

A 3-D Spatiotemporal Model for Remote Sensing Data Cubes

Débora M. Bayer^{ID}, Fábio M. Bayer^{ID}, and Paolo Gamba^{ID}, *Fellow, IEEE*

Abstract—Satellite images from the same scene observed over time can be composed in an image stack, which could be modeled as a 3-D cube. To handle this type of remote sensing data, on the one side, unidimensional dynamical models have been considered, modeling each pixel separately along the time (pixel-based approach), and exploring the temporal correlation. On the other side, 2-D approaches have been considered to process each image at one date, exploring the spatial correlation. In this article, we propose a new 3-D autoregressive (AR) (3-D-AR) model useful for multitemporal image interpretation exploring the correlation in three dimensions altogether. The 3-D-AR model is statistically defined, and a robust parameter estimation method is discussed. The tools for filtering, forecasting, and detecting anomalies are also introduced. A Monte Carlo simulation study is performed to evaluate the finite signal length performance of the robust estimation and its sensitivity to outliers. The proposed model is applied to a multitemporal normalized difference vegetation index (NDVI) image stack for filtering, prediction, and anomaly detection purposes. The numerical results show the importance of the proposed 3-D-AR model for spatiotemporal remote sensing data interpretation.

Index Terms—3-D model, anomaly detection, filtering, remote sensing, robust estimation, spatiotemporal data.

I. INTRODUCTION

TIME series of satellite images are important sources for satellite data interpretation and Earth monitoring [1]–[3]. This type of data has been increasingly available with high temporal and spatial resolutions [1], offering a wide range of Earth observation (EO) applications. In this sense, the search for robust data sequence interpretation methods has increased in recent technical literature [1], [4], [5], considering several different approaches. These works address problems related to land-use classification [6], [7], change detection [8], [9], filtering [10], [11], and missing information reconstruction [12]–[14], among others.

Manuscript received March 12, 2020; revised May 4, 2020; accepted May 20, 2020. Date of publication June 9, 2020; date of current version January 21, 2021. This work was supported in part by the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Brazil. (*Corresponding author: Débora M. Bayer*)

Débora M. Bayer is with the Department of Sanitary and Environmental Engineering, Federal University of Santa Maria, Santa Maria 97105-900, Brazil (e-mail: debora.bayer@uol.com.br).

Fábio M. Bayer is with the Department of Statistics and LACESM, Federal University of Santa Maria, Santa Maria 97105-900, Brazil (e-mail: bayer@uol.com.br).

Paolo Gamba is with the Department of Electrical, Computer and Biomedical Engineering, The University of Pavia, 27100 Pavia, Italy (e-mail: paolo.gamba@unipv.it).

Color versions of one or more of the figures in this article are available online at <https://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TGRS.2020.2998295

Spatiotemporal satellite data can be mathematically described as a 3-D array or a 3-D cube. Two dimensions are related to spatial information, and the third dimension is time. The technical literature is populated by techniques to handle this type of data. In general, these works consider: 1) unidimensional (1-D) pixel-based approaches [8], [15]–[18], ignoring the spatial correlation, or 2) time-by-time spatial techniques [19]–[21], ignoring the temporal correlation. The most common techniques for modeling 1-D signals are autoregressive (AR) models, such as those in the class of AR moving average (ARMA) models [22], [23]. For 2-D signals, such as images, some spatial ARMA models are explored in the literature [24]–[28]. These models and their statistical inference techniques, including robust estimation [24], [25], [29], are explored in different image processing applications. However, to the best of authors' knowledge, a 3-D AR statistical model, as a generalization of the above-mentioned 1-D and 2-D models to the third dimension, is absent in the literature. Thus, we attempt to fulfill this gap by proposing a parametric 3-D statistical model for multitemporal satellite image interpretation that considers the spatial and time correlations in the same AR framework. We introduce the proposed 3-D-AR dynamical model and discuss parameter estimation. The results show that the proposed model can be used for filtering, gap filling, anomaly detection, and future prediction in a multitude of sensors and EO applications.

As remote sensing data usually contain anomalous values (outliers) [30], mostly due to clouds and atmospheric effects in multispectral data, as well as land contamination or sensor failures, this work introduces a robust method for the estimation of the proposed 3-D-AR model parameters. Specifically, in order to obtain the parameter estimates less influenced by outliers, the weighted least square estimators (WLSEs) [29] are considered. Besides, this work introduces filtering and prediction methods, as well as residuals and anomaly detection techniques. In particular, the anomaly detection method is based on control chart theory, which is a tool used in statistical process control [31]. The residual-based control charts can be used to detect the anomalous areas, which in turn may be due to multiple phenomena, such as deforestation, droughts, fires, clouds, plant phenological cycles, or even sensor failures.

Among these and other possible EO applications of multitemporal images, the applications for normalized difference vegetation index (NDVI) series, such as gap-filling techniques [12], [13], are receiving greater and greater attention in the literature [3], [14], [16], [17], [32]. Other studies based

on NDVI data cubes are about the influence of drought on vegetation or drought detection [33]–[35], land use classification [36], vegetation trend estimation [37], changes in vegetation phenology [38], biomass characterization [39], and relationship between vegetation and soil [40]. This is the main driver why, in this work, a NDVI time series of 33 observations of the Caatinga biome, the so-called “Brazilian savannah,” is considered. In the next section, it will be shown that, by fitting this data set with the proposed 3-D-AR model, it is possible to filter and refine the signal, predict future values, and detect atypical observations. The obtained quantitative results show the great potential of the proposed approach for spatiotemporal data modeling and anomaly detection.

This article is, therefore, organized as follows. Section II introduces the 3-D-AR model and tools for robust parameter estimation, filtering, prediction, and anomaly detection. In Section III, a Monte Carlo simulation study is considered for numerical evaluating of the inference theory. Section IV presents some experimental results considering the above-mentioned Caatinga data cube, whereas conclusions and final considerations are presented in Section V.

II. PROPOSED 3-D MODEL

A multitemporal geometrically and radiometrically corrected remote sensing image stack can be defined as a 3-D cube \mathbf{Y} of size $M \times N \times T$. Let the voxel $Y[m, n, t] \in \mathbf{Y}$ be a random variable, with $(m, n, t) \in \mathbb{Z}^3$, where $m = 1, \dots, M$ and $n = 1, \dots, N$ are the spatial dimensions, and $t = 1, \dots, T$ is the time dimension. Each voxel of the 3-D cube can be written as

$$Y[m, n, t] = \mu[m, n, t] + \varepsilon[m, n, t] \quad (1)$$

where $\mu[m, n, t]$ is the mean of $Y[m, n, t]$, and $\varepsilon[m, n, t]$ are independent random variables with $E(\varepsilon[m, n, t]) = 0$ and $\text{Var}(\varepsilon[m, n, t]) = \sigma^2$.

In order to represent the 3-D cube using a statistically treatable model, we model the mean of each voxel as a function of some parameters, with the following 3-D dynamical general structure

$$\mu[m, n, t] = \mathbf{x}[m, n, t]^\top \boldsymbol{\beta} + \sum_{i, j, k} \phi_{(i, j, k)} y[m - i, n - j, t - k] \quad (2)$$

where $y[m, n, t]$ is an observed value of the random variable $Y[m, n, t]$ at $[m, n, t]$, $\mathbf{x}[m, n, t] = (x_1[m, n, t], \dots, x_r[m, n, t])^\top$ is a r -dimensional vector of covariates (non-random input variables), $\boldsymbol{\beta} = (\beta_1, \dots, \beta_r)^\top$ is the r -dimensional vector of unknown parameters related to covariates, $\phi_{(i, j, k)}$ are the unknown AR parameters of the model, and $(m - i, n - j, t - k)$ belongs to some determined support or neighborhood of the model $\mathcal{N}_{\{m, n, t\}}$. The covariates can be used to model inhomogeneous images, with different types of land use or seasonality patterns, for example. When the intercept is included, we have that $x_1[m, n, t] = 1 \forall (m, n, t) \in \mathbb{Z}^3$.

There are several possibilities to consider for the neighborhood $\mathcal{N}_{\{m, n, t\}}$, and they have been investigated in technical

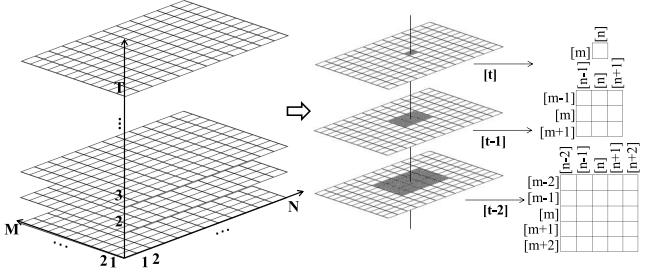


Fig. 1. 3-D-AR(2) model scheme. The voxels at instant t are written as linear combination of the hatched voxels at instants $t - 1$ and $t - 2$.

literature, especially for the 2-D case [24], [27], [41], [42]. Depending on the neighborhood, the model in (2) can be seen as a 3-D convolution operation [43], in a similar way as in the 3-D convolutional neural networks [44]. Motivated by the physical acquisition of the time series satellite images and for parsimonious reason, we define the neighbor structure as the set of past values of $t \in \mathbb{Z}$, i.e., the set of images observed until the current time stamp t . Mathematically, let us define this neighborhood as $\mathcal{N}_{\{m, n, t\}} = \{(i, j, k) \in \mathbb{Z}^3 : k < t\}$. Thus, the results into 3-D-AR structure of order p , called 3-D-AR(p), given by

$$\mu[m, n, t] = \mathbf{x}[m, n, t]^\top \boldsymbol{\beta} + \sum_{k=1}^p \sum_{j=1}^{2k+1} \sum_{i=1}^{2k+1} \phi_{(i, j, k)} y[m - (k + 1) + i, n - (k + 1) + j, t - k] \quad (3)$$

where $p \leq (a - 1)/2$, with $a = \min(M, N)$, and

$$\boldsymbol{\phi} = (\phi_{(1, 1, 1)}, \dots, \phi_{(3, 1, 1)}, \phi_{(1, 2, 1)}, \dots, \phi_{(3, 3, 1)}, \dots, \phi_{(1, 1, k)}, \dots, \phi_{(2k+1, 2k+1, k)}, \dots, \phi_{(1, 1, p)}, \dots, \phi_{(2p+1, 2p+1, p)})^\top$$

is the g -dimensional vector of AR coefficients, with $g = \sum_{k=1}^p (2k + 1)^2$. The considered neighborhood presents a cone shape along with the past values of t considering the “memory” of the signal. The 3-D-AR(2) model (without covariates) scheme is shown in Fig. 1.

