IJEEC

International Journal of Electrical Engineering and Computing
Vol. 5, No. 1 (2021)

Original research paper
UDC 621.311.1.016.2:004.896
DOI 10.7251/IJEEC2101015B

Power Flow Prediction for a Steel Plant with a Dynamic Regression Model

Boris Bizjak

University of Maribor, FERI, Koroška cesta 46, 2000 Maribor, Slovenia

E-mail: boris.bizjak@um.si

Abstract—**A power flow forecast it was shown for an industrial complex consisting of more than 20 different companies. The predominant consumer of electricity in the industrial complex is a steelworks company with an electric arc furnace. A steelworks with an electric arc furnace is a very specific example of an energy consumer. Other companies in the industrial complex are not connected to the steel plant technologically, but they are on the same energy connection. They have a weekly power flow profile significantly different from the steel plant. To calculate the forecast model and perform the forecast of power flows we need only two inputs of data: Historical measurements of power flows and the number of loads of the electric arc furnace in the following days. The first showed a prediction with linear regression. The next model to predict was the seasonal ARIMA model with a regressor, also called a dynamic regression model. The dynamic regression model improved the prediction by 15% compared to linear regression, according to the RMSE measure. This was followed by an improvement in the dynamic regression forecasting model by considering the seasonality 7/5 in the time series. We did this with a model with superimposed noise. With this model, we improved the forecasting by 30% to linear regression. Logically, the filter model of the prediction model also improved, gaining more Lag coefficients and losing a constant. Qualitatively, the result is a forecast of power flow for one month with prediction error MAPE 8% and measure R2 is 0.9.**

Keywords*- arc furnace; ARIMA with regressor; forecasting; model with superimposed noise; power flow; steel plant.*

I. 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 [1]. The goal is to make the forecast with as low an error rate as possible [14]. The motivation is to predict with as low an error rate as possible the amount of energy we need in the next days, which can be useful in reducing the cost of electricity [10],[12]. Secondly, the question is why we would produce more energy than we need and, thus, pollute the environment?

The power flow forecasting project for the steel plant started with the study of the electricity consumption profiles. Typical consumers for a steel plant are an electric arc furnace, a hot rolling mill and mechanical processing. The study of the consumption profile of an electric arc furnace can be done with a physically mathematical model [3] or a stochastic model [4]. Hot rolling mill electricity consumption models are based on SCADA systems (Supervisory Control and Data Acquisition) [5]. Another large group of consumers in the industrial complex has a seasonal consumption profile with two seasonal cycles, the demand for electricity operates on 5 working days, which can be described by Winter`s multiplicative seasonal forecasting model [6].

Our prediction approach, however, was not based on the described partial models, but on a univariate stochastic ARMA [1],[8],[12],[13] model with regressor. Since we did not use the model of the electric arc furnace, we took the impact of this dominant consumer of electricity into account by introducing an explanatory variable, the regressor, the number of loads of the electric arc furnace per day. In the power flow time series addition to the basic 7-day seasonality, we observe an additional two-stage seasonality of 7/5. The aim was to create a simple and transparent forecasting model, with a uniform ARMA methodology with a regressor [1], [8]. It was decided not to use the additional Winter`s multiplicative seasonal model to explain the 7/5 two-stage seasonality. Additional seasonality was covered by the model with superimposed noise [1], [7].

An alternative to the model with superimposed noise could be use Wavelet transform. Wavelet transform is suitable for dynamic signal decomposition. In fact, it is necessary to select the correct wavelets and the appropriate time sampling of the time series. In the case of an unfortunate choice of frequency bands, it can happen that an important signal is between two detailed coefficients and we do not find it [10],[11].

Non-linear prediction models as described in [15],[16] could also be realized for the described system, but tests [12] showed that the use of a non-linear ARIMA model tree is not useful, because they do not have the possibility of using the regressor variable.

II. Characteristics of Electricity Consumers in the Industrial Complex

The steel mill on Fig. 2 uses 70% of all connected energy of the industrial complex for its production (Fig. 1). Steelworks with an electric arc furnace can work all 7 days of the week and at night. The energy consumption of the steelworks follows: Firstly, the steelmaker's commercial orders, secondly, the price of electricity, which is cheaper at night and at weekends.

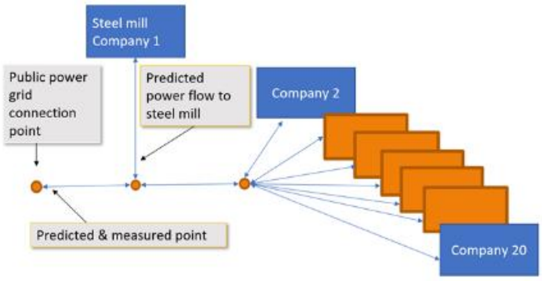

Figure 1. Power flows` branches and block diagram of steel plant

The other 30% of electricity consumption is represented by the other companies, which are located locally in the industrial complex itself. These companies have their own electricity consumption profile 7/5. It should be noted that the block diagram in Fig. 1 is symbolic. The companies in the energy network are not as physically connected as Fig. 1 shows, but only logically. They are physically different, and distributed throughout the industrial energy network.

