Power Flow Forecasting With Low Prediction Error

Boris BIZJAK

UM, FERI, Koroška cesta 46, Maribor E: boris.bizjak@um.si T: +386 41 327 348

Abstract: The goal of the development project was to forecast power flows for a large industrial complex with an hourly step for several days, with as low a prediction error as possible. The primary activity of the industrial complex is the production of steel and technology-related production. However, there are several independent companies on this energy connection who have nothing in common with steel production technology. The biggest consumer is the electric arc furnace and the accompanying technology that has a specific energy profile. The rest of the companies are classic industrial electricity consumers, following the 7/5 working week. We have an interesting mix of electricity consumption profiles, so classic forecasting methods do not produce good results. The article presents an approach to predicting consumption by 24 hours R^2 = 0.93 MAPE = 9.9% and 48 hours R^2 = 0.91 MAPE = 12.3% by combining standard forecasting methods and technical innovation appropriately.

Keywords: forecasting, low prediction error, power flow, steel mill, industrial complex.

Introduction

Predicting the future has been a great motivation for the human mind for centuries. Today's prediction methods are based on the time series theory. Time series harbour a wealth of information. With proper mathematical and statistical processing, they give us prediction models. With forecasting models, we can predict the future. The goal is to make the forecast with as low an error as possible.

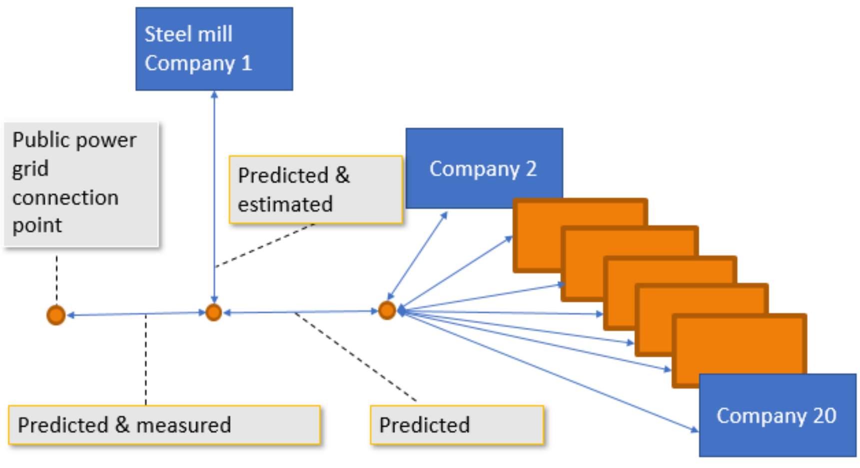

Figure 1: Block diagram of industrial complex and power flows` branches

The motivation is to predict with as low an error as possible the amount of energy we need in the next few days, which can be useful in reducing the cost of electricity. Secondly, the question is why would we produce more energy than we need and, thus, pollute the environment? We present power flow forecasts for three branches of one energy node at the entrance of a large industrial complex. The Root-Mean-Square Error (RMSE) is a measure of the quality of the different forecasting models for the same time series. The classic prediction model for the described case reaches RMSE = 3389. A power flow forecasting model with low prediction error reaches RMSE = 2584, in other words, we have improved the forecast model by over 30%. The prediction step is 1 hour or 15 min. The results can be represented by a forecast error for the next hour on with 6% error, and the next 48 hours with 12% error.

1. Methodology of work

The methodology of the work is based on the time series theory, statistics and machine learning. The forecasting methodology is supported by a database and visualisation is performed through a secure Web server. The work process is basically classical: first, based on the characteristics of the time series of the measured electricity consumption, we choose the appropriate forecasting methodology. The methodology chosen returns a prediction model, but with indeterminate parameters. The parameters of the prediction model are then calculated by the programme before each prediction. Up to this point the methodology is completely normal.

Classical work procedures need to be modified in order to achieve low prediction error. Simplified, for each time series at a branch, it is necessary to determine what is the useful signal and what is the noise. The prediction is then made only for the useful signal, because the noise interferes with the forecast. After the prediction is completed, the noise is again added to the useful signal. That's how we get power flow forecasting with low prediction error. In this way, the basic measuring time series of the measured energy was processed - it is the first input signal. In order to perform the described procedures, we used two different forecast models, which partially cover the dynamic property of consumption in each branch.