A. Robust Estimation

Given a 3-D data cube of satellite observations, the problem is to estimate the vector of parameters $\boldsymbol{\gamma} = (\boldsymbol{\beta}^\top, \boldsymbol{\phi}^\top)^\top$. The traditional LSE could be an option, but it is highly sensitive to outliers [29]. The outliers can be defined as anomalous values with respect to the surrounding pixels (2-D case) [30] or voxels (3-D case). Remote sensing image stack often contains the outliers [30], [45]. In these cases, the robust methods [29], such as WLSE, can be a good choice to obtain inferences less sensitive to outliers.

The WLSE for the proposed model parameters $\boldsymbol{\gamma}$ is obtained on minimizing the weighted sum of squares of the differences between the observed data $y[m, n, t]$ and $\mu[m, n, t]$, i.e., the WLSE is given by

$$\hat{\boldsymbol{\gamma}} = \arg \min_{\boldsymbol{\gamma}} S(\boldsymbol{\gamma})$$

where Θ is the parametric space, and

$$S(\boldsymbol{\gamma}) = \sum_{m=1}^M \sum_{n=1}^N \sum_{t=1}^T w[m, n, t] (y[m, n, t] - \mu[m, n, t])^2 \quad (4)$$

is the weighted residual sum of squares. The weights $w[m, n, t]$ are supposed to be known, usually set to one when $y[m, n, t]$ is not an anomalous value and less than one when there is some evidence that the observation is an outlier. In Section II-A1, few options for the weight value determination are discussed.

By differentiating $S(\boldsymbol{\gamma})$ in (4) with respect to the q th element of the parameter vector $\boldsymbol{\beta}$, for $q = 1, \dots, r$, the following is obtained

$$\frac{\partial S(\boldsymbol{\gamma})}{\partial \beta_q} = -2 \sum_{m=1}^M \sum_{n=1}^N \sum_{t=1}^T (y[m, n, t] - \mu[m, n, t]) \times x_q[m, n, t] w[m, n, t].$$

The derivative with respect to the (i, j, k) th element of $\boldsymbol{\phi}$ is given by

$$\begin{aligned} & \frac{\partial S(\boldsymbol{\gamma})}{\partial \phi_{(i,j,k)}} \\ &= -2 \sum_{m=1}^M \sum_{n=1}^N \sum_{t=1}^T (y[m, n, t] - \mu[m, n, t]) \\ & \quad \times y[m - (k+1) + i, n - (k+1) + i, t - k] w[m, n, t]. \end{aligned}$$

The WLSE $\hat{\boldsymbol{\gamma}}$ of the 3-D-AR(p) model parameters $\boldsymbol{\gamma}$ is thus obtained by equating the derivatives to zero and solving the system. Using similar results from [29], this solution presents the following matrix closed form:

$$\hat{\boldsymbol{\gamma}} = (\mathbf{Z}^\top \mathbf{W} \mathbf{Z})^{-1} \mathbf{Z}^\top \mathbf{W} \mathbf{y}$$

where $\mathbf{y} = (y[1+p, 1+p, 1+p], \dots, y[M-p, 1+p, 1+p], y[1+p, 2+p, 1+p], \dots, y[M-p, 2+p, 1+p], \dots, y[1+p, N-p, 1+p], \dots, y[M-p, N-p, 1+p], y[1+p, 1+p, 2+p], \dots, y[M-p, N-p, T])$, $\mathbf{W} = \text{diag}(w[1+p, 1+p, 1+p], \dots, w[M-p, 1+p, 1+p], w[1+p, 2+p, 1+p], \dots, w[M-p, 2+p, 1+p], \dots, w[M-p, N-p, 1+p], w[1+p, 1+p, 2+p], \dots, w[M-p, N-p, T])$ and \mathbf{Z} is given in (5), as shown at the bottom of this page.

For the variance, the following robust estimate is instead considered:

$$\hat{\sigma}^2 = \frac{1}{J} \sum_{m=1}^M \sum_{n=1}^N \sum_{t=1}^T w[m, n, t] (y[m, n, t] - \mu[m, n, t])^2$$

$$\text{where } J = \left(\sum_{m=1}^M \sum_{n=1}^N \sum_{t=1}^T w[m, n, t] \right) - (r + g).$$

The estimation procedure could be computationally cumbersome for big data cubes, being restrictive for applications in platforms with limited performance. However, a smaller data cube can be used only for parameter estimation. The simulation results (Section III) suggest that the cubes with a dimension about $20 \times 20 \times 30$ present good estimates. Even if the parameter estimates are extracted from a smaller cube, it is then possible to apply the filtering, prediction, and anomaly detection techniques discussed in the following to the whole original 3-D cube.

1) Weight Determination: For the definition of the model and the parameter estimation, no statistical distribution for the noise term ε is assumed. However, for the weight value determination, it is safe to assume that it follows a Gaussian distribution. The Gaussianity, or normality, is a usual distributional assumption for additive noise, which provides a reasonable model for the signals of interest in many situations [29], [46]. Note that if the additive white noise is $\varepsilon[m, n, t] \sim N(0, \sigma^2)$, then $Y[m, n, t] \sim N(\mu[m, n, t], \sigma^2)$.

Based on the Gaussian supposition, an usual form to determinate the weights for weighted estimation methods is given by [47]

$$w[m, n, t] = \begin{cases} F(m, n, t)/\delta, & \text{if } F(m, n, t) < \delta \\ 1, & \text{if } \delta \leq F(m, n, t) \leq 1 - \delta \\ (1 - F(m, n, t))/\delta, & \text{if } F(m, n, t) > 1 - \delta \end{cases} \quad (6)$$

where $F(m, n, t)$ is the cumulative distribution function of the standard normal distribution, and $\delta \in (0, 1)$ is used to delimit the distribution interval for weight in $(1 - 2\delta)\%$. Usual values for δ are 0.01 and 0.001 [47]. When the observed value $y[m, n, t]$ is atypical, implying large or small values of

$$\mathbf{Z} = \begin{bmatrix} x_1[1+p, 1+p, 1+p] & \cdots & x_r[1+p, 1+p, 1+p] & y[1+p, 1+p, p] & y[1+p, 1+p, p-1] & \cdots & y[1+p, 1+p, 1] \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_1[M-p, 1+p, 1+p] & \cdots & x_r[M-p, 1+p, 1+p] & y[M-p, 1+p, p] & y[M-p, 1+p, p-1] & \cdots & y[M-p, 1+p, 1] \\ x_1[1+p, 2+p, 1+p] & \cdots & x_r[1+p, 2+p, 1+p] & y[1+p, 2+p, p] & y[1+p, 2+p, p-1] & \cdots & y[1+p, 2+p, 1] \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_1[M-p, 2+p, 1+p] & \cdots & x_r[M-p, 2+p, 1+p] & y[M-p, 2+p, p] & y[M-p, 2+p, p-1] & \cdots & y[M-p, 2+p, 1] \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_1[1+p, N-p, 1+p] & \cdots & x_r[1+p, N-p, 1+p] & y[1+p, N-p, p] & y[1+p, N-p, p-1] & \cdots & y[1+p, N-p, 1] \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_1[M-p, N-p, 1+p] & \cdots & x_r[M-p, N-p, 1+p] & y[M-p, N-p, p] & y[M-p, N-p, p-1] & \cdots & y[M-p, N-p, 1] \\ x_1[1+p, 1+p, 2+p] & \cdots & x_r[1+p, 1+p, 2+p] & y[1+p, 1+p, 1+p] & y[1+p, 1+p, p] & \cdots & y[1+p, 1+p, p-1] \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_1[M-p, N-p, T] & \cdots & x_r[M-p, N-p, T] & y[M-p, N-p, T-1] & y[M-p, N-p, T-2] & \cdots & y[M-p, N-p, T-p] \end{bmatrix} \quad (5)$$

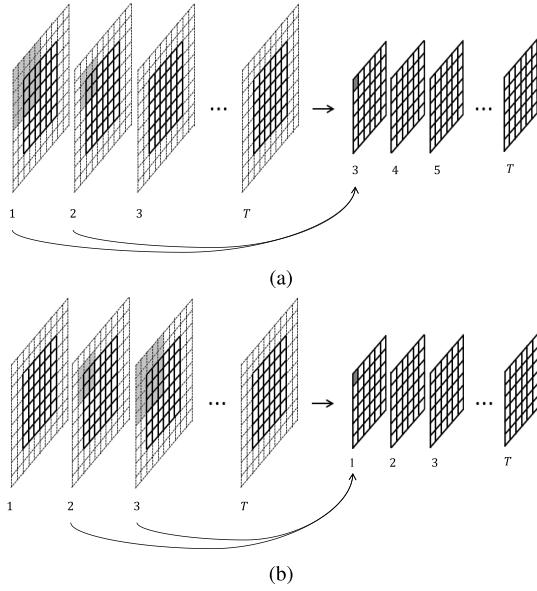


Fig. 2. 3-D-AR(2) model convolution scheme. The observed 3-D cube is represented to the left and the filtered image to the right. (a) Half padding spatial convolution. (b) Back-calculation in time.

$F(m, n, t)$, it is weighted, and the inference distortions due to outliers are minimized. In practice, the unknown parameter in (6) is replaced by its LSE considering $w[m, n, t] = 1 \forall m, n, t$.

