

Finite Rank Series Modeling for Discrimination of Non-stationary Signals

Lina Maria Sepulveda-Cano, Carlos Daniel Acosta-Medina,
and Germán Castellanos-Dominguez

Signal Processing and Recognition Group, Universidad Nacional de Colombia,
Km. 9, Vía al aeropuerto, Campus la Nubia, Manizales, Colombia
lmsepulvedac@bt.unal.edu.co

Abstract. The analysis of time-variant biosignals for classification tasks, usually requires a modeling that may handle their different dynamics and non-stationary components. Although determination of proper stationary data length and the model parameters remains as an open issue. In this work, time-variant signal decomposition through Finite Rank Series Modeling is carried out, aiming to find the model parameters. Three schemes are tested for OSA detection based on HRV recordings: SSA and DLM as linear decompositions and EDS as non-linear decomposition. Results show that EDS decomposition presents the best performance, followed by SSA. As a conclusion, it can be inferred that adding complexity at the linear model the trend is approximate to a simple non-linear model.

1 Introduction

As regards the analysis and processing of biosignals, stochastic modeling has been under continuous formulation for extracting valuable information due to the high non-stationarity inherent of this kind of time series. Mostly, a methodology for analysis of non-stationary time series is based on the assumption that there is a processing time window of such a length that the piecewise stationary-based approach of analysis holds. Although determination of proper stationary data length and the model parameters remains as an open issue [1]. Grounded on piecewise stationary approach, several time-variant linear decomposition techniques had been proposed for non-stationarity characterization: time-frequency representation, smoothing techniques based on ortogonal polynomials, linear splines, wavelets, empirical mode decomposition, and regressive modeling, among others.

Based on generalized spectral representation, a more elaborated approximation, termed *Finite-rank Series Modeling*, can be carried out by decomposing the underlying biosignal into a linear combination of products of harmonics, exponential, or even polynomial series [2], that is, the problem is addressed as modeling multiple time-evolving data streams governed by some linear recurrent formula, and inferring similarities and relationships between these data over prescribed time window lags. Although, a single input signal can be decomposed by

an unknown number of recurrent sequences, which may vary from observation to observation, a strong constrain on this decomposition approach is related with modeling time series holding itself different stochastic structures [3].

This paper considers the extension of Finite Order Series Modeling behind baseline Singular Spectrum Analysis (SSA), explained in [2]. Namely, Dynamic Linear Modeling (DLM) as linear scheme [4], and Exponentially Damped Sinusoids (EDS) as non-linear scheme [5] are also considered. As a demonstrative example, we analyze discrimination of hear rate variability recordings, which are a multi-process exhibiting high non-stationary behavior, for detection of Obstructive Sleep Apnoea [3]. The remainder of the paper is organized as follows: In Section 2.1, the fundamentals of finite rank modeling is briefly introduced. In Section 3 the methodology for parameters estimation is explained. Lastly, in Sections 4 and 5, the effectiveness of each decomposition is illustrated for the considered task detection through cross-validation using a k -nearest neighbors (k -nn) classifier, followed by a discussion of the obtained results.

2 Background

2.1 Basic Finite Rank Series Modeling

Assume a real-valued time series $\mathbf{x} = \{x_n : i, \dots, N\}$, with $\mathbf{x} \in \mathbb{R}^N$, which is mapped into the multidimensional time series set $\tilde{\mathbf{X}} = \{\tilde{\mathbf{x}}_k : k, \dots, K\}$, $\tilde{\mathbf{X}} \in \mathbb{R}^{K \times L}$, termed *Hankel Matrix*, comprising L -lagged vectors $\tilde{\mathbf{x}}_k = [x_{i-1}, \dots, x_{i+L-2}]^\top$; $\tilde{\mathbf{x}}_k \in \mathbb{R}^L$, where $K = N - L$, being $3 \leq L \leq N$ [6]. Furthermore, if the dimension $\dim \mathfrak{L}^{(L)}\{\mathbf{x}\} = \text{span}\{\tilde{\mathbf{x}}_k : \forall k\} = d$, for $0 \leq d \leq L$, then, the series is regarded as L -rank d , noted as $\text{rank}_L\{\mathbf{x}\}$. In turn, when the equality $\text{rank}_L\{\mathbf{x} = d\}$ holds for all suitable L , the series \mathbf{x} has rank d , that is, $\text{rank}\{\mathbf{x}\} = d$. If such d exists, the time series is called a *finite rank series* [2].

