

Very short-term forecasting of power demand of big dynamics objects

Mirosław Parol^{1*}, Paweł Piotrowski¹, and Mariusz Piotrowski²

¹Warsaw University of Technology, Faculty of Electrical Engineering, Warszawa, Poland

²Electrum Sp. z o. o., Białystok, Poland

Abstract. The issue of very short-term forecasting is gaining more and more importance. It covers both the subject of power demand forecasting and forecasting of power generated in renewable energy sources. In particular, for the reason of necessity of ensuring reliable electricity supplies to consumers, it is very important in small energy micro-systems, which are commonly called microgrids. Statistical analysis of data for a sample big dynamics low voltage object will be presented in this paper. The object, in paper author's opinion, belongs to a class of objects with difficulties in forecasting, in case of very short-term horizon. Moreover, forecasting methods, which can be applied to this type of forecasts, will be shortly characterized. Then results of sample very short-term ex post forecasts of power demand provided by several selected forecasting methods will be presented, as well as some qualitative analysis of obtained forecasts will be carried out. At the end of the paper observations and conclusions concerning analyzed subject, i.e. very short-term forecasting of power demand of big dynamics objects, will be presented.

1 Introduction

The issue of very short-term power demand forecasting is gaining more and more importance. In particular, for the reason of necessity of ensuring reliable electricity supplies to consumers, it is very important in small energy micro-systems, which are commonly called microgrids.

The problem of very short-term forecasting of the power demands has been discussed in several publications, e.g. [1-6]. Very comprehensive overview of forecasting methods, taking into account ultra- and short-term time horizon has been included in [7]. In the overview among others some papers dealing with the problems of time horizons being considered, as well as different areas and location types (small city, microgrid, smart building) have been presented. Mainly different kinds of neural networks were used in the papers for forecasting, such as Multi-layer Perceptron (MLP), Support Vector Machine (SVM), Support Vector Regression (SVR) and Self-Organizing Map (SOM).

In the publications being analyzed, the objects of big dynamics in the power demand level are very rarely the subject of very short-term forecasting activities. This is the clue and the key aspect of this paper. Very short-term forecasts will be related to the component

* Corresponding author: mirosław.parol@ee.pw.edu.pl

of microgrid [6] of significant dynamics in the level of received power. One of the main technological components of sewage treatment plant plays the role of an object showing considerable dynamics in the level of power demand. Some controllable and non-controllable loads, together with some generating units are also the part of the microgrid.

In the further part of the paper the results of sample very short-term forecasting processes for the presented component of microgrid, obtained with the use of the following methods (models): naive model, weighted moving average models, auto-regression models, multiple linear regression models, MLP type neural networks and Radial Basis Function type neural networks will be given. The short description of those methods is presented below.

The naive method. In this method it is most often assumed, that the value of the variable being explained (forecasted) by the model in time t will be the same as in time $t - 1$ [8]. In most cases naive method and its results can serve as a reference point in the process of evaluation of other forecasting methods.

Weighted moving average models. The approach making use of these models is based on the following equations set [8]:

$$y_t^* = \sum_{i=t-k}^{t-1} y_i \cdot w_{i-t+k+1}; \quad 0 < w_1 < w_2 < \dots < w_k \leq 1; \quad \sum_{i=1}^k w_i = 1 \quad (1)$$

where: y_t^* – value of the forecast for the variable determined for time t , y_{t-1} – actual value of the variable for which forecasting takes place in time i , k – smoothing factor, $w_{i-t+k+1}$ – weight of the value of the variable being forecast in time i .

Auto-regression model (AR). This model bases on the assumption, that the subsequent values of the variable being forecast are correlated, that is [8]:

$$y_t^* = \varphi_0 + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} \dots + \varphi_p y_{t-p} + e_t \quad (2)$$

where: y_t^* – value of the forecast for the variable determined for time t ; $y_{t-1}, y_{t-2}, \dots, y_{t-p}$ – values of the variable being forecast in time $t - 1, t - 2, \dots, t - p$; $\varphi_0, \varphi_1, \varphi_2, \dots, \varphi_p$ – parameters of the model; p – delay; e_t – error in time t .

