Journal of Applied Statistics
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/cjas20

A wavelet-based time-varying autoregressive model for non-stationary and irregular time series

G. E. Salcedo [a], R. F. Porto [b], S. Y. Roa [a] & F. R. Momo [c]

[a] Department of Mathematics, University of Quindío, Carrera 15 Calle 12N, Armenia, Colombia

[b] Department of Operating Assets Restructuring, Bank of Brazil, Brasília, Brazil

[c] Institute of Science, National University of General Sarmiento, Buenos Aires, Argentina

Version of record first published: 23 Aug 2012

To cite this article: G. E. Salcedo, R. F. Porto, S. Y. Roa & F. R. Momo (2012): A wavelet-based time-varying autoregressive model for non-stationary and irregular time series, Journal of Applied Statistics, DOI:10.1080/02664763.2012.702267

To link to this article: http://dx.doi.org/10.1080/02664763.2012.702267

Journal of Applied Statistics
2012, iFirst article

Taylor & Francis
Taylor & Francis Group

A wavelet-based time-varying autoregressive model for non-stationary and irregular time series

G.E. Salcedo[a]*, R.F. Porto[b], S.Y. Roa[a] and F.R. Momo[c]

[a]*Department of Mathematics, University of Quindío, Carrera 15 Calle 12N. Armenia, Colombia;* [b]*Department of Operating Assets Restructuring, Bank of Brazil, Brasília, Brazil;* [c]*Institute of Science, National University of General Sarmiento, Buenos Aires, Argentina*

(*Received 25 November 2011; final version received 9 June 2012*)

In this work we propose an autoregressive model with parameters varying in time applied to irregularly spaced non-stationary time series. We expand all the functional parameters in a wavelet basis and estimate the coefficients by least squares after truncation at a suitable resolution level. We also present some simulations in order to evaluate both the estimation method and the model behavior on finite samples. Applications to silicates and nitrites irregularly observed data are provided as well.

Keywords: irregularly spaced time series; locally stationary processes; autoregressive model; multiresolution analysis; wavelets

1. Introduction

In time series analysis, most samples are regularly observed over time in which case we say the time series is regular. However, in some real situations it is not possible to obtain equally spaced observations. For these cases, when the time series are said to be irregular, there are some approaches to modeling but most of them focus only on deterministic trends while other approaches consider irregular time series as data with missing observations (see [8,9,15,16,22]) or apply smoothing techniques, as was done by Cipra [7] and Kitagawa [17]. Simple exponential smoothing for irregular data was also suggested by Wright [24] and Cipra [7], and double exponential smoothing is given by Cipra [7]. Parametric models have been proposed by Broersen and Bos [2], who present an algorithm for maximum likelihood estimation of autoregressive moving average models for irregular data. In [16,17,22,23], the authors have used state-space representations to fit continuous-time autoregressions to unequally spaced time series.

*Corresponding author. Email: gsalcedo@uniquindio.edu.co

ISSN 0266-4763 print/ISSN 1360-0532 online

http://dx.doi.org/10.1080/02664763.2012.702267
http://www.tandfonline.com

While these models consider mainly stationary time series, Dahlhaus *et al.* [11] have proposed time-varying autoregressive (tv-AR) models for regular time series that have a locally stationary behavior and used a temporal rescaling in order to facilitate the development of an asymptotic theory. Their models are alike the functional-coefficient autoregressive (FAR) models (see [4,5], for instance), but Dahlhaus' tv-AR models are based on wavelet expansions of the functional parameters with coefficients estimated by ordinary least squares. This technique has been applied by Chiann and Morettin [6] for time-varying linear systems and by Sato *et al.* [21] for vector autoregressive models with parameters varying in time, both based on locally stationary processes. On the other hand, Cai and Brown [3] in the context of regression models, have presented a procedure to properly fit and estimate functional trend models for irregular time series using wavelets.

In this paper we propose an autoregressive model with parameters varying in time in order to model the dynamics of a non-stationary time series irregularly observed over time. The non-stationarity is explained by the functional parameters and the irregularity is explicit in the functional indexes. Our model can be considered as an extension of tv-AR models to the case of irregular time series.

The paper is organized as follows. Section 2 briefly describes wavelet bases and functional wavelet expansions. Section 3 introduces the non-equispaced model and the estimation procedure for autoregressive models varying in time. In Section 4 we give some statistical properties of our estimator. A simulation study is given in Section 5 and two applications are presented in Section 6. Finally, some conclusions are presented in Section 7.

2. Wavelets and wavelet approximations

An orthonormal wavelet basis is generated from dilation and translation of a "father" wavelet ϕ and a "mother" wavelet ψ. These functions are assumed to be compactly supported in $[0, T]$ and ϕ satisfies $\int \phi = 1$. A wavelet is *r-regular* if it has r vanishing moments and r continuous derivatives. A wavelet ψ satisfies the admissibility condition if it is 1-regular.

Let

$$\phi_{j,k}(t) = 2^{j/2}\phi(2^j t - k) \quad \text{and} \quad \psi_{j,k}(t) = 2^{j/2}\psi(2^j t - k), \quad j, k \in \mathbb{Z},$$

where $\phi_{j,k}(t)$ and $\psi_{j,k}(t)$ are the scaling and the wavelet functions, respectively, at level j and translation index k. Thus, $\psi_{j,k}$ has support $[2^{-j}k, 2^{-j}(T + k)]$. Notice that the support of the wavelets are translated by shifts of 2^{-j}.

Also, the periodized wavelets (see [10]) are given by

$$\tilde{\phi}_{j,k}(t) = \sum_{l\in\mathbb{Z}} \phi_{j,k}(t - l) \quad \text{and} \quad \tilde{\psi}_{j,k}(t) = \sum_{l\in\mathbb{Z}} \psi_{j,k}(t - l),$$

for $t \in [0, 1]$. These are the wavelets that we will use in this paper and so the superscript "$\sim$" will be suppressed thereafter. For a given $j_0 \in \mathbb{Z}$, the collection

$$\{\phi_{j_0,k},\ k = 0, \ldots, 2^{j_0} - 1; \psi_{j,k}, j \geq j_0, k = 0, \ldots, 2^j - 1\}$$

constitutes an orthonormal basis of $L_2[0, 1]$, the space of square-integrable functions. Notice that for each level j we have 2^j basis functions.

