Model of electricity consumption in a manufacturing company – time series decomposition

M Michalková¹ and P Ďurčanský²

¹ Department of Applied Mathematics, Faculty of Mechanical Engineering, University of Žilina, Univerzitná 1, 010 26 Žilina, Slovakia
² Department of Power Engineering, Faculty of Mechanical Engineering, University of Žilina, Univerzitná 1, 010 26 Žilina, Slovakia

E-mail: maria.michalkova@fstroj.uniza.sk

Abstract. Electricity has become a commodity capable of being bought, sold or traded. An electricity market is a system enabling purchases, through bids to buy or offers to sell, which can be useful for companies to lower electricity costs by short term trades. This new approach demands new ways of collecting, analysing and predicting not only electricity price, but in first place electricity consumption. The aim of the paper is to model electricity consumption of a manufacturing company. The trend, seasonal and random component, respectively, are identified.

1. Introduction

Time series is a set of observations generated sequentially over time. If the set is continuous, the time series is said to be continuous; if the set is discrete, the time series is said to be discrete [1]. The observations in a discrete time series are usually recorded at equispaced times; based on the frequency of the record we distinguish among long-term time series (data recorded on annual basis), short-term time series (data recorded in time intervals shorter than a year – weekly, monthly, quarterly) and high-frequency time series (data observed daily, hourly or on a finer time scale).

The aim of the time series analysis is dual. First, it is to obtain the structure of data and possible underlying pattern to understand the mechanism of generation of the series. Secondly, to construct the model of the process to help forecast future values of the observed variable in the terms of probability, and to control and optimize the system. Accurate energy forecasting models are necessary for planning and energy optimization. There exists a number of approaches in modelling time series depending on the character of the data. In [2] the adaptive techniques to identify the long-term development of energy consumption in a university campus are compared. The majority of techniques for building the model are now computer-aided, applying machine learning and neural networks. For review of such methods see, for example [3, 4].

The industry sector accounted for 37% (156 EJ) of total global final energy use in 2017. This represents a 1% annual increase in energy consumption since 2010 [5]. The purpose of the paper is to model the energy consumption in a manufacturing company providing hot and cold bending. Generally, processes of forming as well as machining are energy demanding. Moreover, forming processes have lower energy efficiency than machining even though they are much more energy intensive [6]. With growing consumption of electricity caused by growing demand from households, e-mobility and intensification of production in industry, the energy prices are increasing rapidly. Due
to the high cost of electricity, which is the main source for operation of the considered company, it is necessary to look for a cost-effective energy supply solution that meets the needs in terms of primary energy prices. The model of the considered time series might help to understand the character of the energy demand and to help optimize the system in the future. The method of analysis, applied on the time series, is decomposition method. Generally, when decomposing the series, we assume that the time series consists of a several specific components, namely:

- trend component,
- seasonal component,
- cyclic component,
- residual component.

Trend $T_{rt}$ indicates a long term change in the mean level of the series. Periodic components describe a pattern repeating in time. Seasonal component $S_{zt}$ models periodic changes in the time series that appear throughout the year. The cyclic component $C_t$ represents the long-term fluctuations around the trend. Even though this component is considered as periodic, the length of the cycle as well as its intensity may vary during the process. This component mirrors the technological, political, business, production and other changes. The residual component $E_t$ covers the random fluctuations in the series that miss any systematic character. It is assumed that the residual component is a white noise. The trend, seasonal and cyclic component are together called a systematic components because of their deterministic character. The time series does not need to include all of these components; however, the residual component is always present. The decomposition of the time series can take on two different forms:

- additive decomposition

$$y_t = T_{rt} + S_{zt} + C_t + E_t;$$  \hspace{1cm} \text{(1)}$

- multiplicative decomposition

$$y_t = T_{rt} S_{zt} C_t E_t,$$  \hspace{1cm} \text{(2)}$

where $y_t$ denotes the value of the time series in time $t$ [7].