Multiple linear regression (MLR). In the methods basing on MLR for the purpose of determination of the values of explained variable different kinds of explanatory variables are used, including the past values of both the explained variable and explanatory variables

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} \dots + \alpha_n y_{t-n} + \alpha_{n+1} x_{1t} \dots + \alpha_{n+m} x_{mt} + \varepsilon_t \quad (3)$$

where: y_t – explained variable in time t ; x_{1t}, x_{2t}, x_{mt} – explanatory variables in time t ; $\alpha_0, \alpha_1, \alpha_2, \alpha_{n+m}$ – parameters of the model, $n+m$ – the number of explanatory variables, ε_t – error in time t .

Multi-layer Perceptron. It is a unidirectional neural network (trained with supervision) consisting of [9]: input layer, several hidden layers (usually not more than two) and output layer. During the our research neurons in the hidden layer had a non-linear (hyperbolic tangent) activation function, while neurons in the output layer had a linear function.

RBF type network. It is also a unidirectional neural network consisting of input layer, hidden (radial) layer and output layer [9]. Neurons in the hidden layer do not have any weights assigned, but radial base functions instead (during the our research Gaussian type functions were used). Neurons in the output layer had a linear activation function.

The parameters (weights) of individual models were selected using optimization methods (BFGS algorithm for MLP networks, RBFT algorithm for RBF networks and Newton's algorithm for other models for which the values of parameters need to be chosen).

2 Statistical analysis of data

Statistical analysis is based on data from one day from December and one day from June (the only available data). The daily time series includes 8640 periods of 10 seconds. In the data from the December day there is a lack of 169 data. Missing data from the December day was supplemented with information about values neighboring with missing ones. In the data from the June day there is a lack of 578 data. Missing data occur in quite random places of the time series. In a few cases it is the lack of a longer fragment of the time series (several minutes). Data from the December day and the June day were jointly normalized for anonymisation to relative units (1 relative unit is equal to an average value from time series). Table 1 shows selected statistical measures of two mentioned time series of the power demand.

Table 1. Statistical measures of two time series of the power demand. Source: Own elaboration.

Statistical measures	December day	June day
The length of the time series	8472	8063
Max power [p.u.]	6.2690	5.5542
Min power [p.u.]	0.0035	0.0033
The ratio of maximum power to minimum power	1791	1683
Average power [p.u.]	1.0876	0.9079
Median [p.u.]	0.9317	0.8712
Standard deviation [p.u.]	0.7759	0.8745
Variation [-]	0.6020	0.7648
Coefficient of variation [%]	71.3372	96.3149
Skewness [-]	0.2339	0.8347
Kurtosis [-]	-0.3041	0.3512

The time series of the December day has characteristics significantly different from the June day. The time series from the June day has a smaller mean value but a greater variance (over 25%) than the time series from the December day. Probably forecast errors for the June day will be higher than for the December day, assuming that the dominant component is the random component. The minimum values from both time series are almost identical but the maximum value of the power demand is higher for the December day. The Kolmogorov-Smirnov and Lilliefors tests show that the both time series of the power demands do not have a normal distribution. Figure 1 shows the percentage distribution of cases in both time series depending on the power demand.

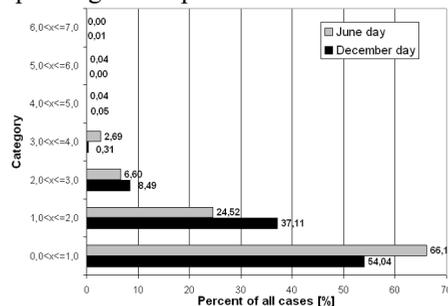


Fig. 1. The percentage distribution of cases in both time series depending on the power demand value. Source: Own elaboration.

The power demand below 1 p.u. is definitely dominant one in both time series. As much as 20% of the power demand from the whole time series are very small values, below 0.05 p.u. (this is at least 20 times smaller than the average value).

Figure 2 shows the daily time series of the power demand for the December day and the June day. Figure 3 shows in turn two parts of the day from December (period from 3:30 to 4:00 am and the period from 9:30 to 10:00 am). Big changes dynamics (rapid jumps) in the power demand is clearly visible.

