Forecasting of daily power consumption dynamics

S Y Petrova
Department of Intelligent Power Supply Networks, Institute of National Technology Initiative, Sevastopol State University, 33 Universitetskaya Street, Sevastopol, 299053 Russian Federation
E-mail: master@sevsu.energynet.ru

Abstract. The purpose of the study is to simulate the time series of power consumption in a power grid segment. It gives us be able to forecast the future state from values of the series according to current and past estimates. We want to test two different approaches: Box-Jenkins method and recurrent neural networks, for detect which of them give more accuracy and speed of prediction calculation. This article discusses the using Box-Jenkins method.

1. Introduction
The power supply continuity is largely determined by the speed of prevention and elimination of emergency situations. One of the promising ways for improving the efficiency of accident elimination is the development of decision support systems used in the maintenance and repair of local region power system. This task can be divided into the following subtasks: a digital description of the topology in the distribution electric network; segmentation of the obtained digital topology into typical auto clusters; the formation of a list of recommendations in each typical auto cluster in the context of the equipment mode; creating an algorithm for forming recommendations.

The clusterization power grid start from idea to split the electricity network into feeder associations (clusters) consisting of ring sections with automatic backup. Normally we segment this configuration into following typical clusters: I-cluster – one ring connected to two sources; A-cluster – two rings connected to two sources; Y-cluster – two rings connected to three sources; H-cluster – three rings connected to four sources.

The purpose of the study is to simulate the time series of power consumption in a power grid segment. It gives us be able to predict the future state from values of the series according to current and past estimates.

The source is the data log of mains load of a recloser exported from feeder's database via TELARM app (Tavrida Electric Automated Relay Manager). This database contains the parameters of the feeder, various libraries, as well as all characteristics of the recloser installed on the feeder. The data log of mains load is presented in the form of a table, with the values of current, voltage, active and reactive power recorded for each of the phases, as well as power factor and possible overflows (figure 1).

2. Results and discussion

2.1. The Box-Jenkins Method
We have a past set of data on the daily dynamics of energy consumption, in the form of active power values, as well as the date and time of taking records \( Y = [y_1, y_2, \ldots, y_{t-1}, y_t] \).
And we need predict the following series active power values $Y' = [y_{t+1}, y_{t+2}, ..., y_{t+k}]$, so that

$$\|Y_{t,k} - Y'_{t,k}\| \rightarrow \text{min}.$$  \hspace{1cm} (1)

We want to test two different approaches: Box-Jenkins method and recurrent neural networks, for detect which of them give more accuracy and speed of prediction calculation. This article discusses the Box-Jenkins method. In this approach, to predict future values of a series, we can use the previous values of the series

$$y_{t+1} = F(y_1, y_2, ..., y_{t-1}, y_t)$$ \hspace{1cm} (2)

and describe this process as combination autoregression and sliding mean into ARIMA [1]

$$\text{arima}(p, D, q) \sim \Delta^D y_t = c + \sum_{k=1}^{p} a_k \Delta^D y_{t-k} + \varepsilon_t + \sum_{k=1}^{q} b_k \varepsilon_{t-k},$$ \hspace{1cm} (3)

where $c$ – is a constant.

Autoregression we describe by simple regression model

$$y_t = \psi + a_1 y_{t-1} + a_2 y_{t-2} + a_3 y_{t-3} + \cdots + \varepsilon,$$ \hspace{1cm} (4)

where $\psi$ – is a constant, $a_i y_{t-i}$ – regression member is the previous member of the series, $\varepsilon$ – is error of point.

Sliding mean uses values of previous errors

$$y_t = \varepsilon_t - b_1 \varepsilon_{t-1} - b_2 \varepsilon_{t-2} - b_3 \varepsilon_{t-3} - \cdots + \mu,$$ \hspace{1cm} (5)

where $\mu$ – is a constant, $b_i \varepsilon_{t-i}$ – is error.

To forecast energy discharge, the Box-Jenkins approach was used in the form of a sequence of steps: data preprocessing, model selection, model estimation, simulation and prediction.