The second input time series is a planned production of the arc furnace in company 1; this is called the predictor. For low prediction error, it is necessary to determine what information the predictor of the production of the arc furnace must contain. The predictor should be as simple as possible so that it can be realised in a project. A periodogram was used to optimise the content of the predictor. The result was a bit surprising: For a prediction with a low error it is only necessary to know the time interval from when and until the production in the furnace will run. So, the predictor only has a value of 0 or 1.

The first step to successful realisation of power flow forecasting is to study the characteristics of the main electricity consumers. There are several companies in the area of the former Ravne Ironworks, employing approximately 3,000 people. The largest company on the site is Metal Ravne, with approximately 900 employees. Metal Ravne is the largest company, and the largest consumer of electricity. The company consists of a steelworks, a rolling mill and electro-smelting under slag. In the steel plant, the basic unit is a 45-tonne electric UHP oven and a vacuum refill kiln for castings of classical ingots. In the Electro-under-slip section under the slag, 36-tonne and 3-tonne ESR devices are in use. At present, it operates in the scale of an industrial complex of 20 companies. The dominant characteristic of energy consumption from the grid is determined by the electric arc furnace, and by companies that are technologically connected to the smelting of iron (Figure 1).

For the predictive work, we used the classical information structure for such examples in the composition: relational database, time series modeller, and HTML visualisation on the WEB intranet server (Figure 2). The time series modeller procedure estimates exponential smoothing, univariate Autoregressive Integrated Moving Average (ARIMA), and multivariate ARIMA models for time series, and produces forecasts. The procedure includes an expert modeller that attempts to identify and estimate the best-fitting ARIMA or exponential smoothing model for one or more dependent variable series automatically, thus eliminating the need to identify an appropriate model through trial and error.

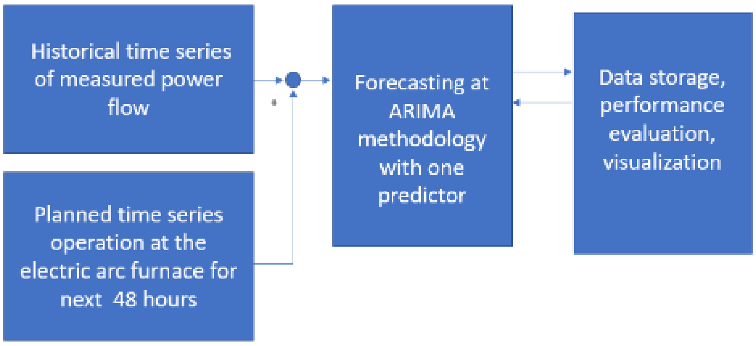

Figure 2: Forecasting information structure

Statistics. Goodness-of-fit measures: stationary R-square, R-square, Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), Maximum Absolute Error (MaxAE), Maximum Absolute Percentage Error (MaxAPE), normalised Bayesian information Criterion (BIC). Residuals: autocorrelation function, partial autocorrelation function, Ljung-Box Q. For ARIMA models: ARIMA orders for dependent variables, transfer function orders for independent variables, and outlier estimates. Also, smoothing parameter estimates for exponential smoothing models.

The Mean Absolute Percentage Error (MAPE), is a measure of prediction accuracy of a forecasting method in statistics. It usually expresses accuracy as a percentage, and is defined by the formula:

$$MAPE = \frac{100}{n}\sum_{i=1}^{n}\left|\frac{Ai-Fi}{Ai}\right|$$

where Ai is the actual value and Fi is the forecast value. Problems can occur when calculating the MAPE value with a series of small denominators. A singularity problem of the form 'one divided by zero' and/or the creation of very large changes in the Absolute Percentage Error can occur, caused by a small deviation in error.

The Root-Mean-Square Error (RMSE) is a frequently used measure of the differences between values predicted by a model and the values observed. The RMSE represents the sample Standard Deviation of the differences between predicted values and observed values. These individual differences are called residuals when the calculations are performed over the data sample that was used for estimation, and are called prediction errors when computed out-of-sample. RMSE is a measure of accuracy to compare forecasting errors of different models for data and not between datasets, as it is scale-dependent. RMSE is sensitive to outliers.

