

UNIVERSIDAD NACIONAL DE CÓRDOBA

FACULTAD DE MATEMÁTICA, ASTRONOMÍA, FÍSICA y COMPUTACION

SERIE “A”

TRABAJOS DE MATEMÁTICA

Nº 119/2017

**Robust BMM estimator in two-dimensional autoregressive
models**

Grisel Maribel Britos - Silvia María Ojeda



Editor: Jorge G. Adrover

CIUDAD UNIVERSITARIA – 5000 CÓRDOBA
REPÚBLICA ARGENTINA

Robust BMM estimator in two-dimensional autoregressive models

Grisel Maribel Britos

Silvia María Ojeda

Abstract

In this paper, we present a new estimator for two-dimensional autoregressive models of first order with three parameters. We compared the performance of our method, called BMM-2D, in cases of non-contaminated and contaminated processes with the classic Least Square estimator and three well known robust estimators. We considered different levels of replacement contamination and varied the observation window sizes. The comparative study was carried out through Monte Carlo simulations. The results show that the new estimator presents a better behavior than the other estimators, both in accuracy and precision, and has a low computational cost. An application to the representation and segmentation of real images is shown.

Keywords— AR-2D models, robust estimators, image processing

1 Introduction

Robust inference techniques appear in a diversity of contexts and applications, though the terms “robust” and “robustness” are quite freely used in the image processing and computer vision literature, not necessarily with the usual statistical meaning. The median and similar order-based filters are basic tools in image processing (Aysal and Barner [5]; Huang and Lee [38]; Palenichka et al. [57], [56]), and in some cases particular attention has been devoted to obtain the distribution of those estimators (Steland [64]). Other resistant approaches have proved to be successful in image restoration (see, for instance, Ben Hamza and Krim [8]; Chu et al. [24]; Koivunen [43]; Marroquin et al. [47]; Rabie [60]; Tarel et al. [66]; Voloshynovskiy et al. [74]; Zervakis and Kwon [81]). A common challenge in these applications is that the number of observations is reduced to a few, typically less than a hundred points. When it comes to image analysis, many robust techniques have been proposed. In this case, the sample size is usually larger than the one available in filters and, frequently, structure and topology do not impose heavy requirements or constraints. In some cases, strong hypotheses are made on the laws governing the observed process (Allende and Pizarro [4]; Brunelli and Messelodi [14]; Bustos et al. [17]; Butler [21]; Dryden et al. [28]; Van de Weijer and Van den Boomgaard [73]); other approaches can be seen in the works by Bouzouba and Radouane [12], Brandle et al. [13], Nirel et al. [51], Sim et al. [62], Tohka et al. [68], Xu [79] and Zervakis et al. [82].

High-level image analysis, or vision, also benefits from the use of robust estimation techniques, as it can be seen in Black and Rangarajan [10], Black et al. [11], Chen et al. [22], Comport et al. [25], Glendinning [30], Gottardo et al. [31], Hasler et al. [37], Kim and Han [42], Li et al. [44], Meer et al. [48], Mirza and Boyer [49], Prastawa et al. [59], Roth [61], Singh et al. [63], Stewart [65], Torr and Zisserman [69] and Wang and Suter [76], [77].

In a wide variety of different situations, such as image analysis, remote sensing and agricultural field trials, observations are obtained on a rectangular 2D lattice or grid. A class of 2D autoregressive processes has been suggested (Whittle [78]) as a source of reasonable models for the spatial correlation in such data (Tjostheim [67]). These models are natural extensions of the autoregressive processes used in time series analysis (Basu and Reinsel [7]).

Consequently, most of the robust techniques developed for parametric models in time series have been implemented for spatial parametric models when the process has been contaminated with innovation or additive outliers (Kashyap and Eom [41]). Since a single outlier can produce bias and large variance in the estimators, most of the proposals are oriented to provide estimators that are more resistant to the presence of contamination.

In the literature, there are at least three classes of robust estimators that have been studied in this context. They are the M, GM and RA estimators. Kashyap and Eom [41] introduced the M estimators for 2D autoregressive models. A recursive image restoration algorithm was implemented by using the robust M estimations to produce a restored image. Later, Allende et al. [1] studied the computational implementation of the Generalized M (GM) estimators for the same class of models. The image restoration algorithm previously developed by Kashyap and Eom [41] was generalized by Allende et al. [2]. The robust Residual Autocovariance (RA) estimators were introduced by Bustos and Yohai [20] in the context of time series, where in the recursive estimation procedure the residuals are cleaned through the application of a robustifying function. An extension of the RA estimators for spatial unilateral autoregressive models and its computational aspects were studied by Ojeda et al. [52]. Monte Carlo simulation studies show that the performance of the RA estimator is better than the M estimator and slightly better than the GM estimator when the model has been contaminated with additive outliers. Besides, Bustos et al. [18] studied the asymptotic behavior of the RA estimator for unilateral autoregressive spatial processes, generalizing the results for 1D time series asymptotic behavior established by Bustos and Fraiman [16]. Although the performance of the M and GM estimators is acceptable under innovation outliers, their asymptotic properties are still open problems.

One of the reasons why the spatial autoregressive model (AR-2D) has been extensively used in image analysis and processing is due to its remarkable ability to represent a variety of real scenarios without the need to use a large number of parameters. However, the robust estimators developed so far for the parameters of the AR-2D model have been constructed only under the assumption of innovative or additive type random noise; there are no proposals for parameter estimation when the model is contaminated from another more general pattern of noise. In this work we define a new class of robust estimators for the AR-2D contaminated models. This class of estimators is robust under replacement contamination that includes additive type contamination. Our proposal is the generalization to the bidimensional case of the BMM estimators, developed by Muler, Peña and Yohai [50] for ARMA time series models. The rest of the paper is organized as follows. In Section 2, the basic definitions are presented. First, some background material on bidimensional autoregressive processes (AR-2D) and model parameter estimators are presented. We also define procedures for generating replacement contamination in such models. In Section 3, we present the new model BIP-AR-2D for spatial processes and the new estimator of the AR-2D model parameters. In Section 4, several Monte Carlo studies are carried out to evaluate the performance of the new estimator against different contamination schemes, compared to the LS, M, GM and RA estimators. Section 5 presents two applications to real images that demonstrate the capabilities of the BMM-estimator to represent, segment and restore contaminated images. Conclusions and future works appear in Section 6. The results of the Monte Carlo studies (Section 4) are shown in the Appendix.

2 Preliminaries

2.1 The spatial ARMA models

In order to represent images using models that can be statistically worked on, three classes of models have been proposed. Whittle [78] studied simultaneous autoregressive (AR) models; Besag [9] introduced conditional autoregressive models. Moving average (MA) models were studied by Haining [36].

Spatial autoregressive moving average (ARMA) processes have also been studied in the context of random fields indexed over \mathbb{Z}^d , $d \geq 2$, where \mathbb{Z}^d is endowed with the usual partial order that is for $s = (s_1, s_2, \dots, s_d)$, $u = (u_1, u_2, \dots, u_d)$ in \mathbb{Z}^d , $s \leq u$ if for $i = 1, 2, \dots, d$, $s_i \leq u_i$. For $a, b \in \mathbb{Z}^d$, such that $a \leq b$ and $a \neq b$, we define $S[a, b] = \{x \in \mathbb{Z}^d | a \leq x \leq b\}$ and $S\langle a, b \rangle = S[a, b] \setminus \{a\}$.

A random field $(Y_s)_{s \in \mathbb{Z}^d}$ is said to be a spatial ARMA(p, q) with parameters $p, q \in \mathbb{Z}^d$ if it is weakly stationary and satisfies the equation

$$Y_s - \sum_{j \in S\langle 0, p \rangle} \phi_j Y_{s-j} = \varepsilon_s + \sum_{k \in S\langle 0, q \rangle} \theta_k \varepsilon_{s-k}, \quad (1)$$

where $(\phi_j)_{j \in S\langle 0, p \rangle}$ and $(\theta_k)_{k \in S\langle 0, q \rangle}$ denote, respectively the autoregressive and moving average parameters

with $\phi_0 = \theta_0 = 1$, and $(\varepsilon_s)_{s \in \mathbb{Z}^d}$ denotes a sequence of independent and identically distributed centered random variables with variance σ^2 . Notice that if $p = 0$, the sum over $S(0, p]$ is supposed to be zero and the process is called spatial moving average MA(q) random field. Similarly if $q = 0$ the process is called spatial autoregressive AR(p) random field. The ARMA random field is called causal if it has the following unilateral representation:

$$Y_s = \sum_{j \in S[0, \infty]} \phi_j \varepsilon_{s-j}, \quad (2)$$

with $\sum_j |\phi_j| < \infty$. Similar to the time series case, there are conditions on the (AR or MA) polynomials for them to be stationary and invertible respectively. Let $\Phi(z) = 1 - \sum_{j \in S(0, p]} \phi_j z^j$ and $\Theta(z) = 1 - \sum_{j \in S(0, q]} \theta_j z^j$, where $z = (z_1, z_2, \dots, z_d)$ and $z^j = z_1^{j_1} z_2^{j_2} \dots z_d^{j_d}$. A sufficient condition for the random field to be causal is that the AR polynomial $\Phi(z)$ has no zeros in the closure of the open unit disc D^d in \mathbb{C}^d . For example if $d = 2$, the process is causal if $\Phi(z_1, z_2)$ is not zero for any z_1 and z_2 which simultaneously satisfy $|z_1| < 1$ and $|z_2| < 1$ [40].