B. Filtering and Prediction

After parameter estimation, a filtered 3-D cube can be obtained. The filtered signal is given by the estimated values of $\hat{\mu}[m, n, t]$, given by the 3-D dynamical structure in (3) evaluated at the WLSE \hat{y} . Accordingly

$$\begin{aligned} \hat{\mu}[m, n, t] &= \mathbf{x}[m, n, t]^T \hat{\beta} + \sum_{k=1}^p \sum_{j=1}^{2k+1} \sum_{i=1}^{2k+1} \hat{\phi}_{(i,j,k)} \\ &\quad \times \{y[m - (k+1) + i, n - (k+1) + j, t - k]\}_f \end{aligned} \quad (7)$$

where

$$\{y[m, n, t]\}_f = \begin{cases} \hat{\mu}[m, n, t], & \text{if } F(m, n, t) < \delta \\ y[m, n, t], & \text{if } \delta \leq F(m, n, t) \leq 1 - \delta \\ \hat{\mu}[m, n, t], & \text{if } F(m, n, t) > 1 - \delta. \end{cases} \quad (8)$$

Note that in (8), when $y[\cdot, \cdot, t]$ is an outlier, it is not considered for the estimate of $\hat{\mu}[\cdot, \cdot, t+k]$, with $k = 1, \dots, p$. In these cases, instead of using the anomalous values $y[m, n, t]$, the estimated values $\hat{\mu}[m, n, t]$ are used.

The filtered scheme in (7) presents some problems at the border of the 3-D cube. To avoid these issues, a half padding procedure [48] in the spatial dimensions is used [Fig. 2(a)], as well as back-calculation values for the first p times [Fig. 2(b)]. In other words, the images at times $t+p, \dots, t+1$ are used to predict the mean values $\hat{\mu}[\cdot, \cdot, t]$, with $t = 1, \dots, p$. In the scheme shown in Fig. 2, considering $p = 2$, we need p additional lines and columns in each special slice. To this

aim, the border adjacent lines and columns are repeated to the outside of the image.

As usual in dynamical models, the fit model can be used to forecast the future values. With the proposed 3-D-AR model, it is possible to predict the future images at instants $t = T+1, \dots, T+H$, where H is the forecast horizon. Accordingly, the mean response estimate at $T+h$, with $h = 1, \dots, H$, can be obtained as

$$\begin{aligned} \hat{\mu}[m, n, T+h] &= \mathbf{x}[m, n, T+h]^T \hat{\beta} + \sum_{k=1}^p \sum_{j=1}^{2k+1} \sum_{i=1}^{2k+1} \hat{\phi}_{(i,j,k)} \\ &\quad \times \{y[m - (k+1) + i, n - (k+1) + j, T+h-k]\}_p \end{aligned} \quad (9)$$

where

$$\{y[\cdot, \cdot, T+h-k]\}_p = \begin{cases} \hat{\mu}[\cdot, \cdot, T+h-k], & \text{if } k < h \\ \{y[\cdot, \cdot, T+h-k]\}_f, & \text{if } k \geq h. \end{cases}$$

Even for predicting purposes, the border values are managed in the same way as discussed above for the filtering procedure.

C. Residual Analysis and Anomaly Detection

Residuals measure the deviation between observed voxels $y[m, n, t]$ and corresponding estimated mean values $\hat{\mu}[m, n, t]$. A residual analysis is used as the assessment criteria for the goodness-of-fit of the fit 3-D-AR(p) model. For this model, we consider the standardized residual given by

$$r[m, n, t] = \frac{y[m, n, t] - \hat{\mu}[m, n, t]}{\hat{\sigma}}. \quad (10)$$

If the model is correctly specified, the standardized residuals are approximately Gaussian distributed with zero mean and unit variance, i.e., $r[m, n, t] \sim N(0, 1)$. Thus, a visual inspection of the residual histogram can be used to check the accurate fit. We expect the symmetric distribution and most residuals around zero in the interval $(-3, 3)$.

The formulation of the 3-D-AR(p) model enables the anomaly detection by considering residuals' monitoring based on the classical theory of control charts [31]. The control charts were originally developed to monitor industrial processes; however, their use has increased in different fields, including signal processing and remote sensing [49]–[51]. Specifically, in the framework of this research, the 3-D-AR residual-based control charts may be used to detect the anomalous areas in 3-D data cubes.

For anomaly detection, the residuals are subjected to a Shewhart control chart [52]. As the approximated value of residuals variance is one, this control chart detects if $r[m, n, t]$ is outside the control limits $\pm L$, where L is the distance between the mean of the residuals (i.e., zero) and the control limits, expressed in standard deviation units. An usual value for L is three [31], [49], [51] so that the residuals remain within $(-3, 3)$ about 99.7% of times. If the residual value is beyond this range, we assume that some anomalous change occurred due to some exogenous factor. Thus, for each residual voxel $r[m, n, t]$, it is verified if $-3 < r[m, n, t] < 3$, and

a binary cube is created with value 1 assigned to the residual voxel outside the interval (anomaly detected) and 0 to the rest of the image (no change).

After the control chart application, some bidimensional morphological operations can be carried out in each temporal “slice” of the 3-D binary cube. The dilation–erosion operations can be used to remove detections considered too small, such as isolated voxels, and to merge the ones that are part of the same target. The binary cube obtained after the latter procedure is called “detected signal.”

Thus, we propose the following algorithm to implement anomaly detection.

- 1) Choose the 3-D-AR model of order p . We suggest starting with the smallest order ($p = 1$) and then increasing p and checking the goodness-of-fit of each model.
- 2) Consider some exogenous variable or a deterministic pattern that can be included as covariates $\mathbf{x}[m, n, t]$ in (3).
- 3) If the cube size is big, choose a smaller dimension to be used for parameter estimation only. If the data size is small, then all the data cube can be used for estimation.
- 4) Fit the 3-D-AR(p) model and obtain the WLSE of the parameters, $(\hat{\beta}, \hat{\phi}, \hat{\sigma}^2)$, as discussed in Section II-A.
- 5) Based on the estimated parameters $(\hat{\beta}, \hat{\phi}, \hat{\sigma}^2)$, compute the filtered (in-sample) signals in (7).
- 6) By using the observed data and the filtered signals, compute the residuals in (10).
- 7) For each residual voxel $r[m, n, t]$, verify if $-3 < r[m, n, t] < 3$. Create a binary cube with value 1 if the residual voxel is outside of the interval $(-3, 3)$ and 0 otherwise.
- 8) Apply dilation–erosion morphological operations on the binary cube.

III. SIMULATION STUDY

In this section, we evaluate the performance of the point estimators of the 3-D-AR model parameters through a Monte Carlo simulation. We assess the estimation performance with and without outliers, measuring the sensitivity of the estimators in the presence of anomalous observations. All implementations and simulations were carried out using the R language [53].

The synthetic spatiotemporal images were generated as a 3-D cube following the 3-D-AR structure given by (1) and (3), where the error terms $\varepsilon[m, n, t]$ in (1) are simulated from zero mean normal distribution. A 3-D-AR(1) with one deterministic covariate was simulated with the following parameter values: $\beta_1 = 0.06$, $\phi_{1,1,1} = 0.19$, $\phi_{2,1,1} = 0.07$, $\phi_{3,1,1} = 0.21$, $\phi_{1,2,1} = 0.03$, $\phi_{2,2,1} = -0.02$, $\phi_{3,2,1} = 0.02$, $\phi_{1,3,1} = 0.15$, $\phi_{2,3,1} = 0.06$, and $\phi_{3,3,1} = 0.17$. Two dispersion parameter values are considered, namely: $\sigma = 0.24$ and $\sigma = 1$. For brevity and similarity of results, just the results for $\sigma = 1$ are shown. The parameter values are based on the fit 3-D-AR(1) model for the actual NDVI times series in Section IV. To mimic a seasonality pattern in the signal, the covariate is considered as $x_1[\cdot, \cdot, t] = \cos(2\pi t/12)$, with $t = 1, \dots, T$. The dimensions of the 3-D cubes were set to $20 \times 20 \times T$, with $T \in \{10, 20, 30\}$.

We set $\delta = 0.01$ for weights’ determination, which proved a suitable choice based on previous experiments and simulations.

In order to evaluate the LSE and its robust version WLSE, the mean, bias, percentage relative bias (RB%), and mean square error (MSE) of the estimators were computed, based on 500 replications of the 3-D signal. The RB is defined as the ratio between the bias and the true parameter value times 100, being how much, in percentage, the mean value of the estimator is far from the real parameter value. We can say that better is an estimator if smaller is RB and MSE. For this evaluation, we considered two approaches: 3-D signal generated with outliers and without outliers. Specifically, to include outliers, we added the value four in 5% of the voxels in randomized positions. This percentage of outliers is also used in [24] for robust estimation analysis in 2-D models. The results without outliers are shown in Table I and with outliers in Table II.

Table I shows the estimation theory, showing small RB for all estimators in all cube dimensions. For $T = 10$, the biggest absolute value of RB is less than 6%, and for $T = 30$, the RB is lower than 4%. The MSE of the estimators is also small, showing their efficiency (i.e., small variance). Table II shows the results in the presence of outliers. We can see that the robust estimation for σ is effective, reducing the RB about three times with respect to the non-robust LSE method. Similar results are found for β_1 . For instance, consider the case with $T = 30$, the RB of LSE for σ is equal to 42.4434%, whereas for WLSE is only 15.0134%. Considering the estimators for β_1 , we have RB = -11.6206% for LSE and less than 1% for WLSE. For the parameters $\phi_{(i,j,k)}$, the RB is more variable. For the AR parameters, the biggest RB for LSE when $T = 10$ is -176.7198%, whereas for the WLSE is -112.0949%, i.e., a reduction about 47%. By considering the mean of absolute values of RB for all parameter estimators, we have 50.23%, 44.91%, and 41.23% for the LSE, respectively, for $T = 10, 20, 30$, and 31.14%, 26.77%, and 25.63% for the WLSE for the same T values. It is important to highlight that the accurate inferences of the σ are the most important ones for the 3-D-AR model applicability. A distorted estimate of σ causes a wrong penalization in the weight determination in (6) and in the filtering procedure in (8). When the overestimated values of σ are included, outliers will be considered as regular observations. We also can note that, in general, the WLSE presents smaller MSE values than LSE.