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(Ai-Fi)^2}{n}}$$

R squared in statistics, the coefficient of determination, denoted R^2, is the proportion of the variance in the predictable variable Fi that is from the actual value Ai:

$$SS_{tot} = \sum_{i=1}^{n}(Ai-\overline{Ai})^2 \qquad SS_{res} = \sum_{i=1}^{n}(Ai-Fi)^2 = \sum_{i=1}^{n}e_i^2$$

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}}$$

2. Forecasting of power flow extremes for 7 days

The first forecasting power flows task predicted of hourly maximum (minimum) at 7 days. An expert modeller has two inputs: Power flow time series and independent variable (predictor). The predictor was the total number of fillings at the electric arc furnace on working days. The inputs to the forecasting expert modeller were calculated easily using aggregate functions from the basic time series (Figure 3). For the predictor, we used the total number of fillings at the electric arc furnace during a day, that is, between midnight and midnight. The administrative production plans were not adequate for the predictor, as they do not take into account the actual production time, which is usually at night, or continuously for 24 hours over the weekend. Therefore, the predictor was calculated from the measured power flows at the common energy connection.

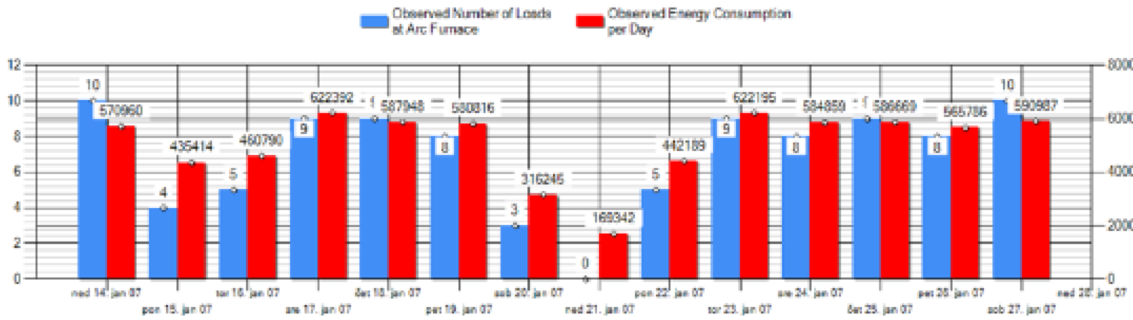

Figure 3: Observed day max and predictor time series

Table 1: The power flow forecast for the day hour maximum

Day Plus	R^2	Št. napovedi	AVG meritev	STDEV meritev	AVG napovedi	STDEV napovedi	MAPE	MAE	RMSE
+ 5	0,9	220	465283,7	144625,5	466093,3	127076,2	13.7 %	36185,9	47861,64
+ 4	0,9	220	464618,6	144370	465873,3	127780	13.7 %	36249,2	48236,38
+ 3	0,9	220	464847	144030,4	468237,4	128608,9	14.1 %	36687,9	49321,13
+ 2	0,9	220	464743,2	144476,5	468588,5	131015,7	13.6 %	35612,9	48418,05
+ 1	0,9	220	465477,1	144846,2	466832,2	136577,5	11.4 %	32210,9	42786,6

The forecasting simulation for the 220 real cases, hour maximum per day, give as a statistic $R^2 = 0.9$. That means that 90% of all of the variance in power flow extremes can be explained. Statistically good prediction results are obtained at high power flows. For small values the MAPE statistic was poorer, due to the nature of MAPE valuation - divide by low value. It should be clarified that the predicted values and the measured values have practically the same mean value and the same Standard Deviation, this is a good indication that the measurements and forecast are very close. (Table 1).

3. Daily load forecasting

The next task was forecasting daily power flows in an hourly step. In reality, the second task is more challenging than the first, since here we have to predict at least 24 steps (better 48 steps), and the prediction error increases with the number of prediction steps. Also, in this task, the expert modeller has two inputs: Power flow time series and predictor time series. We used calculated hourly averages or sum. The total number of fillings at the arc furnace in a working day is not a good predictor for the second task, and the exact time of filling in the arc furnace is still required. The first forecasting try was calculated with a simple predictor, which used the 1-hour step predictor estimated from the hour time series. The predictor has a value of 1 if the hourly average power consumption is > 4,500, and 0 if the hourly average power consumption is < 4,500. In principle, the formula for the calculation is the same as in the case before, except that we added time to the predictor. This predictor was 1 throughout the operation of the electric arc furnace, that is, it did not follow the typical three phase production cycle at the arc furnace.

