Open Access Journal

Journal of Power Technologies 95 (4) (2015) 265–272

journal homepage:papers.itc.pw.edu.pl

Short-term management of hydro-power systems based on uncertainty model in electricity markets

Mansour Hosseini Firouz[a], Noradin Ghadimi [b,*]

[a] *Department of Engineering, Ardabil Branch, Islamic Azad University, Ardabil, Iran;*
[b] *Young Researchers and Elite club, Ardabil Branch, Islamic Azad University, Ardabil, Iran;*

Abstract

There are several methods for generating scenarios in stochastic programming. With extensive historical data records, one possibility is to represent the probability distribution of the uncertain data using a statistical model suitable for sampling. This method is especially useful for handling uncertain data that develops over time by means of time series analysis. In this paper a time series model relevant to the short-term management of hydro-power systems is proposed. This further illustrates the abilities of the models to capture developments in uncertain data over time. To demonstrate the validity of this model, results from the Nordic power exchange—Nord Pool—and a Norwegian power plant are presented.

Keywords: Feature Selection; Hybrid Forecast Engine; Price Forecast

1. Introduction

In the short-term management of a hydro-power plant, the uncertainty of the future surroundings poses a major challenge. In many respects, uncertainty of the inflows to the plant reservoirs is essential. Moreover, uncertainty with respect to electricity demand has in the past been of vital importance in the traditional setting. Demand uncertainty, however, has become less prominent with the deregulation of electricity markets, as a plant is no longer obligated to satisfy demand, but can resort to market exchange. Exchange through short-term markets has called for profit maximization, and market price uncertainty has become highly relevant. Reservoir inflow uncertainty mainly stems from non-anticipated precipitation and snow melt. Market price uncertainty is driven by demand and supply [1–3].

Demand uncertainty is mostly caused by temperature unpredictability and unpredicted customer behavior, whereas supply uncertainty may be due to unexpected failures. Nevertheless, a common feature of inflows and prices is that current observations show strong dependencies on past observations, and therefore the stochastic processes of inflows and prices can be handled by means of time series analysis. The time series analysis serves to gain insight into the empirical time series, to model the underlying stochastic processes and the development of data over time in particular, and to understand future data values. For modeling uncertainty in electricity prices and water streamflows, there are several frameworks in the fields of engineering, economics and statistics. Engineering approaches include bottom-up models for power systems and neural networks for the modeling of hydrological processes. Statistical methods most notably embrace time series analysis, possibly combined with other statistical tools. Finally, econometric approaches to analyzing time series have mostly been developed to analyze top-down models such as aggregated models for electricity prices [4–7].

One of the most common time series models is the ARMA model and its variants; for instance the inte-

*Corresponding author

Email addresses: mansourhosseinifirouz1@gmail.com (Mansour Hosseini Firouz), noradin.ghadimi@gmail.com (Noradin Ghadimi)

grated ARMA model, the fractional integrated ARMA model, the seasonal ARMA model, the vector ARMA model, the transfer model and the contemporaneous model, the ARCH and the GARCH models, some of which will be discussed in this paper.

2. The Proposed Price Forecast Strategy

2.1. The proposed data model

The available historical price series data to forecast the 24 hourly prices of day d is denoted by Ph: h=1,...,T. This series includes historical data up to hour 24 of day $d-1$. The value of T ranges usually from 168 (1 week) to 1,344 (2 months). The ARIMA technique works as follows. Use a specific ARIMA model of each one of the constitutive series to forecast its 24 future values for day d. An ARIMA model is then used to forecast hours $T+1$ to $T+24$ for each one of the constitutive series a_h, b_h, c_h, and d_h; resulting is estimated series a_h^{est}, b_h^{est}, c_h^{est}, and d_h^{est}; h. The standard statistical methodology to construct an ARIMA model includes the following.

Step 0 A class of models is formulated assuming certain hypotheses.

Step 1 A model is identified for the series considered.

Step 2 The parameters of the model are estimated.

Step 3 If the hypotheses of the model are validated, the procedure continues in Step 4; otherwise, the procedure continues in Step 1 to refine the model.

Step 4 The model is used for forecasting.

These steps are briefly explained below.

