MULTI-CHANNEL SSA AND BOOTSTRAP TECHNIQUES APPLIED TO NATURAL INFLOW ENERGY TIME SERIES IN BRAZIL

Paula Medina Maçaira
Pontifical Catholic University of Rio de Janeiro, PUC - Rio
Department of Industrial Engineering
paulamacaira@gmail.com

Fernando Luiz Cyrino Oliveira
Pontifical Catholic University of Rio de Janeiro, PUC - Rio
Department of Industrial Engineering
cyrino@puc-rio.br

Reinaldo Castro Souza
Pontifical Catholic University of Rio de Janeiro, PUC - Rio
Department of Electrical Engineering
reinaldo@ele.puc-rio.br

ABSTRACT

In Brazil, the decisions about the optimal dispatch of hydroelectric and thermoelectric plants are highly dependent on the time series Natural Inflow Energy (NIE). Reliable and accurate long term forecasts and simulations for the NIE series are crucial due to the importance of energy long term planning. To develop such forecasts, the proposed approach relies on the decomposition of NIE series from the South subsystem into signal and white noise using the Multi-channel Singular Spectrum Analysis (MSSA) technique. After that, synthetic series are generated from the noise using the Bootstrap procedure, forecasts 60 months ahead for each series using the traditional Holt Winters are calculated to arrive at the final predictions obtained by taking the average of the individual series forecasts. When compared to the results obtained by applying directly Holt Winters, the proposed methodology proved to be very efficient, especially for lead time of more than 36 months.

KEYWORDS. Natural Inflow Energy. Multi-Channel Singular Spectrum Analysis. Bootstrap.

Main Area: SIM - Simulação, EN - PO na Área de Energia, EST - Estatística.

1. Introduction

The system of production and transmission of electricity in Brazil have a strong predominance of hydroelectric power plants. From the official published statistics, in 2012, 86% of the total energy was generated by hydroelectric sources, 12.4% from thermal origin and only 1.6% from renewable sources (ONS, 2012). Due to the complexity of generation, production and distribution of the Brazilian electrical energy sector, the National Interconnected System (NIS) is composed of utilities from South, Southeast, Midwest, Northeast and part of the Amazon, being called subsystems. Only 1.7% of the country's electricity production capacity is outside the NIS, in small isolated systems, located mainly in the Amazon region.

Planing the Brazilian energy sector means, basically, making decision about the dispatch of hydroelectric and thermoelectric plants, with the risk of financial losses or rationing for wrong decisions, as happened in the 2001, rationing that affected three of the four subsystems. Thus, the planning of the energy operation involves determining targets for the generation of hydroelectric and thermoelectric plants for each stage along the study horizon, given the demand for electricity, the operating constraints of the plants and the restrictions of the electrical system (Pereira, 1989). From the above, it is quite clear that operating a system like the Brazilian one is a rather complex task given its extension and its uniqueness in the world.

One of the most important characteristic of a generation systems with hydro predominance is the strong dependence on the stream inflow regimes. Such, the uncertainly associated with the energy planning requires proper and coherent stochastic modelling and forecast of the hydrological series (and energy). The models that generate hydrological scenarios in order to optimize the performance of the operation system has as stochastic variable, the Natural Inflow Energy (Oliveira et al., 2015), generated by a Periodic Autoregressive model PAR(p), fitted to the historical data (Terry et al., 1986; Maceira and Damázio, 2006).

The Brazilian Research Center of Electrical Energy (CEPEL) developed a chain of models for planning and operating the NIS with platforms that perform long term forecasts (five years ahead), short term (one year ahead) and very short term (fourteen days ahead). Such models are used as the base chain of models that forecast the NIE series (CEPEL, 2004; de Deus, 2010).

This article proposes the joint use of Multi-channel Singular Spectrum Analysis (MSSA), Bootstrap and Holt Winters as an alternative approach to forecast up to sixty months ahead (five years) for the NIE series for the South Subsystem, which is the most irregular one. It is important to mention that this article was completely based on Maçaira et al. (2015).

The paper is organized as follows: section two describes the methods used for the forecasting exercise; in section three it is presented a descriptive analysis of the Natural Inflow Energy series, followed by section four where the results from the application of the proposed methodology are shown. Finally, in section five are drawn some conclusions and directions for further researches on this topic.