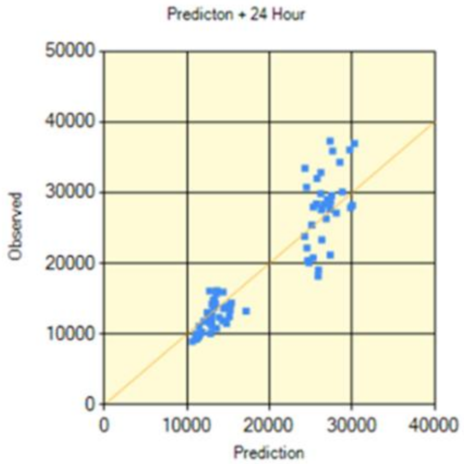

Figure 4: Daily load forecasting with simple predictor

Statistically the results of the forecasts, due to the simple predictor, results in the MAPE statistic 13% and R^2 0.84. The prediction/observed chart shows that such a prediction is not applicable, as the real forecasts, wail points, are hosted in two unrelated clouds.

3.1 Daily load forecasting with low prediction error

3.1.1 Independent variable - predictor

We assumed that the main culprit for the poor prediction was the overly simple predictor, which describes the events surrounding the arc furnace poorly, since the time series modeller was not obvious. In consultation with energy experts in the field, we realised that it is necessary first to understand the technological process of smelting and related electricity consumption. The furnace melting studies gave us the main ideas and guidelines on how to design a predictor to make it easy to put into practice, and to describe the energy developments in an electric arc furnace sufficiently well. The arc furnace and related technological processes are the dominant consumers of electricity. The melting process [1] is always carried out in an arc furnace with reduced voltage, since the conditions for burning the arc are poor in the cold cartridge; the ignition of the arc is carried out in such a way that the graphite electrode is lowered to the cartridge until it touches it, and until contact is reached with the other electrodes with the cartridge. At the discharge of the electrode, then an electric arc is triggered - like the firing of the arc during manual arc welding. Because of this, the current size changes from the short-circuit current through the rated power to the zero current at the end of the arc. We say that the arc furnace is operating restlessly at the beginning of melting. Due to the formation of the first melt at the bottom of the furnace, the conditions for burning the arc are improved due to good ionisation conditions, so we increase the voltage of the arc gradually and the power of melting to the full power: This is always the largest when melting the cartridge when there is already a melt on the bottom of the furnace. We say that we are melting with a hidden arc, which radiates at full power in the crater, which the boulder has drilled into the plunged insert of old iron. In the further heating of the melt or in maintaining its temperature, the power of the furnace is significantly lower. The characteristics of the electric arc must be different in this situation, since the arc can now freeze to the walls and the furnace vane.

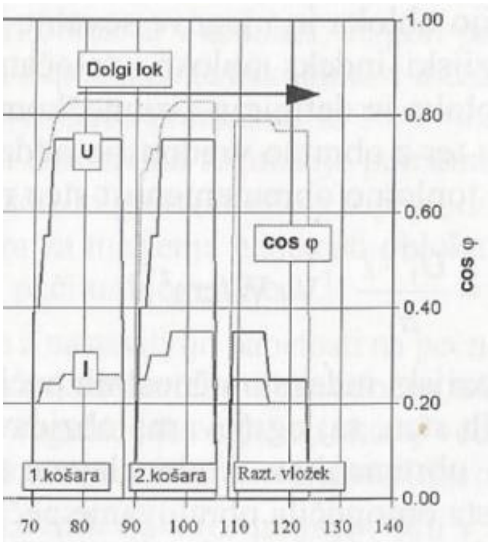

Figure 5: Power flow during melting process [1]

Old iron smelting takes place in three phases, which means three jumps and falls in energy consumption (Figure 5). In fact, these ups and downs in geometric shape are constantly the same, with amplitudes and duration changing slightly. These changes are within reasonable limits, so the principle maintains. The course depends on the composition of the old iron melting in the arc furnace and what kind of steel we want as a final product.

We can’t construct an exact predictor for each batch, since the energy consumption profile during and after melting depends on several parameters, and each batch is unique in principle. We decided on a reasonably simplified universal form of the predictor with a logical amplitude of 0 and 1. The x axis has a 15 min time step. In the phase of testing the adequacy of the predictor for a good forecast, we did not activate the Production Planning Department in the ironworks, but, in the development phase, the predictor was calculated from the energy measurements at the main energy connection. As we said, we did this on a 15 min time axis, and the amplitude was determined by the if statement: The value is 1 if the consumption is greater than 6,000 and 0 if the consumption is less than 6,000. The predictor step line in the chart has a grey colour. The step lines are characterised by the ups and downs of consumption for the smelting technology of the individual batch (Figure 6). An alert reader with a sense of geometry can already be convinced by comparing power flow during steel production and the predictor for forecasting with low prediction error on the proper construction of the predictor.

