

# DAILY LOAD FORECASTING AN INDUSTRIAL COMPLEX

**Boris BIZJAK<sup>1</sup>**

## Introduction

The aim of the forecast is to predict the load profile as accurately as possible, which means minimizing the cost of electricity an the industrial complex. There are load forecasting applications at an industrial plant with an electric arc furnace in the City of Ravne, Slovenia. In this paper we discuss the daily hourly and daily 15-min load forecasting of load for the industrial complex using ARIMA methodology with one predictor. For predicting the daily load, we need to make a forecast of +24 steps in the case of the 1-hour prediction interval and for the announcement 24 hours forward, we should perform +96 steps of the predictions in the case of the 15-minutes prediction interval.

The industrial complex consists of 20 companies and they have a common energy connection. From the 20 companies, the largest electricity consumer is a Steel Mill. The presented technological structure of the Steel Mill production had strong influence of the daily power flow profile at the common electric terminal. It followed a presentation of the electric arc furnace power flow during steel production, since it is the largest consumer of electricity at a Steel Mill. The story of predicting in the article is due to trading in electricity. In Slovenia, it is currently traded in hours' time intervals and not at 15-minute time intervals. In future it is expected to switch to 15-minute trading intervals. This is also followed by chapters in the paper: First, daily load forecasting with a 1-hour step, followed by a prediction of a daily load with a 15-minute step. The basis for the forecasting is the time series of energy consumption data, which is a time series from 15 min values.

For the described type of industry complex we always started the forecasting projects with linear regression. This is a robust method, but one that has a weakness in the limited level of a confidence interval in the forecast. In the case of the 1-hour prediction step, by linear regression, the independent variable of production explains 74% of the variance in the load, which is highly significant, and the F-test says we can trust  $\beta_0$  and  $\beta_1 > 99.9\%$ . Additionally, energy consumption at the industrial plant had a seasonal day, hours feature, or seasonal hours, minute feature. Therefore, we improved forecasting performance - statistical indicators, using ARIMA methodology by approximately 15%. The ARIMA model, with one predictor, explains 85% ( $R^2$  0.85) of the variance in the load, with MAPE 13%.

In the case the daily load forecasting at 15-minutes of forecasting interval, we must change the predictor from the type 1 to type 2. Next is a simple linear regression, in the case of the 15-min prediction step, predicted at MAPE 36 % and  $R^2$  is 0.79 for a one-day prediction. Finally forecasting model was the ARIMA methodology with one dominant predictor. The forecasting result was excellent: MAPE 17 % and  $R^2$  0.9 for "one day plus" at the 15-min step.

The presented load forecasting model needs at the learning phase two input historic time series: Energy consumption and real production data. We get production data (predictor) from the Production Planning Department. We wanted to use a simple and robust predictor, wich not only has a theoretical value for forecasting, but also a practical utility to be used in practice. So, the predictor has a value of 1 when the furnace is running and 0 when the furnace is not in operation. For predictor type 1, the time resolution at the predictor is 1 hour, and for predictor type 2 15 min. The predictor also follows production with different dynamics. Type 1 has a constant value of 1 for all the duration interval of the arc furnace working (for example, for 8 hours), and predictor type 2 follows the arc furnace cycle more dynamically with fluctuation between 1 and 0.

## INDUSTRIAL KOMPLEX AND STEEL PRODUCTION

The beginnings of the ironworks in Koroška go back to the year 1620 [[https://sl.wikipedia.org/wiki/Zelezarna\\_Ravne](https://sl.wikipedia.org/wiki/Zelezarna_Ravne)]. After 1774, the first forges and nails' production started to operate along the river Meža and, thus, the expansion of industrial iron production. Železarna Ravne

is a common name for several companies located at the City of Ravne, Slovenia, which once featured under this name. In 1992, several production and service companies emerged in Ravne from the single company Železarna Ravne. Today, 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 [<http://sij.metalravne.com/sl/>] with approximately 900 employees. Metal Ravne is the largest, and the largest consumer of electricity. The company consists of a steelwork, a rolling mill and electro-smelting under a 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.