Applications of spatial ARMA processes including the analysis of yield trials in the context of incomplete block designs [32, 26], and the study of spatial unilateral first-order ARMA models [7] have been developed. Other extensions of the theory developed for time series to spatial ARMA models can be found in [33, 23, 6, 71, 19].

As an example, consider a particular case of model (1) when $d = 2$ and $p = (1, 1)$. This model is called a first-order autoregressive process. Note that $S = \langle (0, 0), (1, 1) \rangle = \{(1, 0), (0, 1), (1, 1)\}$ and the model is of the form

$$Y_{i,j} = \phi_1 Y_{i-1,j} + \phi_2 Y_{i,j-1} + \phi_3 Y_{i-1,j-1} + \varepsilon_{i,j} \quad (3)$$

where to simplify the notation it took $\phi_1 = \phi_{1,0}$, $\phi_2 = \phi_{0,1}$ and $\phi_3 = \phi_{1,1}$. In equivalent form, (3) can be expressed as

$$Y_{i,j} - \phi_1 Y_{i-1,j} - \phi_2 Y_{i,j-1} - \phi_3 Y_{i-1,j-1} = \varepsilon_{i,j}$$

or in a compact form as

$$\Phi(B_1, B_2) Y_{i,j} = \varepsilon_{i,j} \quad (4)$$

where B_1 and B_2 are the backward operators given by $B_1 Y_{i,j} = Y_{i-1,j}$, $B_2 Y_{i,j} = Y_{i,j-1}$ and in (4), $\Phi(B_1, B_2) = (1 - \phi_1 B_1 - \phi_2 B_2 - \phi_3 B_1 B_2)$. In the case that $\Phi(B_1, B_2)$ has inverse, we can write equation (2) as

$$Y_{i,j} = \Phi^{-1}(B_1, B_2) \varepsilon_{i,j}$$

The correlation structure of a process like (3) was investigated by Basu and Reinsel [7]. They obtained conditions to guarantee the existence of the stationary representation of the model (3) as in (2). In that case, the use of a multinomial expansion for $\Phi^{-1}(B_1, B_2)$ implies the convergent representation

$$Y_{i,j} = \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \sum_{r=0}^{\infty} \lambda_{klr} \varepsilon_{i-k-r, j-l-r} \quad (5)$$

where $\lambda_{klr} = \frac{(k+l+r)!}{k!l!r!} \phi_1^k \phi_2^l \phi_3^r$ with $k, l, r \in \mathbb{N} \cup \{0\}$ are the coefficients of this multinomial expansion.

In the literature there exist several prediction windows that can be considered in the definition of a spatial ARMA process. For example Kashyap and Eom [41] used a finite subset N_1 of the non-symmetrical half plane Ω_- given by

$$\Omega_- = \{(i, j) : (i = 0 \text{ and } j < 0) \text{ or } (i < 0 \text{ and } j \text{ is arbitrary})\}.$$

A complete treatment of the prediction window for spatial ARMA models and examples can be found in Guyon [35] and Bustos et al. [19].

A great variety of texture can be generated through the two-dimensional AR models. Figure (1) shows textures generated with an AR-2D process; (a) and (b) with two parameters, (c) and (d) with

three parameters. The increase in the number of parameters in the model also increases the diversity of possible textures but in contrast, the calculations become more complex. In this paper we worked with the AR 2D model with three parameters as in (3).

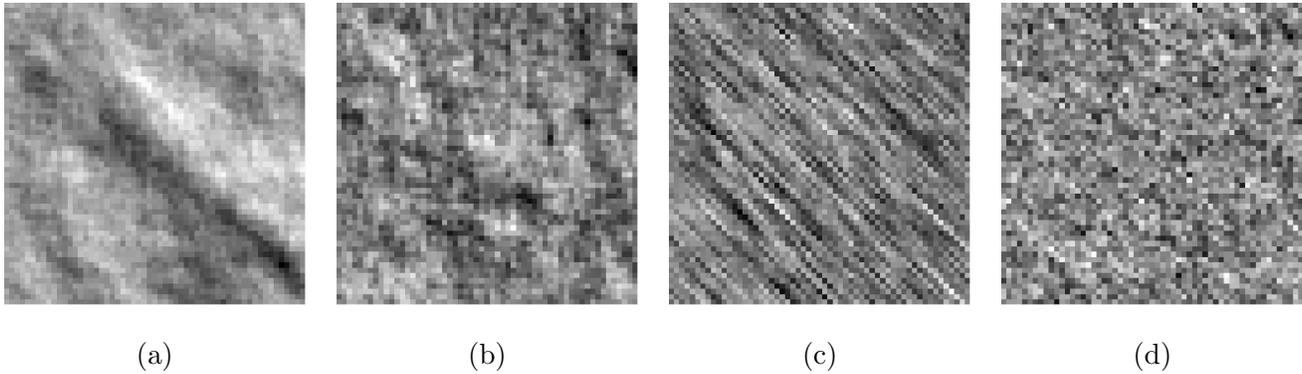


Figure 1: Autoregressive Processes. (a) $\phi_1 = 0.5, \phi_2 = 0.4999, \phi_3 = 0$; (b) $\phi_1 = 0.5, \phi_2 = 0.4, \phi_3 = 0$; (c) $\phi_1 = 0.01, \phi_2 = 0.01, \phi_3 = 0.8$; (d) $\phi_1 = 0.15, \phi_2 = 0.17, \phi_3 = 0.2$.

2.2 Types of contamination in AR-2D processes

Maronna et al. describe in [46] (Chapter 8) some probability models for time series outliers, including additive outliers (AOs), replacement outliers (ROs) and innovation outliers (IOs). Let X be a wide-sense stationary “core” unidimensional process of interest, and let V be a stationary outlier unidimensional process, which is non-zero in a fraction α of the time, i.e., $P(V_t = 0) = 1 - \alpha$. Under an AO model, instead of X_t one observes

$$Y_t = X_t + V_t$$

where the processes X and V are assumed to be independent of one another. The AO model was originally introduced by Fox [29] in 1972 for time series, who called them Type I outliers. Fox attributed such outliers to a “gross-error of observation or a recording error that affects a single observation”.

On the other hand, RO models have the form

$$Y_t = (1 - Z_t)X_t + Z_tW_t$$

where Z is a zero-one unidimensional process with $P(Z_t = 0) = 1 - \alpha$, and W is a “replacement” unidimensional process that is not necessarily independent of X . The RO models contain AO models as a special case in which $W_t = X_t + V_t$, and Z_t is a Bernoulli process.

The IOs are a highly specialized form of outlier that occur in linear processes such as AR, ARMA and ARIMA models in time series. IO models were first introduced by [29], who termed them Type II outliers. For simplicity, we illustrate IOs here in the special case of a first-order autoregression model. A stationary first-order AR model is given by

$$X_t = \phi X_{t-1} + U_t$$

where the innovative unidimensional process U is i.i.d. with zero mean and finite variance, and $|\phi| < 1$. An IO is an outlier in the U process. IOs are obtained, for example, when the innovative process has a zero-mean normal mixture distribution

$$(1 - \alpha)N(0, \sigma_0^2) + \alpha N(0, \sigma_1^2) \tag{6}$$

where $\sigma_1^2 \gg \sigma_0^2$. More generally, we can say that the process has IOs when the distribution of U has a heavy tailed.

In this section we generalized the notion of replacement outliers for spatial processes. Let Y be a stationary process, for example an AR-2D process as (3) and let Z be the observed process. It is said that Z process behaves like a two-dimensional Replacement Outlier model (RO) if it is given by

$$Z_{i,j} = (1 - \xi_{i,j}^\alpha)Y_{i,j} + \xi_{i,j}^\alpha W_{i,j} \quad (7)$$

where ξ^α is a zero-one process such that $P(\xi_{i,j}^\alpha = 1) = \alpha$ and $P(\xi_{i,j}^\alpha = 0) = 1 - \alpha$ and W is a replacement process that is not necessarily independent of Y . The fraction α is positive and small. A particular case of the RO models is the two-dimensional Additive Outlier model (AO), in which

$$W_{i,j} = Y_{i,j} + \nu_{i,j},$$

ν is a stationary process independent of Y and ξ^α is a Bernoulli process. This type of contamination is very important for satellite image processing; for example, it is present in optical images such as those from Landsat satellites. When W does not follow the pattern of AO, we say that the contamination process follows a Strictly Replacement Outlier model (SRO).

Similar to the case of time series, an innovative outlier is defined in an AR-2D process such as the one in (3) as an outlier type that affects $\varepsilon_{i,j}$ innovations. For example, let Y be a stationary first-order AR-2D model given by

$$Y_{i,j} = \phi Y_{i-1,j} + \varepsilon_{i,j}$$

where the innovations $\varepsilon_{i,j}$ are i.i.d. with zero mean and finite variance, and $|\phi| < 1$. An IO is an outlier in the ε process. IOs are obtained when the process of innovations has a zero-mean normal mixture distribution as in (6).

2.3 Robust Parametric Estimation

It is well known that the LS estimators are very sensitive to the outliers (Martin [45]). This fostered the introduction of several alternative estimators to attenuate the impact of contaminated observations on the estimators. Most of these proposals are natural extensions of robust estimators studied in time series.