Generally, a nonzero series \mathbf{x} is governed by the following linear sequence of dimension not exceeding $d \geq 1$ if

$$x_{i+d} = \sum_{k=1}^d w_k x_{i+d-k}, \quad (1)$$

for a certain set $\{w_i \in \mathbb{R} : i = 1, \dots, d\}$ with $w_d \neq 0$, and $0 \leq i \leq N - d$.

Since $\text{rank}_L\{\mathbf{x}\} = \text{rank}_L\{\tilde{\mathbf{X}}\} = \text{rank}_L\{\tilde{\mathbf{X}}\tilde{\mathbf{X}}^\top\}$, the orthonormal system of eigenvectors $\{\mathbf{u}_l : l = 1, \dots, d\}$ (corresponding to d positive eigenvalues of the matrix $\text{rank}_L\{\tilde{\mathbf{X}}\tilde{\mathbf{X}}^\top\}$) constitutes the left singular vectors of the singular value decomposition of the matrix $\tilde{\mathbf{X}}$. This concrete model of Finite Rank Series is known as *Singular Spectral Analysis* (SSA) [7].

Suppose a sufficiently large N of the series \mathbf{x} , under assumption that $d < \min(L, K)$, so, taking into account Eq. (2), the L -lagged vectors $\{\tilde{\mathbf{x}}_k\}$ satisfy the following vector recurrent equation:

$$\tilde{\mathbf{x}}_{i+d} = \sum_{k=1}^d w_k \tilde{\mathbf{x}}_{i+d-k}, \quad (2)$$

The coefficient set in Eq. (2) can be found as follows:

$$\mathbf{w} = \frac{1}{1 - \mu^2} \sum_{i=1}^d u_{di} \mathbf{u}_d^\nabla, \tag{3}$$

where ∇ is the truncation operator, $\mathbf{w} = [w_d, \dots, w_1]^\top$ is the coefficient vector, $\mathbf{u}_d^\nabla = [u_{1i}, \dots, u_{di}]$ is the i -th eigenvector, and $\mu^2 = \sum_{i=1}^d u_{di}^2$ is the *verticality coefficient*.

2.2 Extended Finite Rank Series Modeling

For sake of generalization, model (2) can be rewritten in terms of an *autoregressive model*, as follows:

$$x_{i+d} = \mathbf{w} \mathbf{x}_{i-k-1}, \tag{4}$$

where $\mathbf{x}_{i-k-1} = [x_{i-k-1} \dots x_{i+d-k}]^\top$ is the *autoregressive vector* and $\mathbf{w} \in \mathbb{R}^{d \times 1}$ is the corresponding *parameter vector*, given by: $\mathbf{w} = [w_1, \dots, w_d]$; that is assumed to be constant in time. However, model (2) can be also extended for representing time-varying autoregressive model of order d , termed TVAR(d), if

$$x_i = \sum_{k=1}^d w_{k,i} x_{i-k}, \tag{5}$$

where w_i is the autoregressive parameter at i -th time sample.

An alternative structure of (5) is the Dynamic Linear Model (DLM). Thus, it is important to be able to embed an TVAR model in the DLM form in order to distinguish, and ultimately infer, the process structure from contaminating noise [8]. Particularly, a DLM describing an TVAR process, at moment time i , has the form of Eq. (4) and is given by [4]:

$$x_i = \mathbf{1}^\top \mathbf{H}_i \mathbf{W}_i \mathbf{x}_{i-1} + \delta_i, \tag{6}$$

where $\mathbf{1} = [1, \dots, 1] \in \mathbb{R}^{d \times 1}$, $\mathbf{H}_i = \text{diag}\{\mathbf{U}_i^\top \mathbf{f}\} \mathbf{U}_i^{-1}$, with $\mathbf{H}_i \in \mathbb{R}^{d \times d}$, $\mathbf{f} = [1, 0 \dots 0]^\top \in \mathbb{R}^{d \times 1}$ is the *vector of regressors*, δ a zero-mean normal innovation vector, and $\mathbf{W}_i \in \mathbb{R}^{d \times d}$ is the parameter matrix expressed as:

$$\mathbf{W}_i = \mathbf{U}_i \mathbf{A}_i \mathbf{U}_i^{-1} = \begin{bmatrix} w_{1,i} & w_{2,i} & \dots & w_{d-1,i} & w_{d,i} \\ 1 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & 0 \\ 0 & 0 & \dots & 1 & 0 \end{bmatrix}, \quad i = 1 + d, \dots, N - d \tag{7}$$

being $\mathbf{U}_i \in \mathbb{R}^{d \times d}$ the eigenvector matrix and $\mathbf{A}_i \in \mathbb{R}^{d \times d}$ the eigenvalues of \mathbf{W}_i , respectively, at moment time i .