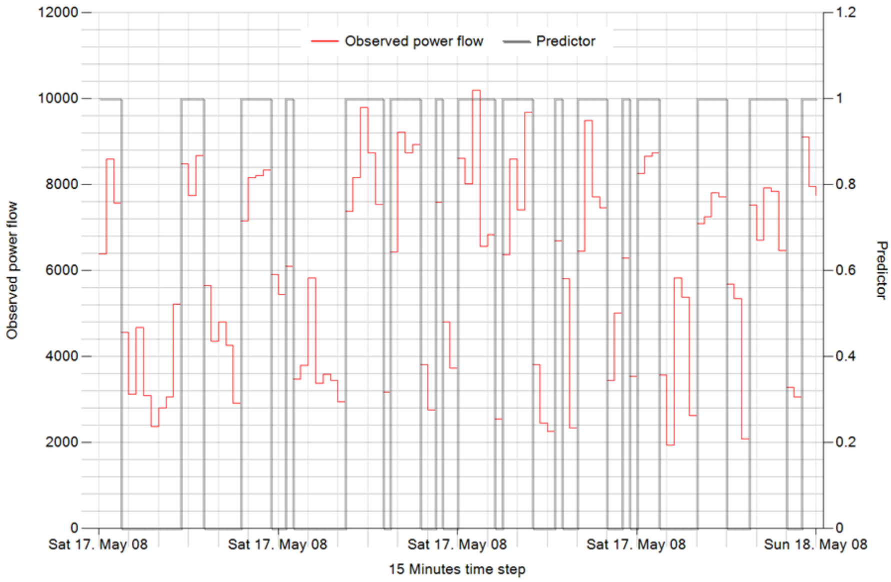

Figure 6: Predictor for forecasting with low prediction error - 15 min time step

With the periodogram we checked the shape of the predictor and whether we really captured the dominant influence of the electric oven. The spectral plots were used to identify periodic behaviour in the time series. Instead of analysing the variation from one time point to the next, it contrasts the variation of the series as a whole into periodic components of different frequencies. Smooth series have stronger periodic components at low frequencies; random variation ("white noise") spreads the component strength over all frequencies. From the spectral figures, we can see that we have a good predictor, since it has a frequency spectrum similar to the frequency spectrum of the time series of common power flows. There is a slight difference in the shape of the curve, the frequency curve of the measurements is only shifted more along the y axis, which is understandable, since the predictor runs in amplitudes from 0 to 1 and the power flow measurements have a maximum amplitude of 40,000.