An usual measure to evaluate the sensitivity of an estimator in the presence of outliers is the sensitivity curve (SC) [29]. In the case of this work, for a simulated signal following the 3-D-AR model, the SC of the estimator $\hat{\gamma}$ of an arbitrary parameter γ is defined by

$$\text{SC}(\hat{\gamma}) = G \cdot (\hat{\gamma}(y_1, \dots, y_{G-1}, y) - \hat{\gamma}(y_1, \dots, y_{G-1}, y^{\text{out}}))$$

where $G = M \times N \times T$ is the signal length, $\hat{\gamma}(y_1, \dots, y_{G-1}, y)$ is the estimator without outliers, and $\hat{\gamma}(y_1, \dots, y_{G-1}, y^{\text{out}})$ is the estimator with the replacement of the observation y by an outlier y^{out} . The SC displays the sensitivity of an estimator when an outlier is added to a signal. Thus, the magnitude of the outlier takes the values in a determined range of values, and the SC is plotted and analyzed. This curve provides a valuable intuitive information about the estimators’ sensitivity [29].

TABLE I
MONTE CARLO SIMULATION RESULTS FOR POINT ESTIMATION FOR THE 3-D-AR(1) MODEL WITHOUT OUTLIERS

Parameter	Value	Mean	Bias	RB (%)	MSE	Mean	Bias	RB (%)	MSE
		LSE				WLSE			
$(M, N, T) = (20, 20, 10)$									
σ	1.00	1.0008	0.0008	0.0767	0.0002	0.9633	-0.0367	-3.6695	0.0015
β_1	0.06	0.0610	0.0010	1.6397	0.0010	0.0601	0.0001	0.2259	0.0011
$\phi_{(1,1,1)}$	0.19	0.1883	-0.0017	-0.8852	0.0003	0.1873	-0.0028	-1.4456	0.0003
$\phi_{(2,1,1)}$	0.07	0.0701	0.0001	0.0943	0.0003	0.0702	0.0002	0.3248	0.0003
$\phi_{(3,1,1)}$	0.21	0.2101	0.0001	0.0238	0.0003	0.2085	-0.0015	-0.6987	0.0003
$\phi_{(1,2,1)}$	0.03	0.0286	-0.0014	-4.6092	0.0003	0.0287	-0.0013	-4.3417	0.0003
$\phi_{(2,2,1)}$	-0.02	-0.0198	0.0002	-0.8828	0.0003	-0.0193	0.0007	-3.6241	0.0003
$\phi_{(3,2,1)}$	0.02	0.0212	0.0012	5.8926	0.0003	0.0212	0.0012	5.8824	0.0003
$\phi_{(1,3,1)}$	0.15	0.1501	0.0001	0.0754	0.0003	0.1494	-0.0006	-0.3801	0.0003
$\phi_{(2,3,1)}$	0.06	0.0586	-0.0014	-2.3744	0.0003	0.0587	-0.0013	-2.1423	0.0003
$\phi_{(3,3,1)}$	0.17	0.1689	-0.0011	-0.6430	0.0003	0.1680	-0.0020	-1.1596	0.0003
$(M, N, T) = (20, 20, 20)$									
σ	1.00	1.0004	0.0004	0.0414	0.0001	0.9624	-0.0376	-3.7580	0.0015
β_1	0.06	0.0590	-0.0010	-1.7180	0.0004	0.0587	-0.0013	-2.1234	0.0004
$\phi_{(1,1,1)}$	0.19	0.1895	-0.0005	-0.2864	0.0001	0.1881	-0.0019	-0.9977	0.0001
$\phi_{(2,1,1)}$	0.07	0.0697	-0.0003	-0.4655	0.0001	0.0697	-0.0003	-0.4602	0.0001
$\phi_{(3,1,1)}$	0.21	0.2102	0.0002	0.1133	0.0001	0.2087	-0.0013	-0.6267	0.0001
$\phi_{(1,2,1)}$	0.03	0.0304	0.0004	1.1728	0.0001	0.0307	0.0007	2.3485	0.0001
$\phi_{(2,2,1)}$	-0.02	-0.0199	0.0001	-0.5265	0.0001	-0.0193	0.0007	-3.5997	0.0001
$\phi_{(3,2,1)}$	0.02	0.0201	0.0001	0.4959	0.0001	0.0204	0.0004	1.9233	0.0001
$\phi_{(1,3,1)}$	0.15	0.1501	0.0001	0.0754	0.0001	0.1490	-0.0010	-0.6907	0.0001
$\phi_{(2,3,1)}$	0.06	0.0591	-0.0009	-1.5099	0.0001	0.0593	-0.0007	-1.2179	0.0001
$\phi_{(3,3,1)}$	0.17	0.1685	-0.0015	-0.9027	0.0001	0.1674	-0.0026	-1.5422	0.0002
$(M, N, T) = (20, 20, 30)$									
σ	1.00	1.0001	0.0001	0.0059	0.0001	0.9623	-0.0377	-3.7720	0.0015
β_1	0.06	0.0592	-0.0008	-1.3012	0.0002	0.0590	-0.0010	-1.6525	0.0002
$\phi_{(1,1,1)}$	0.19	0.1897	-0.0003	-0.1622	0.0001	0.1884	-0.0017	-0.8663	0.0001
$\phi_{(2,1,1)}$	0.07	0.0696	-0.0004	-0.5937	0.0001	0.0698	-0.0002	-0.2575	0.0001
$\phi_{(3,1,1)}$	0.21	0.2094	-0.0006	-0.3034	0.0001	0.2076	-0.0024	-1.1367	0.0001
$\phi_{(1,2,1)}$	0.03	0.0300	0.0000	-0.0240	0.0001	0.0301	0.0001	0.1733	0.0001
$\phi_{(2,2,1)}$	-0.02	-0.0202	-0.0002	0.8672	0.0001	-0.0198	0.0002	-0.9826	0.0001
$\phi_{(3,2,1)}$	0.02	0.0196	-0.0004	-2.1816	0.0001	0.0198	-0.0003	-1.2679	0.0001
$\phi_{(1,3,1)}$	0.15	0.1498	-0.0002	-0.1172	0.0001	0.1489	-0.0011	-0.7074	0.0001
$\phi_{(2,3,1)}$	0.06	0.0588	-0.0012	-1.9831	0.0001	0.0589	-0.0011	-1.8132	0.0001
$\phi_{(3,3,1)}$	0.17	0.1698	-0.0002	-0.1313	0.0001	0.1686	-0.0014	-0.8092	0.0001

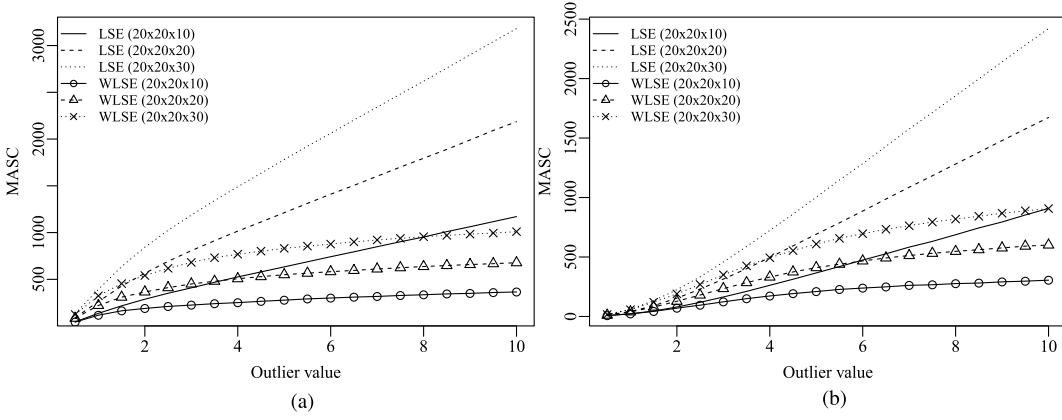


Fig. 3. Sensitivity results for two different variance scenarios. (a) $\sigma = 0.24$. (b) $\sigma = 1.00$.

As we have several parameters, two scenarios, and three cube dimensions, for better graphical analysis, we propose an unified measure called the mean absolute value of SC (MASC). The MASC is defined as the mean of the absolute values of SC of all estimators of the model. The MASC for two scenarios is shown in Fig. 3, considering 5% of outliers with their values ranging from 0.5 to 10. The visual results

show that the WLSE is less sensitive to outliers than LSE for all outlier values in all cases.