Model Selection (Step 0) . The proposed general ARIMA model takes the form

$$\phi(B)Ph = c + \theta(B)\varepsilon_h \quad (1)$$

where Ph is the price at hour h and ε_h is the error term. Polynomials $\phi(B)$ and $\theta(B)$ are functions of the back-shift operator B (observe that B^s_{Ph}=$Ph-s$). That is, $\phi(B)Ph = Ph - \phi_1 P_{h-1} - \phi_2 P_{h-2} - ... - \phi_{nF} P_{h-nF} -$, and $\phi_k(k = 1, ..., nF)$ are polynomial coefficients, and $\theta(B)\varepsilon_h = \varepsilon_h - \theta_1\varepsilon_{h-1} - \theta_2\varepsilon_{h-2} - ... - \theta_{nT}\varepsilon_{h-nT}$, and $\theta_k(k = 1, \ldots, nT)$ are polynomial coefficients.

The number of terms of the polynomial functions $\phi(B)$and $\theta(B)$, nF and nT, respectively, depends on the time series under analysis [8].
Note that including factors of the form allows one to take into account appropriately the seasonality effects. Finally, certain hypotheses on the error terms are needed to ensure the effectiveness of the predictions.

Model identification (Step 1). The target of this step is to identify which polynomial parameters should be estimated. The initial selection is based on the observation of the autocorrelation and partial autocorrelation plots [9]. Further refinement of the selection is based on physical knowledge and on engineering judgment.

Polynomial parameter estimation (Step 2). Once the parameters of the polynomials different from 0 have been identified (through plot observation, physical knowledge and engineering judgment), these parameters should be estimated. The estimation procedure is based on available historical data. Good estimators are usually found assuming that the data constitute observations of a stationary time series and maximizing the likelihood function with respect to the polynomial parameters. Good estimations can be obtained using commercial software, such as [10].

Validation of model hypotheses (Step 3). In this step, a diagnosis check is used to validate the model assumptions. If the estimated model is appropriate, then, the residuals (actual prices minus predicted prices) should behave in a manner consistent with the model. Residuals must satisfy the requirements of a white noise process: zero mean, constant variance, uncorrelation and normal distribution. If the hypotheses on the residuals are validated, then the corresponding model can be used to forecast prices and this step concludes successfully. Otherwise, the residuals contain a certain structure that should be analyzed to refine the model, and the procedure continues in Step 1. To refine the model a careful inspection of the autocorrelation and partial autocorrelation plots of the residuals is advisable.

Actual prediction (Step 4). In this step, the corresponding model from Step 2 is used to predict future values of prices, typically 24 hours ahead. It should be noted that prediction quality deteriorates as the predicted hour increases, i.e., the error of the estimate of hour 24 is typically greater than the error of the estimate of hour 1.