Robust estimators have been defined for models containing a finite number of parameters. Here, we use model (3) to describe the well-known robust estimators; however, a more general treatment for AR and MA models can be found in Kashyap and Eom [41], Allende, Galbiati and Vallejos [2], Ojeda, Vallejos and Lucini [55], Vallejos and Garcia-Donato [71], and Bustos, et al. [18].

Notice that model (3) can be rewritten in the linear model form:

$$Y_{i,j} = \boldsymbol{\phi}^T Z_{i,j} + \varepsilon_{i,j},$$

where $\boldsymbol{\phi}^T = (\phi_1, \phi_2, \phi_3)$ is a parameter vector and $Z_{i,j}^T = (Y_{i-1,j}; Y_{i,j-1}; Y_{i-1,j-1})$. To obtain the LS estimator of $\boldsymbol{\phi}$, we need to minimize, with respect to $\boldsymbol{\phi}$, the function:

$$\sum_{i,j} [\varepsilon_{i,j}(\boldsymbol{\phi})]^2,$$

where

$$\varepsilon_{i,j}(\boldsymbol{\phi}) = Y_{i,j} - \boldsymbol{\phi}^T Z_{i,j} = Y_{i,j} - \phi_1 Y_{i-1,j} - \phi_2 Y_{i,j-1} - \phi_3 Y_{i-1,j-1} \quad (8)$$

Similarly, the class of M estimators for 2D autoregressive processes (Kashyap and Eom, 1988)[41] is defined by the minimization of the function

$$Q(\boldsymbol{\phi}, \sigma) = \sum_{i,j} \left[\rho \left(\frac{Y_{i,j} - \boldsymbol{\phi}^T Z_{i,j}}{\sigma} \right) + \frac{1}{2} \right] \sigma, \quad (9)$$

Likewise, if $\hat{\sigma}$ is an estimate of σ , the M-estimators are defined by minimizing the function

$$M_{nm}(\boldsymbol{\phi}) = \frac{1}{(n-1)(m-1)} \sum_{i=2}^n \sum_{j=2}^m \rho \left(\frac{\varepsilon_{i,j}(\boldsymbol{\phi})}{\hat{\sigma}} \right) \quad (10)$$

This estimator is robust for innovation outliers, when the function ρ is a differentiable function, convex, symmetric with respect to the origin, with bounded derivative, and such that $\rho(0) = 0$. However, the M estimators are very sensitive when the process is contaminated with additive outliers. This suggested the introduction of other robust estimators capable of lessening the effects of additive outliers. Allende, Galbiati and Vallejos [2] developed the generalized M estimators (GM) for spatial AR processes. A GM estimator of ϕ is the solution to the problem of minimizing the non-quadratic function defined by:

$$Q(\phi, \sigma) = \sum_{i,j} l_{ij} t_{ij} \left[\rho \left(\frac{Y_{i,j} - \phi^T Z_{i,j}}{l_{ij} \sigma} \right) + \frac{1}{2} \right] \sigma,$$

where ρ is as in (9), t_{ij} and l_{ij} are weights corresponding to the respective $Z_{i,j}$.

Alternatively, Ojeda, Vallejos and Lucini [55] introduced the robust autocovariance (RA) estimators for spatial autoregressive processes. This estimator was first introduced by Bustos and Yohai [20] for models used in time series.

Let Y be a zero mean AR-2D process described by the equation:

$$Y_{m,n} - \sum_{k,l \in S(0,p]} \phi_{k,l} Y_{m-k,n-l} = \varepsilon_{m,n}, \quad (11)$$

where $\varepsilon_{m,n}$ is a sequence of i.i.d. random variables with $Var(\varepsilon_{m,n}) = \sigma^2$. Let us assume that Y is observed on a strongly causal squared window of order M , $W_M = \{(k,l) \in S : 0 \leq k, l \leq M\}$, where S is an infinite strong causal prediction neighborhood. Let us define $W_M \setminus S(0,p] = \{(m,n) \in W_M : [(m,n) - p] \in W_M\}$. The residual of order (m,n) in ϕ of Y is

$$r(m,n) = \begin{cases} - \sum_{k,l \in T'} \phi_{k,l} Y_{m-k,n-l}, & (m,n) \in (W_M \setminus S(0,p]) \\ 0, & \text{otherwise,} \end{cases} \quad (12)$$

where $T' = S(0,p] \cup \{(0,0)\}$ and $\phi_{0,0} = -1$. In particular, for $p = (1,1)$ we define the coefficients

$$p_\phi(k,l,r) = \frac{(k+l+r)!}{k! l! r!} \phi_1^k \phi_2^l \phi_3^r,$$

then the RA estimator $\hat{\phi}$ of ϕ is defined by the following equations

$$\sum_{k,l,r=0}^{\infty} p_\phi(k,l,r) \sum_{(m,n) \in (W_M \setminus S(0,(1,1)])} \eta \left(\frac{r(m,n)}{\hat{\sigma}}, \frac{r(m-i-k-r, n-j-l-r)}{\hat{\sigma}} \right) = 0, \quad (13)$$

$$\sum_{(m,n) \in (W_M \setminus S(0,(1,1)])} \psi \left(\frac{r(m,n)}{\hat{\sigma}} \right) = 0, \quad (14)$$

where $\hat{\sigma}$ is estimated independently by

$$\hat{\sigma} = Med(|r(m,n)| : (m,n) \in (W_M \setminus S(0,(1,1))))/0.6745, \quad (15)$$

η is a continuous, bounded and odd function in two variables and $0.6745 = Med(|W|)$, where W is a standard normal random variable. Two possible choices for η have been suggested in the literature (see Bustos et al. [19])

3 A new approach

3.1 BIP-AR 2D models

A new class of bounded nonlinear AR-2D models is presented in this work: the bounded innovation propagation AR-2D model (BIP-AR 2D). This model arises from the need to estimate the best possible parameters of a autoregressive central model when a contaminated process is observed. The BIP-AR 2D

model is a two-dimension generalization of the model presented for time series by Muler et al. [50].

Given a stationary and invertible AR-2D model like in (3), it supports a stationary representation as in (5). We define the BIP-AR 2D auxiliary model as:

$$Y_{i,j} = \sum_{(k,l,r) \in D} \lambda_{klr} \sigma \eta \left(\frac{\varepsilon_{i-k-r,j-l-r}}{\sigma} \right) + \varepsilon_{i,j} \quad (16)$$

with $D = \{(k, l, r) \in \mathbb{N}_0^3\} / \{(0, 0, 0)\}$, $\varepsilon_{i,j}$'s are i.i.d. random variables with symmetric distribution and $\eta(x)$ is an odd and bounded function. σ is a robust M-scale of $\varepsilon_{i,j}$, which coincides with the standard deviation if $\varepsilon_{i,j}$ are normal, and is defined as the solution of the equation $E(\rho(\varepsilon_{i,j}/\sigma)) = b$.

Note that (16) can also be written as

$$Y_{i,j} = \sum_{\{k \geq 0, l \geq 0, r \geq 0\}} \lambda_{klr} \sigma \eta \left(\frac{\varepsilon_{i-k-r,j-l-r}}{\sigma} \right) - \sigma \eta \left(\frac{\varepsilon_{i,j}}{\sigma} \right) + \varepsilon_{i,j} = \sigma \Phi^{-1}(B_1, B_2) \eta \left(\frac{\varepsilon_{i,j}}{\sigma} \right) - \sigma \eta \left(\frac{\varepsilon_{i,j}}{\sigma} \right) + \varepsilon_{i,j}$$

and multiplying both members by $\Phi(B_1, B_2)$, we get

$$\Phi(B_1, B_2) Y_{i,j} = \sigma \eta \left(\frac{\varepsilon_{i,j}}{\sigma} \right) - \sigma \Phi(B_1, B_2) \eta \left(\frac{\varepsilon_{i,j}}{\sigma} \right) + \Phi(B_1, B_2) \varepsilon_{i,j}$$

which is equivalent to

$$\begin{aligned} Y_{i,j} &= \phi_1 Y_{i-1,j} + \phi_2 Y_{i,j-1} + \phi_3 Y_{i-1,j-1} + \\ &\sigma \phi_1 \eta \left(\frac{\varepsilon_{i-1,j}}{\sigma} \right) + \sigma \phi_2 \eta \left(\frac{\varepsilon_{i,j-1}}{\sigma} \right) + \sigma \phi_3 \eta \left(\frac{\varepsilon_{i-1,j-1}}{\sigma} \right) + \\ &\varepsilon_{i,j} - \phi_1 \varepsilon_{i-1,j} - \phi_2 \varepsilon_{i,j-1} - \phi_3 \varepsilon_{i-1,j-1} \end{aligned} \quad (17)$$

3.2 BMM estimator for AR-2D processes

In time series, Muler et al. [50] introduce the MM-estimators for ARMA models based in the definition of MM-estimate for regression proposed by Yohai ([80]) with the difference that the residuals are calculated as in the BIP-ARMA model instead of just as in the pure ARMA model. The idea of MM-estimators in regression is to compute a highly robust estimator of the error scale in a first stage, and this estimated scale is used to calculate an M-estimator of the regression parameters in a second stage. However, in time series this differs somewhat because an MM-estimate is not enough to guarantee robustness.