Such an orthogonal wavelet basis has an associated multiresolution analysis on $[0, 1]$ that enables one to analyze the data through a number of resolution scales. Let V_j and W_j, $j \in \mathbb{Z}$, be the closed linear subspaces generated by $\{\phi_{j,k}, k = 0, \ldots, 2^j - 1\}$ and $\{\psi_{j,k}, k = 0, \ldots, 2^j - 1\}$, respectively. Then,

(i) $\cdots \subset V_{j_0-1} \subset V_{j_0} \subset V_{j_0+1} \subset \cdots \subset V_j \subset \cdots$,
(ii) $\overline{\cup_{j=-\infty}^{\infty} V_j} = L_2[0,1]$,
(iii) $V_{j+1} = V_j \oplus W_j$,
(iv) $W_j \perp V_j$.

Denote the usual inner product by $\langle \cdot, \cdot \rangle$. For a given square-integrable function $f(t)$, $t \in [0,1]$, we have that the wavelet transform is given by

$$c_{j_0,k} = \langle f, \phi_{j_0,k} \rangle \quad \text{and} \quad d_{j,k} = \langle f, \psi_{j,k} \rangle, \tag{1}$$

where $c_{j_0,k}$ and $d_{j,k}, j \geq j_0$, and $k = 0, \ldots, 2^j - 1$, are the wavelet coefficients of the coarse and the details scales, respectively.

So, for a given $j_0 \geq 0$, the function $f(t)$ can be expanded, in L_2 norm sense, into an infinite wavelet series as

$$f(t) = \sum_{k=0}^{2^{j_0}-1} c_{j_0,k}\phi_{j_0,k}(t) + \sum_{j=j_0}^{\infty} \sum_{k=0}^{2^j-1} d_{j,k}\psi_{j,k}(t), \quad t \in [0,1], \tag{2}$$

i.e., the wavelet transform decomposes the function into different resolution components.

In practice, $j_0 = 0$ and the expansion in Equation (2) is approximated by the finite summation

$$f(t) = c_{0,0}\phi_{0,0}(t) + \sum_{j=0}^{J-1} \sum_{k=0}^{2^j-1} d_{j,k}\psi_{j,k}(t) = \sum_{j=-1}^{J-1} \sum_{k=0}^{2^j-1} d_{j,k}\psi_{j,k}(t), \quad t \in [0,1], \tag{3}$$

where $d_{-1,0} = c_{0,0}$ and $\psi_{-1,0}(t) = \phi_{0,0}(t)$, and J is chosen based on the expected smoothing degree of the function f.

"Daublets", "Symmlets" and "Coiflets" are the most used wavelet bases and were introduced by Daubechies in [12]. These wavelets are orthogonal and have compact support. Other useful wavelet bases are "Morlet", "Mexican hat" and "Shanon" (see [18] for example).

3. Model and estimation procedure

Consider the irregular autoregressive model of known order $p \geq 1$

$$X_{t_i} = f_1(t_i)X_{t_{i-1}} + f_2(t_i)X_{t_{i-2}} + \cdots + f_p(t_i)X_{t_{i-p}} + \epsilon_{t_i}, \tag{4}$$

for non-equispaced observations $\{X_{t_i}\}$, $i = 1, 2, \ldots, T = 2^N$, $N \in \mathbb{N}$, $0 < t_1 < t_2 < \cdots < t_T < 1$ and independent and identically distributed errors $\epsilon_{t_i} \sim N(0, \sigma_\epsilon^2)$. We could alleviate the requirement that the sample size T is a power of 2. However, as will be clear ahead, our estimation procedure uses the values of $\psi_{j,k}(i/T)$, for $i = 1, \ldots, T$, which are computationally easier to obtain when this requirement is met. Since using sample sizes that are a power of 2 is convenient for us and it is usual in the wavelet literature, we decided to keep it up throughout the paper. The functions $f_l(t_i)$ are intended to capture the non-stationarity and the data irregularity. Our main problem consists in estimating all the p functions $f_l(\cdot)$ and the variance error σ_ϵ^2 from a finite set of irregular observations $\{X_{t_i}, i = 1, 2, \ldots, T\}$.

Besides its ability of fitting some non-stationary time series, model (4) can be physically motivated by measurements X_{t_i}, taken at time or location t_i, that are simultaneously dependent on the continuous time $t \in (t_{i-1}, t_{i+1})$ (as indicated by the notation) and on previous measurements

$X_{t_{i-1}}, \ldots, X_{t_{i-p}}$. This can be seen as a sort of mix between deterministic continuous-time models $f_1, \ldots, f_p$ and a stochastic discrete-time model based on $X_{t_{i-1}}, \ldots, X_{t_{i-p}}$. Consider, for instance, that the indexes t_i are locations of waste water treatment plants on a river. Water quality measurements made at one station clearly depend on water treatments measured in previous stations and on the length of the river upstream of its previous plant. This happens because the water treated by one plant is expected to show a decreasing quality pattern for an increasing length from the plant, mainly due to chemical dilution, new waste discharges, rain and other physical causes (see, for instance, [20]).

Our estimation procedure is based on least-squares estimation of the coefficients of a previous wavelet expansion for the functions $f_l(\cdot), l = 1, \ldots, p$. Since standard wavelet procedures are well suited for equispaced samples, wherein t_i are equally spaced on [0, 1], non-equispaced samples should not in general be directly treated as equispaced. In order to deal with this aspect, we suppose that $t_i = H^{-1}(i/T)$, $i = 1, 2, \ldots, T$, where H is a mapping cumulative density function H on [0, 1] (see [2]).

Consequently,

$$f_l(t_i) = f_l\left(H^{-1}\left(\frac{i}{T}\right)\right), \quad i = 1, 2, \ldots, T, \; l = 1, \ldots, p, \tag{5}$$

i.e. each value $f_l(t_i)$ from an unequally spaced point is mapped to a value of the composite function $f_l \circ H^{-1}$ at the equally spaced point i/T. Notice that the points t_i are assumed to be fixed, not randomly drawn from H.

Let the functions $g_l = f_l \circ H^{-1}$ be such that $f_l(t_i) = g_l(i/T)$, meaning that $g_l(i/T)$ is the equispaced representation of $f_l(t_i)$ mapped through the function H.