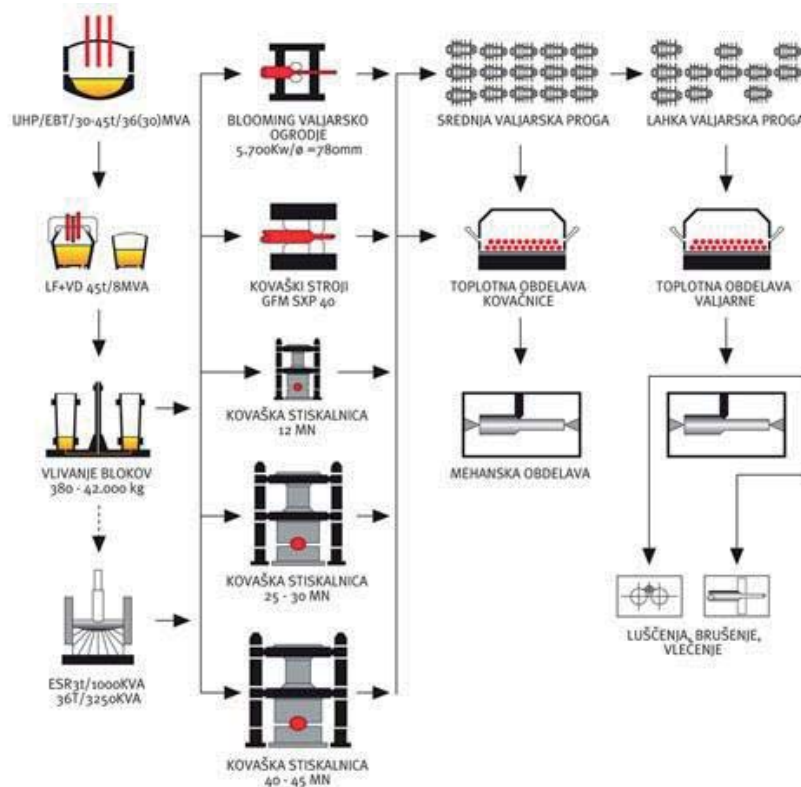


Figure 1: Metal Ravne technological scheme [<http://sij.metalravne.com/sl/>].

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 ionization 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.

For today's long-arched furnace and water-cooled panel, walls and water-cooled vaults are characterized by high transformer secondary voltages and low electrode currents and working  $\cos \Phi$  0.82. The power flow at the time of steel production is shown in Figure 2.

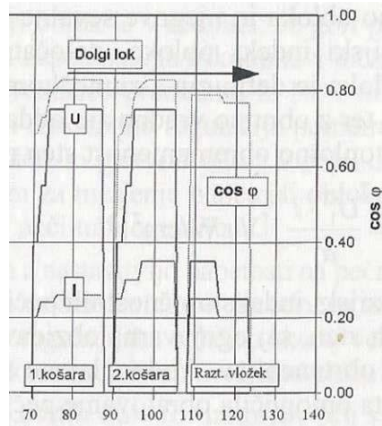


Figure 2: Power flow during steel production [1].

## MARKET AND TRADE WITH ELECTRICITY

By opening the electricity market, electricity became a commodity. Since July 2007, the market has been completely open in Slovenia, which means that all customers can choose their electricity supplier freely. Electricity market participants are producers, traders and suppliers that supply electricity to customers. Electricity is transmitted from power plants to customers via transmission and distribution networks, for which the system network operators are responsible.

The electricity market is split into the wholesale market and the retail market. Wholesale contracts are traded by closed-off contracts, and on the retail market by open contracts. Closed contracts on the supply of electricity are those for which the vendor and the buyer agree on the exact quantities of electricity that are available at a given time interval. Payment in this case is made based only on these arrangements. This also applies when the actual quantities of electricity supplied are different from those specified in closed contracts. Closed contracts are the basis for the compilation of timetables for balance groups, therefore any deviation of quantities delivered from those agreed in closed contracts is subject to the calculation of deviations in accordance with the rules.