The beginnings of the ironworks in Koroška -Slovenia go back to the year 1620. After 1774, the first forges and nails` production started to operate along the river Meža, and, thus, the expansion of industrial iron production. In 1992, several production and service companies emerged in Ravne from the single company. Today, there are several companies around the former Ravne Ironworks, employing approximately 3,000 people. The largest company on the site is Metal Ravne. Metal Ravne is the largest company and the largest consumer of electricity. The company consists of a steelworks, a rolling mill and electro-smelting of 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 (Fig. 2).

To understand the energy consumption profile in a steel plant it is good to understand the melting process in the electric arc furnace. The melting process [3] 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.

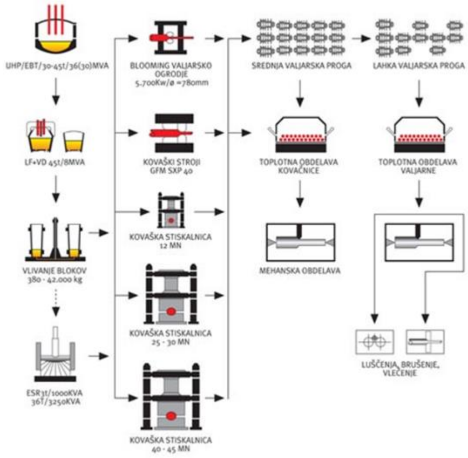

Figure 2. Steel mill technological scheme, Source: https://sij.metalravne.com/

At the discharge of the electrode, an electric arc is then 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 increased 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.

III. Linear Regression Model

Power consumption is measured at 15-minute intervals. We addressed the daily summary consumption in our study (Fig. 3). We would like to predict a complete load at the industry plant for a month in advance.

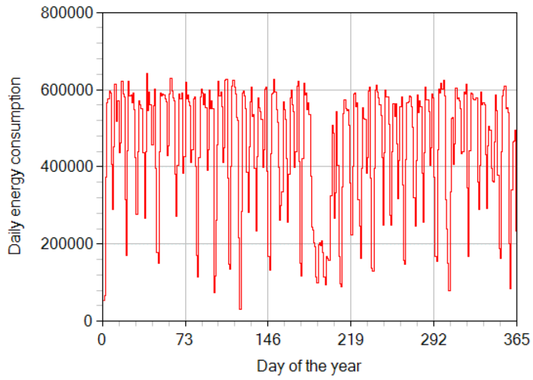

Figure 3. Energy consumption by days in the year

Given the described structure and characteristics of energy consumers in the industrial complex, we can conclude quickly that the electric arc furnace is the main consumer of energy. This is also indicated by Fig. 4, the correlation between the number of loads at the arc furnace per day and common power flow.

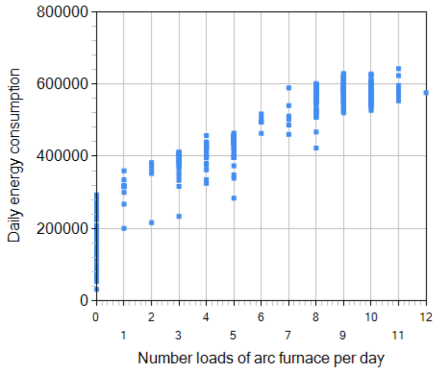

Figure 4. Correlation between the number of loads at the arc furnace per day and power flow

Thus, the first logical attempt at a prediction model is linear regression. An independent predictor variable is the number of loads at the electric arc furnace per day.

Inspections of Fig. 4 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 along a straight line. It is reasonable that the mean of random variable Y is related to x by a straight-line relationship:

$$Y = \beta_0 + \beta_1 x + \varepsilon \tag{1}$$

Where β_0 is a regression coefficient intercept, β_1 is a regression coefficient slope, x is a regressor or predictor variable and Y is a criterion variable. ε is a 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 the best fit to the data. The German scientist Karl Gauss proposed estimating the parameters β_0 and β_1 to minimize the sum of the squares of the vertical deviations. We call this criterion for estimating the regression coefficients the method of least squares [2]. Next, we explain only the major steps to estimating β_0 and β_1, such as definitions:

$$\widehat{\beta_1} = \frac{S_{xy}}{S_{xx}} \tag{2}$$

$$\widehat{\beta_0} = \bar{y} - \widehat{\beta_1}.\bar{x} \tag{3}$$

$\widehat{\beta_0}$…. 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.