For both time series (December day and June day), the occurrence of periods of different length with a very similar power demand is characteristic. These are values from very low to medium power demand (up to 1 p.u. for the June day) and large power demand (up to 2 p.u. for the December day). These periods are usually divided by shorter periods of much bigger power demand (up to 2 orders of magnitude larger).

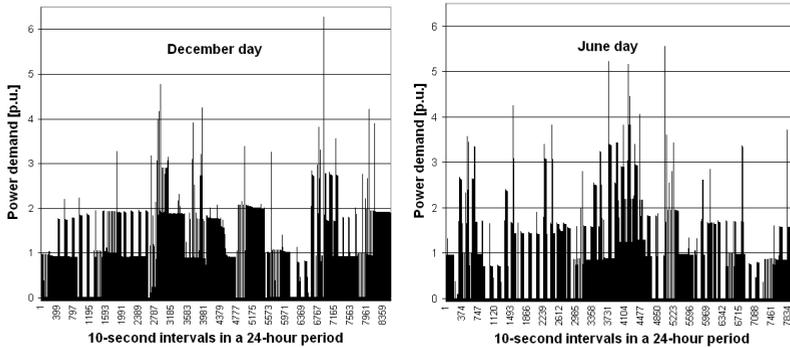


Fig. 2. The daily time series of the power demand for the December day and the June day. Source: Own elaboration.

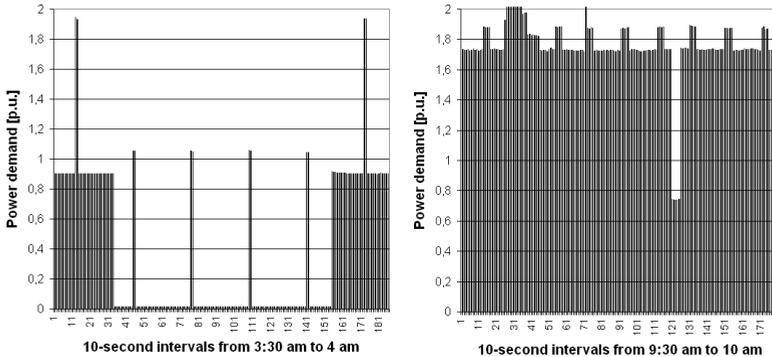


Fig. 3. Two selected 30-minute periods of the power demand in the December day. Source: Own elaboration.

The seasonality problem of the time series of the power demand was not analyzed due to the lack of more amount of data. For the December day, it was examined whether there was daily variation in the power demand in 1-hour periods (see Fig. 4). For each hour the average 10-second power demand was calculated. A polynomial illustrating the best daily variation trend was selected. The biggest value of the determination coefficient R^2 equal to 0.2671 was obtained for the polynomial degree 5 (polynomials from degree 2 to 6 were analyzed). The value of the determination coefficient is very low. The function of polynomial degree 5 explains only less than 27% of the variability of the explanatory variable. The polynomial shape shows that probably (the analysis concerned the only one day) the energy consumer has 2 periods of increased power demand during the day. The first peak is from 6 am to 14 pm. The second peak is from 22 pm to 24 pm. The power

demand lower than the average value (night valley) occurs from 1 am to 4 am and from 5 pm to 7 pm (evening valley).

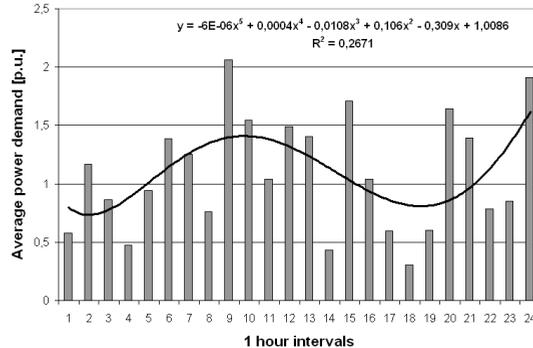


Fig. 4. The daily variation of the power demand in 1-hour periods. Source: Own elaboration.

For the December day, the autocorrelation coefficient dynamically decreases from the value 0.946 (1 period back) to the value 0.300 (90 periods back, i.e. 15 minutes) (see Fig. 5). A surprising and difficult to explain is the temporary increase in the value of the autocorrelation coefficient up to the value of 0.439 (150 periods back, i.e. 25 minutes). After 240 periods back (40 minutes), the value of the autocorrelation coefficient is equal to almost zero.