Thus, the equispaced model equivalent to Equation (4) is given by

$$X_{t_i} = g_1\left(\frac{i}{T}\right)X_{t_{i-1}} + g_2\left(\frac{i}{T}\right)X_{t_{i-2}} + \cdots + g_p\left(\frac{i}{T}\right)X_{t_{i-p}} + \epsilon_{t_i}. \tag{6}$$

Using only one wavelet basis to expand each functional parameter $g_l(\cdot)$, model (6) can be represented as

$$X_{t_i} = \sum_{l=1}^{p}\left[\sum_{j=-1}^{J-1}\sum_{k=0}^{2^j-1} d_{j,k}^{l}\psi_{j,k}\left(\frac{i}{T}\right)\right]X_{t_{i-l}} + s_{t_i} + \epsilon_{t_i}, \quad i = 1, 2, \ldots, T, \tag{7}$$

where

$$s_{t_i} = \sum_{l=1}^{p}\left[\sum_{j\geq J}\sum_{k} d_{j,k}^{l}\psi_{j,k}\left(\frac{i}{T}\right)\right]X_{t_{i-l}}$$

is the error due to the truncation in the resolution level $J - 1$. In practice, one usually chooses J searching from all the available resolution levels.

If we have p additional observations $X_{t_i}, i = 0, -1, \ldots, -p + 1$, which can be used as initial values, then the model (7) can be represented in matrix form as

$$\mathbf{X} = \Psi\mathbf{D} + \mathbf{s} + \boldsymbol{\epsilon}, \tag{8}$$

where

$$\mathbf{X} = (X_{t_1}, X_{t_2}, \ldots, X_{t_T})', \quad \mathbf{s} = (s_{t_1}, s_{t_2}, \ldots, s_{t_T})', \quad \boldsymbol{\epsilon} = (\epsilon_{t_1}, \epsilon_{t_2}, \ldots, \epsilon_{t_T})',$$

$$\mathbf{D} = (d_{-1,0}^{1}, d_{0,0}^{1}, \ldots, d_{\Delta,0}^{1}, d_{\Delta,1}^{1}, \ldots, d_{\Delta,2^{\Delta}-1}^{1}, \ldots, d_{-1,0}^{p}, d_{0,0}^{p}, \ldots, d_{\Delta,0}^{p}, d_{\Delta,1}^{p}, \ldots, d_{\Delta,2^{\Delta}-1}^{p})',$$

with $\Delta = J - 1$, and the transpose of the vector $\mathbf{a}$ is denoted by $\mathbf{a}'$. Also, consider the matrix $\Psi = (\Psi^1 \vdots \Psi^2 \vdots \dots \vdots \Psi^p)$, where

$$\Psi^l = \begin{bmatrix} \psi_{-1,0}\left(\frac{1}{T}\right)X_{t_{1-l}} & \psi_{0,0}\left(\frac{1}{T}\right)X_{t_{1-l}} & \dots & \psi_{\Delta,2^\Delta-1}\left(\frac{1}{T}\right)X_{t_{1-l}} \\ \psi_{-1,0}\left(\frac{2}{T}\right)X_{t_{2-l}} & \psi_{0,0}\left(\frac{2}{T}\right)X_{t_{2-l}} & \dots & \psi_{\Delta,2^\Delta-1}\left(\frac{2}{T}\right)X_{t_{2-l}} \\ \vdots & \vdots & \ddots & \vdots \\ \psi_{-1,0}\left(\frac{T-1}{T}\right)X_{t_{T-1-l}} & \psi_{0,0}\left(\frac{T-1}{T}\right)X_{t_{T-1-l}} & \dots & \psi_{\Delta,2^\Delta-1}\left(\frac{T-1}{T}\right)X_{t_{T-1-l}} \\ \psi_{-1,0}\left(\frac{T}{T}\right)X_{t_{T-l}} & \psi_{0,0}\left(\frac{T}{T}\right)X_{t_{T-l}} & \dots & \psi_{\Delta,2^\Delta-1}\left(\frac{T}{T}\right)X_{t_{T-l}} \end{bmatrix},$$

for $l = 1, 2, \dots, p$.

In model (8), the wavelet coefficients are now the parameters of interest. The ordinary least-squares estimator of $\mathbf{D}$ is given by

$$\begin{aligned} \hat{\mathbf{D}} &= (\Psi'\Psi)^{-1}\Psi'\mathbf{X} \\ &= (\Psi'\Psi)^{-1}\Psi'(\Psi\mathbf{D} + \mathbf{s} + \boldsymbol{\epsilon}) \\ &= (\Psi'\Psi)^{-1}(\Psi'\Psi)\mathbf{D} + (\Psi'\Psi)^{-1}\Psi'\mathbf{s} + (\Psi'\Psi)^{-1}\Psi'\boldsymbol{\epsilon} \\ &= \mathbf{D} + (\Psi'\Psi)^{-1}\Psi'\mathbf{s} + (\Psi'\Psi)^{-1}\Psi'\boldsymbol{\epsilon} \\ &= \mathbf{D} + T_1 + T_2, \end{aligned}$$

where Ψ' is the transpose of Ψ and the term T_1 represents the bias due to the truncation in the resolution level $J - 1$. Hence, from

$$\hat{\mathbf{D}} = (\hat{d}^1_{-1,0}, \hat{d}^1_{0,0}, \dots, \hat{d}^1_{\Delta,0}, \hat{d}^1_{\Delta,1}, \dots, \hat{d}^1_{\Delta,2^\Delta-1}, \dots, \hat{d}^p_{-1,0}, \hat{d}^p_{0,0}, \dots, \hat{d}^p_{\Delta,0}, \hat{d}^p_{\Delta,1}, \dots, \hat{d}^p_{\Delta,2^\Delta-1})'$$

we can estimate each value $g_l(i/T) = f_l(t_i)$, and consequently every observation X_{t_i} through

$$\hat{X}_{t_i} = \sum_{l=1}^{p}\left[\sum_{j=-1}^{J-1}\sum_{k=0}^{2^j-1}\hat{d}^l_{j,k}\psi_{j,k}\left(\frac{i}{T}\right)\right]X_{t_{i-l}} = \sum_{l=1}^{p}\hat{g}_l\left(\frac{i}{T}\right)X_{t_{i-l}} = \sum_{l=1}^{p}\hat{f}_l(t_i)X_{t_{i-l}}, \tag{9}$$

$i = 1, \dots, T$. The error variance can be estimated as the sample variance of the residuals $\hat{\epsilon}_{t_i} = X_{t_i} - \hat{X}_{t_i}$.