The wholesale market participants are electricity producers and traders who trade in electricity for resale. As a rule, wholesale customers do not participate in the wholesale market, as it would be difficult to reconcile their predictable outflow with open contracts. For this reconciliation, the supplier shall arrange for the final customer, who concludes such closed contracts on the wholesale market for different periods of time, enabling him to cover the consumption of final customers with minimal deviations from the schedules. The wholesale market also covers all forms of cross-border trade in electricity, as contracts on cross-border electricity supplies are always closed. Wholesale trading takes place in two forms, i.e. on a regulated market (Stock Exchange) and bilaterally. In Slovenia, most wholesale deals are bilateral, that is, with the agreement between the seller and the buyer. Indices of prices achieved on the electricity exchange operating in the area are usually used as the price criterion on the wholesale market in each area or country, indices of prices achieved on the electricity exchange operating in the area are usually used.

## FORECASTING

We cannot save electricity, so the forecast of consumption, production and losses is very important. With well-planned and well-made predictions, we can save the costs of balancing in the electricity market. Electricity forecasting can be divided into three categories: Short-term forecasting (from one hour to one week), medium-term forecasting (from one week to one year), long-term forecasting (more

than one year). Each measuring site, equipped with 15-minute measurements, can choose electricity products on the electricity market, which are basically divided into belt, trapezoid and night energy, separately for workdays, weekends and holidays. Considering historical measurements and customer forecasts, a precisely calculated price can be obtained. By doing so, we can plan costs accurately, and we also can simulate the cost, depending on the ability to adjust the consumption of consumption in hours when energy is the cheapest on the market. We have the choice of the following products offered: Band energy, trapezoidal energy, night energy and hourly energy.

Irrespective of the changed economic conditions and the various external influences, electricity consumption in Slovenia has been constantly increasing over the years. Predicting future power consumption and load is a complex and far from a simple task. The fact is that in any case it is not possible to capture and mathematically process all the factors that influence the future use of energy and loads. These sizes and their laws can be determined only in simplified statistical form. Various statistics are used to evaluate the performance of models of forecasts, simulations and forecasts. We decided to use in the article only MAPE, MAE, RMSE and  $R^2$  to comparison between individual solutions.

**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:

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

where  $A_i$  is the actual value and  $F_i$  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.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (A_i - F_i)^2}{n}}$$

**R squared** in statistics, the coefficient of determination, denoted  $R^2$ , is the proportion of the variance in the predictable variable  $F_i$  that is from the actual value  $A_i$ :

$$SS_{tot} = \sum_{i=1}^n (A_i - \bar{A})^2 \quad SS_{res} = \sum_{i=1}^n (A_i - F_i)^2 = \sum_{i=1}^n e_i^2$$

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

## FORECASTING DAILY POWER CONSUMPTION PER HOUR INTERVAL

Due to the way of trading, the forecast of the daily flow of electricity flows requires a forecast for at least 24 hours, even better in 48 steps. We want to predict power flow on the common energy supply of the industrial complex Ravne. The structure of total energy consumption for the Ravne industrial complex is the sum of 20 different companies. The biggest 75% of the energy consumers are the Metal Ravne Steel Mill. We presented the technological concept of the Metal Ravne Steel Mill in the introduction. The

dominant energy consumer in the steel industry is the UHP electric arc furnace, the second largest consumer is the oven for overheating LF + VD and then the other technological line of the steel mill, which follows the melting process of old iron. The remaining 25% of consumption, on a common energy connection, is represented by other companies, with a typical spending profile of 7 days a week, of which 5 are working days. Change of 25% of consumption over a day is relatively low, and therefore it can be predicted very well. We managed to verify this with a separate prediction when the steel works were under repair for a month and only other companies were operating.

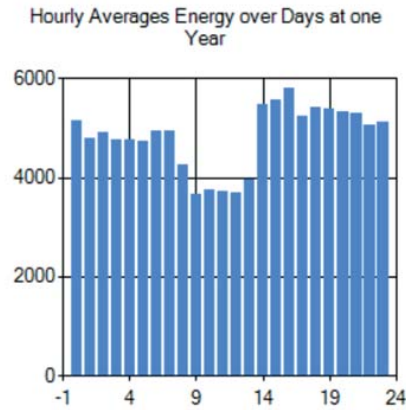


Figure 3: Daily load curves.