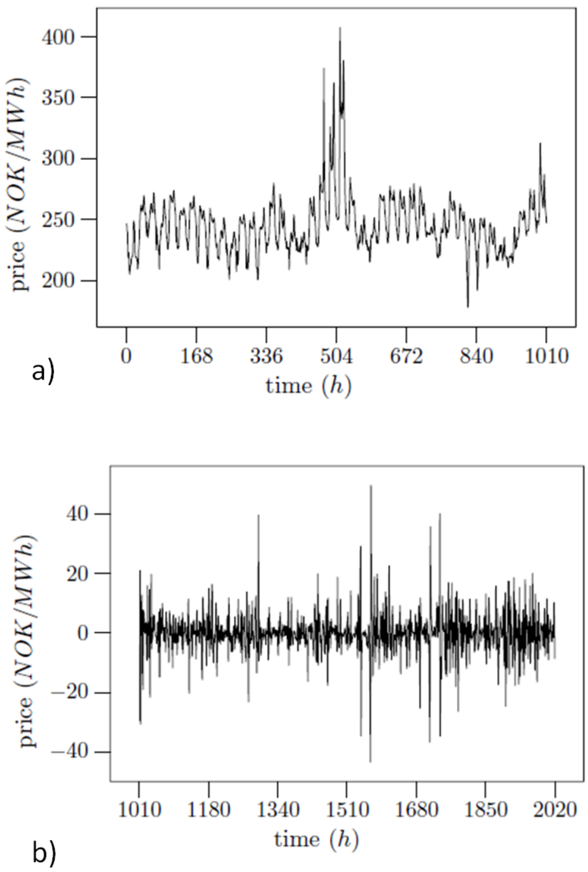

Figure 1: Hourly day-ahead market prices

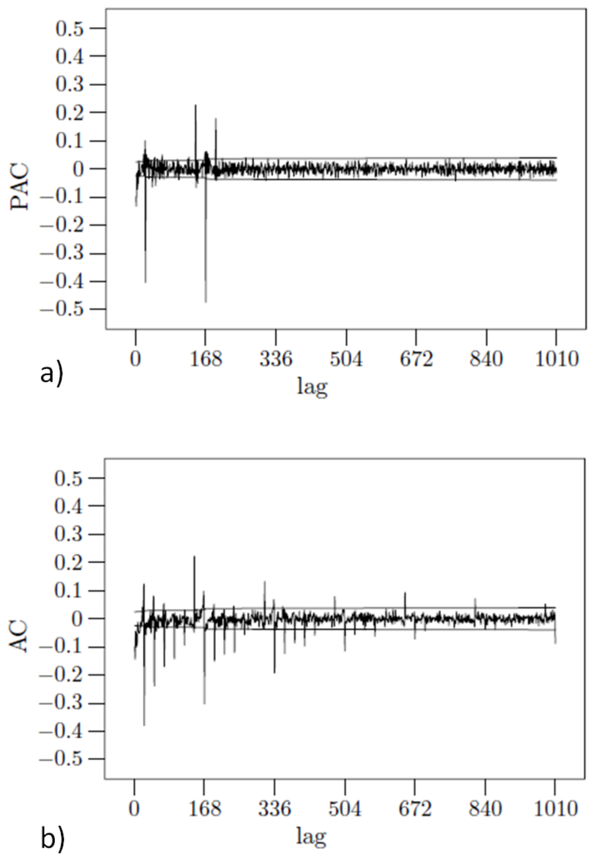

Figure 2: Autocorrelation functions for day-ahead market prices

2.2. Day-ahead market prices

The recent tendency to restructure and deregulate electricity markets has stimulated most power producers to shift the objective from cost minimization and demand satisfaction to profit maximization alone, which has made the behavior of electricity prices of utmost importance. In the short term, power producers may participate in spot markets, such as the day-ahead market for disposing of physical production. To participate in the day-ahead market, bids must be submitted a day in advance, as the name suggests. The market cannot clear and the clearing prices of the following day cannot be announced until afterwards. The clearing prices, also referred to as day-ahead market prices, are therefore uncertain at the time of bidding [11, 12].

Model identification: To identify an appropriate model, the first step is to detect seasonalities. As day-ahead market prices are partly driven by electricity demand, which exhibits a daily and a weekly pattern, both daily and weekly periodic behavior is to be expected. This is supported by the fact that the day-ahead market clears every day of the week except at the weekend. The periodicities are visible from Fig. 1a. The class of proposed models provides a base for identifying a model.

A non-constant mean indicates non-stationarity of the time series data. This is further justified by empirical autocorrelations that decay very slowly to zero. Factors $(1-B)$, $(1-B^{24})$ and $(1-B^{168})$ are included in order to stabilize the mean, $(1-B^{24})$ and $(1-B^{168})$ to remove seasonality, and the process of differences can then be accepted as being stationary, cf. Fig. 1b. Experiments were made with a logarithmic transformation to stabilize the variance. However, the best results were obtained without this transformation. The original process can therefore be described by a SARIMA model. Inspecting the empirical autocorrelation and partial autocorrelation functions, the order of the process of differences can be determined. The functions are shown in Figs 2. The autocorrelations corresponding to the lags 1, 2, 3, ... give an indication of an ARMA (2, 2) process. The autocorrelations of the lags 24, 48, 72, ... show evidence of an MA (2) process, as the lagged-24k autocorrelations are zero for $k>2$ and the partial autocorrelation function is

Table 1: Maximum likelihood estimates for day-ahead market prices

Parameter	ϕ_1	ϕ_2	γ_1	γ_2	γ_{24}	γ_{25}
Estimate	0.30	0.34	0.36	0.55	-0.05	0.73
Parameter	γ_{26}	γ_{47}	γ_{48}	γ_{49}	γ_{168}	σ
Estimate	−0.07	0.05	0.17	0.036	0.99	5.12