4. Some theoretical properties

Consider the model (8) and let $\hat{\mathbf{D}} = (\Psi'\Psi)^{-1}\Psi'\mathbf{X}$ be the estimator of the coefficients vector $\mathbf{D}$ that gives rise to the predictor $\hat{X}_{t_i}$ of Equation (9), and to the residuals $\hat{\epsilon}_{t_i} = X_{t_i} - \hat{X}_{t_i}$, for $i = 1, \dots, T$. In this case, conditional to the observed values $\mathbf{X}^+ = (X_{t_{-p+1}} = x_{t_{-p+1}}, \dots, X_{t_T} = x_{t_T})$,

(i) the estimator $\hat{\mathbf{D}}$ is biased as

$$E(\hat{\mathbf{D}}|\mathbf{X}^+) = \mathbf{D} + (\Psi'\Psi)^{-1}\Psi'\mathbf{s},$$

(ii) independent of the resolution level J chosen for truncation, the covariance matrix of $\hat{\mathbf{D}}$ is the same, such that

$$\mathrm{Var}(\hat{\mathbf{D}}|\mathbf{X}^+) = \sigma^2(\Psi'\Psi)^{-1},$$

(iii) the expected value of the residuals is equal to

$$E(\hat{\epsilon}_{t_i}|\mathbf{X}^+) = s_{t_i} - \Psi_i(\Psi'\Psi)^{-1}\Psi'\mathbf{s},$$

(iv) the variance of the residuals is equal to

$$\mathrm{Var}(\hat{\epsilon}_{t_i}|\mathbf{X}^+) = \sigma^2 - \sigma^2\Psi_i(\Psi'\Psi)^{-1}\Psi_i',$$

(v) the expectation of the predicted value $\hat{X}_{t_i}$ is equal to

$$E(\hat{X}_{t_i}|\mathbf{X}^+) = \Psi_i\mathbf{D} + \Psi_i(\Psi'\Psi)^{-1}\Psi'\mathbf{s},$$

(vi) the variance of the predicted value $\hat{X}_{t_i}$ is close to the variance of the respective residual so that

$$\mathrm{Var}(\hat{X}_{t_i}|\mathbf{X}^+) = \sigma^2 - \mathrm{Var}(\hat{\epsilon}_{t_i}|\mathbf{X}^+),$$

where Ψ_i is the vector formed by the line i of the matrix Ψ, for $i = 1, \ldots, T$. These items are proved by simple linear algebra.

Notice that, for fixed T, these properties are not as good as we usually desire. In fact, the biases of model (8), cited at (i), (iii) and (v) can be substantial with small samples, and the analyst must consider this in practice, when searching for the most appropriate level J, from all the available resolution levels. However, under some assumptions on the functions $g_l(\cdot)$, the asymptotic properties are better than the finite sample properties. To see this, note that the functions $g_l(\cdot)$ lie in the set

$$\mathcal{F}^{(l)} = \left\{ f(t) : f(t) = \sum_{k=0}^{2^{j_0}-1} c^l_{j_0,k}\phi_{j_0,k}(t) + \sum_{j=j_0}^{\infty}\sum_{k=0}^{2^j-1} d^l_{j,k}\psi_{j,k}(t) \right\},$$

where

$$\sup_k |c^l_{j_0,k}| < \infty,$$

and

$$\left(\sum_{j=j_0}^{\infty} \left[2^{js_l p_l} \sum_{k=0}^{2^j-1} |d^l_{j,k}|^{p_l} \right]^{q_l/p_l} \right)^{1/q_l} < \infty,$$

with $s_l = r_l + \frac{1}{2} - 1/p_l$, for some fixed $j_0 \in \mathbb{Z}$ and coefficients $c^l_{j_0,k}$ and $d^l_{j,k}$, where $l = 1, \ldots, p; k = 0, \ldots, 2^j - 1$, and $j \geq j_0$. The class $\mathcal{F}^{(l)}$ is closely related to Besov classes with the same parameters r_l, p_l and q_l, where $r_l \geq 1$ denotes the degree of smoothing and p_l, q_l specify the norm to measure the smoothness, with $1 \leq p_l, q_l \leq \infty$ (see [11]).

To ensure sufficient regularity, assume also that $\tilde{s}_l > 1$, where $\tilde{s}_l = r_l + \frac{1}{2} - 1/\min\{p_l, 2\}$, so that (see [11,13])

$$\sup_{g_l \in \mathcal{F}^{(l)}} \left(\sum_{j=J}^{\infty}\sum_{k=0}^{2^j-1} |d^l_{j,k}|^2 \right) = O(2^{-2J\tilde{s}_l}).$$

In this case, if J is chosen such that $2^{-2J\tilde{s}_l} = O(T^{-2r_l/(2r_l+1)})$ (which according to Dahlhaus *et al.* [11] holds when $2^{J-1} \leq T^{1/2} \leq 2^J$), then the loss due to truncation is also of order $T^{-2r_l/(2r_l+1)}$. Thus, the term T_1 is such that

$$\begin{aligned} \|T_1\|_2 &= \|(\Psi'\Psi)^{-1}\Psi'\mathbf{s}\|_2 = E^{1/2}(|(\Psi'\Psi)^{-1}\Psi'\mathbf{s}|^2) \\ &= O_p((2^{-J\min\{\tilde{s}_l\}} + T^{-1/2}2^{-J\min\{r_l-1/2-1/(2p_l)\}})\sqrt{\log(T)}) \\ &= O_p(T^{-1/2-\tau(J)}), \end{aligned}$$

for some $\tau(J) > 0$. Consequently, $\sqrt{T}\|T_1\|_2 = O_p(T^{-\tau(J)}) = o_p(1)$, and the estimator $\hat{\mathbf{D}}$ enjoys some desired asymptotic properties such as unbiasedness.

In order to achieve good asymptotic properties, we have roughly assumed that the function $g_l \in B^{r_l}_{p_l,q_l}$, the cited Besov functional class with parameters r_l, p_l and q_l, where $r_l > 1 + 1/\min\{p_l, 2\} - \frac{1}{2}$, for $l = 1, \ldots, p$. Although we could prefer assumptions on the functions f and H or H^{-1}, the assumptions made on g_l are sufficient to guarantee good asymptotic results on a wide range of function classes for f and H as well, for practical purposes. To see this, first notice that, since $B^{r_l}_{p_l,q_l} \subset B^{r_l}_{p_l,\infty}$, then $g_l \in B^{r_l}_{p_l,\infty}$. If we also assume that $r_l > 1 + 1/p_l$ and that H^{-1} has a local Lipschitz behavior, i.e. $H^{-1} \in B^{r_l}_{p_l,q} \cap L_\infty$, for any $1 \leq q \leq \infty$, where L_∞ denotes the class of Lipschitz functions, then our assumptions imply that f belongs to a locally Sobolev class of functions that contains all the locally Besov functional classes with parameters r_l, p_l and q [1].

