Prediction of Hydropower Ratio from Total Energy Generation in Romania

Angela Neagoe [1], Eliza-Isabela Tică [1], Bogdan Popa [1], Florica Popa [2]

[1]University Politehnica of Bucharest, Splaiul Independenței no. 313, Bucharest, Romania
[2]University VALAHIA of Târgoviște, Aleea Sinaia, no. 13, Târgoviște, Romania

***Abstract* – In a previous paper, it was made a statistical analysis for hydropower ratio from total energy production in Romania over the period 2006 – 2016 for monthly values, and the main characteristics for this set of observations were determined. Based on monthly recorded data and on mathematical model, this paper presents a method for the prediction of this random variable using the second order autoregressive (AR2) model.**

The reconstructed data are compared with recorded values for the last three years of the analyzed time series and, for the next two years, a forecast was made, whose results are compared with real values published in the annual reports of Hidroelectrica, for 2017, and the Transelectrica Website, for 2018.

***Keywords* – statistical analysis, time series, AR2 model, hydropower generation, confidence interval.**

1. Introduction

Analysis and prediction using time series achieved by chronological recording of measured data analysis are used in many fields, among which meteorological and hydrological ones [1], [2], [3].

The fact is known that forecasting inflows in a reservoir can be performed with this kind of models.

DOI: 10.18421/TEM91-55
https://dx.doi.org/10.18421/TEM91-55

Corresponding author: Bogdan Popa,
University Politehnica of Bucharest,Bucharest, Romania
Email: bogdan.popa@upb.ro

Received: 10 October 2019.
Revised: 30 January 2020.
Accepted: 06 February 2020.
Published: 28 February 2020.

Moreover, if the reservoir is used to supply water to a hydropower plant, the production of energy can be such a forecasted variable as well.

A forecasting cannot be done with high precision due to the other factors which influence the energy production, even if there is a strong relation between the inflows into a reservoir and the energy production in the related hydropower plant.

There are many papers published related to hydropower engineering, regarding the great interest in forecasting and optimizing energy production from water use. With this kind of forecast the development scheme can be improved, for example for a small hydropower plant [4].

Dimitrieva [5] published an extensive study regarding forecasting methods, statistical and machine learning type like regressions, neural networks (NN) etc., methods conceived based on time series analysis. In [6], Zang uses the decomposition method to eliminate the trend and seasonal components, to improve the forecast realized with NN. Abrahart, in [7], made a comparison among NN and ARIMA models which was applied for hydrological data. In order to optimize water, use in [8], a model used for forecasting inflows is presented. Poornima Devi in [9] used fuzzy time series method which can offer good forecast even for incomplete datasets.

In a previous paper of the authors, [10], it was made a statistical analysis for hydropower ratio from total energy production in Romania over the period 2006 – 2016. The timestep is one month and the main characteristics for this set of observations were determined.

Considering the hydropower ratio from the total energy production in Romania as a random variable which can be predicted, the present paper presents a method for the prediction of this variable using the AR2 model. Based on this method, monthly data for the last three years, part of the analyzed data, are estimated, a forecast for average data (from 30 runs) including the confidence interval for the next two years is presented.

2. The time series data

In Romania, hydroelectricity covers, in average, one third of the total energy consumption depending on hydrological characteristic of the year.

The time evolution of the ratio of energy generated in hydropower plants from the total energy produced in Romania, noted further with X, is presented in Figure 1. It can be observed a strong upward gradient between the 1985 and 2000, then, starting with year 2006, it fluctuates around an average of 27.11% (as determined by the previous analysis of the dataset, [1]). Thus, the time series has been set up only for the monthly values recorded for the period 2006 – 2016, the monthly recorded values for this period has been published in the annual reports of the Hidroelectrica available online (hidroelectrica.ro).

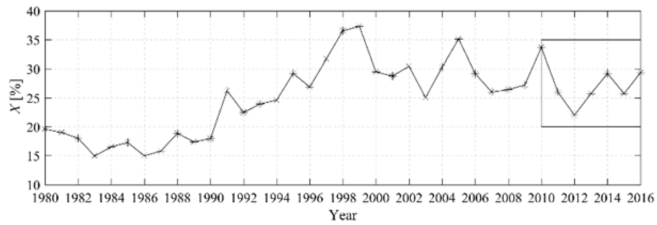

Figure 1. Hydropower ratio from total energy production in Romania (X), over the period 1980 – 2016

In order to realize a forecasting model, for the ratio of energy produced in hydropower plants from the total energy produced in Romania, *X*, in a previous paper the main features of the data set were determined and analyzed [1].