Table 2: Weekly forecast errors for day-ahead market prices

Week	41	42	43	44	45
MPE	0.71	0.33	0.08	0.12	0.15
MAPE	3.11	1.34	2.54	2.44	2.32
MSE	65.66	23.45	54.34	43.67	45.65
Week	46	47	48	49	50
MPE	0.23	0.42	0.45	0.22	0.16
MAPE	1.34	1.44	1.76	2.56	2.64
MSE	31.55	42.56	34.56	38.76	45.80

Table 3: Weekly descriptive statistics for day-ahead market prices

	Week	41	42	43	44	45
Simulation	Mean Value	222.5	234.6	243.3	256.4	231.4
	Std. dev.	16.4	15.7	14.6	16.7	11.2
Real	Mean Value	223.2	221.0	225.2	234.5	231.3
	Std. dev.	11.2	10.3	12.4	13.2	13.4
	Week	46	47	48	49	50
Simulation	Mean Value	225.3	226.4	231.4	243.5	265.7
	Std. dev.	13.2	14.2	12.6	11.2	11.6
Real	Mean Value	221.2	224.2	225.6	226.7	226.1
	Std. dev.	12.4	10.8	11.3	11.5	12.6

exponentially decreasing. Finally, the autocorrelations of the lags 168, 336, 504, ... are indications of $anMA\ (1)$ process, as the lagged-168k autocorrelations are zero for $k>1$ and the partial autocorrelation function is exponentially decreasing. Although not clearly visible in Fig. 2a, the ARMA (2,2) process causes the autocorrelation function to peak in the neighborhood of the lag 24, 48, 72, ... and the MA (2) process causes it to peak at the lags 144, 192, 312, 360, 480, 528. The initial proposal of a model is

$$(1-\phi_1 B-\phi_2 B^2)(1-B)(1-B^{24})(1-B^{168})\varepsilon_h = \\ (1-\gamma_1 B-\gamma_2 B^2)(1-\gamma_{24}B^{24}-\gamma_{48}B^{48})(1-\gamma_{168}B^{168})\varepsilon_h, t\in Z$$

Inspecting the autocorrelation function and the partial autocorrelation function of the residuals, the model can be further refined to

$$(1-\phi_1 B-\phi_2 B^2)(1-B)(1-B^{24})(1-B^{168})\varepsilon_h = \\ (1-\gamma_1 B-\gamma_2 B^2)(1-\gamma_{23}B^{23}-\gamma_{24}B^{24}-\gamma_{25}B^{25} \\ -\gamma_{47}B^{47}-\gamma_{48}B^{48}-\gamma_{49}B^{49})(1-\gamma_{168}B^{168})\varepsilon_h, t\in Z$$

The model is an extension of those of the preceding sections. It is, nevertheless, sufficiently general to include the main characteristics of day-ahead market prices.

3. Numerical Results

The model is an extension of those of the preceding sections. It is, nevertheless, sufficiently general to include the main characteristics of day-ahead market prices. Parameter estimation: Parameter estimates are obtained by the use of maximum likelihood estimation. Estimates based on data of the model identification period can be found in Table 1.

To validate the model, the assumption of a white noise process on the residuals must be confirmed. A plot of the residuals is given in Fig. 3a. It should be clear that the mean value can be assumed to be zero and that the variance appears to be constant. Furthermore, the autocorrelation and partial autocorrelation functions of the residuals, cf. Figs. 3b and 3c, are both close to zero as is the case for a white noise process. The Ljung-Box statistics back up the fit of the model.

Before using forecasts and simulations as tools for planning purposes, both procedures are suitable for further validation of the model. Out-of-sample tests are performed and, hence, the model is tested on the data of the validation period by forecasting and simulating into this period.

The forecast errors of the validation period, i.e., weeks 41–50 or hours 6721–8400, are reported in Table 2. The mean percentual error (MPE), mean absolute percentual error (MAPE) as well as the mean square error (MSE) are displayed on a weekly basis. It should be remarked that the estimation of the validation period has been conducted on a 24-hour basis using an adaptive approach, cf. Contreras et al. (2003) and Nogales et al. (2002). The estimation of the first 24 hours, i.e., hours 6721–6744, is based on data of hours 1–6720.