2.2. Forecasting Procedure

The purpose of data preprocessing is mapping the data to a stationary form. Figure 1 shows initial data source time dependence of power consumption.

![Time Dependence Power Consumption](image)

**Figure 1.** Time Dependence Power Consumption (Reactive Volt-Ampere per Hour).

Data preprocessing is divided into mini - blocks: trend identification, stationarity test and seasonal identification. That give us understanding what is present in the data. For recognition stationary form in data we visual analyzing calculated hourly average for a more than half-year period and RMS deviation (figure 2). That figures shows us non-stationary form of source data, because RMS and mean of $P$ is inconstant.
We must apply several filters to source and delete following components:

\[ y_t = T_t + S_t + C_t \]  

(6)

where \( T_t \) – is a trends; \( S_t \) – seasonal fluctuation; \( C_t \) - business cycle.

Several methods are used for this: trend removal and exponential, linear, polynomial, seasonal components removal. For example, to remove exponential trend from the source data to bring data into a stationary form we used (7).

\[ T_t = \exp(at) \sim y_t - T_t = \log(\exp(at)) = at \]

(7)

The next we show a few stationary testing:

- Test Kwiatkowski, Phillips, Schmidt and Shin if \( H = 0 \) the series is stationary;
- Leybourne test - McCabe if \( H = 0 \), the series is stationary;
- Philips test - Perron if \( H = 0 \), the roots of the characteristic polynomial lie inside the unit circle;
- Advanced Dickey - Fuller if \( H = 0 \), the roots of the characteristic polynomial lie inside the unit circle.

The most common one is the Dickey-Fuller test [2].

\[ t = \frac{\bar{\alpha} - \alpha}{\hat{\sigma}_\alpha}, \]

(6)

\[ y_t = \alpha y_{t-1} + \varepsilon_t, \]

(7)

\( H_0: \alpha = 1, \)

\( H_1: \alpha < 1, \)

\[ \lim_{t_1} t_1 = \frac{W^2(t) - 1}{\frac{1}{n} \int_0^1 W^2(s) \, ds}. \]

(8)

If \( H < 1 \), then the series is still stationary, therefore we need to take the difference again or do something else to bring the data to a stationary form. But the use of multiple differences can make the data stationary, but at the same time we refuse part of the data, and in fact compress it and lose some dimensions.
An alternative approach is to local the trends and then subtract these trends from the source data. Dimension remains the same, but the trend is singled out. In our case, we singled out a polynomial of the third degree and as a result obtained $H = 0$, which says that the data are stationary (figure 3).

![Figure 3](image_url)

**Figure 3.** After applied test Dickey – Fuller to data, $H=0$.

There are two more elements that we have to locking for in the data: seasonal trends and non-cyclical trends. To identify the seasonal trend, we use the Fourier transform. The definition of a one dimensional continuous function, denoted by $f(x)$, the Fourier transform is defined by:

$$F(u) = \int_{-\infty}^{\infty} f(x) \exp(-i2\pi ux) \, dx$$

The next step is count the Fourier component, that is, we get the number of elements $N$ in the array. But $N$ complex Fourier components in the frequency domain count as $2N$ real numbers if we consider both real and imaginary components. The identity above constrains them so that there are only $N$ independent real values involved.

Then we calculate the power – Fourier transform, multiply by the complex conjugate and divide by the number of Fourier component. We use the inverse tangent with four quadrants for numeric and symbolic arguments.

To isolate business cycles in data, we used the Hodrick-Prescott filter [3].

$$\sum_{t=1}^{T} (y_t - s_t)^2 + \lambda \sum_{t=2}^{T-1} \left( (s_{t+1} - s_t) - (s_t - s_{t-1}) \right)^2 \rightarrow \min$$

It consists of two parts: The first part is to fit our original row as close as possible to the new row, and the second part is responsible for ensuring that the row obtained after filtration would be as smooth as possible. That is, on the one hand, we minimize the difference between these series, on the other hand, we maximize smoothness. Parameter $\lambda$ is applied for different periods, the value of this parameter for monthly data was obtained empirically and equal 14400.