For the year p, from the time horizon of *N* years, and for the month τ of the year *p*, the hydropower ratio from total energy production in Romania is defined as:

$$X_{p,\tau} = \frac{E_{h\,p,\tau}}{E_{t\,p,\tau}},\ p = \overline{1,N},\ \tau = \overline{1,12}, \tag{1}$$

in which: $E_{h\,p,\tau}$ is the monthly hydropower output and $E_{t\,p,\tau}$ is the monthly total output in Romania.

The classic steps of statistical analysis have been achieved for detecting and removing components of the trend in the average, $T_{m,\tau}$, and in the variance, $T_{s,\tau}$, of the original data set. In Figure 2 the sketch of the decomposition process is represented.

Also, one year periodicity, the average and standard deviation of the study period have been calculated (μ_τ and σ_τ) in the previous paper [1].

With the variable with no tendency, $Z_{p,\tau}$, can performed a prediction model by adding trends and other determined features in the time series analysis ($\mu_\tau, \sigma_\tau, \varepsilon$), furthermore, the independent stochastic component (ξ) is added.

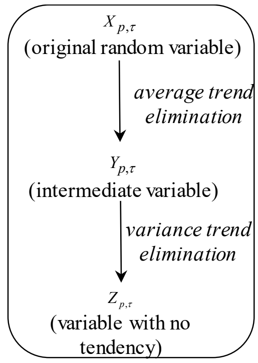

Figure 2. The sketch of the decomposition process of the original random variable

3. Mathematical modelling of the time series

The original variable based on specified components founded out in the process of decomposition [1] can be written in the form:

$$X_{p,\tau} = T_{mp,\tau} + T_{sp,\tau}\left(\mu_\tau + \sigma_\tau \varepsilon_{p,\tau}\right), \tag{2}$$

in which ε means the stochastic time-dependent component and is modelled by an AR2 model:

$$\varepsilon_{p,\tau} = \alpha_{1,\tau-1}\varepsilon_{p,\tau-1} + \alpha_{2,\tau-2}\varepsilon_{p,\tau-2} + \sqrt{1-\left(\alpha_{1,\tau-1}^2 + \alpha_{2,\tau-2}^2 + 2\alpha_{1,\tau-1}\alpha_{2,\tau-2}\rho_{1,\tau-1}^2\right)} \cdot \xi_{p,\tau} \tag{3}$$

in which:

α_1 and α_2 represent the autoregressive coefficients and they are functions of periodic autoregressive coefficients ρ_1 and ρ_2, [1];

ξ is the independent stochastic component for which it was proposed and tested a normal probability density function [1].

To improve the results of the time series modeling, in this paper the tendency on average was modeled using a Fourier series with the relation:

$$T_{m,p} = a_0 + a_1 \cos(wt) + b_1 \sin(wt), \tag{4}$$

where:

t represents the time, in years,

a_0, a_1, b_1 and w are coefficients determined using Matlab cftool so as to adjust as accurately as possible the real values of the annual averages of the analyzed variable.

In the mathematical model characteristics determined were used in the analysis of the time series with available data, 132 values (11 years x 12 months), the monthly values of the last 3 years of the data set were reconstituted (2014, 2015 and 2016). For each year there were synthetically generated 30 monthly values of the analyzed variable, so that the *n* = 30 values of each month constitute a sample with average $\bar{x}$.

Based on the *n* values, it can be defined a confidence interval for each month, where the real variable will be probably placed [11]:

$$\bar{x} - t_c \frac{s}{\sqrt{n}} < X < \bar{x} + t_c \frac{s}{\sqrt{n}}, \quad (5)$$

in which $t_c = t_{1-\frac{\alpha}{2}}$ is the quantile of Student-t distribution with $\nu = n - 1$ degrees of freedom, for which the distribution function is equal to $F(t_c) = 1 - \frac{\alpha}{2}$.

At the significance threshold $\alpha = 0.5$ it is expected that the actual achieved values will be in the confidence interval thus constructed with a probability of 95%.

The mathematical model was used then to predict monthly values of the variable for the next two years, 2017 and 2018, for which the actual values were not included in the statistical analysis based on which the model was built and calibrated.

