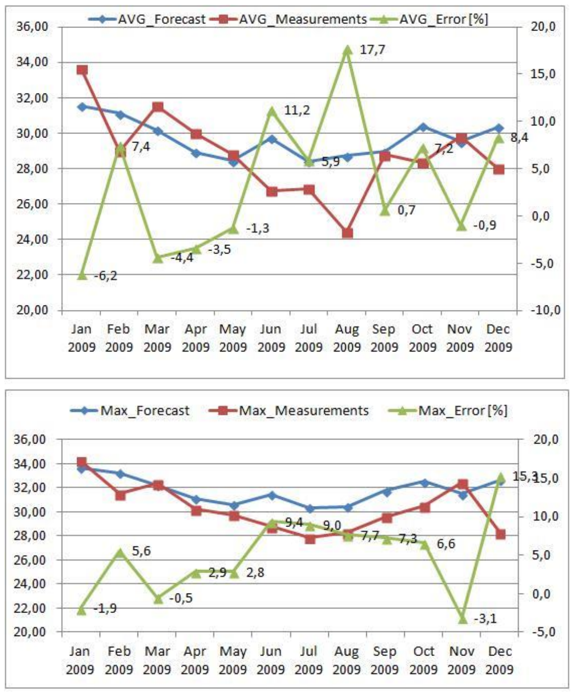


Figure 13: Observed and forecast for twelve months - node 2 Krško.

Figure 14: Load forecast error at node 2.

Table 8: Long-term load prediction statistic

Prediction value description	Model type	MAPE	Forecasting period
Sum month value of P at node 1	Winters Additive	5,4 %	12 months
Monthly hour extreme values of P at node 1	Simple Seasonal	7,1 %	12 months
Average month value of P at node 2	Simple Seasonal	6,2 %	12 months
Monthly hour extreme value of P at node 2	Simple Seasonal	6 %	12 months

Conclusion

Forecasting is a mathematical realisation of the old folk say history repeats itself. In this paper we show that history, in our cases in the energy sector, reiterates per hour (hourly maxima) daily for 5 days, 7 days, month and year. Of course, this does not say just by using coffee extracts pots, but with the help of modern statistics and mathematics, which predict the most likely value, with the addendum that the 95% of the predicted values of the population located in the interval of the confidence limit. Shown in the article they could alternatively be calculated by using the mechanical model, which in other words means a physical modelling of the observed system. Machine models for the cases described here are likely to be complex and lengthy calculations and the results are quite authentic. Model predictions we have for each observed point analysis calculated separately. The models forecast may change over time, depending on the characteristics of the observed phenomenon. Statistical methods require verification of validity of the method of input data - assumption testing, data considerations, which can sometimes be a big part of the mathematical solution in the cartridge.

REFERENCES:

[1] JOE H. CHOW, FELIX I. WU, JAMES A. MOMOH, APPLIED MATHEMATICS FOR RESTRUCTURED ELECTRIC POWER SYSTEMS, 2005, Springer

[2] GEORGE E. P. BOX GWILYM M. JENKINS GREGORY C. REINSEL. Time Series Analysis Forecasting and Control, 2014, Wiley

[3] Janez Bratina, Elektroobločna peč, Ravne na Koroškem, 1994.

[4] Douglas C. Montgomery, George C. Runger, Applied Statistic and Probability for Engineers, Wiley, 2003.

[5] Boris Bizjak, Jure Bizjak, Jože Voršič, Primeri kratkoročnega in dolgoročnega napovedovanja pretokov moči, CIGRE-CIRED 2011, Ljubljana

[6] Boris Bizjak, Napovedovanje porabe električne energije za železarski kompleks, ERK 2009, Portorož

[7] Jamie MacLennan, Zhao Hui Tang, Bogdan Crivat. Data mining with Microsoft SQL Server 2008, Wiley

[8] C.Meek, D.M. Chickering, D.Heckerman, Autoregresive Tree Models for Time-Series Analysis, Microsoft Research

[9] Bo Thiesson, David Maxwell Chickering, David Heckerman, Christopher Meek, ARMA Time-Series Modeling with Graphical Models, Microsoft Research

[10] Leland B. Jackson, Digital filters and Signal Processing, Kluwer academic publishers, 2002, Fifth Printing