The steel plant does not operate statistically on a random basis, but over the day according to the orders and the price of electricity. Since its consumption is 75% of the entire energy of the industrial complex, it has the greatest influence on the consumption profile on the total energy supply of the industrial complex. The issue is the production time; real production will be moved to the late afternoon, next to the evening (cheap electricity), over midnight, until 9:00 AM the next day (Figure 3). Occasionally, when there is a large volume of orders, the steel plant operates for 24 hours continuously, regardless of the daily price of the electric power unit over a working day. Due to cheap electricity, Saturdays and Sundays also operate for 24 hours. Once a year, the steel works carry out a repair that lasts about a month. Of course, the electric arc furnace does not work then.

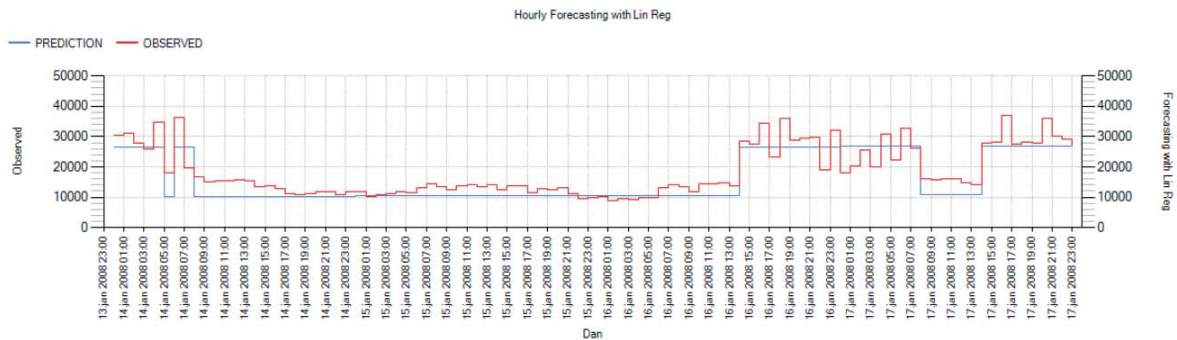


Figure 4: Forecasting daily power consumption per hour with linear regression - predictor type 1.

In the time series of electricity consumption, the outliers are also noticed; these are values when the consumption of electricity escapes from the standard Gaussian distribution and is close to zero. Such values are expressed on May 1 and December 31. These two cases were not addressed specifically in the forecast itself, but, in any case, they have a negative impact on forecasting performance statistics.

For the described type of industry complex we started forecasting continuously with linear regression. This is a robust method that has a weakness in the limited level of confidence in the forecast. In the case of the 1-hour prediction step, with linear regression. The independent variable of production data



at the arc furnace now explains 74% of the variance in the load, which is highly significant, and the F-test says we can trust  $\beta_0$  and  $\beta_1 > 99.9\%$  (Table 1).

We calculate the 1-hour step predictor from the 15-minute time series of consumed electricity using SQL aggregate functions. The predictor has a value of "1" if the hourly average power consumption is  $> 4500$  and "0" if the hourly average power consumption is  $< 4500$ . The predictor's time is coincident with the progress of the blue step line in Figure 4. The forecast itself has a minimum value of 10,493, which means small hourly consumption is a big MAPE error (low value error), which is a weakness in MAPE statistics. Therefore, the values shown in Table 1 are limited to hourly power consumption  $> 3500$ .

MAPE	MAE	RMSE	R square	Number of predictions	AVG observed	STDEV observed	AVG prediction	STDEV prediction
24.58 %	3667,01	4346,34	0,74	8543	19694,7	8594,4	19285,4	8059,7

Table 1 : Observed and prediction at linear regression - predictor type 1.

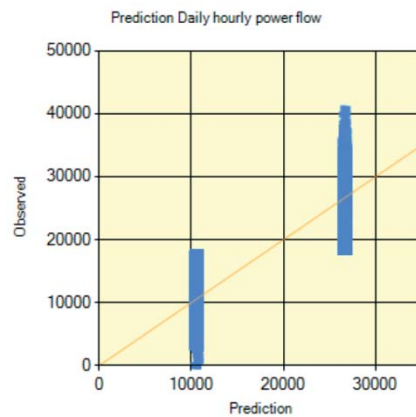


Figure 5: Daily load forecasting at linear regression - predictor type 1

Slight differences between the AVG observed and AVG predictions (Table 1) mean that the regression line using the least squares method was calculated as needed.

$$[\text{Load forecasting}] = 10,493 + 16,203 \cdot [\text{Predictor\_type\_1}] + \varepsilon$$

Figure 5 show a typical chart of a prediction with linear regression. A forecast with linear regression could be improved by change the predictor's timing or amplitude, which means that the on / off principle would be more often adapted to actual production or that the amplitude would follow a production value with not only 0 or 1, but, for example, "0.9", "1", "1.1". We did not do this because we believe that such a predictor cannot be expected from the Production Planning Department. MAPE 25%, R square 0.74 from linear regression will be improved with ARIMA metrology and one predictor.