For the December day, the partial autocorrelation coefficient is statistically significant (5% significance level) only for the first three values and several single values distant from each other (see Fig. 5). Very big values of the partial autocorrelation coefficient for the first three periods back (in turn: 0.9464, 0.9114 and 0.8999) and quite big negative value equal to -0.3590 for 33 periods back, i.e. 5 minutes and 30 seconds, is noteworthy.

The Table 2 shows for the December day the values of Pearson linear correlation coefficients between 10-second power demand and considered explanatory variables. All correlation coefficients are statistically significant (5% level of significance). It is noteworthy that the sum of the power demand values from the period $t-1$ with the weight 0.75 and the period $t-2$ with the weight 0.25 (AWE) has maximum correlation coefficient from all potentially explanatory variables. Moreover the 24-hour profile of the variation of power demand (FUNC) seems to be definitely much better explanatory variable than the hour of power demand (HOU).

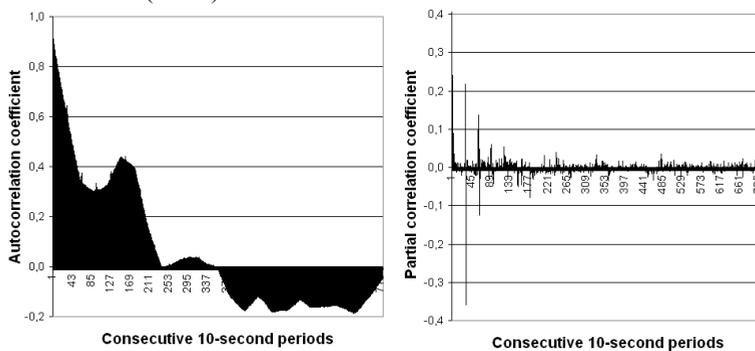


Fig. 5. Autocorrelation function (ACF) and partial autocorrelation function (PACF) of time series in the December day. Source: Own elaboration.

Table 2. The values of Pearson linear correlation coefficients between 10-second power demand and considered explanatory variables. Source: Own elaboration.

The code of variable	Considered explanatory variables	Pearson linear correlation coefficient
POW1	10-second power demand in period $t-1$	0.9464
POW2	10-second power demand in period $t-2$	0.9114
POW3	10-second power demand in period $t-3$	0.8999
POW4	10-second power demand in period $t-4$	0.8882
POW5	10-second power demand in period $t-5$	0.8773
POW6	10-second power demand in period $t-6$	0.8657
POW7	10-second power demand in period $t-7$	0.8539
POW8	10-second power demand in period $t-8$	0.8423
FUNC	24-hour profile of the variation of power demand - the value of the polynomial function of degree 5 of power demand in a given hour	0.3168
AVE3	The average value of power demand from periods $t-1$, $t-2$ and $t-3$	0.9402
AVE2	The average value of power demand from periods $t-1$ and $t-2$	0.9418
AWE	The sum of the power demand values from the period $t-1$ with the weight 0.75 and the period $t-2$ with the weight 0.25	0.9474
HOU	The hour (number from the range 1..24) of power demand	0.0070

3 Comparative analysis of very short-term forecasting methods for power demand of big dynamics objects

Comparative analysis was performed using data from the December day. All data were divided into sets for parameters (weights) estimation of a given model (80% of randomly selected data due to ensuring maximum representativeness of the process) and data for testing the quality of "ex post" forecasts (20% of randomly selected data).

In order to have a broader, multi-aspect view of the quality of individual forecasting models, four measures of "ex-post" forecasts quality were used: MAPE error, SOS error, maximum percentage error and Pearson's linear correlation coefficient.

In the case of MLP networks, the SOS error was minimized, while in other models the minimization of the SOS error, MAPE and the maximization of the linear correlation coefficient were tested.

The selection of the most favorable explanatory variables for the MLP network was made by using the "top-down" method. Undesirable variables (not providing additional information to the model) were eliminated stepwise. The selection of explanatory variables for elimination is supported by the sensitivity analysis of neural networks. In the case of multiple regression models, information about the most advantageous sets of explanatory variables previously found for the MLP network was used.