4. Results and discussion

The trend, *T*, of the annual average values of the analyzed variable, *X*, was determined in the form of a Fourier series with equation (4), for which the following form was found out:

$$T_{m,p} = 27.11 - 1.528 \cdot \cos(1.72 \cdot t) + 2.895 \cdot \sin(1.72 \cdot t). \quad (6)$$

The recorded data for the annual hydropower ratio from total energy production in Romania, *X*, are represented in Figure 3, from which it can be observed that the approximation with the Fourier series is satisfactory/acceptable, anyway it is much more appropriate than the first degree polynomial proposed in the previous work [1].

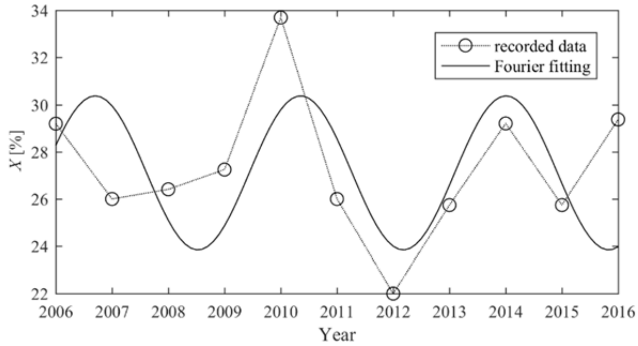

Figure 3. Annual average values of the original variable approximated with the Fourier series

The monthly stochastic component for the 11 years of the time horizon is determined with the equation 18 from [1], and it is represented in Figure 4.

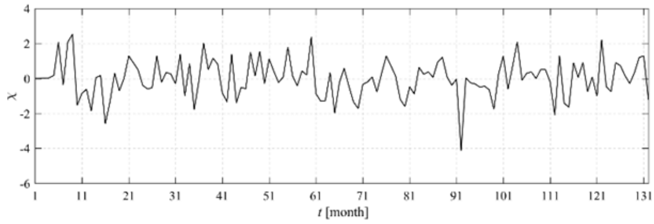

Figure 4. Independent monthly stochastic component for the 11 years of the time horizon

The parameters of the normal probability density function for the independent stochastic component ξ are determined with equations from [1] and results: $\mu_\xi = 0.004$ and $\sigma_\xi = 1.0068$.

In Figure 5 it can be observed the positioning of the average monthly values calculated for 30 runs, for the year 2014, comparing with the real recorded values. It can be noticed that, for this year, the generated values were close enough to the real ones.

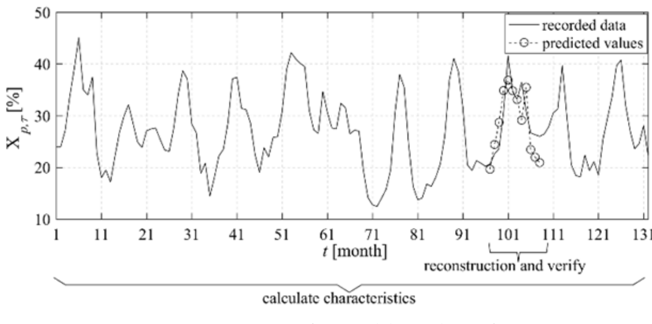

Figure 5. Reconstruction of monthly values for year 2014

For the stochastic variable there are randomly generated monthly values, using the values for the previous last two months of the previous year for α_1 and α_2 for the calculation of ε and using the characteristics of the 11-year series included in the analysis, the values of the variable X for the following months were calculated. In this manner, the values of the last three years of the time series (2014-2016) were generated. For each year, based on the generated values, the confidence intervals were determined, which are plotted beside the recorded values in Figures 6, 7 and 8.

Compared to the confidence interval, five of the recorded data were placed out, three of the values are at the limit, and four values are enclosed in this range (Figure 6).

The same procedure was applied for 2015 and 2016. For 2015 the results were better than for 2014 (Figure 7), only four of the values were being placed outside the confidence interval, while for 2016, five values (April, July, August, September, October) are placed outside the confidence interval (Figure 8).