Moving the time window 24 hours, the estimation of the next 24 hours, i.e., hours 6745–6768, is based on data of hours 24–6744 etc. The forecast errors are seen to be rather small and therefore the model is suitable for forecasting. Descriptive statistics for a simulation of 1000 samples are shown in Table 3 along with the same information on the real observations of the validation period. Concerning the preservation of the descriptive statistics, the mean value is well preserved, whereas the standard deviation is generally over-estimated. Forecast and simulation: Starting from the end of the validation period, forecasts and simulations can be generated further into the future. As an example of a short-term use of the procedures, hourly prices are forecast and simulated a week ahead, i.e., into week 51. A plot of the forecasts and the real observations can be found in Fig. 4a. Moreover, the forecast and its confidence intervals are plotted in Fig. 4b. Examples of a few simulated sample paths are displayed in Fig. 4c.

3.1. Reservoir inflows

As water inflows to hydro reservoirs are often highly uncertain, the forecasting or simulation of future values is crucial in the short-term planning of hydro-power plants. In the management of reservoirs, the scheduling of water releases depends among other things on the current reservoir levels. Water releases in turn affect unit commitment and production level decisions, which makes knowledge of future reservoir inflows very valuable. Early applications of the ARMA framework to forecast and simulate stream-flows can be found in [8]. For the ARMA time series analysis of water inflows to reservoirs, the data consists of hourly observations from the same hydro-power plant located in the vicinity of Trondheim, Norway and run by the company TrønderEnergi. The data dates back to 2004, which is again divided into a model identification period of 40 weeks and a model validation period of 10 weeks. Typically, even smaller hydro-power plants consist of more than one reservoir. Therefore, it is highly relevant to consider multiple reservoir inflow series. Inflows to different reservoirs may stem from the same stream or from different streams. Here, we consider two reservoir inflow series from two different streams and initially handle the series individually. We show a model for one of the inflow series should be fitted, as fitting the other may be done in a similar fashion. The first inflow series corresponds to a reservoir named Samsjøen, the second inflows series to H°aen.

Model identification: Consider the inflow to Samsjøen. The data does not immediately disclose any obvious short-term seasonalities of the reservoir inflows. Hence, the starting point of model identification is stationarity. Highly non-constant mean value and variance reveal non-stationarity of the time series data, which is further verified by slowly decreasing empirical autocorrelations. By experimenting with logarithmic transformation and differences, the inclusion of factor $(1-B)$ was found most suitable in obtaining stationarity. In particular, the autocorrelation functions decreased more quickly without the logarithmic transformation. Hence, an appropriate model should be found within the class of ARIMA models. The original empirical time series and the series of differences are displayed in Fig. 5.

The model is validated by testing the assumption of a white noise process on the residuals and is confirmed by the behavior of the autocorrelation and partial autocorrelation functions as well as by the Ljung-Box statistics, none of which are displayed here.

Out-of-sample testing: Further validation of the model is based on forecasts and simulations. Again, we do out-of-sample testing. Forecast errors are shown in Table 4. We find forecasts to be useful for high inflow weeks, i.e., weeks 41 and 45–50. However, for low inflow weeks, i.e., weeks 42–44, the forecasts are rather poor. The descriptive statistics of the simulations, cf. Table 5, show that the mean is more or less preserved, whereas the standard deviation is highly overestimated. Forecast and simulation: With the validated model at hand, short-term forecasts and simulations of hourly inflows can be made. Plots of the forecast and the real data as well as the forecast and its corresponding confidence interval a week ahead are shown in Fig. 6a and 6b. Examples of a few simulated sample paths are shown in Fig. 6c.