Table 3 shows characteristic of all tested prognostic models. The description of neural network models (MLP/RBF) is as follows: number of neural network inputs, number of neurons in the hidden layer and number of outputs. For the models of neural networks, only the models with the most favorable structure are given in the Table 3 (the number of neurons hidden in a fairly wide range was tested). In addition, for neural network models, the validity of each explanatory variable calculated on the basis of the sensitivity analysis is given in the descriptive field (column) of the given explanatory variable. Validity (number) is the quotient of the neural network error for which the given variable has been replaced in all sets by its average value by the neural network error calculated for the whole given set of explanatory variables. In Table 3 the explanatory variables probably undesirable in a given model are marked in italics. The most important explanatory variable in a given model is marked as bold.

In the case of the naive model (NAIVE), multiple regression models (M.REGR), auto-regression models (AR) and moving average weighted models (MOV.A), the values of parameters of this models assigned to each explanatory variables are given in the descriptive fields (columns) of explanatory variables. Variables with codes FUNC, POW1,

POW2, POW3, POW4 and POW5 create the proper set of explanatory variables for MLP model. The set using additionally variables with AWE, AVE3 and ACE2 codes results in a similar quality of forecasts. The use of both described sets of explanatory variables for multiple regression models has also proved to be the most advantageous. Definitely the most important explanatory variable in almost all models is the power value from the *t*-1 period (POW1).

Table 3. Characteristic of all tested prognostic models. Source: Own elaboration.

The model code / description / error function	Explanatory variables in the model												
	FUNC	HOU	AWE	AVE3	AVE2	POW1	POW2	POW3	POW4	POW5	POW6	POW7	POW8
NAIVE / /	-	-	-	-	-	1.0000	-	-	-	-	-	-	-
MLP1 / 13-10-1 / SOS	1.0166	<i>1.0031</i>	3.7508	1.1682	1.5094	3.3204	2.4477	1.3586	1.0103	1.0230	1.0134	<i>1.0025</i>	<i>1.0049</i>
MLP2 / 10-8-1 / SOS	1.0388	-	3.9480	1.4689	1.0976	3.9268	2.7330	1.5004	<i>1.0077</i>	1.0360	<i>1.0056</i>	-	-
MLP3 / 7-6-1 / SOS	1.0152	-	3.9812	1.4850	1.5011	3.8727	3.0880	1.6301	-	-	-	-	-
MLP4 / 6-6-1 / SOS	1.0203	-	-	1.7375	2.8798	6.2419	2.6974	1.7820	-	-	-	-	-
MLP5 / 4-4-1 / SOS	1.0136	-	-	-	-	10.4000	1.3030	2.7197	-	-	-	-	-
MLP6 / 6-6-1 / SOS	1.0151	-	-	-	-	11.3303	1.2782	1.5836	1.0129	1.0336	-	-	-
MLP7 / 7-6-1 / SOS	1.0172	-	-	-	-	11.2997	1.6829	1.9435	1.0155	1.0274	<i>1.0089</i>	-	-
MLP8 / 9-8-1 / SOS	1.0256	-	3.5493	1.1639	1.5764	4.3138	2.7429	1.5676	1.0142	1.0227	-	-	-
RBF / 6-8-1 / SOS	<i>0.0024</i>	-	-	-	-	0.0708	<i>0.0698</i>	<i>0.0698</i>	<i>0.0604</i>	<i>0.0722</i>	-	-	-
M.REGR1 / / cor.coef.	0.0324	-	0.2713	0.1139	0.1083	0.4342	-0.2176	0.1252	0.0352	0.0509	0.0435	-	-
M.REGR1a / with constant / cor.coef	0.0324	-	0.2713	0.1139	0.1083	0.4342	-0.2176	0.1252	0.0352	0.0509	0.0435	-	-
M.REGR2 / / SOS	0.0303	-	0.2726	0.1145	0.1089	0.4363	-0.2186	0.1258	0.0354	0.0511	0.0437	-	-
M.REGR3 / / MAPE	0.0004	-	0.3676	0.1043	0.1738	0.5615	-0.2139	-0.034	0.0000	0.0001	0.0000	-	-
M.REGR4 / / cor.coef.	0.0204	-	-	-	-	0.4422	-0.0337	0.1022	0.0211	0.0488	-	-	-
M.REGR5 / / SOS	0.0317	-	-	-	-	0.7374	-0.0562	0.1705	0.0352	0.0814	-	-	-
M.REGR6 / / MAPE	0.0004	-	-	-	-	0.9590	-0.0003	0.0003	-0.0001	0.0001	-	-	-
AR(3)1 / p=3 / cor.coef	-	-	-	-	-	0.7907	-0.0471	0.2471	-	-	-	-	-
AR(5)1 / p=5 / cor.coef	-	-	-	-	-	0.7619	-0.0576	0.1763	0.0367	0.0849	-	-	-
AR(5)2 / p=5 / SOS	-	-	-	-	-	0.7445	-0.0542	0.1750	0.0368	0.0887	-	-	-
AR(5)3 / p=5 / MAPE	-	-	-	-	-	0.9889	-0.0001	0.0001	0.0000	0.0001	-	-	-
MOV.A1 / k=3 / cor.coef	-	-	-	-	-	0.3333	0.3333	0.3333	-	-	-	-	-
MOV.A2 / k=3 / cor.coef	-	-	-	-	-	0.7750	0.0000	0.2250	-	-	-	-	-
MOV.A3 / k=4 / cor.coef	-	-	-	-	-	0.7477	0.0000	0.1471	0.1052	-	-	-	-
MOV.A4 / k=5 / cor.coef	-	-	-	-	-	0.7331	0.0000	0.1454	0.0436	0.0779	-	-	-