The paper is organized as follows: in the section 2 the modeled data are described. Section 3 includes the model of systematic components – the trend and the seasonal one, respectively. In section 4 the residual analysis is conducted, and the ARMA model of residuals is identified. In section 5 the criteria for considering the forecast fit of the model to testing dataset are calculated. All of the calculations in the paper are conducted in software Matlab.

2. Data

Data generating the considered time series represent an energy consumption in (kW) recorded on an hourly basis and then averaged per day. The data were collected in the period from 1st March 2016 to 1st May 2017. The time series totally counts of 427 daily observations. The data are divided into two sets: the train set consists of the first 378 observations and the test set includes the rest of the time series. The descriptive statistics of the whole time series is given in the table 1.

| Table 1. Descriptive statistics of energy consumption time series. |
|---|---|---|---|---|---|---|---|---|---|
| Min | Max | Mean | Variance | Lower Quantile | Median | Upper Quantile | Skewness | Kurtosis |
| 31.410 | 820.573 | 327.452 | 43 290.66 | 97.937 | 354.753 | 504.641 | 0.025 | -1.183 |

Before the decomposition process is started, we analyse the time series graphically. In the figure 1 the time series of energy consumption from 1st March 2016 to 13th March 2017 (corresponding to the train set) is plotted. As we can see, the series appears to be of strong periodic nature. There is no evident increasing or decreasing trend in the series. Thus we may assume that the trend is constant.
3. Model of systematic components

In this section we identify the deterministic components of the time series.

3.1. Seasonal component

Among the variety of methods for modelling periodic behaviour of the time series, the analysis in spectral domain is chosen. Due to this approach we may identify all of the significant periods that occur in the data.

Thus, when modelling the seasonal component, we assume that it could be expressed as a sum of sine and cosine functions:

\[ S_{zt} = \sum_{i=1}^{k} (\alpha_i \cos(\omega_i t) + \beta_i \sin(\omega_i t)), \quad t = 1, 2, ..., n \]  

(3)

where \( \omega_i, i = 1, 2, ..., k \) are frequencies from the interval \((0, \pi)\) corresponding to \( k \) periodic components in the series, \( \alpha_i, \beta_i, i = 1, 2, ..., k \) are amplitudes.

![Figure 1. The time series of energy consumption from 1st March 2016 to 13th March 2017.](image)

In order to find the seasonal component of the time series in the form (3), we need to identify the periodicities occurring in the series. For finding the frequencies we apply Fisher’s periodicity test.

Let the time series is given in the form:

\[ y_t = \theta_t + \epsilon_t, \quad t = 1, 2, ..., n, \]  

(4)

where \( \theta_t \) denotes the deterministic component of the series and \( \epsilon_t \sim N(0, \sigma^2) \). We test the null hypothesis that there is no periodic activity:

\[ H_0: \theta_1 = \theta_2 = \cdots = \theta_n = 0. \]  

(5)
The test statistics is based on values of the periodogram:

\[ I(\omega) = \frac{1}{4\pi} \left[ a^2(\omega) + b^2(\omega) \right], \quad 0 \leq \omega \leq \pi, \quad (6) \]

where

\[ a(\omega) = \frac{2}{n} \sum_{t=1}^{n} y_t \cos(\omega t), \quad (7) \]
\[ b(\omega) = \frac{2}{n} \sum_{t=1}^{n} y_t \sin(\omega t). \quad (8) \]

Fisher’s test uses values of periodogram (6) for Fourier frequencies:

\[ \omega_j = \frac{2\pi j}{n}, \quad j = 1, 2, ..., m, \quad (9) \]

where \( m = \left\lfloor \frac{n-1}{2} \right\rfloor \) is the integer part of sample size \( n \). The values of the periodogram are normalized to the form:

\[ X_j = \frac{I(\omega_j)}{\sum_{i=1}^{m} I(\omega_i)}. \quad (10) \]