Forecasting model design is the process where we start with a coarse model and, through more iterations and reflections, get closer to the best fit model. The learning time interval from the observation time series, which the ARIMA forecasting model parameter raises  $(\mu, \varphi_1, \varphi_2, \dots, \varphi_p, \delta_a^2$  and  $\mu, \theta_1, \theta_2, \dots, \theta_q, \delta_a^2)$ , should not be too long or too short. Forecasting models are alive, because they learn permanent parameters from the measurement data. We learned the forecast model in sliding mode with 8,760 historical data (One Year). As an independent predictor, we used the observed number of loads at the electric arc furnace. Basic parameters for ARIMA are determined automatically by Machine Learning.

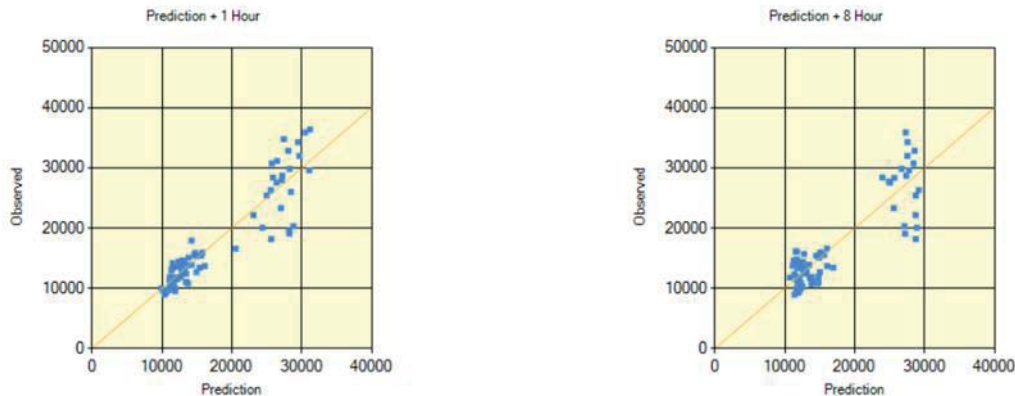
Energy consumption at the industrial plant had a seasonal day, hour (7/24) feature and one dominant prediction. The forecasting results in Table 2 show much better: MAPE 11% and R2 0.86 at "24 hour - first step" prediction at a half year of prediction. For "24 hour - 24 step" forecasting the MAPE statistic was 13% and R<sup>2</sup> 0.84.

The model for our example is the linear standard ARIMA with notation (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.