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

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

n…. the number of observations in a sample.

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

$$e_i = y_i - \hat{y}_i \tag{6}$$

e_i………...is called the residual.

$$SS_E = \sum_{i=1}^{n} e_i^2 \tag{7}$$

SS_E……… error sum of squares.

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

$$\widehat{\delta^2} = \frac{SS_E}{n-2} \tag{8}$$

$\widehat{\delta^2}$…. 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:

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

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

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

An important part of assessing the adequacy of a linear regression model is testing the statistical hypotheses about the model parameters and constructing certain confidence intervals. Hypothesis testing in simple linear regression is discussed and presents the methods:

$$H_0: \beta_0 = \beta_{0,0} \tag{11}$$

$$H_1: \beta_0 \neq \beta_{0,0} \tag{12}$$

Test statistics for the intercept β_0:

$$T_o = \frac{\widehat{\beta_1} - \beta_{0,0}}{SE(\widehat{\beta_0})} \tag{13}$$

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. In an F-test the test statistics have an $F_{1,n-2}$ distribution. The analysis of variance identity is as follows:

$$F_0 = \frac{SS_R}{\frac{SS_E}{n-2}} = \frac{MS_R}{MS_E} \tag{14}$$

$$SS_R = \sum\nolimits_{i=1}^{n} (\hat{y}_i - \bar{y})^2 \tag{15}$$

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

$$f_0 > f_{\alpha,1,n-2} \qquad (16)$$

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 behaves in a linear or first-order manner:

$$e_i = y_i - \hat{y}_i \quad i=1,2\ldots n \qquad (17)$$

e_i……residual

$$d_i = \frac{e_i}{\sqrt{\hat{\delta}^2}} \quad i=1,2 \ldots n \qquad (18)$$

d_i …. standardized residual.

We may also standardize the residuals by computing. If the errors are normally distributed, approximately 95 % of the standardized 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.

A widely used measure for a regression model is coefficient of determination following the ratio of the sum of squares R^2 $(0 \le R^2 \le 1)$:

$$R^2 = \frac{SS_R}{SS_T} = 1 - \frac{SS_E}{SS_T} \qquad (19)$$

$$SS_T = SS_R + SS_E \qquad (20)$$

SS_T ……total corrected sum of squares,

TABLE I. MODEL SUMMARY LINEAR REGRESSION

R^2	**Std. Error of the Estimate**	**Change Statistics**				
		R Square Change	F Change	df1	df2	Sig. F Change
.91	49318.90	.91	2279	1	231	.000

In Table 1 independent variable loads of the arc furnace per day are explained with 91 % of variance ($R^2 = 0.91$) in the power flow, which is highly significant, as indicated by the F-value of 2279. That F-test says we can trust β_0 and β_1 > 99,9 %.

TABLE II. ANOVA TEST LINEAR REGRESSION

Model	**Sum of Squares**	**df**	**Mean Square**	**F**	**Sig.**
Regression	5542770416189	1	5542770416 189	2279	.000
Residual	561874316138	231	2432356346		
Total	6104644732328	232			

An examination of the t-test indicates that if the number of loads at the electric arc furnace contribute to the electric load the t-test says we can trust $\beta 0$ > 99,9 % and $\beta 1$ > 99,9 %. The Standard Deviation of the sampling distribution is called the standard error.

TABLE III. COEFFICIENTS` LINEAR REGRESSION

Model	**Unstandardized Coefficients**		**t**	**Sig.**	**95.0% Confidence Interval for B**	
	B	Std. Error			Lower Bound	Upper Bound
Constant	209480	5767	36	.000	198116	220843
Loads	40072	839	48	.000	38418	41726

The equation of the linear regression model is:

[Load Forecasting] = 209,480 + 40,072 [Number of Loads at the Electric Arc Furnace per Day] $+\varepsilon$ (21)

IV. ARIMA MODEL WITH REGRESSOR

A stochastic model that can be extremely useful in the representation of certain practically occurring series is the autoregressive model [1]. 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 $z_t, z_{t-1}, z_{t-2}, z_{t-3}, \ldots$ Also let $\tilde{z}_t = z_t - \mu$ be the series of deviations from μ. Then

$$\tilde{z}_t = \phi_1\tilde{z}_{t-1} + \phi_2\tilde{z}_{t-2} + \cdot\cdot + \phi_P\tilde{z}_{t-P} + a_t \qquad (22)$$

is called an autoregressive (AR) process of order p. In (22) the variable z is regressed on previous values of itself; hence the model is autoregressive. The model contains p + 2 unknown parameters $\mu, \phi_1, \phi_2, ., \phi_p, \delta_a^2$, which, in practice, must be estimated from the data. The additional parameter is the variance of the white noise process 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 the observed time series, is the finite moving average process. Here we take $\tilde{z}_t$, to be linearly dependent on a finite number q of previous a's. Thus,