In order to evaluate the computational efficiency of the proposed filtering method, we considered the data cubes with dimension $M \times N \times T$, with $M = N = T \in \{20, 30, 40, 50, 60, 70, 80, 90, 100\}$. We run the experiment on a Desktop with 64 bits Intel i7-3770 3.40 GHz CPU core $\times 8$ and

TABLE II

MONTE CARLO SIMULATION RESULTS FOR POINT ESTIMATION FOR THE 3-D-AR(1) MODEL WITH 5% OF CONTAMINATION BY OUTLIERS

Parameter	Value	Mean	Bias	RB (%)	MSE	Mean	Bias	RB (%)	MSE
		LSE		WLSE					
$(M, N, T) = (20, 20, 10)$									
σ	1.00	1.4503	0.4503	45.0278	0.2030	1.1723	0.1723	17.2298	0.0300
β_1	0.06	0.0220	-0.0380	-63.2600	0.0031	0.0803	0.0203	33.8512	0.0022
$\phi_{(1,1,1)}$	0.19	0.1493	-0.0407	-21.4019	0.0020	0.1330	-0.0570	-30.0043	0.0035
$\phi_{(2,1,1)}$	0.07	0.0764	0.0064	9.1617	0.0004	0.0622	-0.0078	-11.1530	0.0003
$\phi_{(3,1,1)}$	0.21	0.1615	-0.0485	-23.0763	0.0027	0.1438	-0.0662	-31.5249	0.0046
$\phi_{(1,2,1)}$	0.03	0.0456	0.0156	52.1222	0.0006	0.0313	0.0013	4.1884	0.0002
$\phi_{(2,2,1)}$	-0.02	0.0153	0.0353	-176.7198	0.0016	0.0024	0.0224	-112.0949	0.0008
$\phi_{(3,2,1)}$	0.02	0.0424	0.0224	112.1523	0.0009	0.0284	0.0084	41.7725	0.0004
$\phi_{(1,3,1)}$	0.15	0.1269	-0.0231	-15.4218	0.0009	0.1119	-0.0381	-25.3856	0.0017
$\phi_{(2,3,1)}$	0.06	0.0691	0.0091	15.1483	0.0004	0.0558	-0.0042	-7.0293	0.0003
$\phi_{(3,3,1)}$	0.17	0.1377	-0.0323	-19.0187	0.0014	0.1219	-0.0481	-28.3235	0.0026
$(M, N, T) = (20, 20, 20)$									
σ	1.00	1.4307	0.4307	43.0663	0.1856	1.1553	0.1553	15.5266	0.0242
β_1	0.06	0.0386	-0.0214	-35.6977	0.0009	0.0661	0.0061	10.2427	0.0006
$\phi_{(1,1,1)}$	0.19	0.1474	-0.0426	-22.4239	0.0020	0.1300	-0.0600	-31.5658	0.0037
$\phi_{(2,1,1)}$	0.07	0.0749	0.0049	7.0190	0.0002	0.0598	-0.0102	-14.5984	0.0002
$\phi_{(3,1,1)}$	0.21	0.1610	-0.0490	-23.3115	0.0025	0.1424	-0.0676	-32.1729	0.0047
$\phi_{(1,2,1)}$	0.03	0.0443	0.0143	47.6948	0.0004	0.0289	-0.0011	-3.7793	0.0001
$\phi_{(2,2,1)}$	-0.02	0.0141	0.0341	-170.5289	0.0013	-0.0004	0.0196	-97.8746	0.0005
$\phi_{(3,2,1)}$	0.02	0.0387	0.0187	93.5068	0.0005	0.0236	0.0036	17.8860	0.0001
$\phi_{(1,3,1)}$	0.15	0.1248	-0.0252	-16.8087	0.0008	0.1079	-0.0421	-28.0717	0.0019
$\phi_{(2,3,1)}$	0.06	0.0678	0.0078	13.0552	0.0002	0.0528	-0.0072	-11.9727	0.0002
$\phi_{(3,3,1)}$	0.17	0.1344	-0.0356	-20.9169	0.0014	0.1177	-0.0523	-30.7497	0.0028
$(M, N, T) = (20, 20, 30)$									
σ	1.00	1.4244	0.4244	42.4434	0.1802	1.1501	0.1501	15.0134	0.0226
β_1	0.06	0.0530	-0.0070	-11.6206	0.0003	0.0604	0.0004	0.6480	0.0004
$\phi_{(1,1,1)}$	0.19	0.1470	-0.0430	-22.6519	0.0020	0.1289	-0.0611	-32.1579	0.0038
$\phi_{(2,1,1)}$	0.07	0.0740	0.0040	5.6693	0.0001	0.0580	-0.0120	-17.1931	0.0002
$\phi_{(3,1,1)}$	0.21	0.1582	-0.0518	-24.6558	0.0028	0.1402	-0.0698	-33.2163	0.0050
$\phi_{(1,2,1)}$	0.03	0.0430	0.0130	43.2898	0.0003	0.0276	-0.0024	-8.0704	0.0001
$\phi_{(2,2,1)}$	-0.02	0.0127	0.0327	-163.6502	0.0012	-0.0019	0.0181	-90.6746	0.0004
$\phi_{(3,2,1)}$	0.02	0.0380	0.0180	89.8006	0.0004	0.0220	0.0020	9.9429	0.0001
$\phi_{(1,3,1)}$	0.15	0.1220	-0.0280	-18.6528	0.0009	0.1054	-0.0446	-29.7237	0.0021
$\phi_{(2,3,1)}$	0.06	0.0662	0.0062	10.4065	0.0001	0.0513	-0.0087	-14.5143	0.0001
$\phi_{(3,3,1)}$	0.17	0.1349	-0.0351	-20.6508	0.0013	0.1177	-0.0523	-30.7899	0.0028

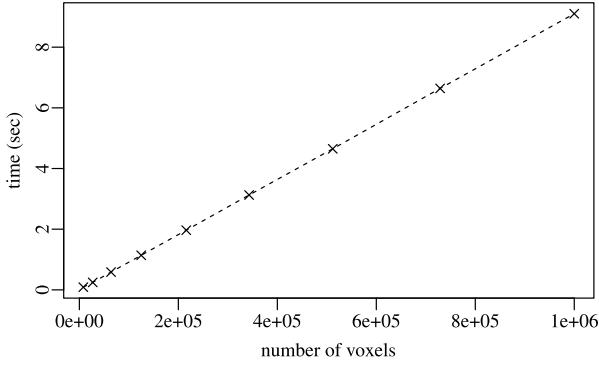


Fig. 4. Execution time for the proposed filtering method versus the number of voxels of the 3-D cube data.

16-GB RAM. Fig. 4 shows the execution times for each 3-D cube size (i.e., number of voxels). These times are average values based on 100 replications. Note that the computational cost increases linearly with the cube size, indicating that the proposed filtering method is compatible with the use of platform with limited performances.

IV. MULTITEMPORAL NDVI DATA INTERPRETATION

In order to show the applicability of the proposal for modeling, predicting, and detecting anomalies in 3-D remote

sensing data cubes, in this section, the 3-D-AR model will be applied to a spatiotemporal NDVI data set. As mentioned earlier, the multitemporal data used in this work are composed of 33 100 × 100 NDVI images observed between February 2018 and June 2019 over the Paraíba State, Brazil. Fig. 5 shows the location of the studied area. The NDVI data for this location are composed over approximately 15 days' periods with 250 m of spatial resolution, where each image can be a composition of a different number of images. We considered the periods from day 1 to day 10 and from day 16 to day 25 of each month. These NDVI images are from the National Aeronautics and Space Administration's (NASA) Earth Observing System (EOS), and they are a part of Earth Resources Observation and Science (EROS) Moderate Resolution Imaging Spectroradiometer (MODIS) collection. The MODIS NDVI V6 collection is based on the Aqua MODIS sensor. Fig. 6 shows the optic images of the scene in four different seasons of the year in the considered period, and Fig. 7(a) shows NDVI images in the same months.

Paraíba State is in the Northeast of Brazil, a region with an atypical rainfall distribution (300–2000 mm) for an equatorial region. It presents three climate types: coastal humid, tropical, and semiarid tropical [54], [55]. In the considered area, the climate dominant type is semiarid. The rainfall is also not

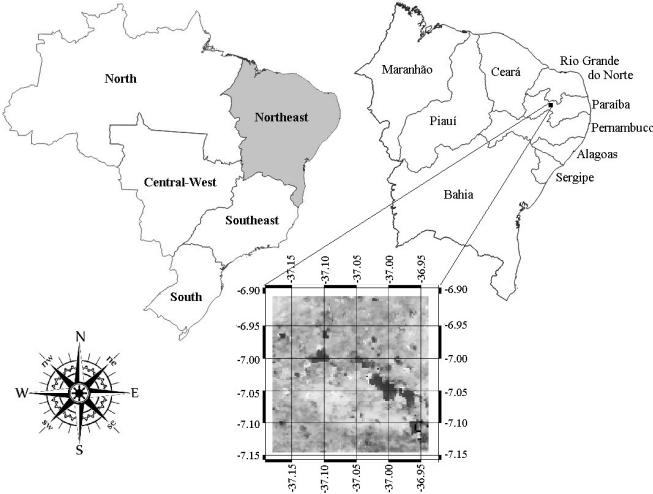


Fig. 5. Aerial view of considered area.

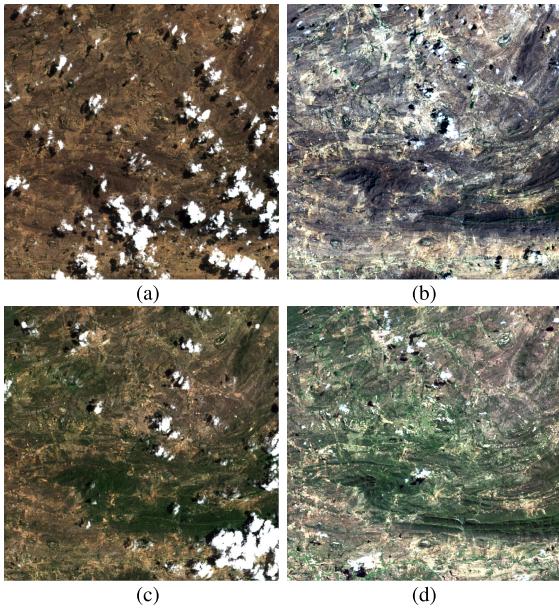


Fig. 6. Optical Landsat-8 (RGB; 432) image shows the study area in different periods of the year. (a) August 2018. (b) November 2018. (c) February 2019. (d) May 2019.

spatially and temporally homogeneous [54], implying a strong seasonality. The seasonal pattern can be seen from optical images in Fig. 6 since May is the period with more vegetation and November presents less vegetation. The phenology is reflected in the vegetation index [see Fig. 7(a)], influenced by the high correlation between NDVI and precipitation [55]. This is in accordance with the dynamic climatic mechanisms responsible for rains, which occur in the period from February to March [54].

Due to the high intrannual variability, this semiarid climate is susceptible to drought [40], [54], and vegetation must adapt to this condition [56]. The Caatinga biome is, therefore, composed of xerophytic, deciduous, semiarid thorn scrub, and forest [56], and it is unique in Brazil [56]. The Caatinga is poorly studied when compared with Cerrado and Amazon biomes [34], for example. The application of a dynamic model

capable of capturing the NDVI behavior is definitely important for different purposes, such as land cover and vegetation change detection, biomass estimation, and eco-hydrological modeling.