So, a TVAR model in Eq. (5) can be rewritten by means of vector recurrent expression (2), namely:

$$x_i = \sum_{k=1}^d h_{k,i} a_{k,i} \mathbf{u}_{k,i} x_{i-k} + \delta_i, \tag{8}$$

where $\mathbf{h}_i \in \mathbb{R}^d$ is the principal diagonal of $\text{diag}\{\mathbf{U}_i^\top \mathbf{f}\}$ and $\mathbf{a}_i \in \mathbb{R}^d$ is the eigenvalue vector of \mathbf{W}_i .

Lastly, the model (2) can be represented as a linear combination of *Exponentially Damped Sinusoidals* (EDS), when the signal is modeled as a finite sum of discrete-time exponentially damped complex sinusoids, as follows [5]:

$$x_i = \sum_{k=1}^d a_k \exp(j\phi_k) \exp((-h_k + j2\pi f_k)i\Delta n) \tag{9}$$

The amplitudes a_k , damping factors h_k , phases ϕ_k , frequencies f_k , and the model order d are free parameters of the non-linear model. Through Prony analysis exposed in [9], the non-linear problem stated in Eq. (9) can be decoupled by solving two introduced sets of linear equations, given in least square sense. Thus, model (9) is rewritten by using Eq. (2), as follows:

$$x_{i+d} = - \sum_{k=1}^d w_k x_{i+d-k}, \tag{10}$$

where the the parameter vector \mathbf{w} is given as [10]:

$$\mathbf{w} = \sum_{k=1}^d \sigma_k \mathbf{u}_k \mathbf{v}_k^\top, \tag{11}$$

being σ the principle singular vector of $\tilde{\mathbf{X}}$. $\{\mathbf{v}_k : k = 1, \dots, d\}$ constitutes the right singular vectors of the singular value decomposition of the matrix $\tilde{\mathbf{X}}$. At the end, the EDS model can be expressed as:

$$x_{i+d-k} = \mathbf{w} \mathbf{x}_{i+d-1}, \tag{12}$$

3 Experimental Set-Up

The methodology is divided into five steps: *i*) Preprocessing *ii*) Signal Decomposition and *iii*) Classification, where a simple k -nearest neighbors (k -NN) classifier is used.

3.1 Database and Preprocessing

This collection holds $M = 70$ electrocardiographic recordings (modified lead V2) from PhysioNet [11], each one including a set of reference annotations added

every minute of the recording indicating either the presence or absence of apnoea during each segment of time. The recordings were subdivided in three groups: apneic patients, with more than 100 minutes in apnea, borderline patients, with total apnea duration more than 5 and less than 99 minutes and control or normal patients, with less than 5 minutes in apnea. These recordings were obtained from men and women between the ages 27 through 63 years old, with weights ranging from 53 to 135 kg. From the database, 25 recordings were used as a training set for the classification algorithms. A second group with 25 recordings was used as a test set to measure the performance of the algorithms. It must be quoted that OSA diagnosis requires the extraction of HRV time series from each ECG recording, which can be estimated of the *QRS* complex. In this work, complex detection is carried out including linear filtering, followed by a non linear transformation, and adaptive decision rules, as well. Then, the HRV time serie is normalized, termed \mathbf{x}' , as recommended in [3]:

$$\mathbf{x}' = \frac{2(\mathbf{x} - \mathcal{E}\{\mathbf{x}\})}{\max_{\forall n}\{\mathbf{x}\} - \min_{\forall n}\{\mathbf{x}\}}, n \in N$$

3.2 Training Based on Signal Time Decomposition

The HRV signal is highly non-stationary, due to this, is necessary a better representation of the signal. In this work, for signal decomposition, three approaches are considered: the signal is decomposed directly through SSA, EDS and DLM. Tuning of the different schemes of considered signal decomposition throughout this training procedure is carried out by using the average classification accuracy for the automatic OSA detection, which is estimated using a *k-nn* classifier, followed by the well-known cross-validation methodology.