4. Conclusion

In this paper a time series model relevant to the short-term management of hydro-power systems is proposed. The results obtained illustrate the ability of ARMA time series models to forecast and simulate hourly day-ahead electricity prices and reservoir inflows. As indicated, forecasts and simulations are suitable as input for optimization problems such as deter-

Table 4: Weekly forecast errors for reservoir inflows

Week	41	42	43	44	45
MPE	15.47	15.09	14.84	14.88	14.91
MAPE	15.53	13.76	109.3	78.54	14.74
MSE	30.01×10^6	12.32×10^6	18.69×10^6	8.02×10^6	10×10^6
Week	46	47	48	49	50
MPE	0.75	0.94	0.97	0.74	0.68
MAPE	9.57	9.67	9.99	10.79	10.87
MSE	5.21×10^6	16.22×10^6	8.22×10^6	12.42×10^6	19.46×10^6

Table 5: Weekly descriptive statistics for reservoir inflows

	Week	41	42	43	44	45
Simulation	Mean Value	10046.5	6546.3	10067.3	10080.4	76542.4
	Std. dev.	9659.4	9658.7	9657.6	9659.7	9654.2
Real	Mean Value	6547.2	6545	6549.2	6558.5	6555.3
	Std. dev.	3476.2	3475.3	3477.4	3478.2	3478.4
	Week	46	47	48	49	50
Simulation	Mean Value	17012.3	17013.4	17018.4	17030.5	17052.7
	Std. dev.	9686.6	9687.6	9686	9684.6	9685
Real	Mean Value	23985.8	23988.8	23990.2	23991.3	23990.7
	Std. dev.	1354.9	1353.3	1353.8	1354	1355.1

Table 6: Maximum likelihood estimates for reservoir inflows

Parameter	ϕ_1	γ_1	γ_2	γ_{41}	σ^2
Estimate	0.89	1.23	0.45	0.78	680.12

ministic and stochastic planning problems that have future values of data as input. Direct application of the ARMA framework might not successfully capture intermittency and the current results show a moderate performance with respect to hourly stream-flows.

References

[1] M. Shahidehpour, H. Yamin, Z. Li, Market overview in electric power systems, Market Operations in Electric Power Systems: Forecasting, Scheduling, and Risk Management (2002) 1–20.

[2] L. Wu, M. Shahidehpour, A hybrid model for day-ahead price forecasting, Power Systems, IEEE Transactions on 25 (3) (2010) 1519–1530.

[3] P. Kou, D. Liang, L. Gao, J. Lou, Probabilistic electricity price forecasting with variational heteroscedastic gaussian process and active learning, Energy Conversion and Management 89 (2015) 298–308.

[4] R. Hafezi, J. Shahrabi, E. Hadavandi, A bat-neural network multi-agent system (bnnmas) for stock price prediction: Case study of dax stock price, Applied Soft Computing 29 (2015) 196–210.

[5] J. Che, J. Wang, Short-term electricity prices forecasting based on support vector regression and auto-regressive integrated moving average modeling, Energy Conversion and Management 51 (10) (2010) 1911–1917.

[6] W.-M. Lin, H.-J. Gow, M.-T. Tsai, Electricity price forecasting using enhanced probability neural network, Energy Conversion and Management 51 (12) (2010) 2707–2714.

[7] F. J. Nogales, J. Contreras, A. J. Conejo, R. Espínola, Forecasting next-day electricity prices by time series models, Power Systems, IEEE Transactions on 17 (2) (2002) 342–348.

[8] F. Nogales, A. J. Conejo, Electricity price forecasting through transfer function models, Journal of the Operational Research Society 57 (4) (2006) 350–356.

[9] J. Contreras, R. Espinola, F. J. Nogales, A. J. Conejo, Arima models to predict next-day electricity prices, Power Systems, IEEE Transactions on 18 (3) (2003) 1014–1020.

[10] O. Abedinia, N. Amjady, M. Shafie-Khah, J. Catalão, Electricity price forecast using combinatorial neural network trained by a new stochastic search method, Energy Conversion and Management 105 (2015) 642–654.

[11] L. Zhang, P. B. Luh, K. Kasiviswanathan, Energy clearing price prediction and confidence interval estimation with cascaded neural networks, Power Systems, IEEE Transactions on 18 (1) (2003) 99–105.

[12] J.-J. Guo, P. B. Luh, Improving market clearing price prediction by using a committee machine of neural networks, Power Systems, IEEE Transactions on 19 (4) (2004) 1867–1876.