Following the simulation results in Section III and for parameter estimation purpose only, as discussed in Section II-A, we selected a smaller cube with $40 < m < 60$, $40 < n < 60$, and $1 < t < 33$. Besides, a square of synthetic outliers (synthetic anomalous observations) was added to the signal. We arbitrary consider this inclusion at time $t = 14$ in the pixels with $45 < m < 55$ and $45 < n < 55$. We replaced the observed values of these voxels by -0.5 . This negative value can represent, for example, a thick cloud, and is shown in Fig. 7(a) and (c). This synthetic inclusion has three main aims, namely: 1) to evaluate the performance of the robust estimation procedure in the presence of atypical values; 2) to check the filtering performances on a real data set with some exogenous intervention; and 3) to verify the performance of the anomaly detection technique.

For 3-D-AR model application, we followed the algorithm described in Section II-C. For this specific application, to account for seasonality along the determined period, we considered the covariate $x_1[\cdot, \cdot, t] = \cos(2\pi(t + 18)/24)$ that consists of 24 biweekly observations for each year. However, there are some other interesting alternatives for covariates. As discussed in [13], since the Aqua and Terra MODIS sensors have similar design patterns, we can use the Terra MODIS data as a covariate in order to obtain gap-filling results in Aqua sensor data, for example. Moreover, the topography is an important factor that affects the distribution of vegetation cover [57], [58] and can be used as a covariate. As an alternative, another good covariate could be the altitude.

The fit 3-D-AR(1) model has the following estimated parameter values: $\hat{\sigma} = 0.2442$, $\hat{\beta}_1 = 0.0570$, $\hat{\phi}_{(1,1,1)} = 0.1913$, $\hat{\phi}_{(2,1,1)} = 0.0734$, $\hat{\phi}_{(3,1,1)} = 0.2126$, $\hat{\phi}_{(1,2,1)} = 0.0295$, $\hat{\phi}_{(2,2,1)} = -0.0250$, $\hat{\phi}_{(3,2,1)} = 0.0201$, $\hat{\phi}_{(1,3,1)} = 0.1483$, $\hat{\phi}_{(2,3,1)} = 0.0642$, and $\hat{\phi}_{(3,3,1)} = 0.1710$. Fig. 8 shows that the residuals are normally distributed around zero, and most of them are in the interval $(-3, 3)$. Thus, the model seems well adjusted. Fig. 9 shows the observed and filtered images, as well as the detected signal. The filtered images in Fig. 7(b) show that the model is able to capture the main patterns of the observed signal, smoothing variability, and gap-filling the outliers in the squared region. Fig. 7(c) shows that the proposed residual-based anomaly detection technique is effective in detecting anomalies in multitemporal data. After morphological operations, only the included synthetic gap is detected.

For comparison purposes, we also fit unidimensional first-order AR (AR(1)) models pixel by pixel. This approach was suggested by [8] for prediction and change detection applications in multitemporal SAR images. We used the R package *forecast* [59] to fit the AR(1) model. Table III shows the correlation (r), mean absolute percentage error (MAPE), and root-mean-square error (RMSE) between the observed data cube and filtered values, as well as the time consumption of each approach. Regarding the quality measures, the two models presented similar results. The proposed model slightly outperforms the AR model in terms of r and MAPE,

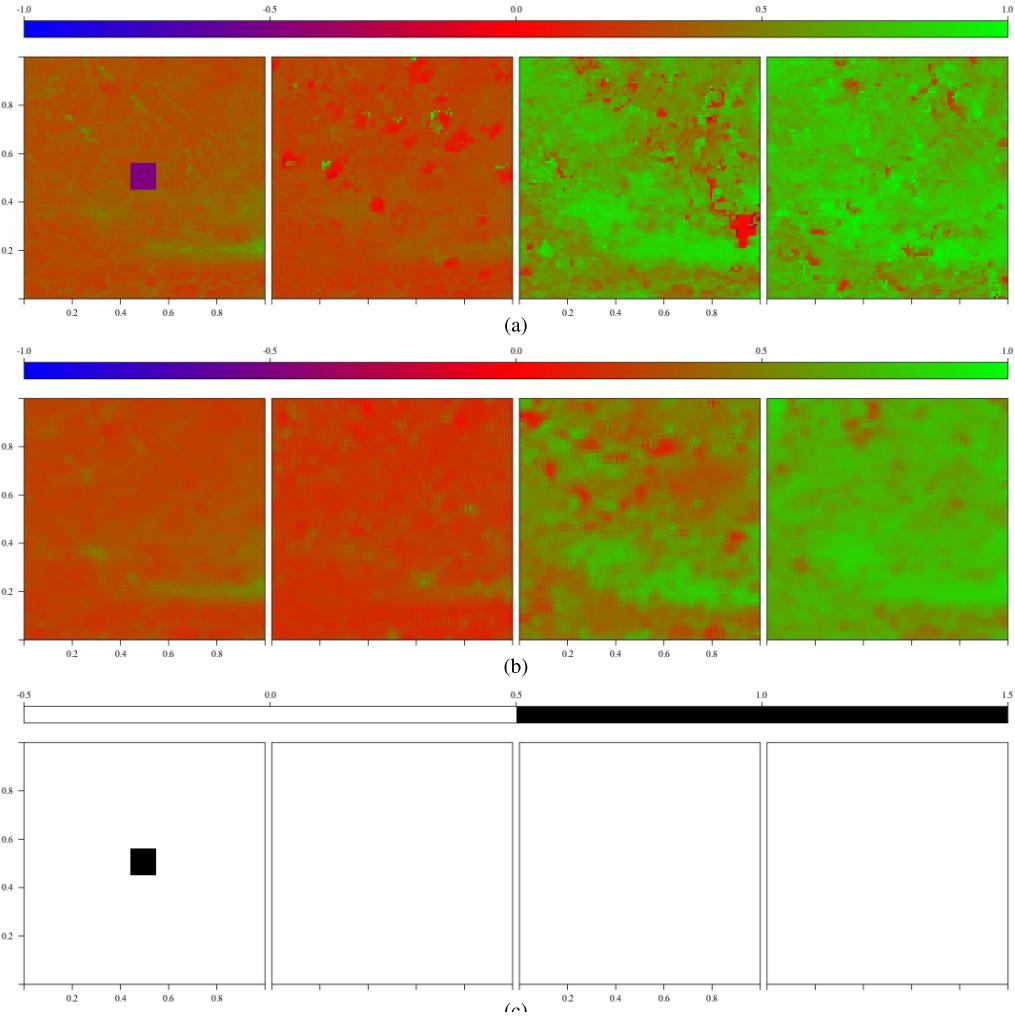


Fig. 7. Quarterly observation images. (From Left to Right) August 2018, November 2018, February 2019, and May 2019. (a) Observed images (with synthetic outliers included). (b) Filtered images. (c) Detected signal.

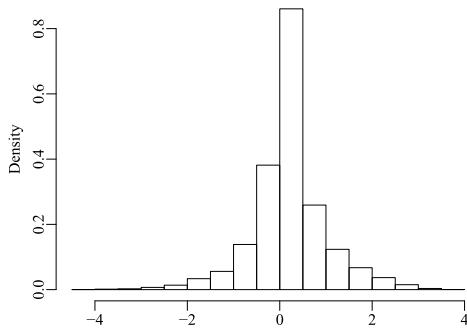


Fig. 8. Histogram of residuals.

TABLE III

COMPARISON MEASURES FOR 3-D-AR AND UNIDIMENSIONAL AR MODELS FIT TO NDVI DATA (BEST FIGURES IN BOLD)

Model	r	MAPE	RMSE	time (sec)
3D-AR (proposed)	0.57	0.54	0.20	41.23
AR (pixel by pixel)	0.56	0.60	0.18	58.36

reaching these results with just 70% of time consumption. We note that using unidimensional pixel-based approaches requires the fitting of estimation and filtering as many models

as the size of the spatial dimension ($M \times N$). Differently, in the proposed approach, we estimate parameters only once, thus being an advantage in bigger data cube applications.

A pixel-based visual analysis of a few spatial points is also presented to verify the accuracy of the filtered and predicted signals based on the proposed 3-D-AR(1) and AR(1) models. Fig. 9 shows the observed NDVI values for pixels [25, 25], [50, 50], and [75, 75] together with corresponding filtered and predicted values for both models. We note that the filtered values by the proposed 3-D model are closer to the observed ones than the 1-D approach. Fig. 9(b) shows that the included outliers do not influence next filtered values, evidencing the robustness of the 3-D-AR model. Differently, the filtered signal from 1-D model is strongly affected by atypical observations. The predicted values for six steps ahead (after vertical line) of the 3-D-AR are more accurate than those of the 1-D model. The predictions based on the proposed model better capture the seasonal behavior of the signal, useful for future decision-making.

Some land cover works are interested to the percentage of vegetation in a certain scene [60], [61]. In this sense, we analyzed the percentages of observed and filtered NDVI values over 0.3, 0.5, and 0.7 thresholds in each image time for

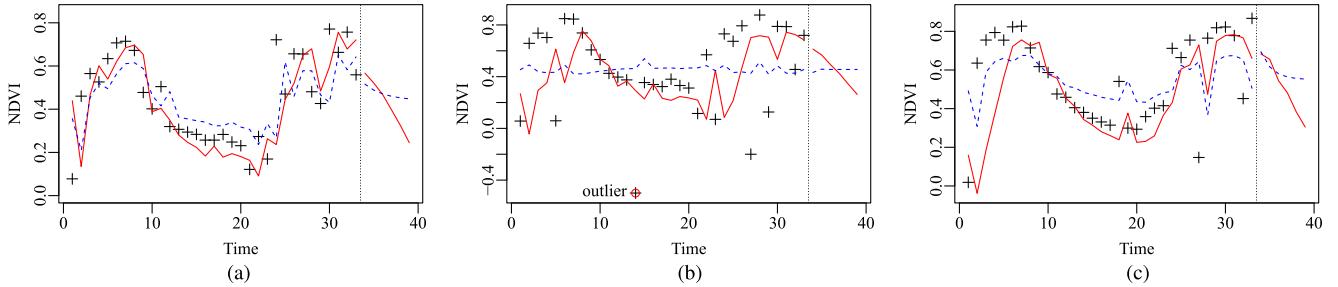


Fig. 9. Observed (points) unidimensional time series for pixels [25, 25], [50, 50], [75, 75], and their filtered values considering the 3-D-AR(1) model (continuous line) and unidimensional AR(1) model (dashed lines). The values after the vertical line are out-of-sample predicted values. (a) Pixel [25, 25]. (b) Pixel [50, 50]. (c) Pixel [75, 75].