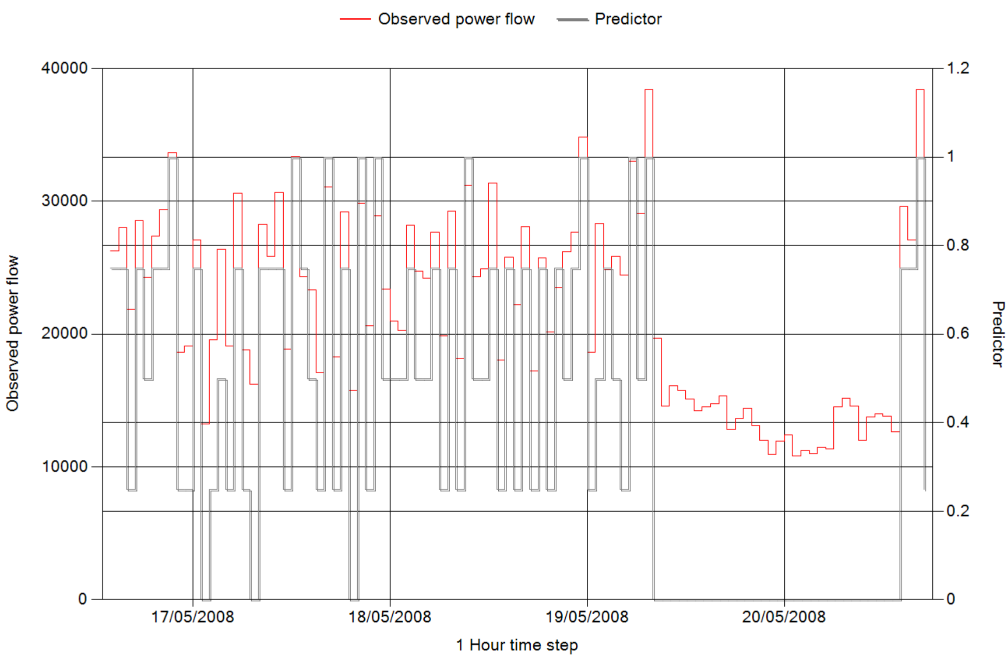

Figure 7: Final predictor – hour time step

3.1.2 Forecasting model

As we made predictions on an hour scale, the predictor had to be calculated into an hour scale, with hourly average aggregate function. Thus, we get a new predictor on the hour scale, which now had four values: 0.25, 0.5, 0.75, 1 (Figure 7). This was followed by the computation of a forecast model with hourly power flow time series and a new predictor. The model for our example is the linear standard ARIMA with notation (0,1,1) (1,0,1), where 0 is the order of autoregression, 1 is the order of differentiation, and 1 is the order of moving-average, and (1,0,1) are their seasonal counterparts. Model statistics are very good, approaching the magic limit to predict power flows with 95% probability

The model can be estimated with statistics, or we can draw a chart. Model statistics are calculated for a 1-step virtual prediction, and, of course, for the same time series from which the model originated. Stationary R-squared explains that there is a seasonality in the measurement data, which the forecast model did not explain properly. The R-squared value tells us that the variability of the time series is well covered by the prediction model. Statistically, the prediction model misses the measurement data at 9%. The model's residual does not have autocorrelation, so the autocorrelation of time series is captured adequately in the model. The time series from which the model was created was selected without outliers (Table 2). The "observed/model fit prediction" dot chart fits in well with the ideal 45 degree line. It is encouraging that the points clouds in the chart are continuous, and that the points cloud is a narrow ellipse. The narrow cloud of the ellipse is due to the good correlation between the model and the measurements (Figure 8). A continuous cloud means that the 4-level predictor is sufficiently segmented, since, for the 2-level predictor, we had a point cloud with a torn ellipse (Chapter 2). The statistics of the forecasting model are usually better than the statistics of the real forecasting itself. We want to predict 48-steps with an hourly step, so we are interested in the power flow over the next 48 hours. Of course, the first step to the final solution is a good model; a realistic prediction must be verified by the next phase, that is, by simulating the prediction on the real measurement data.

Table 2: The prediction model fit statistic

Model	Number of Predictors	Stationary R-squared	R-squared	RMSE	MAPE	MAE	MaxAPE	MaxAE	Ljung-Box Q(18) Statistics	DF	Sig.	Number of Outliers
Poraba-Model_1	1	.914	.947	2081	9.217	1579.641	285.055	9948	21.905	15	.110	0

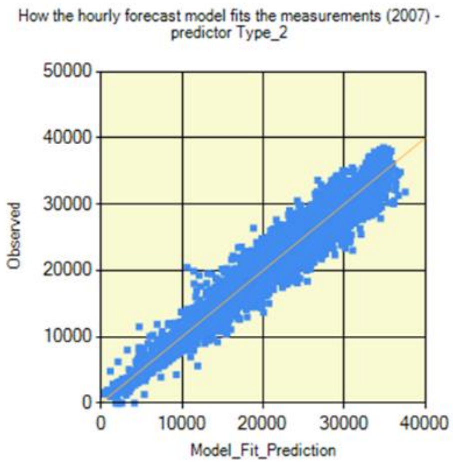

Figure 8: Model fit for forecasting with low prediction error

3.1.3 Procedures for improving forecasts

To simulate the forecast, we can use a fixed model as we have just calculated. However, it can be recalculated before each prediction, which means that we adjust to any changes in the time series from which we learn. The learning time series has a constant length, which means moving the start and end of the time series one step to the right. The 1-step simulation prediction statistics are very close to the model statistics. Forecasts for 24 hours already have a 15% prediction error. However, the statistics of the 48-step simulation forecast were even worse. The relative error started approaching 19%. Statistics are a professional numerical indicator of the quality of forecasts, but the observed/prediction graphs are also very illustrative, and are an excellent tool for developing reflection and improving forecasts. A careful observation of the 48-step forecast chart makes it clear that the forecast is (statistically) poor, due mainly to the forecast of low power consumption values (Figure 9). The question arises how to improve the prediction, as it seems that we have chosen the predictor well, and also the expert modeller has determined the parameters of the prediction model optimally. We will explain how to improve our forecasts below.