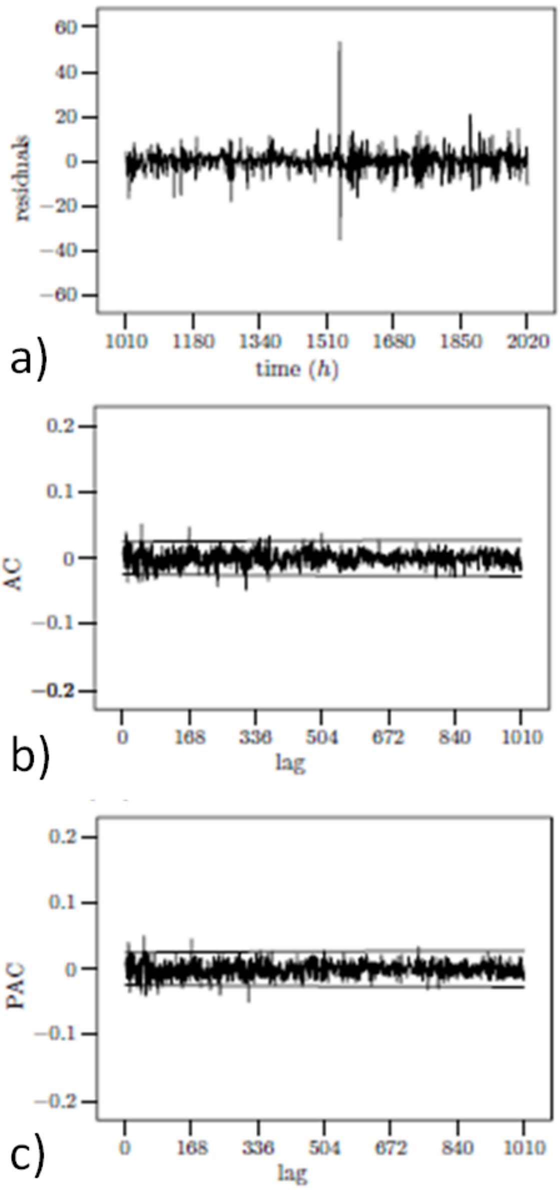

Figure 3: Residuals of day-ahead market prices

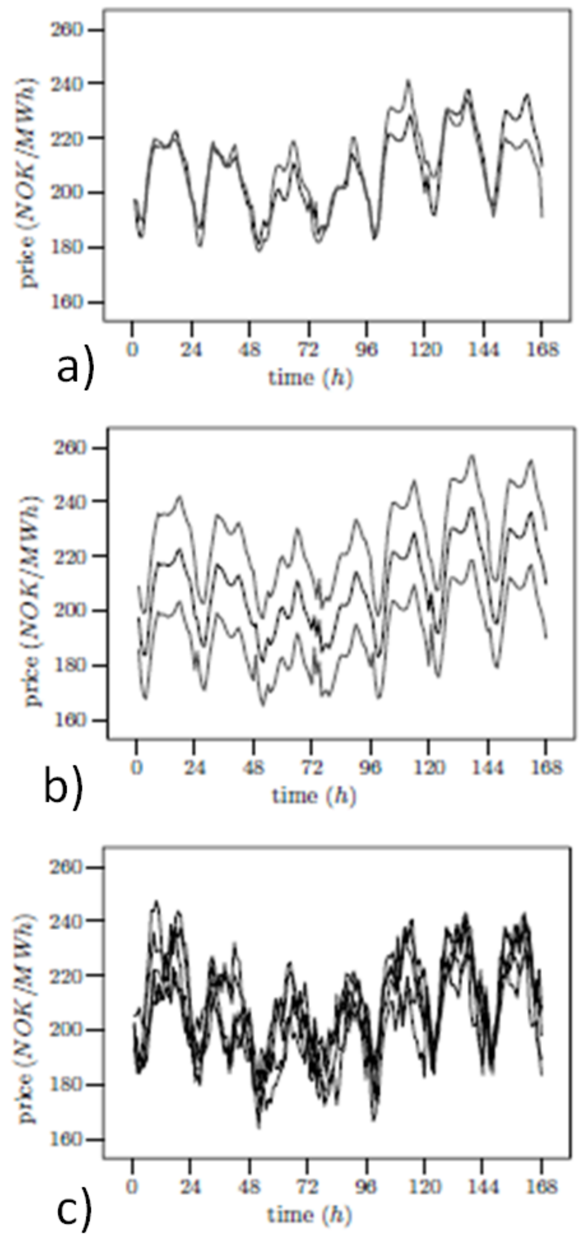

Figure 4: Forecast and simulated day-ahead market prices