The hypothesis \( H_0 \) is rejected at the significance level \( \alpha \) if:

\[ W = \max_{j=1, \ldots, m} X_j \geq g_\alpha, \quad (11) \]

where \( g_\alpha \) is a critical value of this test (the way to calculate the critical values of Fisher’s test can be found in [8]). In case of rejection of the null hypothesis, we find as well the frequency of the periodic component [7]. However, this way only the highest value of periodogram can be identified. In case of compound periodicity (the existence of a number of periodicities), Whittle [9] suggested the following modification of the test:

When the null hypothesis is rejected for the maximal value \( X_j \) of the sampled periodogram, this value is omitted and the Fisher’s test is repeated for the rest of the frequencies. This way we continue until the null hypothesis is not rejected, i.e. there are no periodic components in the time series.

The sample periodogram of the time series is in figure 2. The frequencies identified via the Fisher’s test are given in table 2.

![Sample periodogram](image.png)

\[ \text{Figure 2. The sample periodogram.} \]
Table 2. Frequencies of the seasonal component for the energy consumption time series.

| j  | \( \omega_j \) | period T (days) |
|----|----------------|-----------------|
| 54 | 0.8976         | 7               |
| 108| 1.7952         | 3.5             |
| 1  | 0.0166         | 378             |
| 2  | 0.0332         | 189             |
| 4  | 0.0665         | 94.5            |
| 8  | 0.1330         | 47.25           |
| 11 | 0.1828         | 34.36           |
| 3  | 0.0499         | 126             |
| 162| 2.6928         | 2.33            |

As we see from the periodogram in figure 2 and from results of Fisher’s test in table 2, the most dominant period is that of 7 days – a week. The second dominant period is of 3.5 days – this one may represent the alteration of weekdays and weekend.

When all of the significant frequencies are identified, we can find the estimates \( a_i, b_i \) of the amplitudes \( \alpha_i, \beta_i \) in (3). The estimates of coefficients are calculated using the ordinary least square method (OLS) – in the form (7), (8) for significant frequencies. The estimates of the amplitudes are in the table 3.

Table 3. Estimates of the amplitudes \( a_j, b_j \) for frequencies \( \omega_j \).

| i  | \( \omega \) | \( \alpha \)  | \( \beta \)  | \( \text{p-value} \) |
|----|-------------|--------------|--------------|-----------------|
| 0  | 0.8976      | -33.818      | 187.150      | (9.00e-5)       |
| 1  | 1.7952      | 75.998       | 39.502       | (2.77e-17)      |
| 2  | 0.0166      | 36.353       | 41.918       | (2.64e-5)       |
| 3  | 0.0332      | 52.985       | -15.192      | (1.50e-9)       |
| 4  | 0.0665      | -29.757      | -43.391      | (5.52e-4)       |
| 5  | 0.1330      | 40.071       | 9.242        | (3.82e-6)       |
| 6  | 0.1828      | -4.2698      | 39.210       | (0.62)          |
| 7  | 0.0499      | 23.525       | 26.830       | (6.16e-3)       |
| 8  | 2.6928      | 33.990       | -6.424       | (8.30e-5)       |
| 9  |             |              |              |                 |

Since not all of the periodical components need to consist of both the sine and cosine element, we test each coefficient whether it is statistically significant. We test the hypothesis:

\( H_0: \) the coefficient is equal to 0,

\( H_A: \) the coefficient is not equal to 0.

The hypothesis \( H_0 \) is rejected at significance level \( \alpha \) if \( p - value < \alpha \).

Omitting those sine and cosine elements, that are not statistically significant at significance level \( \alpha = 0.05 \) (their \( p-value \) is higher than 0.05), the estimate of the seasonal component of the time series is in the form:

\[
S_t = \sum_{i=1}^{9} (a_i \cos(\omega_i t) + b_i \sin(\omega_i t)), \quad t = 1,2,\ldots,378
\]

with coefficients \( a_i, b_i \) from table 3 (\( a_7 = b_4 = b_6 = b_9 = 0 \)). The seasonal component (upper figure) and the deseasoned time series (lower figure) are in figure 3.
Figure 3. The seasonal component of the time series (upper); the deseasoned time series (lower).