2. Methodology

As mentioned above, to achieve the proposed objectives, the following techniques are used in the paper: Multi-channel Singular Spectrum Analysis (MSSA),Bootstrap and Holt Winters. This approach was inspired in the works of Bergmeir et al. (2014); Dantas and Oliveira (2014), where the last one uses the decomposition know as Seasonal-Trend Decomposition Using Loess (STL), to model a wind power time series and the Moving Blocks Bootstrap technique applied to the noise from the fitted model. Twenty nine new series were generated and 24-step-ahead forecasts were obtained for each one of these simulated series plus the original one. The final 24-step-ahead forecast were obtained by averaging these forecasts generated by the 30 series. Such procedure produced more accurate forecasts than those obtained by modelling only the original series. Some modifications were introduced in the methodology described in the implementation of the present approach. The steps to implement the proposed procedure are described below:

1. Decompose the original series into signal and noise using MSSA technique;
2. Apply the Bootstrap to the noise series, generating new P noise series;
3. Add to the P new noise series the signal, resulting in P synthetic series;
4. Fit, to each one of the P synthetic series, the Holt-Winters method and generate h-steps-ahead forecasts; and
5. Make the simple average of the P forecasts and compare them with the ones obtained by the model fitted to the original series.

The main inspiration of this paper has been cited previously, but it is worth mentioning two important papers that have successfully integrated the SSA technique and some other methodology. Rahmani (2014) joint SSA with Bootstrap, to provide precise estimation of underlying coefficients of LRF (Linear Recurrent Formula). And Cassiano et al. (2013), that jointly use SSA, PCA (Principal Component Analysis) and ARIMA model to produce hydroelectric energy forecasting.

In what follows a brief description of each of the techniques used in the proposed implementation are briefly addressed.

2.1. Multi-channel Singular Spectrum Analysis (MSSA)

By denoting Y_t as an observation of the random variable Y at time t then $Y_1, ..., Y_N$ is a time series with N observations, it can also be considered one of many possible realization of a stochastic process. In general, a time series can be described as:

$$Y_t = f(t) + \epsilon_t \quad (1)$$

where $f(t)$ is the signal and ϵ_t is the noise, i.e. an independent and identically distributed random variable with zero mean and constant variance σ^2. Therefore, in order to produce good forecasts, it is essential that the proposed models are able to separate the signal and the noise components. The former can be decomposed into seasonal, trend, cycle and other factors.

The main objective of the MSSA technique is to decompose the original series in the sum of a few interpretable and independent components, such as trend, periodic or quasi-periodic and noise. MSSA is a non parametric technique that incorporates elements of classical time series, multivariate statistics, multivariate geometry, dynamic systems and signal processing. This technique is a multidimensional extension (or multivariate) of the Singular Spectrum Analysis (SSA). The birth of SSA is usually associated to the paper of Broomhead and King (1986) and a complete description of the theory and practice of the technique can be found in Hassani (2007); Hassani and R. (2013).

To understand the MSSA technique it is first necessary to understand the process performed by SSA and after that extend the idea to the multivariate case. Thus, SSA consists of two complementary stages: decomposition and reconstruction, both including two separate steps. In the first step of the first stage, SSA transforms the one-dimensional time series Y_t into multi-dimensional series $X_1, ..., X_K$ with vectors $X_i = (y_i, ..., y_{i+L-1})'$, called vectors of lag, where $K = N - L + 1$. The only parameter of this step is the window length L, an integer such that $2 \leq L \leq N/2$ and the result is the trajectory matrix $X = [X_1, ..., X_K] = (x_{ij})_{i,j=1}^{L,K}$. In the second step of the first stage it is carried out the decomposition of the trajectory matrix in singular values and its representation as the sum of bi-orthogonal elementary matrix of rank 1. Denote $\lambda_1, ..., \lambda_L$ as the eigenvalues of XX' in descending order of magnitude and $U_1, ..., U_L$ the orthonormal system of eigenvectors of the matrix XX' corresponding to the eigenvalues. By defining $V_i = X'U_i/\sqrt{\lambda_i}$, then the Singular Value Decomposition (SVD) of the trajectory matrix can be written as:

$$X = X_1 + ... + X_d \quad (2)$$

where $X_i = \sqrt{\lambda_i} U_i V_i'$, $(i = 1, ..., d)$ and the collection $(\sqrt{\lambda_i}, U_i, V_i)$ are called eigentriple of matrix X at time i. If all the eigenvalues have multiplicity 1, then the expansion is defined uniquely.