In the same way that residues of AR-2D model exist (8), there are residues obtained from BIP-AR 2D model:

$$\begin{aligned} \varepsilon_{i,j}^b(\phi, \sigma) &= Y_{i,j} - \phi_1 Y_{i-1,j} - \phi_2 Y_{i,j-1} - \phi_3 Y_{i-1,j-1} - \\ &\sigma \left(\phi_1 \eta \left(\frac{\varepsilon_{i-1,j}^b(\phi, \sigma)}{\sigma} \right) + \phi_2 \eta \left(\frac{\varepsilon_{i,j-1}^b(\phi, \sigma)}{\sigma} \right) + \phi_3 \eta \left(\frac{\varepsilon_{i-1,j-1}^b(\phi, \sigma)}{\sigma} \right) \right) + \\ &\phi_1 \varepsilon_{i-1,j}^b(\phi, \sigma) + \phi_2 \varepsilon_{i,j-1}^b(\phi, \sigma) + \phi_3 \varepsilon_{i-1,j-1}^b(\phi, \sigma) \end{aligned} \quad (18)$$

for all $i, j \geq 2$. With this residue the objective function that must be minimized to obtain the M-estimator of the parameters under a model BIP-AR 2D is defined:

$$M_{nm}^b(\phi) = \frac{1}{(n-1)(m-1)} \sum_{i=2}^n \sum_{j=2}^m \rho \left(\frac{\varepsilon_{i,j}^b(\phi, \hat{\sigma})}{\hat{\sigma}} \right) \quad (19)$$

where $\hat{\sigma}$ is a robust estimate of σ .

One way of robustly estimating the scale was introduced in 1964 by Huber ([39]) as follows: Let a sample $\mathbf{u} = (u_1, \dots, u_n)$, with $u_i \in \mathbb{R}$, an M-estimate of scale $S_n(\mathbf{u})$ is defined by any value $s \in (0, \infty)$ satisfying

$$\frac{1}{n} \sum_{i=1}^n \rho\left(\frac{u_i}{s}\right) = b \quad (20)$$

where ρ is a continuous and non-constant function, non-decreasing in $|x|$ and symmetric around zero as well. To make the M-scale estimate consistent with the standard deviation when the data are normal, it requires that $E(\rho(x)) = b$ under the standard normal distribution. Taking $b = \max(\rho)/2$, we get a maximum breakdown point of 0.5.

With all this, we can define the new BMM-2D estimator by following the two steps given below:

First Step: At this stage, an estimate of σ is obtained. For this purpose, two σ estimates are considered: one using an AR-2D model, another using a BIP-AR 2D model, and choosing the smallest of them. Let ρ_1 a continue, non-constant, non-decreasing in $|x|$, bounded and symmetric function and such that: $b = E(\rho_1(u)) \Rightarrow b/\max(\rho_1) = 0.5$. This guarantees that for a normal random sample, the M-scale estimator s based on ρ_1 converges to the standard deviation and the breakdown point of s is 0.5. Put

$$\mathcal{B}_{0,\xi} = \{\phi = (\phi_1, \phi_2, \phi_3) \in \mathbb{R}^3 : |z| \geq 1 + \xi \text{ holds for all the roots } z \text{ of } \Phi(B_1, B_2)\}$$

Let us call $\mathcal{B} = \mathcal{B}_{0,\xi}$ for some small $\xi > 0$.

Then, we define an estimate of $\phi \in \mathcal{B}$:

$$\hat{\phi}_S = \arg \min_{\phi \in \mathcal{B}} S_{nm}(\epsilon_{nm}(\phi))$$

and the corresponding estimate of σ :

$$s_{nm} = S_{nm}(\epsilon_{nm}(\hat{\phi}_S)) \quad (21)$$

where $\epsilon_{nm}(\phi) = (\epsilon_{22}(\phi), \epsilon_{32}(\phi), \dots, \epsilon_{n2}(\phi), \epsilon_{23}(\phi), \dots, \epsilon_{2m}(\phi), \dots, \epsilon_{nm}(\phi))$ with $\epsilon_{ij}(\phi)$ as in (8) and S_{nm} is the M-estimate of scale based on ρ_1 and b defined as in (20).

Let us describe now the estimate corresponding to the BIP-AR model. Define $\hat{\phi}_S^b$ by the minimization of $S_{nm}(\epsilon_{nm}^b(\phi, \hat{\sigma}(\phi)))$ over \mathcal{B} . The value $\hat{\sigma}(\phi)$ is an estimate of σ computed as if ϕ were the true parameters and the $\epsilon_{i,j}$'s were normal. Then, since in this case the M-scale σ coincides with the standard deviation of $\epsilon_{i,j}$, from (16) we have:

$$\sigma^2 = \frac{\sigma_Y^2}{1 + \kappa^2 \sum_{k,l,r \geq 0} \lambda_{klr}^2}$$

where $\kappa^2 = \text{Var}(\eta(\frac{\epsilon_{i,j}}{\sigma}))$ and $\sigma_Y^2 = \text{Var}(Y_{i,j})$. Let $\hat{\sigma}_Y^2$ a robust estimate of σ_Y^2 and $\kappa^2 = \text{Var}(\eta(Z))$ where $Z \sim N(0, 1)$. Then, we define

$$\hat{\sigma}^2 = \frac{\hat{\sigma}_Y^2}{1 + \kappa^2 \sum_{k,l,r \geq 0} \lambda_{klr}^2(\phi)}$$

The scale estimate s_{nm}^b corresponding to the BIP-AR-2D model is defined by

$$\hat{\phi}_S^b = \arg \min_{\phi \in \mathcal{B}} S_{nm}(\epsilon_{nm}^b(\phi, \hat{\sigma}(\phi)))$$

and

$$s_{nm}^b = S_{nm}(\epsilon_{nm}^b(\hat{\phi}_S^b, \hat{\sigma}(\hat{\phi}_S^b)))$$

where $\epsilon_{nm}^b(\phi) = (\epsilon_{22}^b(\phi, \hat{\sigma}(\phi)), \epsilon_{32}^b(\phi, \hat{\sigma}(\phi)), \dots, \epsilon_{n2}^b(\phi, \hat{\sigma}(\phi)), \epsilon_{23}^b(\phi, \hat{\sigma}(\phi)), \dots, \epsilon_{2m}^b(\phi, \hat{\sigma}(\phi)), \dots, \epsilon_{nm}^b(\phi, \hat{\sigma}(\phi)))$ with $\epsilon_{ij}^b(\phi, \hat{\sigma}(\phi))$ defined as in (18).

Our estimate of σ is

$$s_{nm}^* = \min(s_{nm}, s_{nm}^b)$$

Second Step: We consider a bounded function ρ_2 that satisfies the same properties as ρ_1 but also $\rho_2 \leq \rho_1$. This function is chosen such that the corresponding M-estimator is highly efficient under normal errors. Given the objective functions defined in (10) and (19) with scale obtained in the first step (s_{nm}^*):

$$M_{nm}(\phi) = \frac{1}{(n-1)(m-1)} \sum_{i=2}^n \sum_{j=2}^m \rho_2 \left(\frac{\varepsilon_{i,j}(\phi)}{s_{nm}^*} \right) \quad (22)$$

and

$$M_{nm}^b(\phi) = \frac{1}{(n-1)(m-1)} \sum_{i=2}^n \sum_{j=2}^m \rho_2 \left(\frac{\varepsilon_{i,j}^b(\phi, s_{nm}^*)}{s_{nm}^*} \right) \quad (23)$$

The corresponding M-estimators of the parameters for each function are:

$$\hat{\phi}_M = \arg \min_{\phi \in \mathcal{B}} M_{nm}(\phi)$$

and

$$\hat{\phi}_M^b = \arg \min_{\phi \in \mathcal{B}} M_{nm}^b(\phi)$$

Then, we define the BMM-estimator 2D as:

$$\hat{\phi}_M^* = \begin{cases} \hat{\phi}_M, & \text{si } M_{nm}(\hat{\phi}_M) \leq M_{nm}^b(\hat{\phi}_M^b); \\ \hat{\phi}_M^b, & \text{si } M_{nm}(\hat{\phi}_M) > M_{nm}^b(\hat{\phi}_M^b) \end{cases}$$

4 Monte Carlo Results

The aim of this section is to analyze the performance of the new BMM estimator to estimate the parameters in the model (3) compared to the LS, M, GM and RA estimators. We performed several experiments. Each experiment is based on different Monte Carlo studies. We set the parameter values of (3) as:

$$Y_{i,j} = 0.15Y_{i-1,j} + 0.17Y_{i,j-1} + 0.2Y_{i-1,j-1} + \varepsilon_{i,j} \quad (24)$$

We performed our study under five different conditions of the model (Cases); in Case I, the model was non-contaminated, while in Cases II, III, IV and V, the model was contaminated according to (7):

- Case I) Non-contaminated model like in (24), where ε is a normal distribution process with $Var(\varepsilon_{i,j}) = 1$ and $E(\varepsilon_{i,j}) = 0$ for all i, j .
- Case II) AO, where the ν process is independent of the Y process and follows a normal distribution with zero mean and variance 50.
- Case III) SRO, where the replacement process W follows a t-student distribution with 2.3 f.d.
- Case IV) SRO, where the replacement process W is another autoregressive process, independent of the Y process, with parameters $\tilde{\phi}_1 = 0.1$, $\tilde{\phi}_2 = 0.2$ and $\tilde{\phi}_3 = 0.3$.
- Case V) SRO, where the replacement process W is a white noise process with normal distribution with zero mean and variance 50.