$$\tilde{z}_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \cdot\cdot - \theta_q a_{t-q} \qquad (23)$$

is called a moving average (MA) process of order q. It contains q + 2 unknown parameters $\mu, \theta_1, \theta_2, \ldots, \theta_q, \delta_a^2$, which, in practice, must be estimated from the data. We will define an ARMA (p, q) model with no regressor:

$$\tilde{z}_t = \phi_1\tilde{z}_{t-1} + \phi_2\tilde{z}_{t-2} + \cdot + \phi_2\tilde{z}_{t-2} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \cdot\cdot - \theta_q a_{t-q} \qquad (24)$$

B is defined to perform the following operation: It causes the observation that it multiplies to be shifted backwards in time by 1 period. That is, for any time series $\tilde{z}_t$ and any period t:

$$(B)\,\tilde{z}_t = \tilde{z}_{t-1} \qquad (25)$$

With the *B* operator ARMA can be written:

$$\phi(B)\,\tilde{z}_t = \theta(B)\,a_t \qquad (26)$$

where $\phi(B) = 1 - \phi_1 B - \ldots - \phi_p B^p$ and $\theta(B) = 1 - \theta_1 B - \ldots - \theta_q B^q$.

An ARMA model can be considered as a special type of regression model, so it is straightforward in principle to extend an ARMA model to incorporate information provided by leading indicators and other exogenous variables: You simply add one or more regressors to the forecasting equation. How to include a regressor in ARMA models:

$$\tilde{z}_t = \beta x_t + \phi_1 \tilde{z}_{t-1} + \phi_2 \tilde{z}_{t-2} + \cdot\ \cdot + \phi_P \tilde{z}_{t-P} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \cdot\ \cdot - \theta_q a_{t-q} \qquad (27)$$

where x_t is a regressor at time t and β is its coefficient. While this looks straightforward, one disadvantage is that the regressor coefficient is hard to interpret. The value of β has no effect on $\tilde{z}_t$ when the x_t is increased by one (as it is in classic linear regression). The presence of lagged values of the response variable on the right-hand side of the equation mean that β can only be interpreted conditionally on the value of previous values of the response variable.

If we write the model using backshift operators, the ARMA model with regressor is given by:

$$\phi(B)\tilde{z}_t = \beta x_t + \theta(B)\tilde{z}_t \qquad (28)$$

or

$$\tilde{z}_t = \frac{\beta}{\phi(B)} x_t + \frac{\theta(B)}{\phi(B)} a_t$$

Notice how the AR coefficients get mixed up with both the regressor and the error term. This model can be considered as a special case of transfer function models [1]:

$$\tilde{z}_t = \frac{\beta(B)}{\upsilon(B)} x_t + \frac{\theta\,(B)}{\phi(B)} a_t \qquad (29)$$

This allows for lagged effects of the regressor (via the $\beta(B)$ operator), and for decaying effects of the regressor (via the $\upsilon(B)$ operator). These are called dynamic regression models.

The forecasting model was made from the observed N=234 data records. As an independent predictor we used the number of loads at the electric arc furnace per day. The time series model ARIMA supports an exponential smoothing model and, for our example, used the standard notation (0,0,1)(1,0,0), where 0 is the order of autoregression, 0 is the order of differencing, and 1 is the order of moving-average, and (1,0,0) are their seasonal counterparts. The seasonal part has a periodicity of 7 days. The forecasting model has determined that the load is best described by a seasonal ARIMA model with no order of differencing.

The model statistics in Table 4 provide summary information and goodness-of-fit statistics for the estimated model. First, notice that the model contains one predictor. The stationary R^2 value 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^2 indicate a better fit. A value of 0.956 means that the model does the job of explaining the observed variation in the time series very well.

TABLE IV. MODEL FIT STATISTIC ARIMA WITH REGRESSOR

Predictor	R^2	RMSE	MAPE
1	.956	34636.490	10.557
Ljung-Box Q (18)			
Statistics	DF	Sig.	
18.988	15	.214	

The Ljung-Box statistic, also known as the modified Box-Pierce statistic, provides an indication of whether the model is specified correctly. A significance value less than 0.05 implies that there is a structure in the observed series which is not accounted for by the model. The value of 0.214 shown here is significant, so we can be confident that the model is specified correctly.

In Table 5 all Lags coefficients have a Sig. value less than 0.01; meanings are significant, trust in them is 99%. The only exception is MA Lag 2, which has a confidence of 97%.