Prediction for a Hour	MAPE	MAE	RMSE	R square	Number of predictions	AVG observed	STDEV observed	AVG prediction	STDEV prediction
+ 23	13.35 %	2537,87	3385,74	0,84	80	19424,1	8414,2	19322,5	6859,6
+ 22	12.87 %	2449,85	3322,91	0,85	80	19280,4	8473,2	19118,3	6793,2
+ 21	13.07 %	2406,33	3199,62	0,85	80	18925,5	8301,8	18902,7	6731
+ 20	12.76 %	2341,5	3148,12	0,86	80	18853	8308,4	18738,3	6646,6
+ 19	13.07 %	2376,44	3248,57	0,85	80	18651	8281	18568,3	6571,4
+ 18	13.19 %	2401,27	3299,28	0,84	80	18502,7	8291,2	18396,8	6490,1
+ 17	13.53 %	2418,79	3335,39	0,84	80	18364,2	8347	18216	6399
+ 16	13.87 %	2436,3	3286,32	0,84	80	18109,8	8183	18065,5	6310,6
+ 15	14.29 %	2494,19	3341,5	0,83	80	17894,7	8142,7	17886,7	6236
+ 14	14.6 %	2559,37	3440,3	0,82	80	17726,3	8089,9	17722,2	6204,5
+ 13	14.82 %	2580,3	3446,63	0,81	80	17555,2	7979	17562,3	6188,4
+ 12	15.49 %	2669,45	3536,92	0,8	80	17377,5	7855,8	17404,4	6201,1
+ 11	15.73 %	2599,4	3383,8	0,8	80	17096	7588,6	17206,4	6277,4
+ 10	15.8 %	2634,04	3430,33	0,79	80	16949,6	7487,7	17029,2	6323,4
+ 9	16.06 %	2654,45	3411,6	0,79	80	16772,4	7390,7	16836,5	6397,1
+ 8	15.81 %	2600,55	3393,01	0,78	80	16658,5	7266,1	16707,5	6407,1
+ 7	15.41 %	2490,88	3262,55	0,78	80	16377,2	6908,8	16528,1	6394,8
+ 6	15.35 %	2426,79	3206,3	0,78	80	16247,8	6793,7	16512,8	6455,5
+ 5	15.21 %	2502,01	3418,03	0,76	80	16393,6	7024,9	16538,2	6535,3
+ 4	14.73 %	2417,13	3308,55	0,78	80	16413,1	7034,3	16556,1	6536,5
+ 3	14.12 %	2385,65	3344,71	0,79	80	16677	7318,3	16717,8	6603
+ 2	13.57 %	2309,76	3303,4	0,8	80	16787,1	7396,8	16881,1	6696,5
+ 1	11.02 %	1951,3	2860,87	0,85	80	16924,8	7507,2	17026,4	6948,1
+ 0	11.28 %	2011,82	2903,37	0,86	80	17149,4	7653,9	17211,4	7002,3

Table 2: Forecasting daily power consumption per hour at ARIMA methodology - predictor type 1.

Linear regression (Table 1) has an RMSE constant value of 4,346 for all predictions. With ARIMA methodology, for the same historical time series, RMSE ranges between 2,903 and 3,385, which is a significant improvement (Table 2). So, the model ARIMA + predictor is a better prediction model. Looking at the 24-hour forecasts, we notice that better forecasts are for the values near 1+ and 24+ hours, which shows the 24-hour periodicity of our time series.



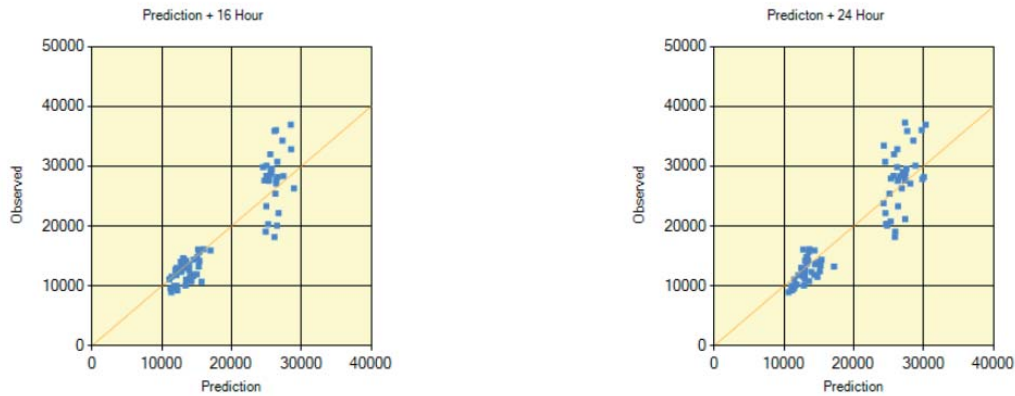


Figure 6: Forecasting daily power consumption per hour with ARIMA methodology - predictor type 1.

Figure 6 shows that we have improved the forecast for low power consumption dramatically. There is also an improvement in high consumption, but it's not as good as you would like. This type of forecast could be improved by improving the timing of the predictor, which means that the on/off principle would be more often adapted to actual production data. This new type of predictor was used at a 15-minute daily-day forecast. We named it predictor type 2.