In each of the five variants of the model (24), the parameters ϕ_1, ϕ_2 and ϕ_3 were estimated by the five procedures presented in the previous sections. In each experiment, 500 simulations of the model were generated, and the mean value, the mean square error (MSE) and the sample variance were computed. For the contaminated models we considered four levels of contamination: 5%, 10%, 15% and 20%. Besides, we performed our study considering different window sizes: 8×8 , 16×16 , 32×32 , and 57×57 . For the calculation of the BMM estimator, a robust estimator of the scale was obtained as in (21). $\rho_1(x) =$

$\rho_2(\frac{x}{0.405})$ was selected according to the same criteria that was taken for the definition of the BMM estimator for time series present in [50], where the function ρ_2 is given by:

$$\rho_2(x) = \begin{cases} 0.5 * x^2, & \text{si } |x| \leq 2; \\ 0.002 * x^8 - 0.052 * x^6 + 0.432 * x^4 - 0.972 * x^2 + 1.792, & \text{si } 2 < |x| \leq 3; \\ 3.25, & \text{si } 3 < |x| \end{cases} \quad (25)$$

The same ρ_2 function was used to calculate the M estimators. In addition, for the implementation of the GM estimator the weights were set:

$$l_{i,j} = 1 \quad \forall i, j$$

$$t_{i,j} = \frac{\psi_H((Y_{i-1,j}^2 + Y_{i,j-1}^2 + Y_{i-1,j-1}^2)/3)}{(Y_{i-1,j}^2 + Y_{i,j-1}^2 + Y_{i-1,j-1}^2)/3}$$

where ψ_H is the following version of the Huber function:

$$\psi_H(x) = \begin{cases} x, & \text{si } |x| \leq 1.5; \\ 1.52, & \text{si } 1.5 < x; \\ -1.5, & \text{si } x < -1.5 \end{cases} \quad (26)$$

Finally the RA estimators were implemented according to the details formulated in [52]. To facilitate the paper reading, only the boxplots of the simulations have been included in the body of the work; the numerical Monte Carlo results are shown in the Appendix.

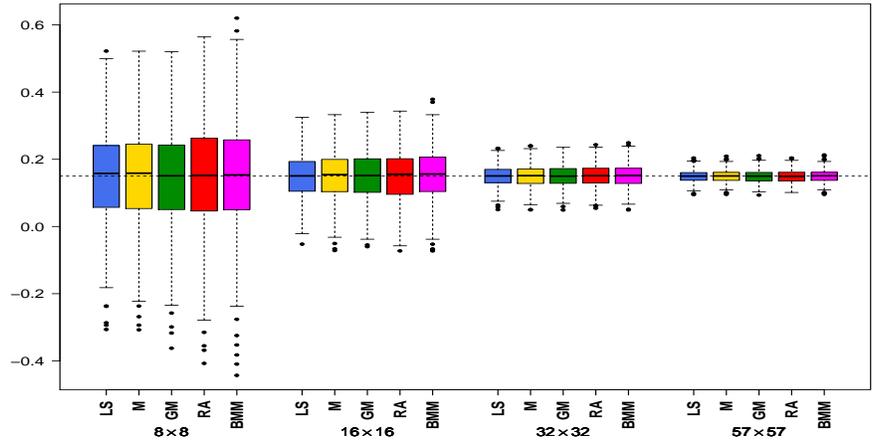
4.1 Experiments

In a first experiment, we studied the performance of the BMM estimator for the non-contaminated model (Case I). All the methods estimated the parameters quite well. Table 3 shows the results obtained for the four different window sizes considered. In Figure 2, the corresponding boxplots are shown. In this case, it is convenient to use the LS method due to its simplicity and calculation speed.

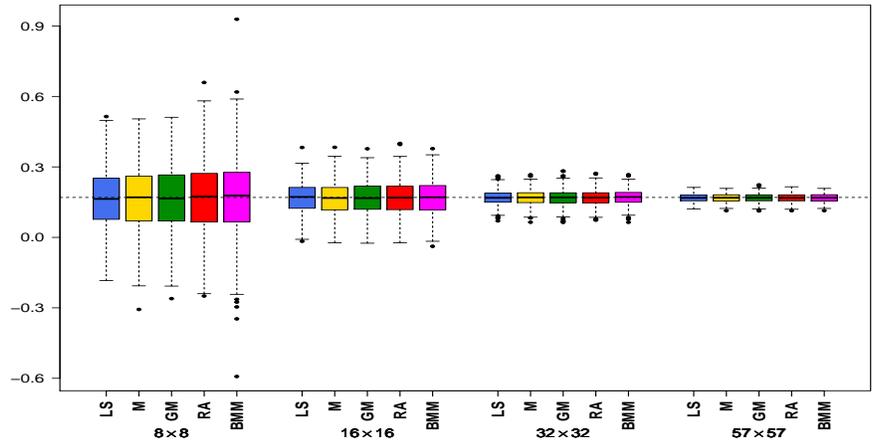
The second experiment was developed in the context of Case II. We analyzed the ability of the BMM method to estimate the parameters of the model, considering a 10% of additive contamination, and for window sizes 8×8 , 16×16 , 32×32 and 57×57 , in comparison with the LS, M, GM and RA methods. Table 4 shows the estimated values for ϕ_1 , ϕ_2 and ϕ_3 , using the different five procedures analyzed. Figure 3 exhibits the corresponding boxplots. For window size 32×32 and 57×57 , it can be seen that the BMM estimator is the best because its values are closer to the parameter than the estimates produced by the other methods mentioned. In addition, BMM estimator has the lowest variance and the lowest MSE. When the window size was 8×8 or 16×16 , the best performance corresponded to the GM and RA estimators; however, the BMM estimation values were similar to the RA and GM estimations. An analogue affirmation is valid to the sample variance and MSE of BMM. We also noted that for any window size, the M estimator had a very small sample variance but their estimations were wrong when compared to the ones in the other methods.

The third experiment also refers to Case II. We set the window size at 32×32 and varied the additive contamination level, considering four levels: 5%, 10%, 15% and 20%. The BMM method was the best in most of the cases studied, followed by the RA estimator. This behavior is deduced from the comparison of the values estimated by the BMM method with the respective estimates obtained by the other procedures. The values of the dispersion measures also point out that the BMM estimator is the most accurate methodology. These results can be seen in Table 5. In addition, from Figure 4 we can note that using any of the five estimators, the parameter ϕ_3 was estimated with less precision than ϕ_1 and ϕ_2 as the percentage of contamination increases. Besides, while ϕ_3 was underestimated by all methods, for all levels of contamination, the RA estimator was the only method that overestimated ϕ_1 and ϕ_2 , independently of the contamination level. The fourth experiment is related to Case III. The process of contamination is a replacement contamination where the replacement process W follows a t-student distribution with

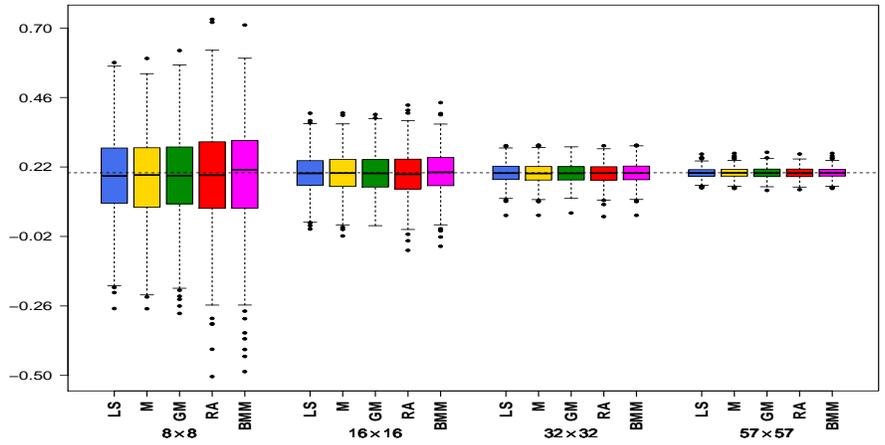
Figure 2: LS, M, GM, RA and BMM estimation boxplots for $\phi_1 = 0.15$ (a), $\phi_2 = 0.17$ (b) and $\phi_3 = 0.2$ (c); model (24) without contamination, varying the window sizes.



(a)



(b)



(c)

2.3 f.d.. The simulations were performed for a 57×57 window. Table 6 and plot 5 show the results. Boxplots show that the BMM method is the best performing estimator, followed by the RA, GM, M and LS methods, in that order. We also noted that in all methods the estimates deteriorate as the level of contamination increases. Additionally, the classic LS estimator presented greater dispersion of the data. The fifth experiment was performed in the context of Case IV, where the replacement process W was an autoregressive process. We set the window size at 32×32 , varying the level of contamination (5%, 10%, 15% and 20%). Table 7 shows these results. Besides, Figure 6 displays the boxplots corresponding to these tables. We can see a pattern similar to the fourth experiment; except that in this case, the variances

seem very much alike. Finally, the sixth experiment was carried out according to Case V. The replacement process of the contamination was a white noise with variance 50. As in the fifth experiment, we set the window size at 32×32 , varying the level of contamination (5%, 10%, 15% and 20%). Table 8 presents the estimated values obtained. The corresponding boxplots are shown in Figure 7. The parameter values ϕ_1 , ϕ_2 and ϕ_3 were underestimated for all methods, excluding the RA estimator that overestimated the values of ϕ_1 and ϕ_2 parameters. The BMM estimator was less affected by the contamination process. The LS and M estimators are less accurate than the GM, RA and BMM estimators. Comparatively, the RA estimator presented the highest variance, whereas the GM estimator, although quite accurate, deteriorates more than the BMM estimator as the level of contamination increases.