3.2. Trend component
After subtracting the seasonal component from the original time series, we obtain the deseasoned series. Now we can proceed to trend modelling. The trend component is usually modelled by a mathematical function, in majority of cases by a polynomial. We can see in figure 3 that the time series does not show any evident trend. We try to fit the linear trend in the form:

\[ T_r = \beta_0 + \beta_1 t \]  \hspace{1cm} (13)

to the deseasoned data. The estimates of coefficients are found by OLS method – table 4.

| Estimate \( \beta_0 \) | Estimate \( \beta_1 \) | p-value  |
|-------------------------|------------------------|----------|
| \( \hat{\beta}_0 \)    | 313.096                | 1.925e-87|
| \( \hat{\beta}_1 \)    | 0.068                  | 0.215    |

We see that the slope of the line is not statistically significant on the significance level \( \alpha = 0.05 \). Thus the trend component of the time series is constant, in the form: \( T_r = 313.1 \).
Figure 4. The constant trend component + seasonal component fitted to the data (upper); the residual component of the time series after eliminating the systematic part (lower).

4. Residual analysis

Residual analysis is an essential part of the time series modelling. After subtracting the systematic components of the series, the residual component needs to fulfil the conditions of white noise:

The probabilistic model of random variables \( \epsilon_t \) is said to be white noise when for \( \forall t \in T \subset R \) [10]:

- \( E(\epsilon_t) = 0 \),
- \( D(\epsilon_t) = \sigma^2 \),
- \( cov(\epsilon_t, \epsilon_{t-k}) = cov(\epsilon_t, \epsilon_{t+k}) = 0 \),
- \( \epsilon_t \sim N(0, \sigma^2) \).

This means that the residuals:

\[
\hat{\epsilon}_t = y_t - \hat{y}_t, \tag{14}
\]

\( y_t \) the original time series, \( \hat{y}_t \) the values estimated by the model, must be mutually uncorrelated and distributed approximately normally about zero with constant variance \( \sigma^2 \).

First of all we check the third condition, i.e. whether the residuals are mutually uncorrelated. For such purpose we estimate the autocorrelation function. The estimates of autocorrelation coefficients are given by [1]:

\[
r_k = \frac{\sum_{t=k+1}^{T} \epsilon_{t-1} \epsilon_{t-k}}{\sum_{t=1}^{T} \epsilon_{t-1}^{2}}, r_0 = 1; \quad k > 1. \tag{15}
\]

To decide whether the residuals of the model are autocorrelated, we apply Bartlett’s test. We are testing the hypothesis:

\( H_0: \) random variables \( \hat{\epsilon}_t \) and \( \hat{\epsilon}_{t-k} \) are linear independent,

\( H_A: \) random variables \( \hat{\epsilon}_t \) and \( \hat{\epsilon}_{t-k} \) are linear dependent.
The null hypothesis is rejected on the significance level \( \alpha \) if:

\[ |r_k| > z_{1-\frac{\alpha}{2}} \sqrt{\frac{1}{n} \left( 1 + 2 \sum_{i=1}^{k-1} r_i^2 \right)}, \]  

(16)

where \( z_{1-\frac{\alpha}{2}} \) are critical values of standard normal distribution [1]. The autocorrelation function of the residuals \( \tilde{\epsilon}_t \) together with the boundaries given by the Bartlett’s test are shown in figure 5.

![Figure 5](image)

**Figure 5.** The sample autocorrelation function alongside with the boundaries of the Bartlett’s test.

We reject the null hypothesis on the significance level \( \alpha = 0.05 \). The residuals are autocorrelated because the coefficient \( r_1 \) exceeds the boundaries given by Bartlett’s test. In such situation the model does not describe the past of the time series adequately and it needs to be supplemented.