Moving now to the second stage, the first step is to split the elementary matrices X_i into several groups, adding matrices inside each group. Let $I = i_1, ..., i_p$ be a group of index $i_1, ..., i_p$, then the matrix X_I correspond to the group I is defined as $X_I = X_{i1} + ... + X_{ip}$. The disaggregation of the set of index $J = 1, ..., d$ into disjoint subsets $I_1, ..., I_m$ corresponds to the representation:

$$X = X_{I1} + ... + X_{Im} \tag{3}$$

For a given group I, the contribution of component X_I in the expansion (2) is measured by the share of the corresponding eigenvalues $\sum_{i \in I} \lambda_i / \sum_{i=1}^{d} \lambda_i$. Finally, in the second step of stage two, the method, called Diagonal Averaging or Hankelization, transform each matrix I into a time series, which is an additive component of the original series Y_t. Applying the procedure in all matrices of (3) a new expansion of X is obtained, where the result is equivalent to the decomposition of the initial series $Y_t = (y_i, ..., y_N)$ as the sum of m series:

$$y_t = \sum_{k=1}^{m} \tilde{y}_t^{(k)} \tag{4}$$

where $\tilde{Y}_t^{(k)} = (\tilde{y}_1^{(k)}, ..., \tilde{y}_N^{(k)})$ corresponds to the matrix X_{Ik}. All description and application of the method are detailed shown in Hassani (2007); Hassani and R. (2013); Golyandina and Korobeynikov (2013).

The extension of the SSA (univariate) explained above to the multivariate (or multidimensional) case (s series), is carried out by the so called MSSA. For that, it suffices to consider in the first step of the first stage that for each time series $F^{(k)}$ the initial procedure form $K = N - L + 1$ vectors of lag $X_j^{(k)} = (f_{j-1}^{(k)}, ..., f_{j+L-2}^{(k)})^T$, $1 \leq j \leq K$. The trajectory matrix of the multidimensional series $(F^{(1)}, ..., F^{(s)})$ must have dimension LxKxs and shape

$$X = [X_1^{(1)} : ... : X_K^{(1)} : ... : X_1^{(s)} : ... : X_K^{(s)}] = [X^{(1)} : ... : X^{(s)}] \tag{5}$$

For the following three steps the same SSA mathematical procedure is applied, resulting in

$$\tilde{F}^{(k)} = \sum_{j=0}^{N-1} \tilde{f}_j^{(k)} \tag{6}$$

with k $(k = 1, ..., s)$ time series reconstructed corresponding to the index group of the signal.

Golyandina and Stepanov (2005) call attention to the fact that often, in the case of two sets, it is more reasonable to solve the problem of extracting a single component (for example, noise reduction), rather than two components. In this case, one is only interested in one group of indices related to the signal, instead of looking for each index in a series, not taking into account the influence of the correlation that may possibly exist.

2.2. Bootstrap

The Bootstrap technique, first developed in Efron (1979), is a method of resampling data that allows the assessment of the variability estimator based on existing data from a single sample. Briefly speaking, this technique consists of a random sample with replacement of the elements of a random sample, generating a new sample, called as "Bootstrap sample", usually with the same size of the original sample. In the context of time series, there are basically two ways to apply: Bootstrap in the residuals and the method called Moving Blocks (Kuensch, 1989).

In this paper we use Bootstrap in the residuals, due to the fact that from the white noise extracted by MSSA it is possible to ensure the hypotheses of independence of them, the required condition for application of the method.

A formal description of the method is: consider $R_1, ..., R_N$ the random noise extracted by MSSA and B the number of residuals series to be generate. B residual series are drawn with replacement from the original residual series, generating B Bootstrap residuals series of size N each.

On the following step, i.e., the generation of synthetic NIE scenarios, one should ensure that all simulated values are positive; as it does not make sense a negative value for the NIE. The initial approach does not guarantee this constraint, so the strategy employed was to resample with replacement the corresponding residuals until a positive value for the NIE is obtained, similar to what Oliveira et al. (2015) has done.

2.3. Holt Winters

Holt Winters' method, first developed by Holt with reprinted version, in Holt (2004), and improved by Winters, in Winters (1960), is a classic method of Time Series. Widely used, it aims to capture the behaviour of the series by separating the components of trend, seasonality and error term. The fitting and forecasting are performed iteratively in steps with the equations of level, slope and seasonal component. The last term can be additively or multiplicatively linked to the trend component. In this paper only the additive version of Holt Winters is used and therefore just the expressions for this method will be presented.

Let y_t be the time series, t, l_t, b_t and s_t be the level, slope and seasonal components, respectively at time t and m the seasonal period of the series, the parameters updating equations are as follows.

$$l_t = \alpha(y_t - s_{t-m}) + (1-\alpha)(l_{t-1} + b_{t-1}) \tag{7}$$

$$b_t = \beta(l_t - l_{t-1}) + (1-\beta)b_{t-1} \tag{8}$$

$$s_t = \gamma(y_t - l_{t-1} - bt - 1) + (1-\gamma)s_{t-m} \tag{9}$$

where the constants α, β and γ are smoothing hyperparameters, all of them defined in the interval $[0, 1]$.