Singular Spectrum Analysis : for SSA tuning, the model order is chosen according to the classification accuracy, as shown in Figure 1(a). Since the lowest order presents a better result, then, order 1 is selected for considered model.

Dynamic Linear Model: the Bayesian information criterion (BIC) is used as order estimator, with the aim to find the recurrent coefficient of the embedded TVAR model, in terms of DLM form. The estimated mean of the order, given by BIC, is 10. Nevertheless, higher and lower orders are tested. The best classification accuracy is presented by the higher order, as seen in Figure 1(b).

Exponentially Damped Sinusoidal: Likewise, the model order is chosen according to the classification accuracy, as shown in Figure 1(c). Since the lowest order presents a better result, then, order 1 is chosen for this model.

4 Results and Discussion

Due to a huge dimension, which is inherent to the signal decomposition training space, the *k-nn* classifier requires a high computational cost with large memory

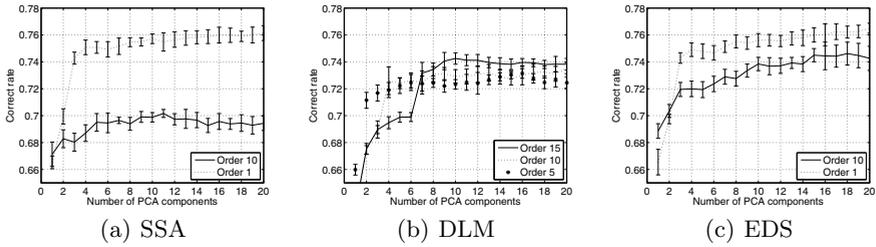


Fig. 1. Tuning of the different schemes of considered signal decomposition

amount. So, strong dimension reduction of such large training feature set should be carried out. Since there is a need for finding groups that are highly correlated (as it is the case with time series decomposition data), the principal component analysis is used throughout this study as unsupervised method to perform dimensionality reduction over the input training set in hand as recommended in [1].

Figure 2 illustrates how the estimated performance changes when ranging from 1 to 20 the number of considered principal components, which are a mutually linear independent combination of original variables. As seen, for the compared cases of signal decomposition, the proposed methodology is able to find differences between control and apneic events.

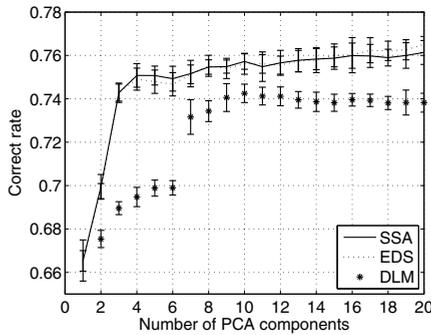


Fig. 2. Accuracy performance depending upon the signal decomposition when varying the number of used principal components

Table 1 summarizes the best performance obtained after the parameter tuning for each approach tested and the comparison between another outcomes related to OSA diagnosis. In the case of the results presented by [12], several approaches for the spectral splitting upon time–frequency planes in the concrete case of OSA detection are studied.

The model based on EDS presents the best performance, with $76.50 \pm 0.37\%$, followed by SSA with $76.13 \pm 0.55\%$. The linear model SSA and the non-linear

Table 1. Performance outcomes for time series analysis using a single tuned k -nn classifier

Approach	Ac	Number of Features
SSA	$76.13 \pm 0.55\%$	20
DLM	$74.25 \pm 0.42\%$	10
EDS	$76.50 \pm 0.37\%$	20
[12] Heuristic Approach	$75.42 \pm 0.88\%$	6
[12] Relevance Maximum variance	$75.22 \pm 0.58\%$	7

model EDS are very similar. That similar behavior can be explained. So, if adding complexity at the linear model SSA, the trend becomes more approximate to a simple non-linear model, as in case of EDS. Both models represent the signal with a lower order, i.e. the decomposition converges to the original signal through the highest principal eigenvalue, associated to the Hankel matrix in the projections (12) and (3). Besides, the order increases, on the opposite, the accuracy decreases; this can be explained because the other eigenvalues are informativeness for the process, adding just noise.