## FORECASTING DAILY POWER CONSUMPTION PER 15 MINUTE INTERVALS

Predictor type 1 for forecasting daily power consumption per 15-minute interval is not an appropriate test with linear regression  $R^2 = 0.68$ , MAPE 26.4% and RMSE 1500. First, not because it has a time scale of 1 hour, and if we look more closely at Figure 8, it can be observed that predictor type 1 can also overtake or delay the start of the production process in an electric arc furnace for 30 minutes. Second, the spectral density of the for 15 minutes power consumption (Figure 7) has a different frequency characteristic than the frequency spectrum of 1-hour power consumption.

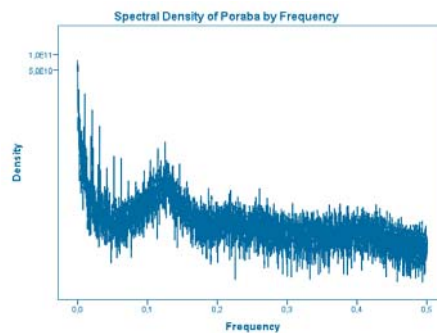


Figure 7: Spectral density for 15 minutes power flow at one year

The spectral density of the for 15 minutes power flow for one year indicates additional periodicity of the signal at frequency 0.12. Given the spectral density fluctuation for 1-hour power flow at one year, we have had monotonic falls from low frequencies to the highest frequencies. This is understandable, since 1-hour average power consumption is calculated from 15-minute measurements with using SQL aggregate functions SUM or AVG, that is in a principle low band pass filter. With another words: 1-hour time series have less frequency content. Therefore, we have introduced a new type 2 predictor with only two amplitude values: 0 and 1. Considering how to improve the forecasting, we had to delve into the production process Metal Ravne (Figure 1) and the operation of an electric arc furnace (Figure 2). Power flow in a UHP arc furnace during steel production is more or less constant. The difference between individual loads at the arc furnace is only in what happens after we have finished melting in an arc furnace. Sometimes overheating occurs and sometimes it does not, which means a short leap in the consumption of electricity to the maximum. Predictor type 2 is more time-divergent than type 1, so it is better to follow the operation of the electric arc furnace and other technological lines of the steel plant. For research we calculate 15 minutes step predictor from the 15-minute time series of consumed



electricity using SQL functions. The predictor has a value of "1" if the 15 minutes power consumption is > 6500 and "0" if the 15 minutes power is < 6500. It is still necessary to check how much such a predictor can be realized in the Production Planning Department. Next is simple linear regression, in the case of a 15 min prediction step, predicted at MAPE 36% and R<sup>2</sup> is 0.79 for one day prediction.

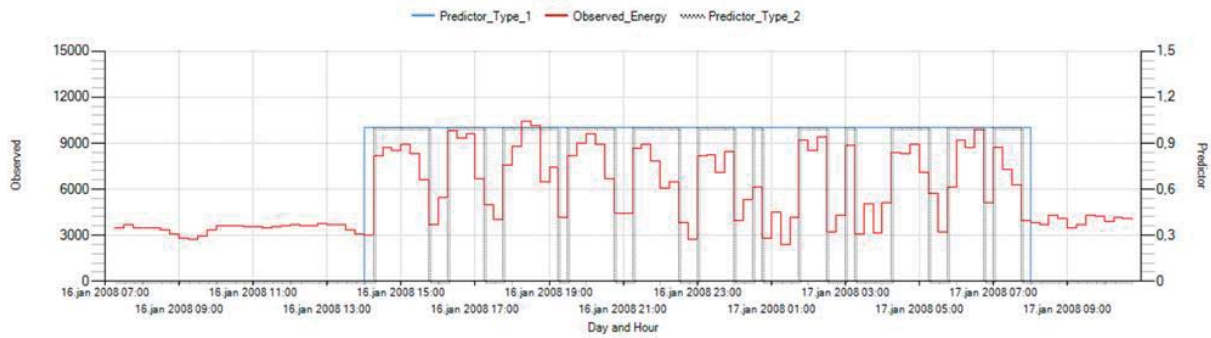


Figure 8: What is the difference between predictor type 1 and predictor type 2 – 15 min interval

For the described type of load profile of the electricity consumption, we started forecasting with linear regression. This is a robust method that has a weakness in the limited level of confidence in the forecast. In the case of the 15 minute prediction step, with linear regression, the independent variable of production data at the arc furnace now explains 79% of the variance in the load, which is highly significant, and the F-test says we can trust  $\beta_0$  and  $\beta_1 > 99.9\%$  (Table 3). With MAPE 36% we are not satisfied.

$$[\text{Load Forecasting}] = 3,130 + 5,023[\text{predictor\_type\_2}] + \varepsilon$$

R_square	MAE	stevilo_napovedi	AVG_meritev	STDEV_meritev	AVG_napovedi	STDEV_napovedi	MAPE	RMSE
0,79	985,9	34996	4824,9	2662,8	4817,8	2372,6	36.39 %	1219,9

Table 3: Observed and predictions at linear regression 15 - min interval - predictor type 2.