The h step ahead forecast at origin t is given by:

$$\hat{y}_{t+h|t} = l_t + h * b_t + s_{t-m+h} \tag{10}$$

3. Descriptive Analysis

Historical data of NIE for the South Subsystem used in this paper are available in the National Operation System (NOS) website and were split into the training set, in which all the techniques are applied, and the validation set, where the comparisons between the forecast and actuals are made. The period of January 1970 to December 2008 is the training period (468 months) and from January 2009 til December 2013 (60 months) the validation period. Figure 1 presents the series of NIE from the South Subsystem for the training period; a quick look at it reveals the presence of outliers and a rather weakly seasonal pattern.

According to Table 1, the time series presents an average value of 10458 MWme, a median of 8482 MWmed and a standard deviation of 7089 MWmed. It is also displayed in this table some others descriptives statistics of the series, such as minimum and maximum values and quantiles estimates. Table 2 presents the same descriptives statistics for the entire series, but now for each month. From this table one can observed that December has the minimum value recorded in the training period and July the maximum value, this occurred in 1985 and 1983, respectively. Another point to observe is that the month of April is the one with the lowest mean and median while the month of October has the highest values for these statistics.

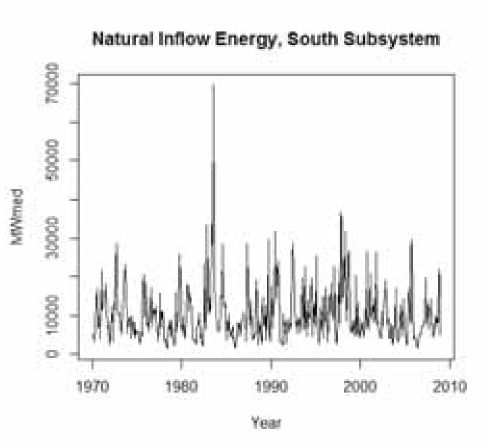


Figure 1: Graphic of Natural Inflow Energy series from South Subsystem

Table 1: Descriptive statistics: Natural Inflow Energy series

Minimum	1442
1st Quartile	5654
Mean	10458
Median	8482
3st Quartile	12982
Maximum	69521
Standard Deviation	7089

Table 2: Descriptive statistics by month: Natural Inflow Energy series

Statistics	Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec
Min.	2202	2196	2575	1798	1455	1854	3120	2592	2797	4124	2762	1442
1st Quart.	5196	5607	4397	4087	3623	5970	7575	6468	7085	7817	6502	5034
Mean	8763	9450	7481	7182	9900	11329	13191	11845	12750	14592	10572	8444
Median	7419	8051	7116	5956	7476	10104	11925	9306	8798	13014	8579	7444
3st Quart.	10556	11736	9180	8641	12815	14594	15159	17240	17219	19982	12762	10929
Max.	25380	24239	19251	31363	33465	31544	69521	31215	30281	36634	34421	18274
St. Dev.	4977	5441	3839	5200	7819	6830	10183	7425	8069	7925	6587	4477

In Figure 2 it is shown the monthplot. From it, one can clearly see the presence of outliers in all months, emphasising once again the complexity of the series.

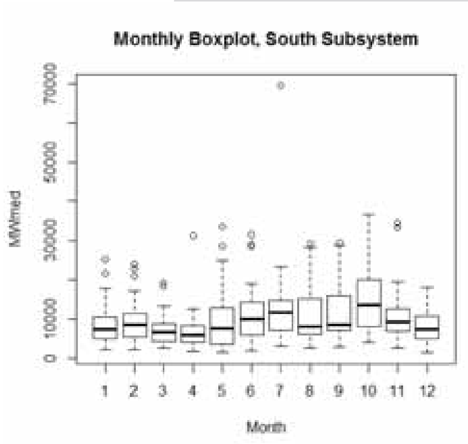


Figure 2: Monthplot of Natural Inflow Energy series

4. Results

Before the implementation of the approach proposed in this paper, four traditional time series methods were applied to the original series: Holt Winters, Seasonal Naive, SARIMA and PAR(p) - Periodic Autoregressive of orders $p_1, ..., p_{12}$ (one for each month). For the SARIMA, the best model, i.e. the one that presents the lower Akaike Information Criterion (AICc) in the fitting exercise, was $(1, 0, 0)$x$(0, 1, 1)_{12}$. For the PAR(p), the selection was based on the results of Oliveira and Souza (2011); the following orders $(1, 1, 4, 5, 1, 5, 4, 1, 1, 1, 6, 1)$ were identifyed for each month of the year. In Figure 3 it is presented the forecasts (red) against the actuals (black), while Table 3 presents the Mean Absolute Percentage Error (MAPE) for the periods of 12, 24, 36, 48 and 60 steps ahead for each one of the fitted models.