Figure 3: LS, M, GM, RA and BMM estimation boxplots for $\phi_1 = 0.15$ (a), $\phi_2 = 0.17$ (b) and $\phi_3 = 0.2$ (c), varying the window sizes. Model (24) with additive contamination 10% level, with a normal noise.

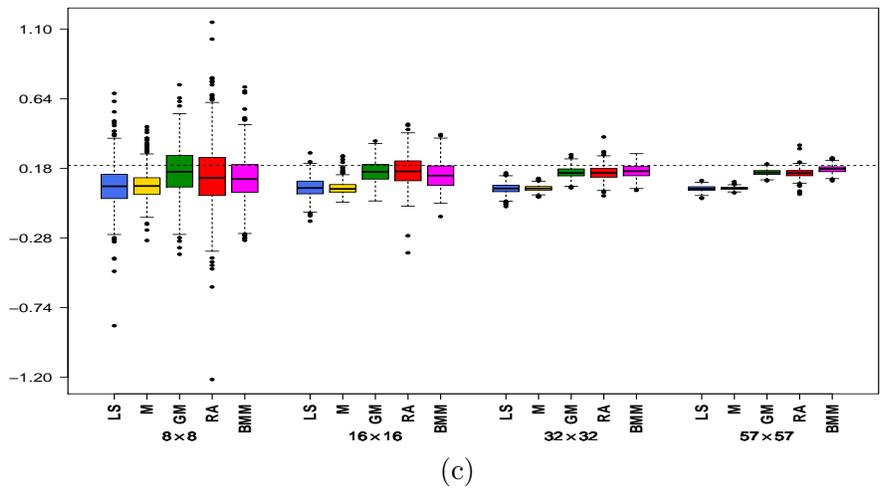
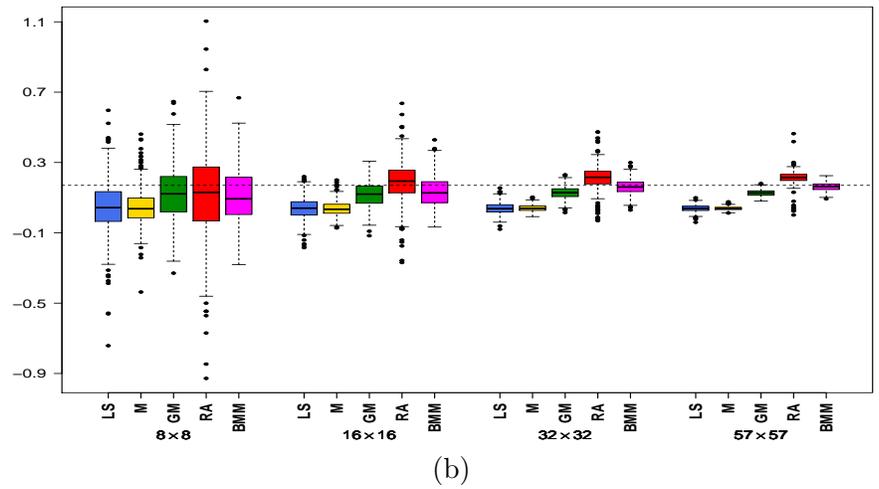
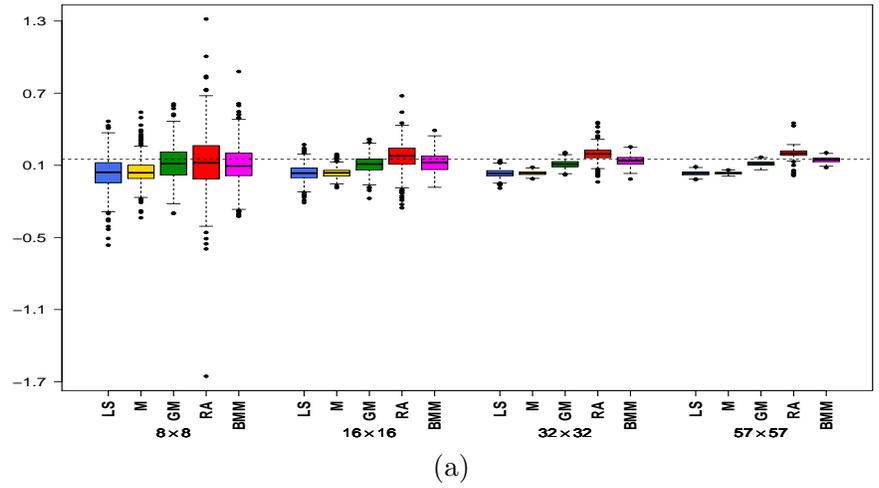
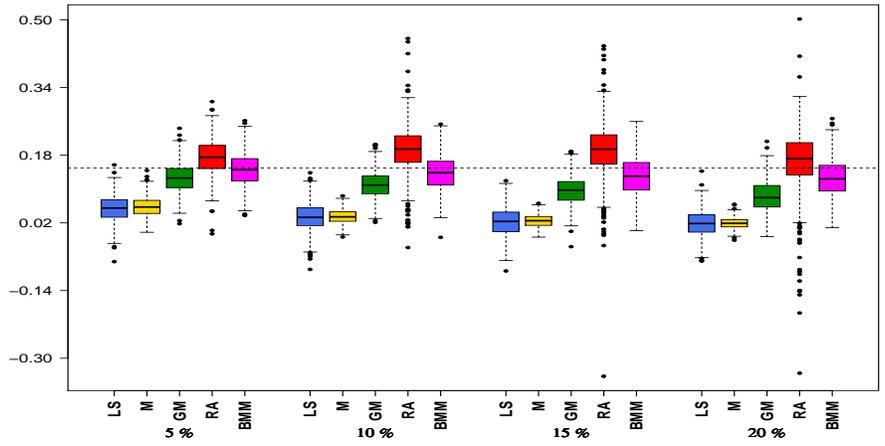
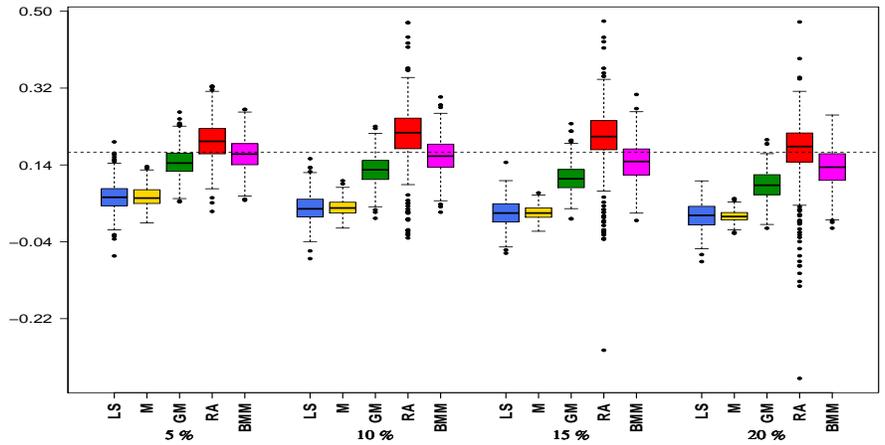


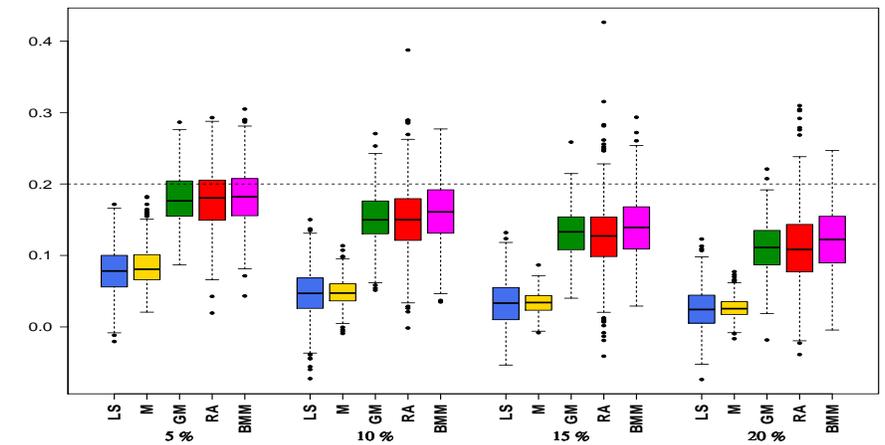
Figure 4: LS, M, GM, RA and BMM estimation boxplots, for $\phi_1 = 0.15$ (a), $\phi_2 = 0.17$ (b) and $\phi_3 = 0.2$ (c) in model 24, with additive contamination, varying the contamination level, with a 32×32 window size. The contamination process is a normal noise, with $\sigma^2 = 50$.



(a)



(b)



(c)

Figure 5: LS, M, GM, RA and BMM estimation boxplots, for $\phi_1 = 0.15$ (a), $\phi_2 = 0.17$ (b) and $\phi_3 = 0.2$ (c) in model 24, with contamination of replacement, varying the contamination level, with a 57×57 window size. The process of contamination follows a t-Student distribution with 2.3 d.f.