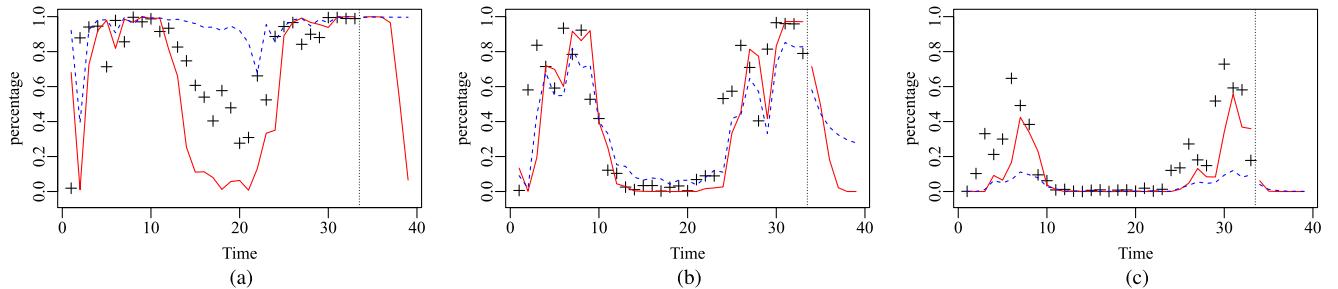


Fig. 10. Percentage of NDVI values greater than 0.3, 0.5, and 0.7 over time in the observed images (points) and filtered/predicted images considering the 3-D-AR(1) model (continuous line) and unidimensional AR(1) model (dashed lines). The values after the vertical line are based on the out-of-sample predicted values. (a) NDVI > 0.3. (b) NDVI > 0.5. (c) NDVI > 0.7.

each model. These results are shown in Fig. 10, and on the contrary to the results from the unidimensional AR model, the proposed model shows very similar percentages between the observed and filtered values. It means, for example, that the predicted percentage of land cover types can be obtained from the predicted NDVI values. These prediction results could be useful for a multitude of applications, such as the input of deterministic and stochastic models, from hydrology to epidemiology.

V. CONCLUSION

This article proposes a 3-D dynamical model for the interpretation of multitemporal remote sensing images. The model is statistically robust, and the tools for parameter estimation, filtering, prediction, and anomaly detection obtained from the models were also discussed. As outliers are often present in remote sensing images, the weighted least square estimation was considered as a robust alternative for parameter estimation. A simulation study was carried out for point estimation evaluation. The simulation results validate the inference procedures.

Since the model is useful to model the signals that can be described as a 3-D cube, an application to an NDVI data cube was performed. The NDVI test shows that the proposed 3-D-AR model is able to model the pattern of the time series, providing accurate filtered signals and predicted values. It was also capable to detect exogenous artifacts such as thick clouds or sensor failures. The numerical experiments show the flexibility and usefulness of the proposed model to model 3-D data cubes. Indeed, the proposed model is general enough to model the other types of 3-D structures, such as hyperspectral

images [62] and 3-D heightmap, where the third dimension is the wavelength in the former and the altitude in the latter.

REFERENCES

- [1] F. Petitjean, J. Inglada, and P. Gancarski, "Satellite image time series analysis under time warping," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 8, pp. 3081–3095, Aug. 2012.
- [2] L. Poggio, A. Gimona, and I. Brown, "Spatio-temporal MODIS EVI gap filling under cloud cover: An example in Scotland," *ISPRS J. Photogramm. Remote Sens.*, vol. 72, pp. 56–72, Aug. 2012.
- [3] L. Xu, B. Li, Y. Yuan, X. Gao, and T. Zhang, "A temporal-spatial iteration method to reconstruct NDVI time series datasets," *Remote Sens.*, vol. 7, no. 7, pp. 8906–8924, Jul. 2015.
- [4] E. F. Lambin and M. Linderman, "Time series of remote sensing data for land change science," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 7, pp. 1926–1928, Jul. 2006.
- [5] T. Hoberg, F. Rottensteiner, R. Queiroz Feitosa, and C. Heipke, "Conditional random fields for multitemporal and multiscale classification of optical satellite imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 2, pp. 659–673, Feb. 2015.
- [6] K.-F. Lin and D. Perissin, "Single-polarized SAR classification based on a multi-temporal image stack," *Remote Sens.*, vol. 10, no. 7, p. 1087, Jul. 2018.
- [7] A. Marinoni, G. C. Iannelli, and P. Gamba, "An information theory-based scheme for efficient classification of remote sensing data," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 10, pp. 5864–5876, Oct. 2017.
- [8] B. G. Palm *et al.*, "Autoregressive model for multi-pass SAR change detection based on image stacks," in *Proc. SPIE*, vol. 10789, Oct. 2018, Art. no. 1078916.
- [9] C. Marin, F. Bovolo, and L. Bruzzone, "Building change detection in multitemporal very high resolution SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 5, pp. 2664–2682, May 2015.
- [10] T. Trang Le, A. M. Atto, E. Trouve, and J.-M. Nicolas, "Adaptive multitemporal SAR image filtering based on the change detection matrix," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 10, pp. 1826–1830, Oct. 2014.
- [11] J. Yuan, X. Lv, and R. Li, "A speckle filtering method based on hypothesis testing for time-series SAR images," *Remote Sens.*, vol. 10, no. 9, p. 1383, Aug. 2018.