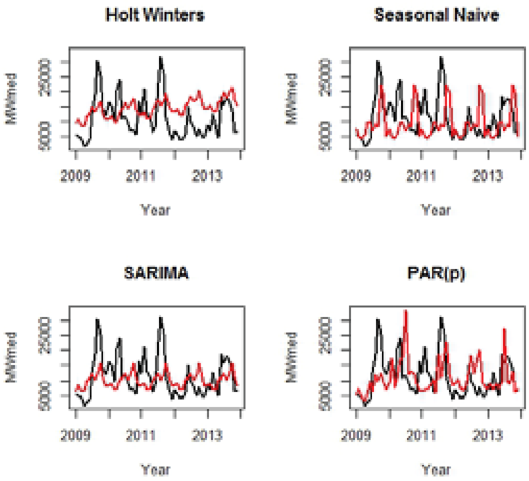


Figure 3: Forecasts obtained for the Natural Inflow Energy

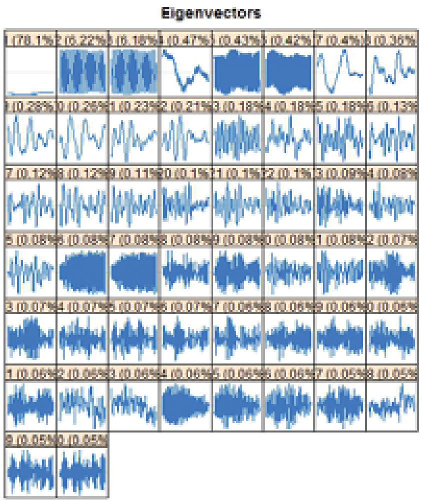
Table 3: MAPE of the forecasts with Holt Winters, Seasonal Naive, SARIMA and PAR(p)

Methods	12 months	24 months	36 months	48 months	60 months
Holt Winters	115,59	88,41	80,5	99,35	94,55
Seasonal Naive	79,46	77,63	73,75	71,51	69,11
SARIMA	88,39	69,5	63,71	65,18	59,59
PAR(p)	55,88	57,52	56,79	64,09	57,91

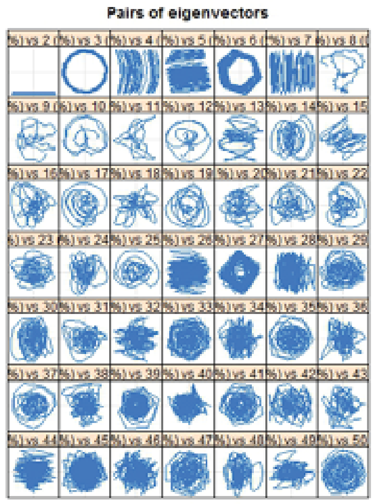
After fitting these four models, 60-step-ahead forecasts were calculate by each one of them and the comparison among them was carried out on the validation set. As can be observed from Table 3 for all forecast lead time, the PAR(p) was the technique that produced lower error measures. That is not a surprise, as it is used in the context of the Brazilian Hydrothermal Dispatch (Gladyshev, 1961). However, as it will be shown later, the worst method (HW) improved its performance substantially when used in conjunction with MSSA and Bootstrap, as proposed in this paper.

To separate the original series into signal and noise, it was used initially the univariate SSA technique, applied only to the South NIE series. However, after numerous tests of window length combinations (L) and number of components, the approach was not successful in extracting white noise. Therefore, it was decided to test the multi-channel approach (MSSA) using also the NIE series of the others subsystems (Southeast/Midwest, Northeast and North) as an aid for the filtering process of the subsystem South series and the results are as follows.

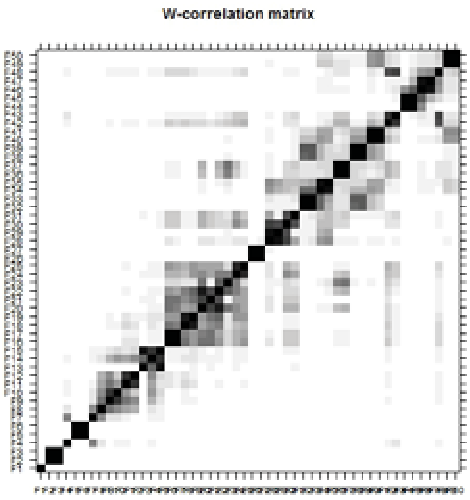
Using the window length of 234, i.e., $N/2$, where N (468), the total size of the training set, the eigenvectors and the pairs of eigenvectors obtained are shown in Figures 4(a) and 4(b), while the eigenvectors correlation matrix W in Figure 4(c).