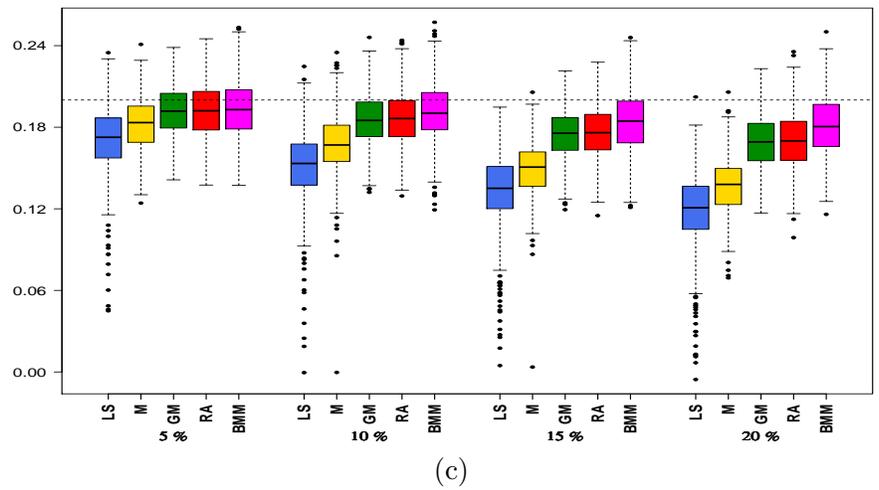
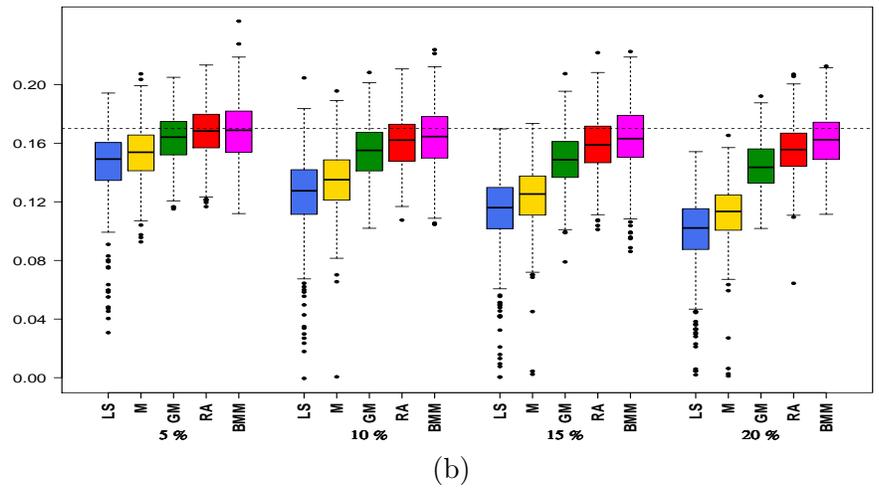
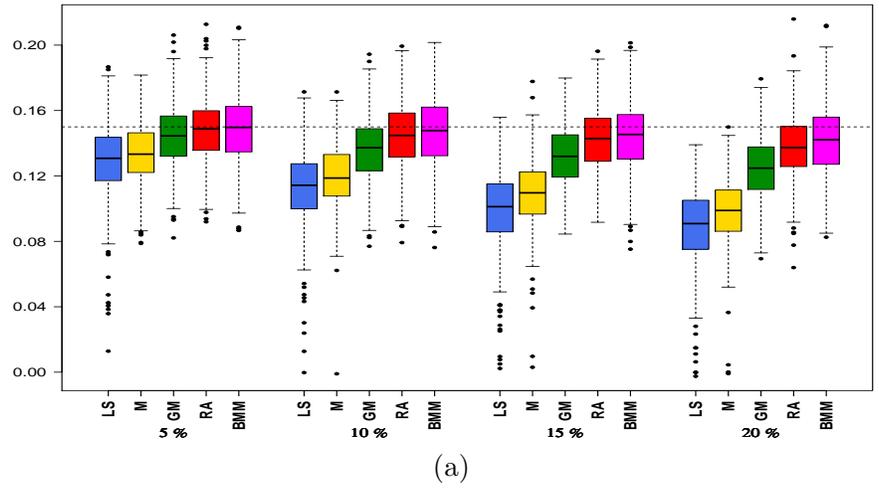


Figure 6: LS, M, GM, RA and BMM estimation boxplots, for $\phi_1 = 0.15$ (a), $\phi_2 = 0.17$ (b) and $\phi_3 = 0.2$ (c) in model 24, varying the contamination level, with a 32×32 window size. The contamination process is of replacement type, by an AR process with $\tilde{\phi}_1 = 0.1$, $\tilde{\phi}_2 = 0.2$ and $\tilde{\phi}_3 = 0.3$ parameters.

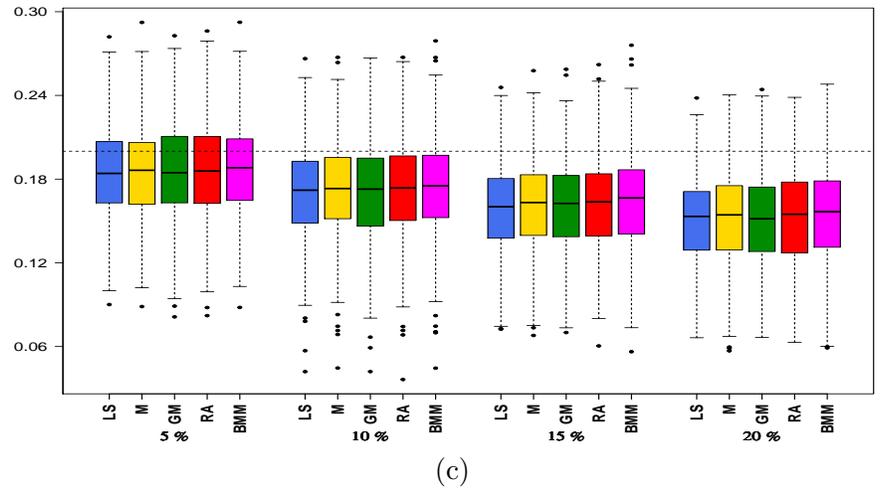
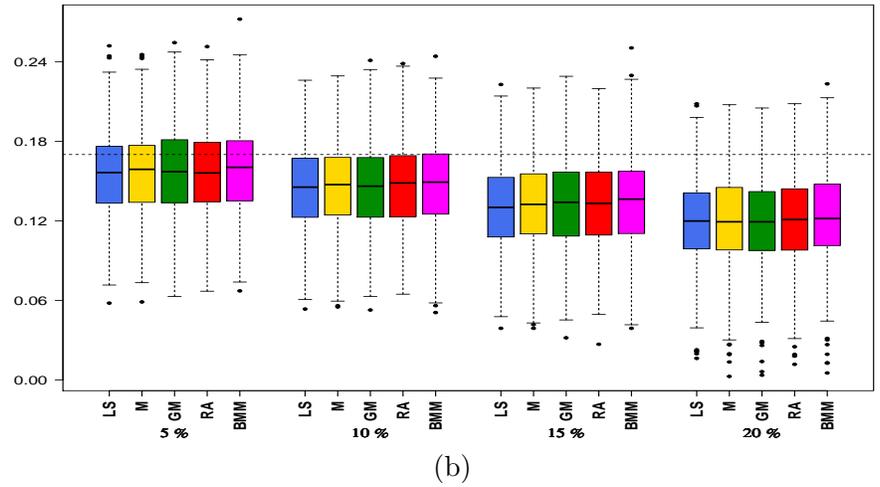
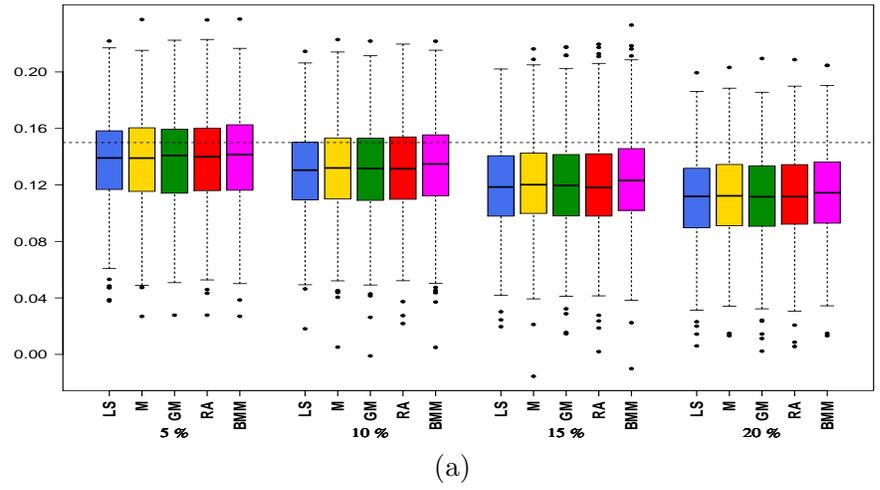
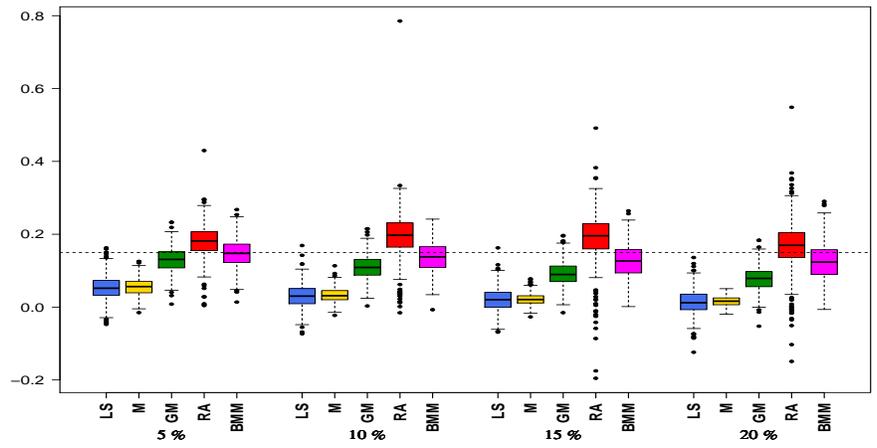
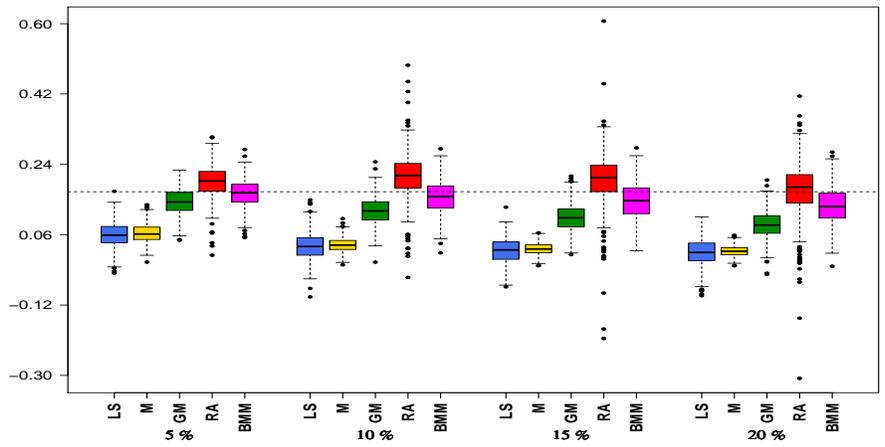