Table 4 presents the results of "ex post" forecasts of the tested forecasting models. Significantly worse results in a given error category are marked in italics. While significantly better results in a given error category are marked as bold. Figure 6 shows the chosen forecast results for 71 10-second periods from hour 3:33.

The obtained forecast errors of the tested models are large (the smallest MAPE error was slightly above 107%), but this is a normal situation for a process with such a large and frequent changes dynamics and a large random component. However, compared to the simplest and the worst naive model, the improvement in the quality of forecasts of the best models is very significant (more than doubling the size of the MAPE error and the SOS error).

It is not possible to unambiguously choose the best model due to the use of four measures of forecast quality with slightly different properties.

The most favorable forecast quality measures have MLP models (the smallest maximum percentage errors, the smallest SOS errors and the largest correlation coefficients). MAPE errors of MLP models (overestimation of power demand values especially for very small values) are significantly bigger than in some other models, e.g. moving average weighted models or autoregressive models. MLP models have the worst forecasts by very low power demand (MAPE errors typically above 200%) (see Fig.6). This is due to the fact that the SOS error is minimized by the neural network during the learning. The difference between the two very small values of power demand (actual value and forecast) is a very small number, however the MLP models tries to minimize the large SOS errors (from its point of

view) occurring for much bigger power demand values (see Fig.6). Multiple regression models and autoregressive models in which MAPE error is minimized have much smaller MAPE errors of forecasts than MLP models. However, in these cases the SOS errors of forecasts of these models are significantly worse than in the MLP models. A very characteristic phenomenon for multiple regression models and autoregressive models are very accurate forecasts of very small power demand values, but this is done at the expense of the forecasts accuracy for bigger power demand values (see Fig.6). On the other hand, what is surprising, the models of the moving average weighed despite the use of the correlation coefficient as the maximized error function have MAPE errors of forecasts quite much smaller than in the MLP models.

Table 4. The results of "ex post" forecasts of the tested forecasting models. Source: Own elaboration.