Finally, we used traditional ARIMA methodology with predictor type 2. The forecasting result was excellent: MAPE 17% and R<sup>2</sup> 0.9 for "one day plus" at 15-minute prediction steps. As previously described, RMSE is a measure of the quality of the forecast model for the same time series. ARIMA has this value the smallest (RMSE 903), so this model is the best. It may be explained that Figure 6 represents a simulation of true prediction, and Figure 9 represents a "prediction model fit". Experience from work tells us that the MAPE, MAE, RMSE, R<sup>2</sup> statistics for predictive simulation and the prediction model fit are very similar, but the charts'prediction/observed can be different.

R_square	MAE	Number of predictions	AVG observed	STDEV observed	AVG prediction	STDEV prediction	MAPE	RMSE
0,89	665,03	35092	4812,8	2669,2	4812	2494,8	17.16 %	902,5

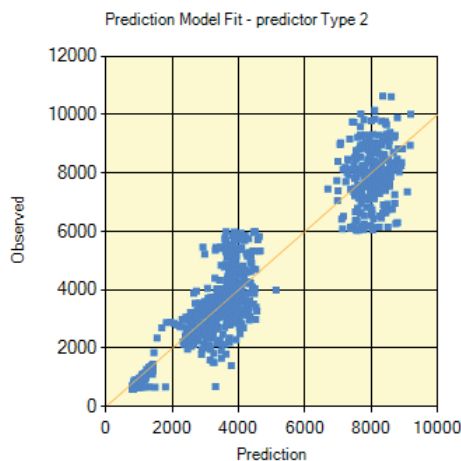


Figure 9: The daily load forecasting model fit at 15 minute interval ARIMA + predictor type 2

For research and development we used standard time series for the years 2007 and 2008, from the Company Petrol Energetika Slovenia.

## CONCLUSION

Production time at the arc furnace is dependent upon the types of input material etc. To learn and to improve a forecasting model we need exact production data. How do we get it? The answer is hidden in the industrial plant energy consumption time series and in discussion with the Production Planning Department. So we research and develop an advanced algorithm, like the forecasting pre-processor, which captures the best predictor data. The tariff policy of electricity has the potential to improve forecasts, for example, by introducing new predictors. Another option to perform prediction is with non-linear prediction models, as described in [2][5] and [6]. The purpose of load flow forecasting is to enhance the performance of the electricity trading, so it would make sense to continue to work towards the integration of forecasting with the tariff policy and the system trading of electricity.

## References

- [1] JANEZ BRATINA, ELEKTROOBLOČNA PEČ, RAVNE NA KOROŠKEM, 1994.
- [2] JAMIE MACLENNAN, ZHAO HUI TANG, BOGDAN CRIVAT. DATA MINING WITH MICROSOFT SQL SERVER 2008, WILEY.
- [3] DOUGLAS C. MONTGOMERY, GEORGE C. RUNGER, APPLIED STATISTIC AND PROBABILITY FOR ENGINEERS, WILEY, 2003.
- [4] GEORGE E. P. BOX GWILYM M. JENKINS GREGORY C. REINSEL. TIME SERIES ANALYSIS FORECASTING AND CONTROL, 2014, WILEY.
- [5] C.MEEK, D.M. CHICKERING, D.HECKERMAN, AUTOREGRESSIVE TREE MODELS FOR TIME-SERIES ANALYSIS, MICROSOFT RESEARCH.
- [6] BO THIESSON, DAVID MAXWELL CHICKERING, DAVID HECKERMAN, CHRISTOPHER MEEK, ARMA TIME-SERIES MODELING WITH GRAPHICAL MODELS, MICROSOFT RESEARCH