### 4.1. ARMA model of residuals

The ARMA models are applied to model the stochastic development of the time series, i.e. to model the random component of the series. The ARMA model is of the form:

\[ w_t = \sum_{i=1}^{p} \varphi_i w_{t-i} + \varepsilon_t - \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} \]  

(17)

where \( \varphi_i \) are autoregressive coefficients, \( \theta_j \) are moving average coefficients and \( \varepsilon_t \) is white noise. Such models allow us to describe the situation when the residuals are determined by their previous values, as well as the situation that the time series reacts to a series of random shocks. Based on the values of autocorrelation coefficients \( r_k \) and partial autocorrelation coefficients \( r_{kk} \) we judge on the structure of the ARMA model. The partial autocorrelation coefficients are given as [1]:

\[ r_{kk} = \frac{r_k - \sum_{j=1}^{k-1} r_{k-j} r_{k-j}}{1 - \sum_{j=1}^{k-1} r_{k-j} r_{k-j}}, \quad r_{11} = r_1; \quad k > 1, \]

\[ r_{kj} = r_{k-1,j} - r_{kk} r_{k-1,j}, \quad j = 1, 2, \ldots, k - 1. \]  

(18)

The autocorrelation coefficients and the partial autocorrelation coefficients are given in figures 6 and 7, respectively.
We decided to model the residuals by AR(18) model. Based on the estimation of coefficients of the model conducted in Matlab we obtain the following results given in table 5.

### Table 5. Estimates of autoregressive coefficients.

| Estimate | AR1  | AR6  | AR8  | AR14 | AR18 |
|----------|------|------|------|------|------|
|          | 0.2728 | 0.1332 | -0.1144 | 0.1048 | -0.1317 |

All of the other coefficients are not statistically significant on significance level $\alpha = 0.05$, i.e. they are equal 0. The estimation of the variance of the random component $\sigma^2 = 11.164$. According to results, we see that the system reacts to its values with delays of 1, 6, 8, 14 and 18 days.

From figure 8 we see that the residuals of AR18 model are not correlated since they do not exceed the boundaries equal to $2\sqrt{\frac{2}{n}}$.

When the condition of uncorrelated residuals is fulfilled, we can test the other conditions. The estimate of the mean of the AR18 model residuals is $\bar{\varepsilon} = 0.00084$ (we may conclude that the mean is zero). The histogram of the residuals of AR18 model is in figure 9. Based on the Anderson Darling test with $p-value = 0.08$ we can conclude that the residuals are normally distributed. The homoscedasticity of residuals (the constant variance) is tested using the following test: The residuals are divided into two groups of the same size and we are testing the hypothesis:

$H_0: \sigma^2_{\varepsilon_1} = \sigma^2_{\varepsilon_2}$,

$H_A: \sigma^2_{\varepsilon_1} \neq \sigma^2_{\varepsilon_2}$.

The test statistics is:

$$F = \frac{s_{\varepsilon_1}^2}{s_{\varepsilon_2}^2} \approx F\left(\frac{n}{2} - 1; \frac{n}{2} - 1\right).$$

(19)
The null hypothesis is rejected on the significance level $\alpha$ if $F < F_{\alpha/2}$ or $F > F_{1-\alpha/2}$. In our case the null hypothesis is not rejected, based on the $p$-value of the test that is 0.0621. Therefore the residuals are homoscedastic.

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure8.png}
\caption{Autocorrelation function of the residuals of AR18 model.}
\end{figure}

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure9.png}
\caption{Histogram of the residuals of AR18 model.}
\end{figure}

Since the conditions on the residual component of the time series decomposition are fulfilled, we may write the final form of the model. That is

$$w_t = y_t - 313.1 - \sum_{i=1}^{9} (a_i \cos(\omega_i t) + b_i \sin(\omega_i t)), \quad a_7 = b_4 = b_6 = b_9 = 0,$$

$$w_t = 0.2728 w_{t-1} + 0.1332 w_{t-6} - 0.1144 w_{t-8} + 0.1048 w_{t-14} - 0.1317 w_{t-18} + \varepsilon_t,$$

$$\hat{\sigma}^2 = 11164$$

(20)

where values of $a_i, b_i, \omega_i$ are from table 2.
5. Forecast adequacy of the model