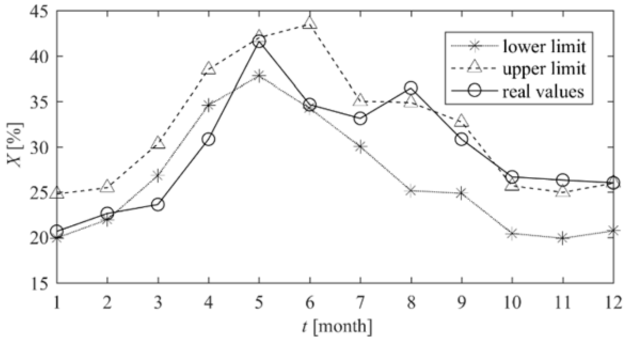

Figure 6. Confidence interval for year 2014

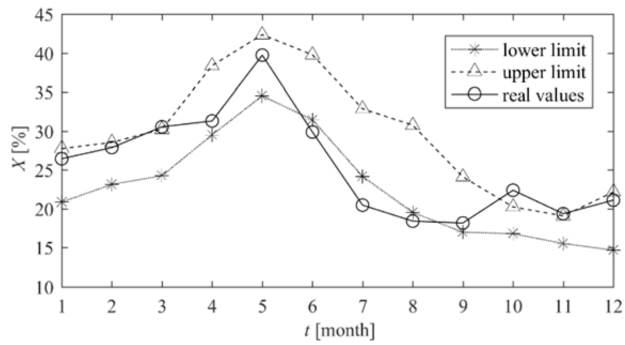

Figure 7. Confidence interval for year 2015

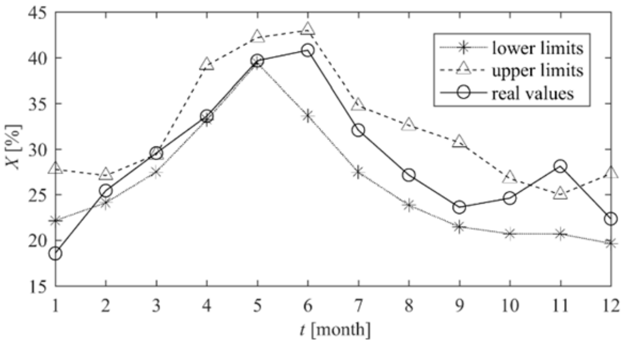

Figure 8. Confidence interval for year 2016

The same procedure was used for generating monthly values for the three years of the analyzed period, 2014-2016, and it was used to predict the monthly evolution of this variable for the next two years, 2017 and 2018.

In Figure 9 it is presented the variation of the variable X for the time horizon, 2006-2016, completed with 2017 (published in annual reports of Hidroelectrica) and 2018 (Transelectrica Website) and the predicted values determined with the mathematical model. It can be noticed that the average of the generated monthly values in 30 runs overestimates this variable in relation to the data recorded for the 12 months of 2017.

In the same manner, as for the period 2014-2016, the confidence limits for predicted values for the next two years, 2017 and 2018, were determined.

From Figures 10 and 11 it can observed that the recorded values were outside the confidence interval in six months of the year 2017, in three cases the deviation was considerable (July, August, September 2017) and in three months for 2018, February, May and July, the deviation was a small one.

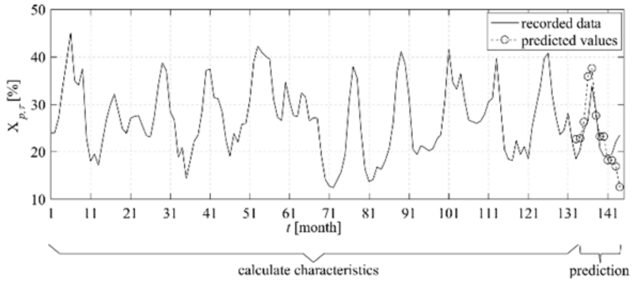

Figure 9. Reconstruction of monthly values for year 2017

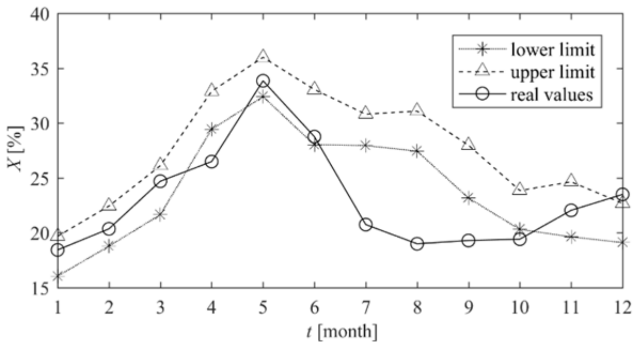

Figure 10. Recorded data and confidence interval for predicted values for year 2017