The model code / description / error function	Error measure			
	MAPE [%]	SOS [(p.u.) ²]	Max. error [%]	Correlation coefficient
NAIVE / /	239.77	249.55	48590	0.8800
MLP1 / 13-10-1 / SOS	196.22	99.67	26347	0.9509
MLP2 / 10-8-1 / SOS	200.65	97.15	29491	0.9522
MLP3 / 7-6-1 / SOS	178.88	98.33	27443	0.9516
MLP4 / 6-6-1 / SOS	187.39	98.23	26919	0.9516
MLP5 / 4-4-1 / SOS	181.64	99.52	26733	0.9510
MLP6 / 6-6-1 / SOS	172.04	98.65	25832	0.9514
MLP7 / 7-6-1 / SOS	187.63	97.97	27061	0.9517
MLP8 / 9-8-1 / SOS	186.17	97.51	27231	0.9520
RBF / 6-8-1 / SOS	2.2*10 ⁷	9.5*10 ¹⁰	4.5*10 ⁸	0.4887
M.REGR1 / / cor.coef.	218.66	109.52	33579	0.9459
M.REGR1a /with constant / cor.coef	474.87	124.69	35786	0.9459
M.REGR2 / / SOS	214.02	109.48	33689	0.9459
M.REGR3 / / MAPE	107.12	124.49	43894	0.9400
M.REGR4 / / cor.coef.	149.99	567.91	20336	0.9460
M.REGR5 / / SOS	213.66	109.32	33925	0.9460
M.REGR6 / / MAPE	107.12	124.48	43900	0.9400
AR(3)1 / p=3 / cor.coef	120.75	111.97	35676	0.9452
AR(5)1 / p=5 / cor.coef	138.65	110.47	34240	0.9460
AR(5)2 / p=5 / SOS	138.47	110.08	33482	0.9460
AR(5)3 / p=5 / MAPE	108.05	123.47	45268	0.9400
MOV.A1 /k=3/ cor.coef	131.86	128.83	25358	0.9367
MOV.A2 /k=3/ cor.coef	120.82	112.31	35474	0.9452
MOV.A3 /k=4/ cor.coef	135.19	111.74	34225	0.9453
MOV.A4 /k=5/ cor.coef	137.63	110.42	33558	0.9459

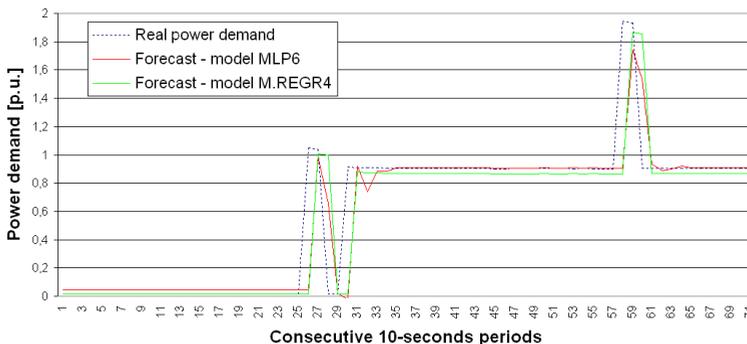


Fig. 6. The chosen forecast results for 71 10-second periods from hour 3:33. Source: Own elaboration.

It is also worth adding that in the autoregression models, the addition of a constant significantly worsened the forecast results. In moving average weighed models, the increase of k (number of back periods in model) reduced the SOS error of forecasts but

simultaneously increased the MAPE error of forecasts. An unexpected phenomenon is also fact that in all models of the moving average weighted after the estimation of the model parameters, the parameter related to the value of the power demand from the period $t-2$ is equal to almost zero.

It seems that from the point of view of the purposes of the forecasts it is more important to obtain greater accuracy of forecasts for the value of power demand other than very small values. Therefore, the measure of the SOS error of forecasts should be considered as a priority in our opinion.

However, if we treat SOS and MAPE errors of forecasts as equivalent numbers (without units) then the most advantageous models of the same quality are: the multiple regression model with the code M.REGR6 and the autoregression model having the code AR(5)3, thus quite simple construction models. In both models, however, the MAPE error is minimized. It is worth adding, which is very characteristic that in the case of both models the most important was the parameter assigned to the value of power demand from period $t-1$. Information on the power demand from periods from $t-2$ to $t-5$ has almost zero importance for these models, which makes these models very similar to the naive model, which has however significantly bigger forecast errors.

Noteworthy is also a very simple moving average weighed model ($k = 3$) due to the relatively small forecast errors of both MAPE and SOS (the sum of both errors as numbers is almost as small as for the models M.REGR6 and AR(5)3).

None of the forecasting models was able to effectively perform the forecasts for single 10-second very large increases in the power demand - the phenomenon is probably entirely random.

Definitely the worst model is the RBF model. In practice, this model is entirely not suitable for this type of forecasts.