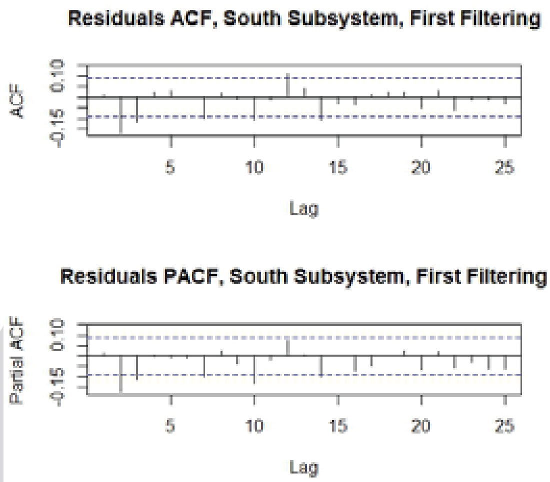
(a) 50 eigenvectors extracted with MSSA



(b) Pairs of eigenvectors extracted with MSSA



(c) Correlation matrix W

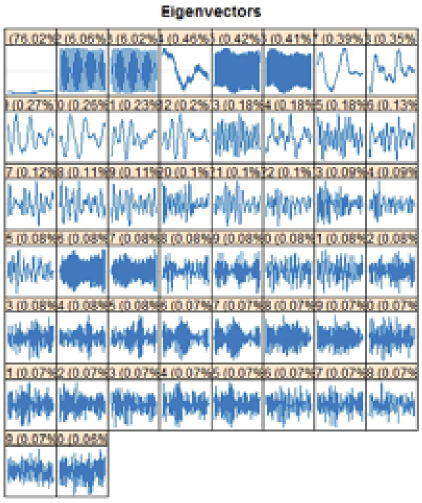


(d) ACF and PACF of the residuals series extracted with MSSA

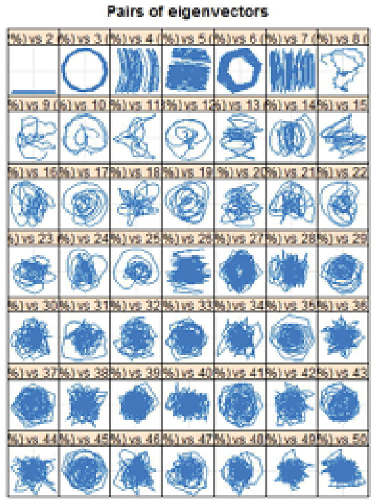
Analysing the Figures 4(a), 4(b) and 4(c) and after several attempts, it was decided that the first 50 components would respond for the signal and the remaining the residuals. As the initial objective of the decomposition is to obtain white noise, Figure 4(d) shows the graphs of the auto-correlations and partial autocorrelations of the obtained residual series. It is possible to see that the residual series still preserve a correlation in the first lags, confirmed by the BDS test (Brock et al., 1987) with p-value less than 0.001.

It is important to mention that no other window length and number of components combination produced white noise to the residuals, so we move to the next step, called re-filtering. That is, the residual series obtained in the previous step were considered input data to the MSSA, and it is expected that the existing pattern in the noise will be captured by the method.

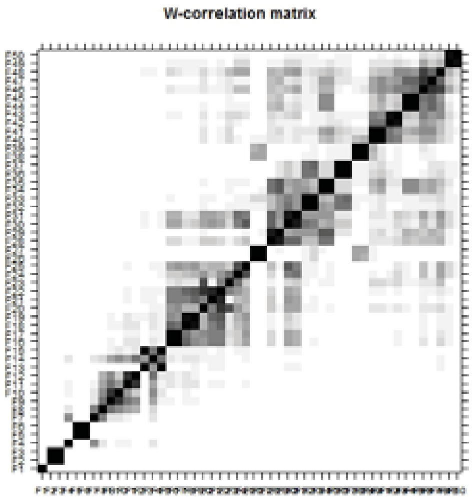
In Figures 4(e), 4(f) and 4(g) are graphically represented the eigenvectors, pairs of eigenvectors and the correlation matrix W obtained after the re-filtering (or second filtering) of the NIE series using, again, 234 (N/2) as window size.