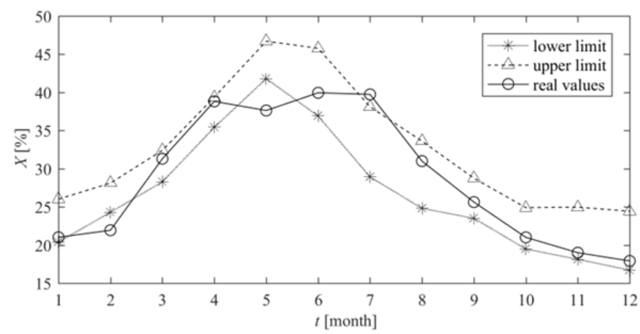

Figure 11. Recorded data and confidence interval for predicted values for year 2018

5. Conclusion

The hydropower generation in Romania is limited by the approximately 6.8 GW installed capacity in hydropower plants and hydrological type of the year. Therefore, the hydropower ratio from total energy production in Romania represents a variable that can be estimated by mathematical modelling based on time series analysis.

The decomposition of the time series was done in a previous paper, [10], and in this paper we built a model for the prediction of monthly values of the hydropower ratio from total energy production in Romania and a method for determine the confidence interval. The time horizon was the period 2006-2016.

For the period 2014-2016 synthetic data developed using the model are in good concordance with the recorded data and well defined by the confidence interval.

For the next two years, 2017 and 2018, we did not elaborate the model, considered forecast, and the situation was a bit different.

For the year 2017, in which the values have been smaller than the lower limit values of the months in study period, the predicted values are far from reality and most of the monthly values are out of the confidence interval.

Of course, the values which deviate substantially from the founded pattern for the time series does not have to be predicted with the methods combined with artificial intelligence. For 2018 the predictions are in good concordance with recorded values and the data are well placed in relation to the confidence interval.

Acknowledgements

The work has been funded by the Operational Programme Human Capital of the Ministry of European Funds through the Financial Agreement 51675/09.07.2019, SMIS code 125125.

References

[1]. Rahayu, S. P., Prastyo, D. D., & Wijayanti, D. G. P. (2017, September). Hybrid model for forecasting time series with trend, seasonal and salendar variation patterns. In *Journal of Physics: Conference Series* (Vol. 890, No. 1, p. 012160). IOP Publishing.

[2]. Hu, J., Wang, J., & Zeng, G. (2013). A hybrid forecasting approach applied to wind speed time series. *Renewable Energy*, *60*, 185-194.

[3]. Wu, C. L., Chau, K. W., & Fan, C. (2010). Prediction of rainfall time series using modular artificial neural networks coupled with data-preprocessing techniques. *Journal of Hydrology*, *389*(1-2), 146-167.

[4]. Hammid, A. T., Sulaiman, M. H. B., & Abdalla, A. N. (2018). Prediction of small hydropower plant power production in Himreen Lake dam (HLD) using artificial neural network. *Alexandria Engineering Journal*, *57*(1), 211-221.

[5]. Dmitrieva, K. (2015) Forecasting of a hydropower plant energy production, Master's Thesis in Computer Science, Halden, Norway.

[6]. Zhang, G.P., & Qi, M. (2005). Neural network forecasting for seasonal and trend time series, *European Journal of Operational Research. 160*, 501-514.

[7]. Abrahart, R. J., & See, L. (2000). Comparing neural network and autoregressive moving average techniques for the provision of continuous river flow forecasts in two contrasting catchments. *Hydrological processes*, *14*(11-12), 2157-2172.

[8]. Dashora, I., Singal, S. K., & Srivastav, D. K. (2015). Streamflow prediction for estimation of hydropower potential. *Water and Energy International*, *57*(11), 54-60.

[9]. Devi, P., Lakshmi, C. V., & Sakthivel, E. (2013). Design of a Fuzzy Time Series Forecasting Model for Hydro Power Generation. *International Journal of Computer Applications*, *74*(16).

[10]. Neagoe, A., Tică, E. I., Diminescu, M. A., Popa, B., & Cismaş, S. C. (2019). Analysis of hydropower ratio from total energy production in Romania. In *E3S Web of Conferences* (Vol. 85, p. 06009). EDP Sciences.

[11]. Popa, R., & Neagoe, A. (2008). Elements of statistics and data processing (in original in Romanian: Elemente de statistică şi prelucrarea datelor), Printech Publishing House, Bucharest.