TABLE V. MODEL PARAMETERS ARIMA WITH REGRESSOR

		Estimate	SE	t	Sig.
Constant		212449	9025.6	23.538	.000
MA	Lag 1	-.664	.067	-9.874	.000
	Lag 2	-.156	.068	-2.314	.022
AR, Seasonal	Lag 1	.415	.063	6.631	.000
Numerator	Lag 0	35755.8	864.3	41.4	.000
	Lag 2	-16258.6	4486.6	-3.624	.000

When we want to compare the absolute values of individual Lags, we must consider that we did not standardize the time series of power flows, nor the regressor. Therefore, the comparison of absolute values for individual lags is difficult. The time series of power flows could be normalized with a maximum power flow, i.e., around 600,000. The time series of the regressor could be standardised with the maximum number of arc furnace loads per day, i.e., 12. When we consider these two values, we can conclude that Numerator Lag 0 and MA coefficient Lag 1 have about 80% contribution to the forecast.

If we look at the ARIMA model parameters, we notice a little more closely the constant 212,449, which is very close in value to the pure linear regression coefficient intercept. Another detail of comparison ARIMA with linear regression is Numerator Lag 0 35,755, which is essentially a linear regression coefficient slope. In models it is generally accepted that we put values in constants that we cannot explain with the model. Our model was disturbed by a constant with a value of 212,449.

V. DYNAMIC REGRESSION AND MODEL OF SUPERIMPOSED NOISE

We got the idea of how to improve the forecast model during a visual inspection of the shape of the annual time series. The electric arc furnace is maintained once a year, so it does not operate, and the technologically related production processes do not produce. In Fig. 5 we see this characteristic shape on the graph of the annual power flow between 187 and 195 days of the year. At that time, the other companies are operating normally. The form of electric consumption has the characteristic: 5 working days and 2 days off, and its amplitude is only 1/3 of the maximum. So, we can conclude that there are two periodicities in the time series, first a period of 7 days and then a two-stage 7/5. The seasonal ARIMA model presented in the previous chapter has a periodicity of 7 days. Periodicity 7/5 could be modeled with the seasonal Winter`s multiplicative model. Winter's multiplicative model was confirmed manually for the mentioned partial signal with R^2 = 0.95 and MAPE 3%, and cannot be interpreted with any combination of ARIMA coefficients. For the seasonal ARIMA with a regressor the signal 7/5 is a noise.

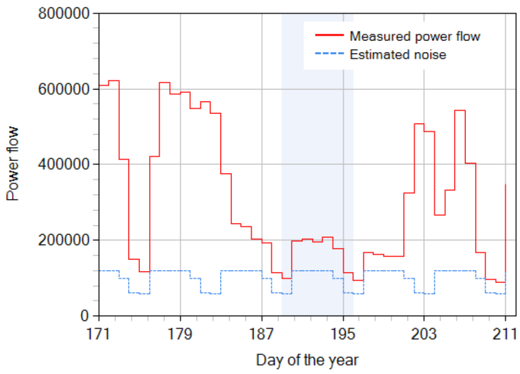

Figure 5. Characteristic power flow when the arc furnace is not in operation

Thus, to improve our predictions, we moved on to the theory of models with superimposed noise. This problem was complicated in practice by the presence of noise N_t, which we assumed corrupts the true relationship between input and output according to

$$\tilde{z}_t = \frac{\beta(B)}{\upsilon(B)} x_t + N_t \tag{30}$$

$$N_t = \frac{\theta(B)}{\phi(B)} a_t \tag{31}$$

where N_t and X_t are independent processes.

Of course, equation (29) from the previous chapter and equation (30) in this chapter are the same. In the previous chapter, we wanted to show how ARIMA and the regressor merged into a single model. In this chapter, however, we are more interested in pure transfer function $\frac{\beta(B)}{\upsilon(B)} x_t$ and a noise $N_t = \frac{\theta(B)}{\phi(B)} a_t$. According to the indications at the beginning of this chapter about additional periodicity 7/5, it is necessary to supplement for our transfer function (30) with additional noise N_t^1

$$\tilde{z}_t = \frac{\beta(B)}{\upsilon(B)} x_t + N_t + N_t^1 \tag{32}$$

Noise N_t^1 represents the power flow of 20 companies in the characteristic 7/5.

Ideally, the common prediction mode for our case would consist of three partial models: Dynamic regression (predictor), ARIMA, and Winter`s multiplicative model. In any case, this would be a complex model that would require a dynamic decomposition of the time series or additional measurements in an industrial energy network.

We opted for a simpler and more efficient path that does not require the Winter`s multiplicative model. The detail for a week of the noise N_t^1 is given in Fig. 6. We assume that the noise is repeated periodically throughout all weeks of the year, for the same days of the week (Fig. 5). It is not really like that, but we take that as like a definition that holds true for a year.