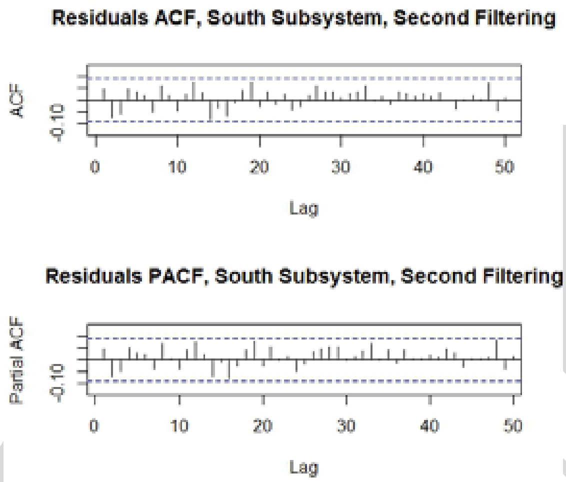
(e) 50 eigenvectors extracted in the second filtering with MSSA



(f) Pairs of eigenvectors extracted in the second filtering with MSSA



(g) Correlation matrix W in the second filtering



(h) ACF and PACF of the residuals series extracted in the second filtering with MSSA

Once again, after a careful analysis of the figures, it was decided that the first 50 components will be part of the signal and the remaining the residual series. Analysing the correlation in the residuals series, it is confirmed, in Figure 4(h) and also by the BDS test (p value = 0.054), that no dependence structure were found between the noise series.

In Figure 4, it is observed the signal obtained with MSSA against the original series and also the noise series.

With the signal and the noise extracted from the original series, the Bootstrap technique was applied to the residual series to generate 199 new series and added to the signal, generating 199 synthetic NIE series for the South Subsystem. This quantity of new series is justified by the fact that this number provide measures of error more stable between the simulation. Adding the original series as the 200th series, it was applied to each one of them the Holt Winters Additive method to forecast 60 steps ahead. The Holt Winters method was chosen because it is a simple technique to be implemented and its additive term comes from the fact that the NIE series presents an additive seasonality. The simple average of the values predicted from each synthetic series was considered as the final forecast for the South Subsystem.

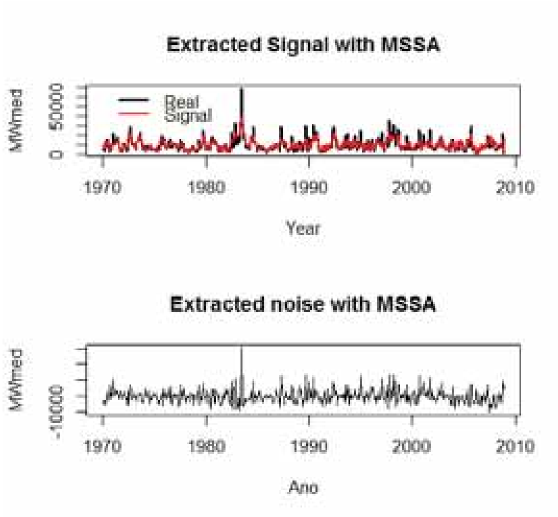


Figure 4: Signal and Noise extracted from the Natural Inflow Energy series with MSSA

The error measures obtained are in Table 4, and an analysis revels that the proposed approach not only improves the MAPE of the Holt Winters method applied to the new series, but also produces more accurate forecasts for longer horizons (36-months-ahead).

Table 4: MAPE of the forecasts with Holt Winters, Seasonal Naive, SARIMA, PAR(p) and MSSA with Bootstrap

Methods	12 months	24 months	36 months	48 months	60 months
Holt Winters	115,59	88,41	80,5	99,35	94,55
Seasonal Naive	79,46	77,63	73,75	71,51	69,11
SARIMA	88,39	69,5	63,71	65,18	59,59
PAR(p)	55,88	57,52	56,79	64,09	57,91
MSSA Bootstrap	77,52	61,39	54,89	47,94	45,79

5. Conclusion

The approach proposed in this paper can be summarize as follows: decompose the NIE time series in its signal and noise components via MSSA technique, generating synthetic series from the application of Bootstrap to its noise components, forecast each time series using Holt Winters Additive method and then perform the simple average to set the final forecast. This approach, aimed to produce more reliable and accurate long-term forecasts, and it was achieved for lead time greater than 36 months, showing a considerable reduction in measurement errors when compared against the methods applied directly to the original series.

As an extension of this work, the authors intend to use other time series models, as the PAR (p), with the aim of reducing the forecasts for both short-term (up to 36 months) and for long-term. Also, for future research we intend to use related climate series, such as El Niño, La Niña and Sunspots, to improve the decomposition process. As a final word, the exercise applied to the South Subsystem NIE series will be applied to the other three subsystems to check whether the same improvement on the MAPE is observed.