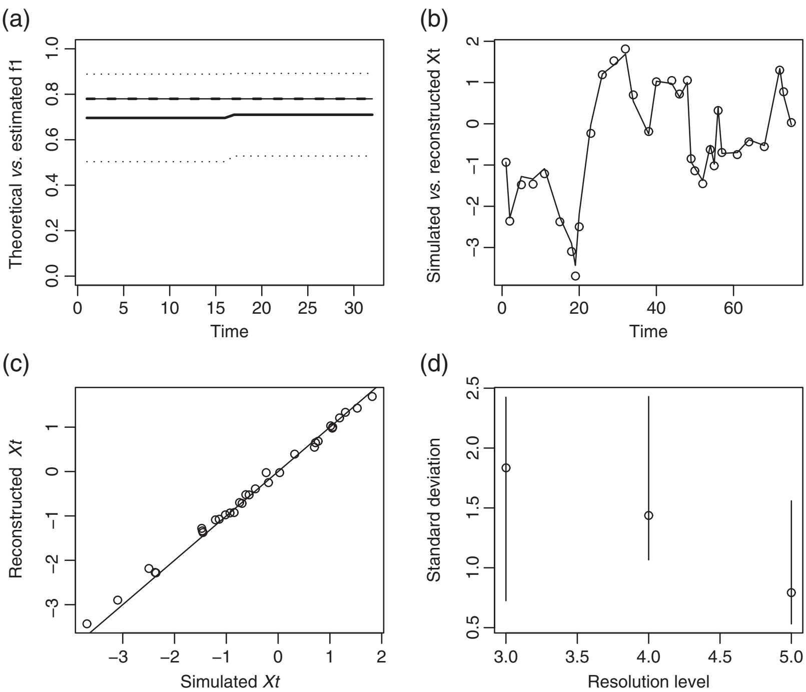

Figure 1. Simulation results for $T = 32$ and a constant function. (a) Theoretical function (dashed), estimated mean function (filled) and the intervals of one standard deviation (dotted). (b) Example of X values (circles) and respective reconstructed values (lines). (c) Simulated and reconstructed X values, using g and its estimate, respectively, in Equation (6), for the example of (b). (d) Residual correlation check for the example of (b) using the mean absolute deviation (std. dev.) from zero of the wavelet coefficients at each resolution level j 177 × 177 mm.

5. Simulation results

In this section we present some simulation results to evaluate the performance of the estimation procedure for irregular series fitted by the proposed model. In order to do this, we have studied the proposed time-varying irregular autoregressive model of known orders $p = 1$ and $p = 2$.

The simulations were done in the R language using the Wavethresh package [19].

5.1 *The design*

For models of order $p = 1$, we have generated 1000 irregular time series of sample length $T = 32$ from a stationary process for the constant function $f_1(t) = 0.78$, as well as from a time-varying process with $f_1(t) = f(t) + 0.1$ if $0 < t \leq T/4$ or $3T/4 < t \leq T$, and $f_1(t) = -0.25$ otherwise with $f(t) = -0.15\cos(2\pi t/T) - 0.35$.

For the irregular autoregressive models of order $p = 2$, we have generated 1000 irregular time series, using two different pairs of functional parameters with $T = 128$. The first pair of functions is given by $f_1(t) = 0.35$ and $f_2(t) = \frac{1}{3}\cos(2\pi t/T) + 0.35$, and the second pair of functions is given by $f_1(t) = \frac{1}{3}\sin(2\pi t/T) + 0.35$ and $f_2(t) = \frac{1}{3}\cos(2\pi t/T) + 0.35$.

From each generated time series and for all values of t_i, $i = 1, 2, \ldots, T$, we have estimated the theoretical autoregressive functions $f_l(t_i)$, $l = 1, 2$, the estimated values (used to build the fitted series)

$$\hat{X}_{t_i} = \hat{f}_1(t_i)X_{t_{i-1}} + \hat{f}_2(t_i)X_{t_{i-2}}$$

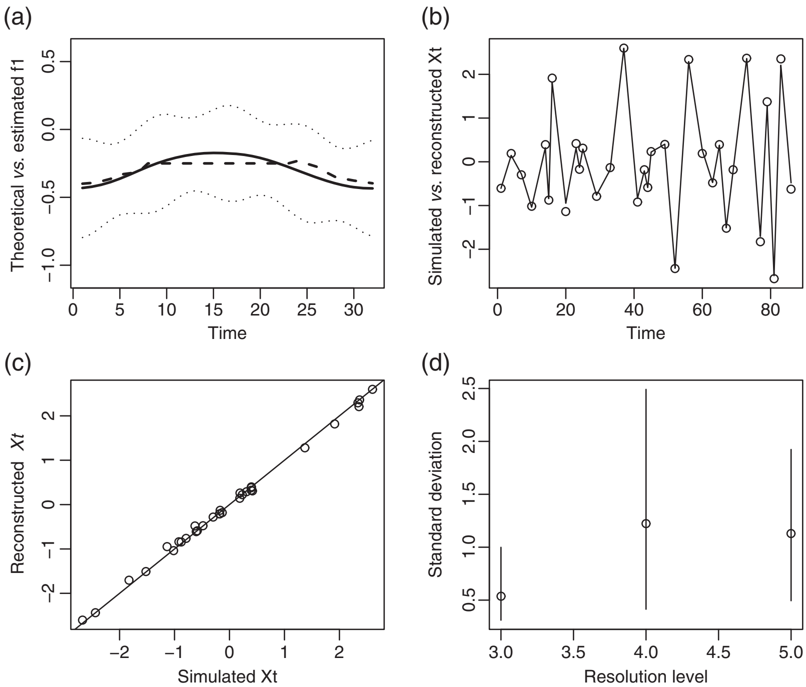

Figure 2. Simulation results for $T = 32$ and a time-varying function. (a) Theoretical function (dashed), estimated mean function (filled) and the intervals of one standard deviation (dotted). (b) Example of X values (circles) and respective reconstructed values (lines). (c) Simulated and reconstructed X values, using g and its estimate, respectively, in Equation (6), for the example of (b). (d) Residual correlation check for the example of (b) using the mean absolute deviation (std. dev.) from zero of the wavelet coefficients at each resolution level j 177 × 177 mm.

and the reconstructed values

$$\tilde{X}_{t_i} = \hat{f}_1(t_i)X_{t_{i-1}} + \hat{f}_2(t_i)X_{t_{i-2}} + \epsilon_{t_i}.$$

Notice that in the reconstructed values the error terms used are the simulated ones; obviously, when $p = 1, f_2(t_i) = 0$ for all t_i values and was not estimated.