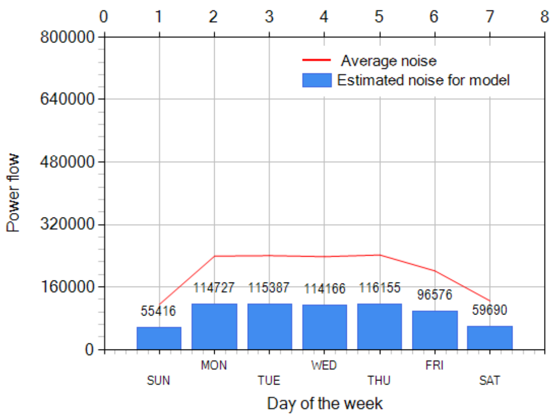

Figure 6. Estimated noise N_t^1 for a week

The noise N_t^1 was subtracted from the original measured time series. When the prediction model was obtained based on a modified time series and a prediction was made, an inverse transformation must be performed for the result of the prediction, i.e., noise N_t^1 is added back for the final prediction. Here, we would also write a dilemma: Whether we are talking about the transformation of the basic time series, or about the subtraction and addition of noise. Basically, an additive signal transformation is used, but our time series did not need a characteristic ARIMA transformation, since the base time series was stationary without it. But, we needed this transformation to improve our forecast.

The use of superimposed noise (32) changed the model greatly. Our time series model ARIMA supports exponential smoothing. Dynamic regression and the model of superimposed noise was using the standard notation ARIMA (0,0,1)(1,0,1), where 0 is the order of autoregression, 0 is the order of differencing, and 1 is the order of moving-average, and (1,0,1) are their seasonal counterparts. We noticed changes in the symbolic notation of the ARIMA model, where the AR part appeared in the seasonal part. The forecasting model determined that a load is best described by the seasonal ARIMA model with no order of differencing. A value of R^2 = 0.97 means that the model does an excellent job of explaining the observed variation in the time series.

TABLE VI. MODEL FIT STATISTIC FOR ARIMA MODEL WITH REGRESSOR AND A SUPERIMPOSED NOISE

Predictor	R^2		RMSE	MAPE
1	.968		29823.6	12.577
Ljung-Box Q (18)				
Statistics	DF	Sig.		
12.653	15	.629		

TABLE VII. MODEL PARAMETERS – ARIMA WITH REGRESSOR AND SUPERIMPOSED NOISE

		Estimate	SE	t	Sig.
MA	Lag 1	-.607	.056	-10.9	.000
AR, Seasonal	Lag 1	.913	.142	6.446	.000
MA, Seasonal	Lag 1	.854	.178	4.788	.000
Numerator	Lag 0	36526	780	46.82	.000
	Lag 1	15507	4419	3.509	.001
	Lag 2	17705	3905	4.535	.000
Denominator	Lag 1	.631	.118	5.342	.000
	Lag 2	.313	.109	2.886	.004

Consideration of superimposed noise caused improvements on the ARIMA model. Table 7 shows that constant 212,449 from the previous model disappeared and had been replaced by 3 new Lags in the transfer function model. The ARIMA MA Lag 1 value coefficient changed very little compared to the previous model. The seasonal part of ARIMA was given a new structure (1,0,1) and new values. The Numerator coefficient at the Lag 0 changed a bit. Numerator Lag 1 was new and Lag 2 coefficients had new positive values. In the Denominator segment were completely new Lag 1 and Lag 2 coefficients.

The use of a modified superimposed noise model in conjunction with the structure of the ARIMA model caused the constant in the forecasting model to disappear. The role of the constant was taken over by the new Lags, so we could get a better filter model of the prediction. As a result, the RMSE also decreased, and the accuracy of the prediction increased.

VI. FORECASTING WITH MODELS

Various statistics are used to evaluate the performance of models of forecasts. We decided to use in the final test only MAPE, RMSE and R^2 to obtain a comparison between individual solutions:

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:

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}} \qquad (33)$$

where $SS_{tot} = \sum_{i=1}^{n}(Ai - \bar{A})^2$

and $SS_{res} = \sum_{i=1}^{n}(Ai - Fi)^2 = \sum_{i=1}^{n} e_i^2$.

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 the same data.

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(Ai - Fi)^2}{n}} \qquad (34)$$

All three models we compared were calculated from the same historical time series N = 234. With the models thus obtained, we predicted the same 30 days in the future.

The linear regression itself is a static method, since there is no time in the equation. We assumed that the linear dependence was valid in the future as well, and so we made a prediction. It is usually the case that the forecast is worse than the model. The linear regression model promises R^2 0.91, but this was not achieved during the 30-day simulation predictions – they gave us R^2 0.78, values of MAPE 12.5% and RMSE 57824.3.