Finally, we test the forecast ability of the model. In order to analyse the fit of the prediction on the testing data, we calculate the following criteria [10]:

- mean error (ME):
  \[ ME = \frac{1}{n} \sum_{i=1}^{n} (y_t - \hat{y}_t), \]

- mean absolute error (MAE):
  \[ MAE = \frac{1}{n} \sum_{i=1}^{n} |y_t - \hat{y}_t|, \]

- root mean square error (RMSE):
  \[ RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_t - \hat{y}_t)^2}, \]

- mean absolute percentage error (MAPE):
  \[ MAPE = \frac{1}{n} \sum_{i=1}^{n} \frac{|y_t - \hat{y}_t|}{y_t} \times 100\% \]

- coefficient of determination \( R^2 \):
  \[ R^2 = \frac{\sum_{i=1}^{n} y_t^2 - n\bar{y}^2}{\sum_{i=1}^{n} y_t^2 - n\bar{y}_t^2} \]

The results of the criteria for model (20) compared to testing set of time series are given in table 6.

| ME   | MAE   | RMSE  | MAPE   | \( R^2 \) |
|------|-------|-------|--------|---------|
| -73.2032 | 112.5378 | 133.8481 | 133.7916 | 0.6497 |

Based on the ME value we can conclude that the model systematically overestimates the real situation. The fit of the model on the test dataset is in figure 10.

![Figure 10. The fit of the model to the test dataset.](image-url)
6. Conclusion
In the paper the model of the time series representing the electricity consumption in a manufacturing company has been derived. Using the decomposition of the time series the deterministic components – trend and seasonal one, respectively, were identified. The trend of the time series is constant, equal to 313.1. The seasonal component consists of nine periodic components which frequencies were found by application of Fisher’s test. The most dominant frequency refers to 7 day period. Since the residual component of the series showed the autocorrelation, the model of residuals based on the Box-Jenkins methodology needed to be done. The autoregressive model AR18 was applied. The forecast ability of the model was tested. Based on the value of ME, the model systematically overestimates the real situation. The model fits the data on 64 % (according to coefficient of determination).

7. References
[1] Box G, Jenkins G, Reinsel G and Ljung G 2016 Time Series Analysis. Forecasting and Control (New Jersey: Wiley) p 709
[2] Lee Y W, Tay K G and Choy YY 2018 Forecasting electricity consumption using time series model International Journal of Engineering and Technology 7 218–223
[3] Deb C, Zhang F, Yang J, Lee S E and Shah K W 2017 A review on time series forecasting techniques for building energy consumption Renewable and Sustainable Energy Reviews 74 902–924
[4] Öğcu G, Demirel OF and Zaim S 2012 Forecasting electricity consumption with neural networks and support vector regression Procedia – Social and Behavioral Sciences 58 1576–1585
[5] Fernandez-Pales A, Levi P and Vass T 2019 Tracking Clean Energy Progress – Industry. Available from: https://www.iea.org/tcep/industry/
[6] Gao M, Huang H, Wang Q, Liu Z and Li X 2018 Energy consumption analysis on sheet metal forming: focusing on the deep drawing processes The International Journal of Advanced Manufacturing Technology 96 3893–3907
[7] Cipra T 2013 Financial econometry (Praha: Ekopress) p 538 (in Czech)
[8] Cipra T 1984 Investigation of periodicity for dependent observations Aplikace matematiky 29 134–142
[9] Whittle P 1952 The simultaneous estimation of time series, harmonic components and covariance structure Trabajos de Estadistica 3 43–57
[10] Rublíková E 2007 Time series analysis (Bratislava: Iura Edition) p 208 (in Slovak)

Acknowledgments
This paper was supported by the Slovak Grant Agency KEGA through the projects No. 040ŽU-4/2018 and KEGA-063ŽU-4/2018 and by the Slovak Grant Agency VEGA through the project VEGA-1/0738/18.