The irregular times t_i were fixed at the beginning of the simulation where $t_1 = 1$ and the fixed intervals $\delta_i = t_{i+1} - t_i$ previously obtained as $\delta_i \sim U(1, 5)$, for $i = 1, 2, \ldots, T - 1$.

For the stationary case we have used the Haar wavelet, denoted by DB1, while for the time-varying models we have used Daublets with eight vanishing moments, denoted here by DB8 (see [12]).

Remark The presence of correlated noise manifests itself as a dependence of the standard deviations σ_j of the wavelet coefficients on the level j (see [14] for details). Figures 1(d) to 4(d) show σ_j (as estimated by the median absolute deviation from zero) versus j, along with 95% confidence intervals (CIs) obtained by a boostrapping procedure. We have just used the three highest resolution levels. The fact that all the CIs overlap suggests that the residuals are uncorrelated as expected from a reasonable estimation procedure. Although not shown, the CIs do not overlap for the original series.

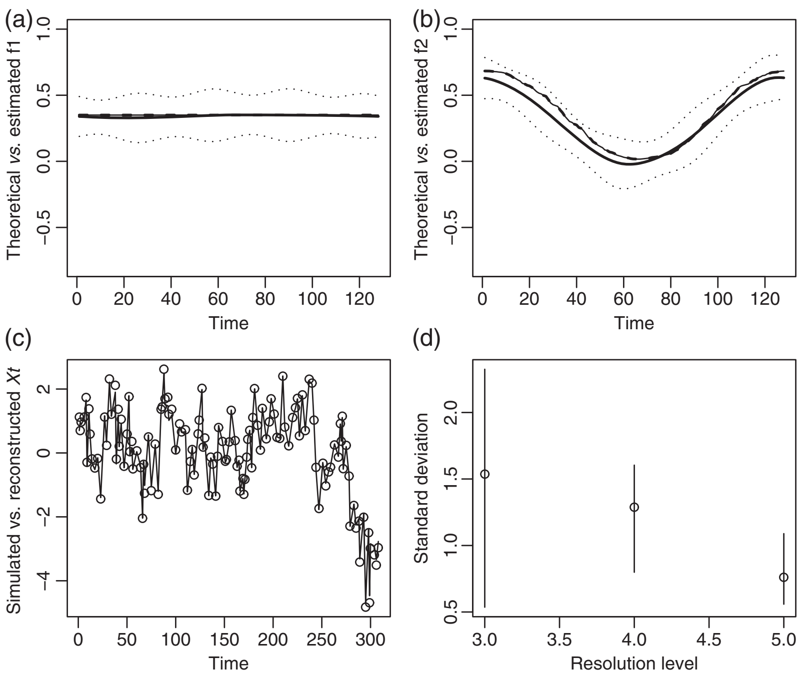

Figure 3. Simulation results for $T = 128$ with the constant and cosine functions. (a) Theoretical function (dashed), estimated mean function (filled) and the intervals of one standard deviation (dotted). (b) Example of X values (circles) and respective reconstructed values (lines). (c) Simulated and reconstructed X values, using g and its estimate, respectively, in Equation (6), for the example of (b). (d) Residual correlation check for the example of (b) using the mean absolute deviation (std. dev.) from zero of the wavelet coefficients at each resolution level j 177 × 177 mm.

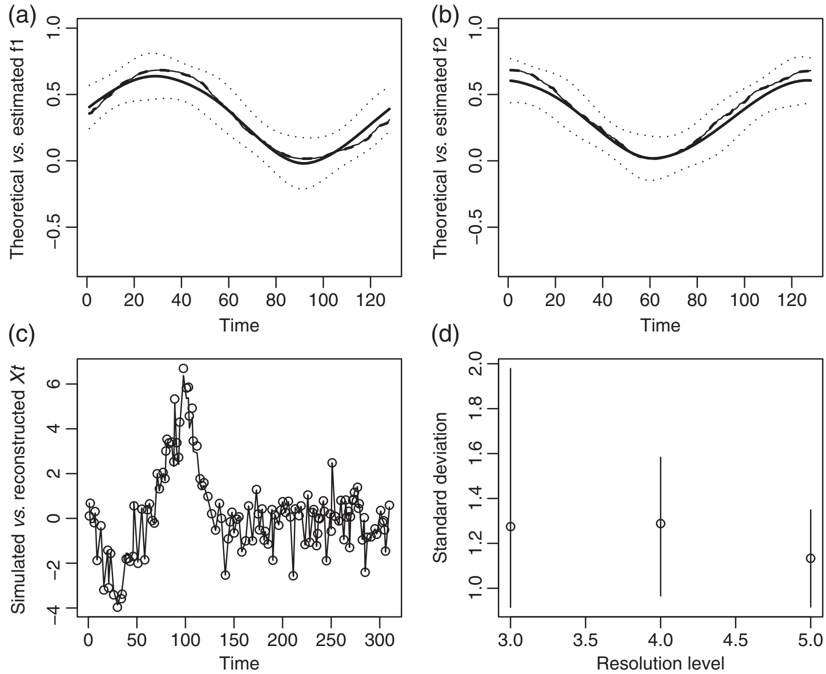

Figure 4. Simulation results for $T = 128$ with the sine and cosine functions. (a) Theoretical function (dashed), estimated mean function (filled) and the intervals of one standard deviation (dotted). (b) Example of X values (circles) and respective reconstructed values (lines). (c) Simulated and reconstructed X values, using g and its estimate, respectively, in Equation (6), for the example of (b). (d) Residual correlation check for the example of (b) using the mean absolute deviation (std. dev.) from zero of the wavelet coefficients at each resolution level j 177 × 177 mm.

5.2 *Results*

Figures 1 and 2 show the results for the simulated series from the stationary and non-stationary irregular autoregressive processes of order $p = 1$ for $T = 32$, respectively. In these cases, we have used $J = 1$ because, compared to $J = 2, 3$, this value resulted in better adjusted curves with narrow intervals in figures (a) and less in scattered figures (c). Figures (a) contain the theoretical function (dashed curve), the mean (at each point t_i) of the 1000 estimated functions (filled curve) and the interval of one standard deviation (dotted curves). Figures (b) contain one example of the simulated irregular series (circles) and the corresponding reconstructed series (lines). Figures (c) show the validation plot containing the simulated versus the reconstructed values $\tilde{X}_{t_i}$ for the example of (b). Figures (d) are residual correlation check plots for the example of (b) using the mean absolute deviation from zero of the wavelet coefficients at each resolution level j.