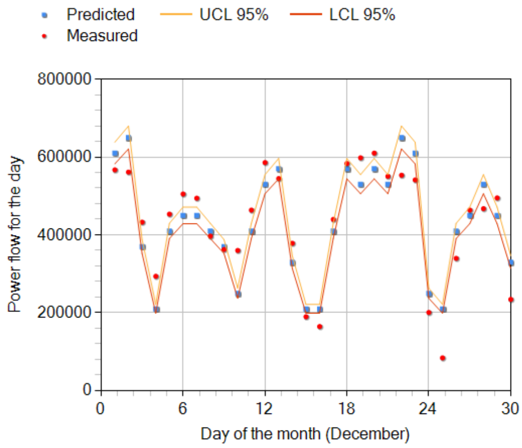

Figure 7. Monthly load forecasting using the linear regression model

There is another problem in linear regression as we do not have a satisfactory confidence interval in which the predictions lie. Table 3 shows a 95% confidence interval for regression coefficient intercept and regression coefficient slope, but this was not the 95% confidence interval for the predicted values. So, it's hard to say with linear regression, the prediction is so much, with the lower limit of the LCL and the upper limit of the UCL. In Fig. 7 we see that most of the measurements lay outside the 95% confidence interval.

TABLE VIII. FORECASTING STATISTICS FOR 30 DAYS IN ADVANCE WITH THREE DIFFERENT FORECAST MODELS

	Linear regression	**ARIMA with regressor**	**ARIMA with regressor an superimposed noise**
RMSE	57824.3	50205.8	40699.8
MAE	50950.1	41452.8	35144.1
MAPE	12.5 %	10.3 %	8 %

STDEV forecasting	131028.7	128506.7	131474.4
AVG forecasting	445618	451552.7	447023
STDEV measured	123688.5	123688.5	123688.5
AVG measured	450870.5	450870.5	450870.5
R^2	0.78	0.84	0.89

In the ARIMA model with regressor, parameters were optimized for a step or two into the future from the historical learning time series itself. In fact, during the optimization of the ARIMA model, we predicted right through the historical time series itself. A 95% confidence interval was calculated for such a short model prediction. For a 30-day forecast, we must check the model with a simulation forecast, and we make a forecast with the model for 30 days ahead. Only in this way can we determine whether the model is useful for medium-term forecasting of 30 days. Fig. 8 shows the forecast for 30 days using the ARIMA model with regressor and superimposed noise. All predictions and measured values lay within the 95% confidence interval. The exception was 24 and 25 days (national and religious holidays). These are two so-called outliers.

Outliers first appeared during model computation and during forecasting. In our time series for one year, we had 7 outliers, on national and religious holidays. Power flows were extremely low on these days. There are different strategies when working with outliers: Replace with an average value, replace with a previous value, or delete. We did not do any of that, which is often a good solution as well. However, it is necessary to mention that they are here, and have an impact on the results.

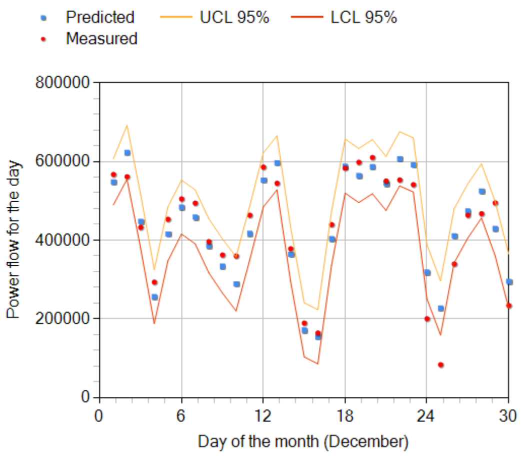

Figure 8. Monthly load forecasting using the ARIMA model with regressor and a superimposed noise

VII. CONCLUSION

For an industrial complex with very specific and different types of electricity consumers, we performed a forecast of power flows for a month into the future. In our case, we had two groups of electricity consumers: The first is a steelworks with an electric arc furnace that can work all 7 days of the week and at night. Secondly, there are companies that work only 5 days a week. The choice of prediction technique is highly dependent on the properties of the measured power flow time series. Our time series was created in a very specific industrial environment and required a regressor a (so-called predictor). At this point, the number of possible choices of forecasting techniques is greatly reduced.

Various authors have described very different models of prediction technique [6],[7],[8],[10] but all acknowledge that ARIMA is a universal model, while other models are specialized for a particular type of time series. ARIMA has the great advantage that it can also include one or more regressors in its filter model. The loads of the electric arc furnace per day were the regressor at our case.