Acknowledgements

The authors would like to thank the R&D program of the Brazilian Electricity Regulatory Agency (ANEEL) for the financial support (R&D project 0387-0113/2013). Finally, the authors would like to thank CNPq (research project 443595/2014-3), FAPERJ (research project E-

26/111.621/2014) and CAPES, which are governmental agencies that have also given support to this research.

References

Bergmeir, C., H. R. J., and B. J. M. (2014). Bagging exponential smoothing methods using stl decomposition and box-cox transformation. working paper 11/14.

Brock, W. A., W. D. Dechert, and J. A. Scheinkman (1987). A test for independence based on the correlation dimension. *Departament of Economics, University of Wisconsin at Madison, University of Houston and University of Chicago*.

Broomhead, D. S. and G. P. King (1986). Extracting qualitative dynamics from experimental data. *Physica D: Nonlinear Phenomena 20*, 217–236.

Cassiano, K. M., L. A. Teixeira Jr, R. M. Souza, M. L. Menezes, J. F. Pessanha, and R. C. Souza (2013). Hydroelectric energy forecast. *International Journal of Energy and Statistics 1*, 205–214.

CEPEL (2004). Reference manual of prediction model for weekly in ows applied to brazilian hydroelectric system, model previvaz. Technical report, Center for Electric Power Research.

Dantas, T. M. and F. L. C. Oliveira (2014). Prediction of wind speed: An approach using bagging holt winters with decomposition stl. In *XLVI Simpósio Brasileiro de Pesquisa Operacional (SBPO)*.

de Deus, F. L. D. (2010). Generation of natural inflow energy scenarios through the adoption of the latin hypercube methodology. Master's thesis, Federal University of Rio de Janeiro.

Efron, B. (1979). Bootstrap methods: Another look at the jacknife. *The Annals of Statistics 7*, 1–26.

Gladyshev, E. G. (1961). Periodically correlated random sequences. *Soviet Mathematics Doklady 2*, 385–388.

Golyandina, N. and A. Korobeynikov (2013). Basic singular spectrum analysis and forecasting with r. *Computational Statistics and Data Analysis 71*, 934–954.

Golyandina, N. and D. Stepanov (2005). Ssa-based approaches to analysis and forecast of multidimensional time series. In *Proceedings of the 5th St. Petersburg Workshop on Simulation*.

Hassani, H. (2007). Singular spectrum analysis: Methodology and comparison. *Journal of Data Science 5*, 239–257.

Hassani, H. and M. R. (2013). Multivariate singular spectrum analysis: A general view and new vector forecasting approach. *International Journal of Energy and Statistics 1*, 55–83.

Holt, C. (2004). Forecasting seasonals and trends by exponentially weighted moving averages. *International Journal of Forecasting 20*, 5–10.

Kuensch, H. R. (1989). The jackknife and the bootstrap for general stationary observations. *The Annal of Statistics 17*, 1217–1241.

Maçaira, P., F. Cyrino Oliveira, and R. C. Souza (2015). Forecasting natural inflow energy series with multi-channel singular spectrum analysis and bootstrap techniques. *International Journal of Energy and Statistics 3*, 15500051–155000517.

Maceira, M. E. P. and J. M. Damázio (2006). Use of the par(p) model in the stochastic dual dynamic programming optimization scheme used in the operation planning of the brazilian hydropower system. *Probability in the Engineering and Informational Sciences 20*, 143–156.

Oliveira, F. L. C. and R. C. Souza (2011). A new approach to identify the structural order of par (p) models. *Pesquisa Operacional 31*, 487–498.

Oliveira, F. L. C., R. C. Souza, and A. L. M. Marcato (2015). A time series model for building scenarios trees applied to stochastic optimisation. *International Journal of Electrical Power & Energy Systems 67*, 315–323.

ONS (2012). National electric system operator: Publications of relevant data 2012, production by generation type.

Pereira, M. V. F. (1989). Optimal stochastic operations scheduling of large hydroeletric systems. *International Journal of Electrical Power & Energy Systems 11*, 161–169.

Rahmani, D. (2014). A forecasting algorithm for singular spectrum analysis based on bootstrap linear recurrent formula coefficients. *International Journal of Energy and Statistics 2*, 287–300.

Terry, L., M. Pereira, T. A. Araripe Neto, L. F. C. A. Silva, and P. R. H. Sales (1986). Coordinating the energy generation of the brazilian national hydrothermal electrical generating system. *Interfaces 16*, 16–38.

Winters, P. R. (1960). Forecasting sales by exponentially weighted moving average. *Management Science 6*, 324–342.