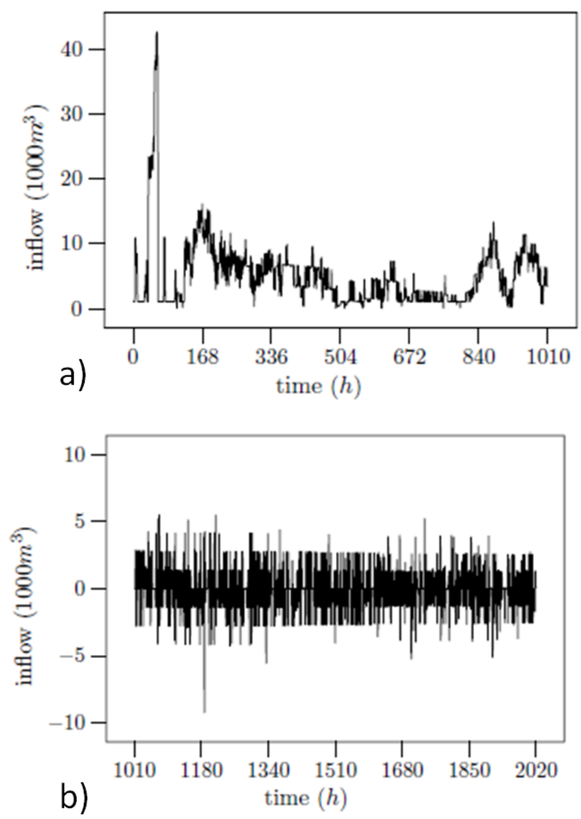

Figure 5: Hourly reservoir inflows

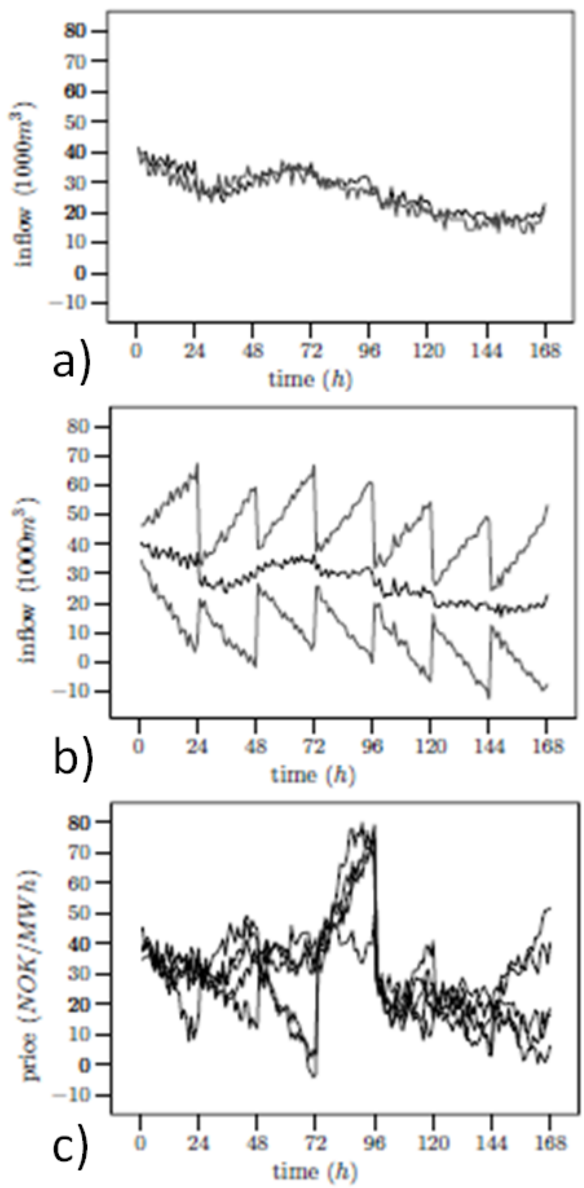

Figure 6: Forecast and simulated reservoir inflows