When we looked at the annual time series of consumption on the total energy connection, we found that the time series was not monotonous. For the first time, we noticed extremely low spending on national and religious holidays, which means that there are time series outliers. Another example of deviation from monotony is the maintenance of an electric arc furnace. During the maintenance of the arc furnace, also, the technological lines related to the melting of iron do not operate. At that time, consumption on the main energy connection follows the characteristic of the seasonal multiplicative Winters model 7/5.

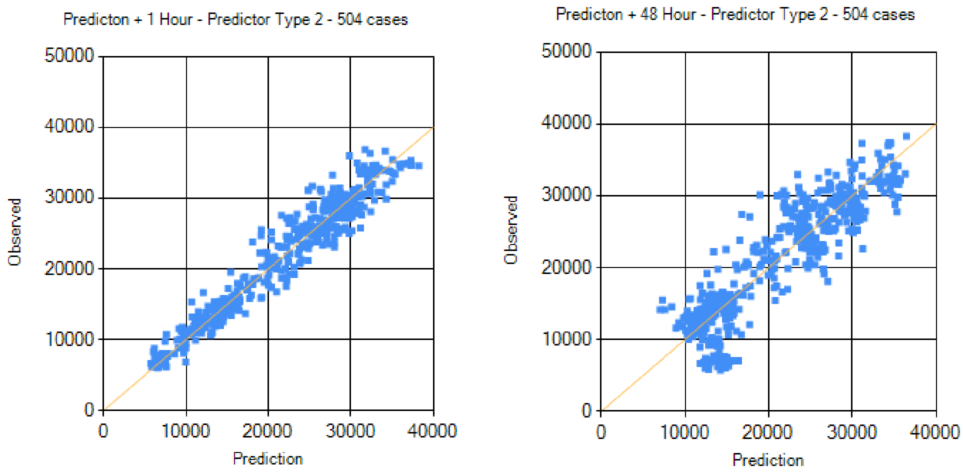

Figure 9: Forecasting simulation, next step to final prediction model

We don't need a predictor to predict with the Winters seasonal model. The Winters' observed/prediction chart is almost a 45 degree line, which means that the model does the prediction very well, in other words, we can predict this fraction of low power flows for a long time (Figure 10). The magnitude of the power flows is between 5,000 and 10,000, and is in amplitude of the problematic area of poor forecasting, as shown by the 48 hour prediction chart at full arc furnace operation (Figure 9). The second conclusion is that the season weeks are repeated throughout the year, whether the arc furnace is or is not in operation. When the arc furnace is in operation, the seasonal power flow 7/5 is added to the dominant power flow caused by the arc furnace technology part.

So far, we have explained that the power flow at the common energy connection can be explained by ARIMA with a predictor and the Winters multiplicative seasonal model. The Winters model explains power flow exactly where ARIMA fails.

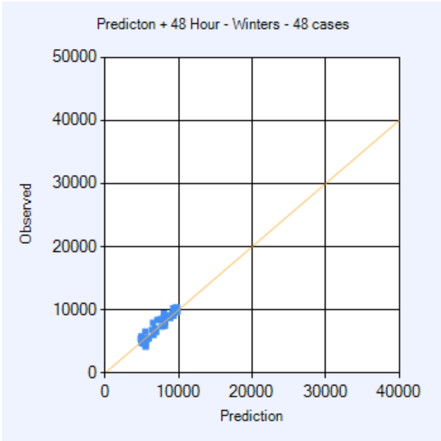

Figure 10: Load forecasting with multiplicative Winters

We came to the idea of subtracting the seasonal data covered by Winters from the total power flow, thus allowing the ARIMA model with the predictor to do its part of the forecast perfectly. When the ARIMA model predicted for 48 hours, we finally add seasonal "Winters data" back to forecast of total consumption. The correlation between the forecast with Winters and the associated weekly measurement data is very strong. In order to perform the described subtraction and summation, we had to determine a "standardised weekly seasonal power flow". We were able to do this because the Winters model is practically the same over a long period, and has a low forecast error. That is how we came to the final forecasting with a low prediction error. The forecasts were made by an expert modeller, and the last described transactions with the data were performed within the framework of a real-time database, using the SQL scripting language.

4 Results - power flow forecasting with low prediction error

Branch A represents a prediction of the power flow from the public grid, the total energy consumption, and has the result: for the next 24 hours: $R^2_{24} = 0.93$ $MAPE_{24} = 9.9\%$; for the next 48 hours: $R^2_{48} = 0.91$ $MAPE_{48} = 12.3\%$ $RMSE_{48} = 2584$. The solution was tested by simulating a forecast of 21 days. We predicted each time point 48 times, first predicting 48 steps away, then 47 steps away, etc. The last prediction is one step ahead of the time point. Thus, for 21 days, we made 504 forecasts (Figure 11, Table 3).