Once we have given a series to a stationary form, now we can simulate it using arima $(p, D, q)$ models and is forecasting of daily power consumption dynamics. Parameters $P$, $D$, $Q$ are those parameters on which the degree of the model depends.

When modeling a series using the ARIMA model (3), you need to understand how many significant members are in it. There are methods for choosing these ones by the autocorrelation function and by the partial autocorrelation function. Because there are also autoregression $AR(p)$ and sliding mean $MF(q)$ components in ARIMA model, then significant members often overlap, therefore makes their choice difficult. This means that MA members and auto regression members can influence each other and, accordingly, either neutralize the influence or create unnecessary effects. Let’s make a list of
candidate models and our task is to choose the model that best fits into the data table 1 and 2. The choice of the model is based on the assessment of statistical characteristics.

### Table 1. ARIMA (2,0,3) Model Seasonally Integrated with Seasonal AR (12) (Gaussian Distribution).

|                | Value | Standard Error | T Statistic | P Value |
|----------------|-------|----------------|-------------|---------|
| Constant       | 0     | 0              | NaN         | NaN     |
| AR \{1\}       | 0.41743 | 0.20923      | 1.9951      | 0.04603 |
| AR \{2\}       | 0.37312 | 0.19222      | 1.9411      | 0.05225 |
| SAR \{12\}     | -0.35671 | 0.05914      | -6.0316     | 1.6232e-09 |
| MA \{1\}       | -0.08295 | 0.20247    | 0.40971     | 0.68202 |
| MA \{2\}       | 0.022238 | 0.13219     | -0.16823    | 0.8664 |
| MA \{3\}       | 0.28413 | 0.076746 | 3.7022     | 0.00021374 |
| Variance       | 2.6823e-05 | 2.1458e-06 | 12.5       | 7.4678e-36 |

Description:
- P: 26
- D: 0
- Q: 3
- Constant: 0
- AR: \{NaN NaN\} at lags \[1 2\]
- SAR: \{NaN\} at lag [12]
- MA: \{NaN NaN NaN\} at lags \[1 2 3\]
- SMA: {}
- Seasonality: 12
- Beta: \[1×0\]
- Variance: NaN

### Table 2. ARIMA (3,0,4) Model Seasonally Integrated with Seasonal AR (12) (Gaussian Distribution).

|                | Value | Standard Error | T Statistic | P Value |
|----------------|-------|----------------|-------------|---------|
| Constant       | 0     | 0              | NaN         | NaN     |
| AR \{1\}       | -0.088388 | 0.20888       | -0.42315    | 0.67218 |
| AR \{2\}       | 0.17562 | 0.13675      | 1.2842      | 0.19906 |
| AR \{3\}       | 0.56979 | 0.16743      | 3.4032      | 0.00066608 |
| SAR \{12\}     | -0.38074 | 0.057775    | -6.5901     | 4.3956e-11 |
| MA \{1\}       | 0.41836 | 0.20649      | 2.026       | 0.042763 |
| MA \{2\}       | 0.37581 | 0.15358      | 2.447       | 0.014406 |
| MA \{3\}       | 0.038531 | 0.13489     | 0.28564     | 0.77515 |
| MA \{4\}       | 0.15485 | 0.099633    | 1.5542      | 0.12015 |
| Variance       | 2.6749e-05 | 2.1769e-06 | 12.288     | 1.0567e-34 |

Description:
- P: 27
- D: 0
- Q: 4
- Constant: 0
- AR: \{NaN NaN NaN\} at lags \[1 2 3\]
- SAR: \{NaN\} at lag [12]
- MA: \{NaN NaN NaN NaN\} at lags \[1 2 3 4\]
- SMA: {}
- Seasonality: 12
- Beta: [1×0]
- Variance: NaN

For forecasting, we choose the second model, since it has the best characteristics.