First attempts to predict daily consumption were by linear regression. This technique is relatively simple. The disadvantage of it is that is a static method and cannot track daily variation in consumption dynamically. It was the basis for comparison with the improved ARIMA models. When we combined linear regression and ARMA methodology, we obtained a dynamic regression model that, with its linear filter structure, adapted dynamically to changes in time series. This model can be considered as a special case of transfer function models. If we compare the predictions with linear regression and predictions using the ARIMA model with a regressor was the second, 15% better model.

In the last prediction model, we considered superimposed noise: Classic forecasting models, each of which works well for a certain form of time series. The problem occurs when the time series signal is complex, meaning it consists of several characteristic signals. At that time, no known classic model gives a satisfactory result. The solution is to decompose or denoise the time series. In our case, we split the time series it into two signals. Because of the above, we opted for the superimposed noise method.

Thus, we obtained a simple model with 8 Lags and a reliable predictive model that offered a 30-day forecast into the future with $R^2 = 0.9$ and RMSE = 40699, which was 30% better than for the linear regression model. Finally, the prediction showed a MAPE error of 8% and a very narrow constant confidence interval UCL and LCL for 30 days. This type of forecast can be extended to a longer period, but we only needed to know how many times a day we would fill the electric arc furnace.

REFERENCES

[1] George E. P. Box Gwilym M. Jenkins Gregory C. Reinsel, Time Series Analysis Forecasting and Control, 2014, Wiley.

[2] Douglas C. Montgomery, George C. Runger, Applied Statistic and Probability for Engineers", 2014, Wiley.

[3] Janez Bratina, Electric arc furnace for steel production (in the Slovene language), Ravne na Koroškem, 1994, Slovenia.

[4] Johannes Dock, Daniel Janz, Thomas Kienberger, Modellierung des Lastprofils eines Elektrolichtbogenofens mittels Markov-Ketten, 16. Symposium Energy Innovation EnInnov2020, 12.-14. Februar 2020 TU Graz, Austria.

[5] Gonzalo Alonso Orcajo, Josué Rodríguez D., Pablo Ardura G., José M. Cano, Joaquín G. Norniella, Rocio Llera T., and Diego Cifrián R., Dynamic Estimation of Electrical Demand in Hot Rolling Mills, IEEE Transactions on Industry Applications, Vol. 52, No. 3, May/June 2016.

[6] James W. Taylor and Patrick E. McSharry, Short-Term Load Forecasting Methods: An Evaluation Based on European Data, IEEE Transactions on Power Systems, Vol 222, No 44, November 2007.

[7] T. Haida; S. Muto, Regression based peak load forecasting using a transformation technique, IEEE Transactions on Power Systems Volume: 9, Issue: 4, Nov. 1994.

[8] R.E.Abdel-Aal A.Z.Al-Garni, Forecasting monthly electric energy consumption in eastern Saudi Arabia using univariate time-series analysis, Energy, Volume 22, Issue 11, November 1997.

[9] Chaoqing Yuan, Sifeng Liu, Zhigeng Fang, Comparison of China's primary energy consumption forecasting by using ARIMA (the autoregressive

integrated moving average) model and GM (1,1) model, Energy, Volume 100, April 2016.
[10] Gholamreza Memarzadeh, Farshid Keynia, Short-term electricity load and price forecasting by a new optimal LSTM-NN based prediction algorithm, Electric Power Systems Research, Volume 192, March 2021.
[11] Boris Bizjak, Measurement of electrical voltage distortion by wavelets transformation (in the Slovene language), ERK 2015, Portorož, Slovenia.
[12] Boris Bizjak, Load Forecasting Applications for The Energy Sector, 14. Symposium Energy Innovation EnInnov2016, 10.-12. February 2016 TU Graz, Austria.
[13] Boris Bizjak, Daily Load Forecasting an Industrial Complex, 15. Symposium Energy Innovation EnInnov2018, 14.-16. February 2018 TU Graz, Austria.
[14] Boris Bizjak, Power Flow Forecasting with Low Prediction Error, 16. Symposium Energy Innovation EnInnov2020, 12.-14. February 2020 TU Graz, Austria.
[15] Jamie McLennan, Zhao Hui Tang, Bogdan Crivat. Data Mining with Microsoft SQL Server, 2011, Wiley.
[16] Bo Thiesson, David Maxwell Chickering, David Heckerman, Christopher Meek, ARMA Time-Series Modeling with Graphical Models, Microsoft Research, Proceedings of the 2004 Siam International Conference on Data Mining

Boris Bizjak was born in Maribor 1957, High Technical School – Maribor Dipl. Eng. (1981). University of Maribor – Faculty of Electrical Engineering and Computer Science
Master of Science with diploma title “Servo drive control with elastically coupled loads” (1990)
He has worked in various companies as a Project Manager. He is currently working at the University of Maribor – Faculty of Electrical Engineering and Computer Science, as a coordinator of practical education.