- [12] H. Shen *et al.*, "Missing information reconstruction of remote sensing data: A technical review," *IEEE Geosci. Remote Sens. Mag.*, vol. 3, no. 3, pp. 61–85, Sep. 2015.
- [13] Q. Zhang, Q. Yuan, C. Zeng, X. Li, and Y. Wei, "Missing data reconstruction in remote sensing image with a unified spatial-temporal-spectral deep convolutional neural network," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 8, pp. 4274–4288, Aug. 2018.
- [14] Y. Julien and J. A. Sobrino, "Optimizing and comparing gap-filling techniques using simulated NDVI time series from remotely sensed global data," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 76, pp. 93–111, Apr. 2019.
- [15] Y. Julien and J. A. Sobrino, "Global land surface phenology trends from GIMMS database," *Int. J. Remote Sens.*, vol. 30, no. 13, pp. 3495–3513, Jul. 2009, doi: [10.1080/01431160802562255](https://doi.org/10.1080/01431160802562255).
- [16] Y. Julien and J. A. Sobrino, "Comparison of cloud-reconstruction methods for time series of composite NDVI data," *Remote Sens. Environ.*, vol. 114, no. 3, pp. 618–625, Mar. 2010.
- [17] G. J. Roerink, M. Menenti, and W. Verhoef, "Reconstructing cloudfree NDVI composites using Fourier analysis of time series," *Int. J. Remote Sens.*, vol. 21, no. 9, pp. 1911–1917, Jan. 2000.
- [18] E. B. Brooks, V. A. Thomas, R. H. Wynne, and J. W. Coulston, "Fitting the multitemporal curve: A Fourier series approach to the missing data problem in remote sensing analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 9, pp. 3340–3353, Sep. 2012.
- [19] C. Zhang, W. Li, and D. Travis, "Gaps-fill of SLC-off landsat ETM+ satellite image using a geostatistical approach," *Int. J. Remote Sens.*, vol. 28, no. 22, pp. 5103–5122, Nov. 2007.
- [20] H. Shen and L. Zhang, "A MAP-based algorithm for destriping and inpainting of remotely sensed images," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 5, pp. 1492–1502, May 2009.
- [21] A. Bugeau, M. Bertalmio, V. Caselles, and G. Sapiro, "A comprehensive framework for image inpainting," *IEEE Trans. Image Process.*, vol. 19, no. 10, pp. 2634–2645, Oct. 2010.
- [22] G. Box, G. M. Jenkins, and G. Reinsel, *Time Series Analysis: Forecasting and Control*. Hoboken, NJ, USA: Wiley, 2008.
- [23] P. J. Brockwell and R. A. Davis, *Time Series: Theory and Methods*. New York, NY, USA: Springer-Verlag, 2013.
- [24] O. Bustos, S. Ojeda, and R. Vallejos, "Spatial ARMA models and its applications to image filtering," *Brazilian J. Probab. Statist.*, vol. 23, no. 2, pp. 141–165, Dec. 2009.
- [25] G. M. Britos and S. M. Ojeda, "Robust estimation for spatial autoregressive processes based on bounded innovation propagation representations," *Comput. Statist.*, vol. 34, no. 3, pp. 1315–1335, Sep. 2019.
- [26] O. H. Bustos, M. Ruiz, S. Ojeda, R. Vallejos, and A. C. Frery, "Asymptotic behavior of RA-estimates in autoregressive 2D processes," *J. Stat. Planning Inference*, vol. 139, no. 10, pp. 3649–3664, Oct. 2009.
- [27] B. Choi and D. N. Politis, "Modeling 2-D AR processes with various regions of support," *IEEE Trans. Signal Process.*, vol. 55, no. 5, pp. 1696–1707, May 2007.
- [28] A. Sarkar, K. M. S. Sharma, and R. V. Sonak, "A new approach for subset 2-D AR model identification for describing textures," *IEEE Trans. Image Process.*, vol. 6, no. 3, pp. 407–413, Mar. 1997.
- [29] A. M. Zoubir, V. Koivunena, E. Ollila, and M. Muma, *Robust Statistics for Signal Processing*. Cambridge, U.K.: Cambridge Univ. Press, 2018.
- [30] A. Alvera-Azcárate, D. Sirjacobs, A. Barth, and J.-M. Beckers, "Outlier detection in satellite data using spatial coherence," *Remote Sens. Environ.*, vol. 119, pp. 84–91, Apr. 2012.
- [31] D. C. Montgomery, *Introduction to Statistical Quality Control*, 6th ed. Hoboken, NJ, USA: Wiley, 2009.
- [32] A.-R. Cho and M.-S. Suh, "Detection of contaminated pixels based on the short-term continuity of NDVI and correction using spatio-temporal continuity," *Asia-Pacific J. Atmos. Sci.*, vol. 49, no. 4, pp. 511–525, Aug. 2013.
- [33] W. T. Liu and F. N. Kogan, "Monitoring regional drought using the vegetation condition index," *Int. J. Remote Sens.*, vol. 17, no. 14, pp. 2761–2782, Sep. 1996.
- [34] H. A. Barbosa, A. R. Huete, and W. E. Baethgen, "A 20-year study of NDVI variability over the Northeast region of Brazil," *J. Arid Environ.*, vol. 67, no. 2, pp. 288–307, Oct. 2006.
- [35] H. A. Barbosa, T. V. Lakshmi Kumar, F. Paredes, S. Elliott, and J. G. Ayuga, "Assessment of caatinga response to drought using meteosat-SEVIRI normalized difference vegetation index (2008–2016)," *ISPRS J. Photogramm. Remote Sens.*, vol. 148, pp. 235–252, Feb. 2019.
- [36] D. A. Mariano *et al.*, "Use of remote sensing indicators to assess effects of drought and human-induced land degradation on ecosystem health in Northeastern Brazil," *Remote Sens. Environ.*, vol. 213, pp. 129–143, Aug. 2018.
- [37] B. J.-B. Zoungrana, C. Conrad, M. Thiel, L. K. Amekudzi, and E. D. Da, "MODIS NDVI trends and fractional land cover change for improved assessments of vegetation degradation in Burkina Faso, West Africa," *J. Arid Environ.*, vol. 153, pp. 66–75, Jun. 2018.
- [38] L. Yu *et al.*, "Monitoring the long term vegetation phenology change in Northeast China from 1982 to 2015," *Sci. Rep.*, vol. 7, no. 1, pp. 1–9, Dec. 2017.
- [39] X. Zhu and D. Liu, "Improving forest aboveground biomass estimation using seasonal landsat NDVI time-series," *ISPRS J. Photogramm. Remote Sens.*, vol. 102, pp. 222–231, Apr. 2015.
- [40] H. L. F. da Silveira, L. S. Galvão, I. D. Sanches, I. B. de Sá, and T. A. Taura, "Use of MSI/Sentinel-2 and airborne LiDAR data for mapping vegetation and studying the relationships with soil attributes in the Brazilian semi-arid region," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 73, pp. 179–190, Dec. 2018.
- [41] P. Whittle, "On stationary processes in the plane," *Biometrika*, vol. 41, nos. 3–4, pp. 434–449, 1954.
- [42] M. S. Bartlett, "Physical nearest-neighbour models and non-linear time-series," *J. Appl. Probab.*, vol. 8, no. 2, pp. 222–232, 1971.
- [43] Y. Sekikawa, K. Ishikawa, and H. Saito, "Constant velocity 3D convolution," *IEEE Access*, vol. 6, pp. 76490–76501, 2018.
- [44] S. Ji, W. Xu, M. Yang, and K. Yu, "3D convolutional neural networks for human action recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 1, pp. 221–231, Jan. 2013.
- [45] H. Ghafarian Malamiri, I. Rousta, H. Olafsson, H. Zare, and H. Zhang, "Gap-filling of MODIS time series land surface temperature (LST) products using singular spectrum analysis (SSA)," *Atmosphere*, vol. 9, no. 9, p. 334, Aug. 2018.
- [46] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Upper Saddle River, NJ, USA: Prentice-Hall, 1993.
- [47] C. Field and B. Smith, "Robust estimation: A weighted maximum likelihood approach," *Int. Stat. Rev.*, vol. 62, no. 3, pp. 405–424, 1994.
- [48] V. Dumoulin and F. Visin, "A guide to convolution arithmetic for deep learning," 2016, [arXiv:1603.07285](https://arxiv.org/abs/1603.07285). [Online]. Available: <http://arxiv.org/abs/1603.07285>
- [49] E. B. Brooks, R. H. Wynne, V. A. Thomas, C. E. Blinn, and J. W. Coulston, "On-the-fly massively multitemporal change detection using statistical quality control charts and landsat data," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 6, pp. 3316–3332, Jun. 2014.
- [50] R. Jr *et al.*, "Iterative change detection algorithm for low-frequency UWB SAR," in *Proc. Anais de XXXIV Simpósio Brasileiro de Telecomunicações*, 2016, pp. 654–658.
- [51] F. M. Bayer, A. J. Kozakevicius, and R. J. Cintra, "An iterative wavelet threshold for signal denoising," *Signal Process.*, vol. 162, pp. 10–20, Sep. 2019.
- [52] W. A. Shewhart, *Economic Control of Quality of Manufactured Product*. New York, NY, USA: MacMillan, 1931.
- [53] R. C. Team, *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing, 2018. [Online]. Available: <https://www.R-project.org/>
- [54] M. Kayano and R. Andreoli, *Tempo e Clima no Brasil*. São Paulo, Brazil: Oficina de Textos, 2009, pp. 213–234.
- [55] H. A. Barbosa and T. V. Lakshmi Kumar, "Influence of rainfall variability on the vegetation dynamics over Northeastern Brazil," *J. Arid Environ.*, vol. 124, pp. 377–387, Jan. 2016.
- [56] I. R. Leal, J. M. C. Da Silva, M. Tabarelli, and T. E. Lacher, "Changing the course of biodiversity conservation in the Caatinga of Northeastern Brazil," *Conservation Biol.*, vol. 19, no. 3, pp. 701–706, Jun. 2005.
- [57] I. V. Florinsky and G. A. Kuryakova, "Influence of topography on some vegetation cover properties," *Catena*, vol. 27, no. 2, pp. 123–141, Aug. 1996.
- [58] B. Wang, G. Zhang, and J. Duan, "Relationship between topography and the distribution of understory vegetation in a *Pinus massoniana* forest in Southern China," *Int. Soil Water Conservation Res.*, vol. 3, no. 4, pp. 291–304, Dec. 2015.
- [59] R. J. Hyndman. (2017). *Forecasting Functions for Time Series and Linear Models R Package Version 8.0* ed. [Online]. Available: <https://pkg.robjhyndman.com/forecast/authors.html>
- [60] R. S. Lunetta, J. F. Knight, J. Edirwickrema, J. G. Lyon, and L. D. Worthy, "Land-cover change detection using multi-temporal MODIS NDVI data," *Remote Sens. Environ.*, vol. 105, no. 2, pp. 142–154, Nov. 2006.

- [61] K. Prabhakara, W. D. Hively, and G. W. McCarty, "Evaluating the relationship between biomass, percent groundcover and remote sensing indices across six winter cover crop fields in Maryland, United States," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 39, pp. 88–102, Jul. 2015.
- [62] Q. Yuan, Q. Zhang, J. Li, H. Shen, and L. Zhang, "Hyperspectral image denoising employing a spatial-spectral deep residual convolutional neural network," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 2, pp. 1205–1218, Feb. 2019.



Fábio M. Bayer received the B.Sc. degree in mathematics from the Federal University of Santa Maria (UFSM), Brazil, in 2006, and the D.Sc. degree in statistics from the Federal University of Pernambuco (UFPE), Recife, Brazil, in 2011.

In 2019, he was a Visiting Researcher with the Telecommunications and Remote Sensing Laboratory, The University of Pavia, Pavia, Italy. He is an Associate Professor with the Department of Statistics, UFSM, where he is also a Researcher with the Santa Maria Space Science Laboratory (LACESM).

His research interests include data science, regression and dynamic models, statistical computing, parametric inference, and statistical signal processing.



Paolo Gamba (Fellow, IEEE) received the Laurea degree in electronic engineering (*cum laude*) and the Ph.D. degree in electronic engineering from The University of Pavia, Pavia, Italy, in 1989 and in 1993, respectively.

He is a Professor with The University of Pavia, where he leads the Telecommunications and Remote Sensing Laboratory. He has been invited to give keynote lectures and tutorials in several occasions about urban remote sensing, data fusion, EO data for physical exposure and risk management. He has authored or coauthored more than 140 articles in international peer-review journals and presented nearly 300 research works in workshops and conferences.

Dr. Gamba served as an Editor-in-Chief for the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS from 2009 to 2013, and as the Chair of the Data Fusion Committee of the IEEE Geoscience and Remote Sensing Society (GRSS) from October 2005 to May 2009. He has been elected in the GRSS AdCom, since 2014, and he is the GRSS President. He has been the Organizer and the Technical Chair of the Biennial GRSS/ISPRS Joint Workshops on Remote Sensing and Data Fusion over Urban Areas from 2001 to 2015. He also served as a Technical Co-Chair for the 2010 and 2015 IGARSS Conferences, in Honolulu (Hawaii), and Milan (Italy), respectively.



Débora M. Bayer received the B.Sc. degree in civil engineering from the Federal University of Santa Maria, Brazil, in 2008, the M.Sc. degree in civil engineering from the Federal University of Pernambuco, Recife, Brazil, in 2010, and the D.Sc. degree in water resources and environmental sanitation from the Federal University of Rio Grande do Sul, Porto Alegre, Brazil, in 2014.

In 2019, she was a Visiting Researcher with the Telecommunications and Remote Sensing Laboratory, The University of Pavia, Pavia, Italy. She has experience in hydrology, water resources, hydrological simulation, statistical hydrology, impact of land use changes on water resources and environmental sanitation. She is an Associate Professor with the Department of Sanitary and Environmental Engineering, Federal University of Santa Maria.