In the case of DLM, despite being a TVAR model, the accuracy is the lowest of the tested decompositions. The criterion for order selection is not coherent with the results of Figure 1(b), i.e. the order mean is not an adequate estimator due to the difference between orders in the training group are over five points. Additionally, the dynamic imposed by the model requires a higher order, near to the signal length, aiming to follow the signal trend as the other models. However, even with the above problem, the decomposition is able to find differences between control and apneic events with an acceptable precision.

5 Conclusions

A methodology for non-stationary signal discrimination is explored that imposes a finite rank modeling for a considered time series signal. The training procedure appraises the time-variant signal decomposition, aiming to find the model parameters. In this work, three approaches are tested: Singular Spectrum Analysis and Dynamic Linear Model as linear schemes, and Exponentially Damped Sinusoids as non-linear scheme. The methodology is proved particularly on biosignal data, namely, for Obstructive Sleep Apnoea syndrome detection. This data collection is chosen due to the fact that the interaction between the sympathetic and parasympathetic systems on the HRV recordings, gives rise to several non-stationary components added to the signal [3].

According to the performance of each time-variant signal decomposition, it can be concluded that the linear model SSA and the non-linear model EDS are very similar. It is possible because adding complexity at the linear model SSA, the trend is more approximate to a simple non-linear model as EDS. Regarding to DLM, the dynamic imposed by the model requires a higher order, near to the signal length, aiming to follow the signal trend as the other models. Therefore,

for OSA detection, it would be of benefit to explore needed enhancement by using more elaborated non-linear models.

Acknowledgments. This research is supported by “*Becas para estudiantes sobresalientes de Posgrado - Universidad Nacional de Colombia*”

References

1. Sepulveda-Cano, L., Gil, E., Laguna, P., Castellanos-Dominguez, G.: Selection of Nonstationary Dynamic Features for Obstructive Sleep Apnoea Detection in Children. *EURASIP Journal on Advances in Signal Processing* 2011, 10 (2011)
2. Golyandina, N., Osipov, E.: The “caterpillar”-ssa method for analysis of time series with missing values. *Journal of Statistical Planning and Inference* 137(8), 2642–2653 (2007)
3. Martínez-Vargas, J., Sepulveda-Cano, L., Travieso-Gonzalez, C., Castellanos-Dominguez, G.: Detection of obstructive sleep apnoea using dynamic filter-banked features. *Expert Systems with Applications* 39, 9118–9128 (2012)
4. West, M., Prado, R., Krystal, A.D.: Evaluation and comparison of eeg traces: Latent structure in nonstationary time series. *Journal of the American Statistical Association* 94, 1083–1095 (1999)
5. De Clercq, W., Vanrumste, B., Papy, J.M., Van Paesschen, W., Van Huffel, S.: Modeling common dynamics in multichannel signals with applications to artifact and background removal in eeg recordings. *IEEE Transactions on Biomedical Engineering* 52, 2006–2015 (2005)
6. Mahmoudvand, R., Zokaeo, M.: On the singular values of the hankel matrix with application in singular spectrum analysis. *Chilean Journal of Statistics* 3, 43–56 (2012)
7. Thomakos, D.: Optimal linear filtering, smoothing and trend extraction for m-period differences of processes with a unit root. Technical report (2008)
8. West, M., Harrison, J.: Bayesian forecasting and dynamic models, 2nd edn. Springer-Verlag New York, Inc., New York (1997)
9. Allu, G.K.: Estimating the parameters of exponentially damped sinusoids in noise. Technical report, Electrical Engineering Department, University of Rhode Island, Kingston, Rhode Island (2003)
10. Li, S., Lv, X., Sun, H., Hu, W.: Multiple scattering centers measurements using the kumaresan-tufts method. In: 2006 IET International Conference on Wireless, Mobile and Multimedia Networks, pp. 1–4 (2006)
11. Mendez, M., Corthout, J., Huffel, S., Matteucci, M., Penzel, T., Cerutti, S., Bianchi, A.: Automatic screening of obstructive sleep apnea from the eeg based on empirical mode decomposition and wavelet analysis. *Physiol. Meas.* 31, 273–289 (2010)
12. Martínez-Vargas, J., Sepúlveda-Cano, L., Castellanos-Domínguez, G.: On determining available stochastic features by spectral splitting in obstructive sleep apnea detection. In: 33rd Annual International Conference of the IEEE EMBS, Boston, Massachusetts USA, August 30-September 3, pp. 6079–6082. IEEE (2011)