Figures 3 and 4 show the results for the simulated series from the non-stationary irregular autoregressive processes of order $p = 2$ for $T = 128$. In these cases, we have used $J = 2$ because compared to $J = 3, 4$, this value resulted in better adjusted curves with narrow intervals in figures (a) and (b) and correlated residuals in figures (d). Figures (a) and (b) contain the theoretical function (dashed curve), the mean (at each point t_i) of the 1000 estimated functions (filled curve) and the interval of one standard deviation (dotted curves) for $f_1(t)$ and $f_2(t)$ functions, respectively. Figures (c) contain one example of the simulated irregular series (circles) and the corresponding reconstructed series (lines). Figures (d) are residual correlation check plots for the example of (c) using the mean absolute deviation from zero of the wavelet coefficients at each resolution level j.

In general, we consider that the performance of the estimation procedure for finite samples is good because, in each case, both the simulated and the reconstructed series match each other reasonably well even for small sample size. Since in figures (d) the CIs overlap we can conclude that the residuals are uncorrelated and we have adequately eliminated the temporal dependence through the model. As expected, the estimation is better for larger values of T.

6. Application

We have applied the proposed model to irregular time series of 32 points of silicates and nitrites that were sampled from the waters of the Beagle Channel in Argentina. This channel separates the Tierra del Fuego from the islands in its south and monitoring its water quality is important, for instance, to farmers in this area. The data were collected from March 2005 to December 2006 at irregularly spaced dates due to weather and operational conditions.

Since the sample size is small, the times are unequally spaced and the series seems to exhibit a non-stationary behavior, we have tried to fit a time-varying irregular autoregressive model of order $p = 1$ to these data. The claimed non-stationarity behavior can be seen in Figures 5(b) and 6(b) for the silicates and nitrites series, respectively.

In order to choose the specific wavelets for the model, we have tried Daublets with $N = 6,8$ and 10 vanishing moments. In each case we have chosen the wavelet with the smallest residual mean square errors (MSE) to proceed with the analysis. We have decided to use $J = 2$ in this application after experimenting with some other values.

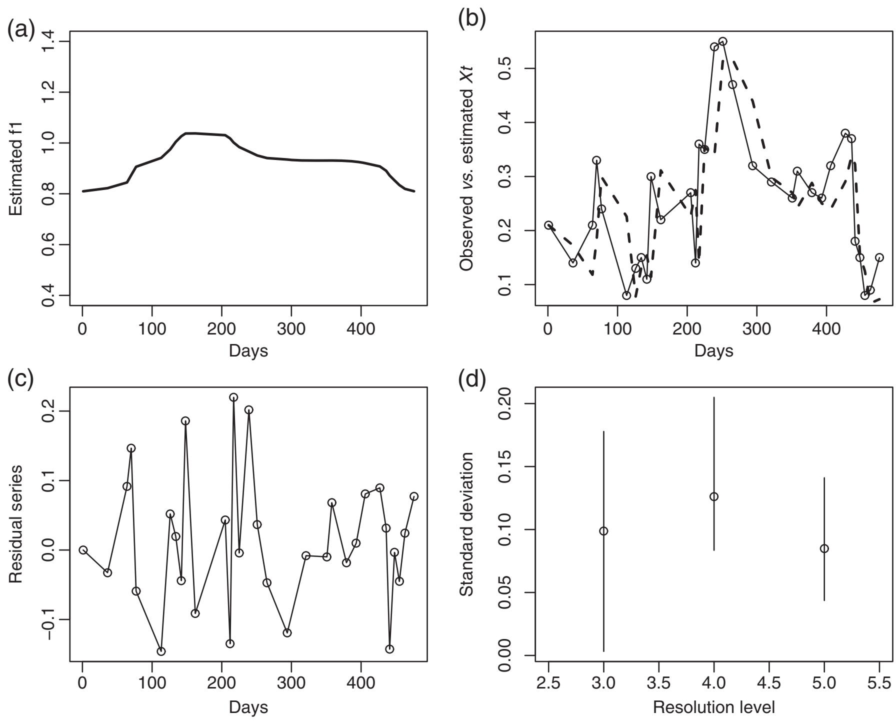

Figure 5. Results of fitting the time-varying irregular autoregressive model of order $p = 1$ to the water silicates series in the Beagle Channel from March 2005 to December 2006. (a) Estimated AR(1) functional parameter with DB8. (b) Observed (solid lines) and fitted time series (dashed lines). (c) Irregular residual series. (d) Residual correlation check of the fitted model 177×177 mm.

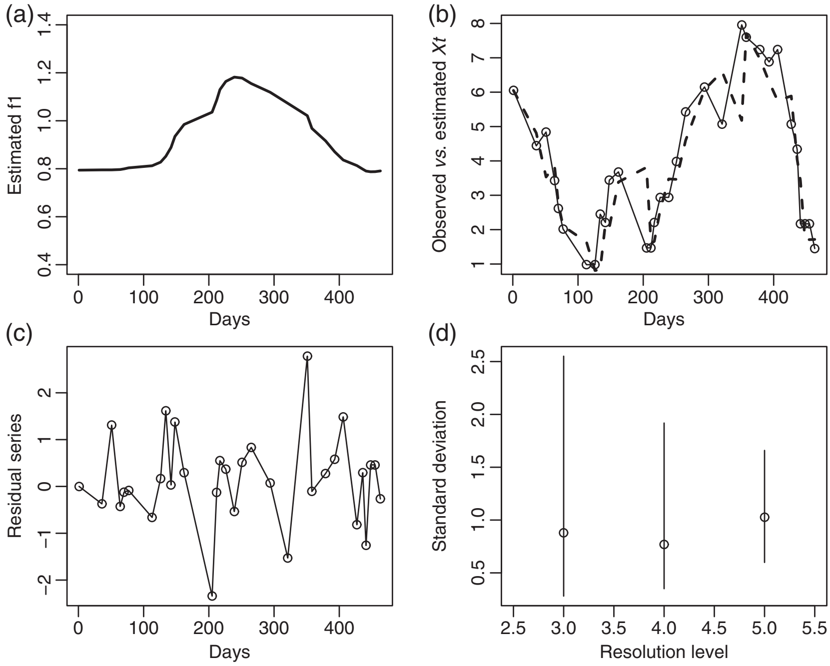

Figure 6. Results of fitting the time-varying irregular autoregressive model of order $p = 1$ to the water nitrites series in the Beagle Channel from March 2005 to December 2006. (a) Estimated AR(1) functional parameter with DB8. (b) Observed (solid lines) and fitted time series (dashed lines). (c) Irregular residual series. (d) Residual correlation check of the fitted model 189 × 181 mm.