3. Conclusion

We got that there are two significant members in the model. At the output, we get the model variables hpi_model already with the adjusted coefficients, which were obtained based on the data that we passed to the model.

After performing the forecast within the sample, we got the following result table 3 and table 4.

**Table 3. ARIMA (2,0,3) Model Seasonally Integrated with Seasonal AR (12) (Gaussian Distribution).**

|   | Value | Standard Error | T Statistic | P Value |
|---|-------|----------------|-------------|---------|
| Constant | 0     | 0              | NaN         | NaN     |
| AR {1}   | 0.41743 | 0.20923        | 1.9951      | 0.04603 |
| AR {2}   | 0.37312 | 0.19222        | 1.9411      | 0.0525  |
| SAR {12} | -0.35671 | 0.05914        | -6.0316     | 1.6232e-09 |
| MA {1}   | -0.082953 | 0.20247       | -0.40971   | 0.68202 |
| MA {2}   | -0.022238 | 0.13219       | -0.16823   | 0.8664  |
| MA {3}   | 0.28413  | 0.076746       | 3.7022     | 0.00021374 |
| Variance | 2.6823e-05 | 2.1458e-06     | 12.5       | 7.4678e-36 |

Effective Sample Size: 229
Number of Estimated Parameters: 7
LogLikelihood: 880.289
AIC: -1746.58
BIC: -1722.54

**Table 4. ARIMA (3,0,4) Model Seasonally Integrated with Seasonal AR (12) (Gaussian Distribution).**

|   | Value | Standard Error | T Statistic | P Value |
|---|-------|----------------|-------------|---------|
| Constant | 0     | 0              | NaN         | NaN     |
| AR {1}   | 0.090771 | 0.66163        | 0.13719     | 0.89088 |
| AR {2}   | 0.08881  | 0.98863        | 0.089832    | 0.92842 |
| AR {3}   | 0.081954 | 0.80356        | 0.10199     | 0.91877 |
| SAR {12} | -0.10805 | 0.068317       | -1.5816     | 0.11373 |
| MA {1}   | 0      | 0.67484        | 0           | 1       |
| MA {2}   | 0      | 1.0431         | 0           | 1       |
| MA {3}   | 0      | 0.72675        | 0           | 1       |
| MA {4}   | 2e-12  | 0.11654        | 1.7162e-11  | 1       |

We checked the heteroscedasticity of the model, that is, the dependence on variance: ARCH described (11), (12) and GARCH described (13). Result shown in table 5.

\[
y_t = \varepsilon_t \sqrt{\alpha_0 + \sum_{i=1}^{q} \alpha_i y_{t-i}^2}
\]

(11)

\[
\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i y_{t-i}^2
\]

(12)

\[
\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i y_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2
\]

(13)
Table 5. GARCH (0,1) Conditional Variance Model (Gaussian Distribution).

|          | Value   | Standard Error | T Statistic | P Value   |
|----------|---------|----------------|-------------|-----------|
| Constant | 2.3678e-05 | 4.7808e-06     | 4.9527      | 7.3175e-07 |
| ARCH {1} | 0.59282   | 0.18975        | 3.1243      | 0.0017825  |

Sample error: 0.0042 is a good result.

So, ARIMA predicts gets good forecasting of daily power consumption dynamics when there is not a lot of data, but we suppose when there is a lot of data, neural networks will give us better results. This one next research.

References
[1] Box G E, Jenkins G M, Reinsel G C and Ljung G M 2015 *Time Series Analysis: Forecasting and Control* 5th Edition (Hoboken, New Jersey: Published by John Wiley and Sons Inc.) p 712
[2] Dickey D A and Fuller W A 1979 Distribution of the Estimators for Autoregressive Time Series with a Unit Root *Journal of the American Statistical Association* **74** (366) 427–431
[3] Arajou F, Areosa M B and Neto J A 2003 r-filters: a Hodrick-Prescott Filter Generalization *Reproduction permitted only if source is stated as follows Working Paper Series* **69** 1–37