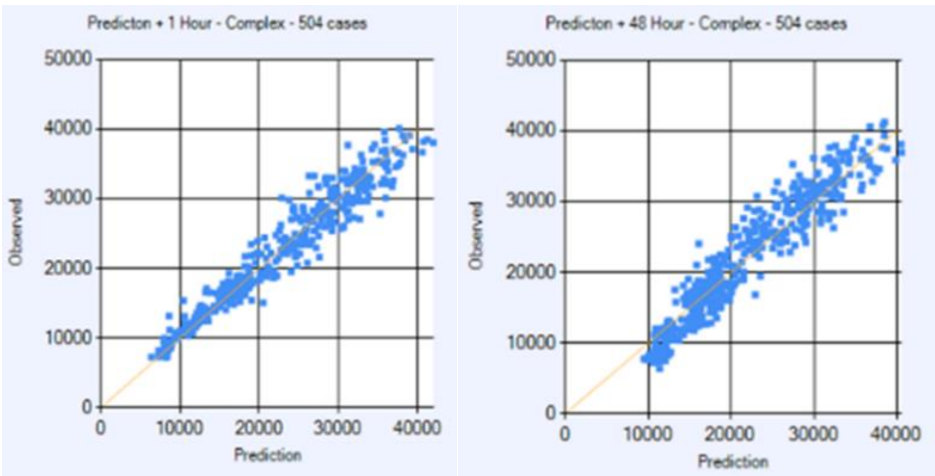

Figure 11: Final result forecasting with low prediction error

Branch B represents the power flow prediction to company 1, the largest consumer of power, with a typical coefficient of determination $R^2_{48} = 0.93$. Branch C represents the flow to all other businesses with $MAPE_{48} = 7\%$. The project timeline can only be carried out in close cooperation with experts on site. A good real-time forecasting system for business can come about with steps to verify and implementation. Verification of the forecasts` model from a data set of the last year. Determine how a steel mill is capable to provide a predictor. Estimate of financial savings using predicted daily power flows in the daily market, virtual trading for 3 months. Definition of the trading worst cases. Finally, signing a software lease and software maintenance contract.

Table 3: Final forecasting statistic

Hour +	R square	MAE	MAPE	RMSE	Number of predictions
1	0,95	1386,93	6.81 %	1965,98	504
2	0,94	1455,65	7.33 %	2060,92	504
3	0,94	1534,3	7.8 %	2106,36	504
4	0,94	1616,43	8.26 %	2149,73	504
5	0,93	1675,69	8.78 %	2224,52	504
6	0,93	1711,12	9.08 %	2238,19	504
7	0,93	1748,47	9.28 %	2268,13	504
8	0,93	1756,52	9.51 %	2280,33	504
9	0,93	1759,68	9.47 %	2280,7	504
10	0,93	1789,77	9.76 %	2317,68	504
11	0,93	1828,01	9.83 %	2360,4	504
12	0,93	1837,03	9.91 %	2364,44	504
13	0,93	1824,9	9.89 %	2330,4	504
14	0,92	1855,75	10.02 %	2387,83	504
15	0,93	1816,1	9.87 %	2330,17	504
16	0,93	1845,7	10.03 %	2352,68	504
17	0,93	1795,34	9.73 %	2276,53	504
18	0,93	1794,97	9.77 %	2286,87	504
19	0,93	1813,72	9.91 %	2295,4	504
20	0,93	1814,17	9.77 %	2322,71	504
21	0,93	1758,78	9.7 %	2262,13	504
22	0,93	1781,22	9.81 %	2309,08	504
23	0,93	1780,48	9.86 %	2296,16	504
24	0,93	1786,35	9.99 %	2298,05	504

References:

[1] JANEZ BRATINA, ELEKTROOBLOČNA PEČ, RAVNE NA KOROŠKEM, 1994.

[2] DOUGLAS C. MONTGOMERY, GEORGE C. RUNGER, APPLIED STATISTIC AND PROBABILITY FOR ENGINEERS, WILEY, 2003.

[3] GEORGE E. P. BOX GWILYM M. JENKINS GREGORY C. REINSEL. TIME SERIES ANALYSIS FORECASTING AND CONTROL, WILEY, 2014.