In Figure 5(a) we show the estimated functional parameter for the model fitted to the silicates series, using the Daublet with $N = 8$ (DB8). Observed (dark lines) and fitted (dashed lines) series are shown in Figure 5(b). We have obtained an MSE $= 0.9389$ for the residual series in Figure 5(c) and p-values greater than 0.10 for different normality tests, after standardization (Kolmogorov–Smirnov, 0.4920; Anderson–Darling, 0.1560; Shapiro–Wilks, 0.5183; Jarque–Bera, 0.2753). Figure 5(d) shows that this model removes the time series dependence.

For the nitrites series, in Figure 6(a) we show the estimated functional parameter also using the Daublet DB8. Observed (dark lines) and fitted (dashed lines) series are shown in Figure 6(b). We have obtained an MSE $= 0.0088$ for the residual series in Figure 6(c). This model removes the time series dependence, see Figure 6(d), and generates normal residuals with p-values greater than 0.10 for different normality tests, after standardization (Kolmogorov–Smirnov, 0.9733; Anderson–Darling, 0.5356; Shapiro–Wilks, 0.3921; Jarque–Bera, 0.7672).

7. Conclusions

In this paper, we have proposed a tv-AR model for irregularly spaced non-stationary time series, similar to wavelet models for locally stationary processes. Irregularity and non-stationarity are included in the model through the autoregressive functional parameters. Estimation is done by least squares of the coefficients of a wavelet expansion of the functional parameters. The model performance in finite samples and its usefulness are illustrated through some simulations and an application to two time series of data collected from a shipping channel.

Acknowledgement

We would especially like to thank PhD. Marcelo Pablo Hernando, Facultad de Medicina, Universidad de Buenos Aires (Argentina) for providing the irregular time series used in Section 6.

References

[1] G. Bourdaud and W. Sickel, *Composition operators of function spaces with fractional order of smoothness*, Kokyutoku Bessatsu B26 (2011), pp. 93–132.
[2] P.M.T. Broersen and R. Bos, *Estimating time series models from irregularly spaced data*, MIEEE Trans. Instr. Meas. (2006).
[3] T.T. Cai and L.D. Brown, *Wavelet shrinkage for nonequispaced samples*, Ann. Statist. 26(5) (1998), pp. 1783–1799.
[4] Z. Cai, J. Fan, and Q. Yao, *Functional-coefficient regression models for nonlinear time series*, J. Am. Statist. Assoc. 95 (2000), pp. 941–956.
[5] R. Chen and R.S. Tsay, *Functional-coefficient autoregressive models*, J. Am. Statist. Assoc. 88 (1993), pp. 298–308.
[6] C. Chiann and P.A. Morettin, *Time domain nonlinear estimation of time varying linear systems*, J. Nonparametric Statist. 17 (2005), pp. 365–383.
[7] T. Cipra, *Some problems of exponential smoothing*, Appl. Math. 34 (1989), pp. 161–169.
[8] T. Cipra, J. Trujillo, and A. Rubio, *Holt–Winter's method with missing observations*, Manag. Sci. 41 (1995), pp. 171–178.
[9] T. Cipra, *Exponential smoothing for irregular data*, Appl. Math. 6 (2006), pp. 597–604.
[10] A. Cohen, I. Daubechies, and P. Vial, *Wavelets on the interval and fast wavelet transforms*, Appl. Comput. Harmon. Anal. 1 (1993), pp. 54–81.
[11] R. Dahlhaus, M.H. Neumann, and R. von Sachs, *Nonlinear wavelet estimation of time-varying autoregressive processes*, Bernoulli 5(5) (1999), pp. 873–906.
[12] I. Daubechies, *Ten Lectures on Wavelets*, CBMS 61, Regional Conference Series in Applied Mathematics, SIAM, Philadelphia, 1992.
[13] D. Donoho and I. Johnstone, *Minimax estimation via wavelet shrinkage*, Ann. Statist. 26 (1998), pp. 879–921.
[14] I.M. Johnstone and B.W. Silverman, *Wavelet threshold estimators for data with correlated noise*, J. Roy. Statist. Soc. Ser. B, Statist. Methodol. 59 (1997), pp. 319–351.
[15] R.H. Jones, *Maximum likelihood fitting of ARMA models to time series with missing observations*, Technometrics 22 (1980), pp. 389–395.
[16] R.H. Jones, *Fitting multivariate models to unequally spaced data*, in: *Time Series Analysis of Irregularly Observed Data, Proceedings*, E. Parzen, ed., Lectures Notes in Statistics, Springer-Verlag, Berlin, 1984, pp. 158–178.
[17] G. Kitagawa, *State space modeling of non-stationary time series and smoothing of unequally spaced data*, in: *Time Series Analysis of Irregularly Observed Data, Proceedings*, E. Parzen, ed., Lectures Notes in Statistics, Springer-Verlag, Berlin, 1984, pp. 189–210.
[18] P.A. Morettin, *Waves and Wavelets*, Editora da Universidade de São Paulo, São Paulo, 1998.
[19] G. Nason, *wavethresh: Wavelets statistics and transforms*, R package version 4.5. Available at http://CRAN.R-project.org/package=wavethresh, 2010.
[20] E. Polus, N. Flipo, C. Fouquet, and M. Poulin, *Geostatistics for assessing the efficiency of a distributed physically-based water quality model: Application to nitrate in the Seine River*, Hydrol. Process. 25 (2011), pp. 217–233.
[21] J.R. Sato, P.A. Morettin, P.R. Arantes, and J.R. Amaro, *Wavelet based time-varying vector autoregressive models*, Comput. Statist. Data Anal. 51 (2006), pp. 5847–5866.
[22] R.H. Shumway, *Some applications of the EM algorithm to analyzing incomplete time series data*, in: *Time Series Analysis of Irregularly Observed Data, Proceedings*, E. Parzen, ed., Lectures Notes in Statistics, Springer-Verlag, Berlin, 1984, pp. 290–324.
[23] R.H. Shumway and D.S. Stoffer, *Time Series Analysis and its Applications with R Examples*, 2nd ed., Springer Verlag, NY, 2006.
[24] D.J. Wright, *Forecasting data published at irregular time intervals using an extension of Holt's method*, Manag. Sci. 32 (1986), pp. 499–510.