4 Conclusions

The issue of very short-term power demand forecasting of big dynamics object was the subject of this paper. First detailed statistical analysis of available measurement data has been carried out. Then results of sample forecasts obtained by means of 6 selected methods (models) have been presented and compared.

Four measures of "ex-post" forecasts quality were used: MAPE error, SOS error, maximum percentage error and Pearson's linear correlation coefficient. The most important explanatory variable in almost all tested models was the value of power level from the $t-1$ period of time.

It is not possible to unambiguously choose the best model due to the use of these four measures of forecast quality with slightly different properties.

The most favorable values of forecast quality measures can be observed for MLP models (the smallest maximum percentage errors, the smallest SOS errors and the largest correlation coefficients). MAPE errors of MLP models are significantly bigger than in some other models. Multiple regression models and autoregressive models have much smaller MAPE errors of forecasts than MLP models. However, in these cases the SOS errors of forecasts of these models are significantly worse than in the MLP models. The models of the weighted moving average have MAPE errors of forecasts much smaller than in the MLP models.

The obtained forecast errors of the tested models are large (the smallest MAPE error was slightly above 107%). However, compared to the simplest naive model, the improvement in the quality of forecasts of the best models is very significant.

If we treat SOS and MAPE errors of forecasts as equivalent numbers (without units) then the most advantageous models of the same quality are: the multiple regression model

with the code M.REGR6 and the auto-regression model having the code AR(5)3. In the case of both models the most important was the parameter assigned to the value of power demand from period $t-1$. A very simple weighted moving average model ($k = 3$) results in relatively small forecast errors of both MAPE and SOS. Definitely the worst model is the RBF model.

None of the forecasting models was able to effectively perform the forecasts for single 10-second long very large increases in the power demand.

Data set used for calculations in this paper have been collected during the realization of the scientific project being a part of the ERA-Net Smart Grid Plus initiative, having a support from the European Union's Horizon 2020 research and innovation programme.

References

1. P. Shamsollahi, K.W. Cheung, Quan Chen, E.H. Germain, A neural network based very short term load forecaster for the interim ISO New England electricity market system, *Proc. of the 2001 Power Industry Computer Applications Conf.*, May 20 - 24, Sydney, Australia, 217 – 222, (2001)
2. M. Parol., P. Piotrowski, Very short-term load forecasting for optimum control in microgrids, *2nd International Youth Conference on Energetics*, Budapest, 4-6 June (2009)
3. M. Parol, P. Piotrowski, Prognozy zapotrzebowania na energię elektryczną o horyzoncie 15 minut na potrzeby sterowania w sieciach niskich napięć ze źródłami generacji rozproszonej, *Przegląd Elektrotechniczny (Electrical Review)*, **86**, 12, 303-309 (2010)
4. A. Khosravi, S. Nahavandi, P. Creighton, D. Srinivasan, Interval Type-2 Fuzzy Logic Systems for Load Forecasting: A Comparative Study. *IEEE Trans. on Power Syst.*, **27**, 3, 1274 – 1282 (2012)
5. P. Piotrowski, *Prognozowanie w elektroenergetyce w różnych horyzontach czasowych*, Prace Naukowe Politechniki Warszawskiej „Elektryka” **144** (Oficyna Wydawnicza Politechniki Warszawskiej, Warszawa, 2013)
6. D. Baczyński, K. Księżyk, M. Parol, P. Piotrowski, J. Wasilewski, T. Wójtowicz, *Mikrosieci niskiego napięcia. Praca zbiorowa pod redakcją M. Parola*, (OWPW, Warszawa, 2013)
7. L. Hernandez, C. Baladrón, J.M. Aguiar, Belén Carro, A.J. Sanchez-Esguevillas, J. Lloret, and J. Massana, A Survey on Electric Power Demand Forecasting: Future Trends in Smart Grids, Microgrids and Smart Buildings, *IEEE Comm. Surv. & Tut.*, **16**, 3, 1460-1495 (2014)
8. M. Cieślak, *Prognozowanie gospodarcze. Metody i zastosowania*. PWN (2001)
9. S. Osowski, *Sieci neuronowe do przetwarzania informacji*. (OWPW, Warszawa, 2006)