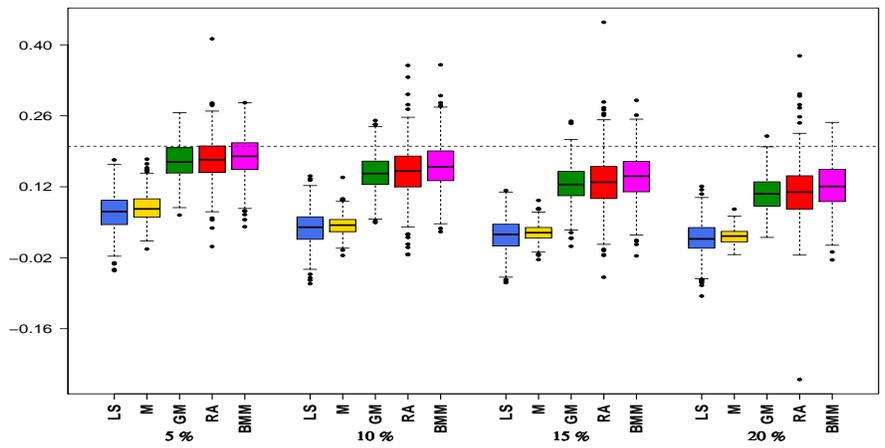
Figure 7: LS, M, GM, RA and BMM estimation boxplots, for $\phi_1 = 0.15$ (a), $\phi_2 = 0.17$ (b) and $\phi_3 = 0.2$ (c) in model 24, varying the contamination level, with a 32×32 window size. The contamination process is of replacement type, with a white noise, of variance 50.



(a)



(b)



(c)

4.2 Computational Time Evaluation

All the computational routines were developed in R and were carried out on the server JupiterAce of FaMAF-UNC. It has 12-cores 2.40GHz Intel Xeon E5-2620v3 processor, with 128 GiB 2133MHz of available DDR4 RAM. Running time as time logarithm of a single simulation to each estimator vs the window size in Case I is presented in Figure 8. Time was expressed in seconds. The graph shows that the computational cost of the RA estimator is the highest; for example, in a 32×32 window size, the RA running time was 43.812 seconds, while BMM, GM, M and LS computational costs were 2.936, 0.552, 0.516 and 0.436 seconds, respectively. This results show that, although RA estimator is one of the major competitors of BMM estimator, due to its accuracy and good asymptotic properties, it exhibits its computational cost as a disadvantage. This makes RA an unattractive estimator for the processing of big size images.

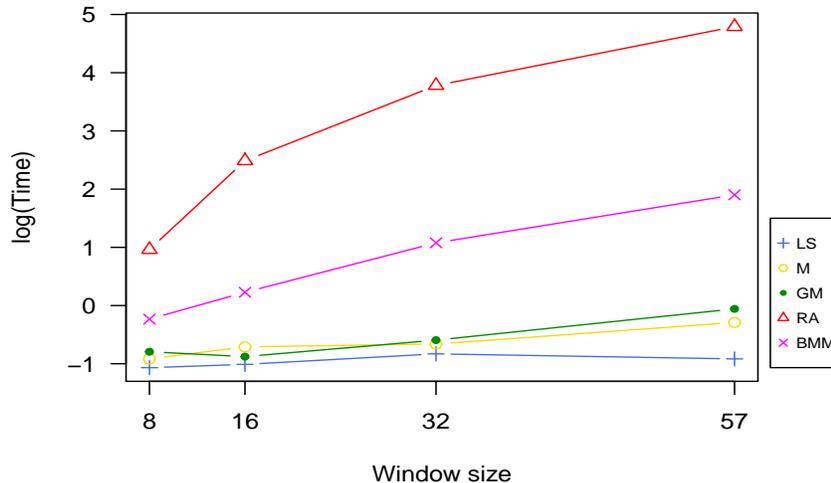


Figure 8: Logarithm of the estimation time (in seconds) when the process has additive contamination of $\sigma^2 = 50$ according to the window size.

5 Application to real images

The analysis of contaminated images is of great interest in several areas of research. For example, the reconstruction of contaminated images is relevant in modeling of images (Allende and Galbiati, [3], Vallejos and Mardesic, [72]), and, in general, the reduction of the noise produced by interferences taking place in the processes of obtaining the physical image and transmitting it electronically plays an important role in the literature (Bustos, [15]).

In Ojeda et al. ([53]), two algorithms for image processing based on the unilateral AR-2D model with two parameters were presented. The foundations of the algorithms are random field theory and robustness for spatial autoregressive processes. The first one produces a local approximation of images, and the second one, is a segmentation algorithm. In this work, we proposed to use a variant of these algorithms using a unilateral AR-2D process with three parameters (model (3)), instead of two parameters as it was originally proposed. We called the modified algorithms as Algorithm 1 and Algorithm 2. We applied them to reconstruction and segmentation of images using the LS, GM and BMM estimators of the parameters in the model (3). Later, we inspected and compared the performance of these estimators in Algorithms 1 and 2 on contaminated images. To compare the images generated by the algorithms and, therefore, the performance of the different estimators, we calculated three indexes used in the literature; the SSIM index [75], the CQ index ([54]), and CQmax index [58]. Next, we present two numerical experiments using the image “Lenna”, which was taken from the USC-SIPI image database <http://sipi.usc.edu/database/>. In Figure 9-(I), the original 512×512 image is shown.

Now, we describe Algorithms 1 and 2: Let $Z = [Z_{m,n}]_{0 \leq m \leq M-1, 0 \leq n \leq N-1}$ be the original image, and let $X = [X_{m,n}]_{0 \leq m \leq M-1, 0 \leq n \leq N-1}$ where, for all $0 \leq m \leq M-1$, $0 \leq n \leq N-1$, $X_{n,m} = Z_{n,m} - \bar{Z}$ and \bar{Z} is the mean of Z . Consider the approximate image Y of Z based on a AR-2D process as in (3). Let $4 \leq k \leq \min(M, N)$. To simplify, from now on we shall consider that the images to be processed (Z and X) are arranged in such a way that the number of columns minus one and the number of rows minus one are multiples of $k-1$; i.e., $Z = [Z_{m,n}]_{0 \leq m \leq M'-1, 0 \leq n \leq N'-1}$ and $X = [X_{m,n}]_{0 \leq m \leq M'-1, 0 \leq n \leq N'-1}$ where $M' = \frac{M-1}{k-1}(k-1)+1$, $N' = \frac{N-1}{k-1}(k-1)+1$. For all $i_b = 1, \dots, \frac{M-1}{k-1}$, and for all $j_b = 1, \dots, \frac{N-1}{k-1}$, we defined the $(k-1) \times (k-1)$ block (i_b, j_b) of the image X by $B_X(i_b, j_b) = [X_{r,s}]_{(k-1)(i_b-1) \leq r \leq (k-1)i_b, (k-1)(j_b-1) \leq s \leq (k-1)j_b}$. The $M' \times N'$ approximated image \hat{X} of X is provided by the following algorithm:

Algorithm 1 Local approximation of images by using AR-2D processes

Require: Original image Z .

Ensure: Approximated image \hat{Z} of the original image Z

- 1: Define X as $X = Z - \bar{Z}$
- 2: Generate block $B_X(i_b, j_b)$
- 3: Compute the estimations $\hat{\phi}_1(i_b, j_b)$, $\hat{\phi}_2(i_b, j_b)$, $\hat{\phi}_3(i_b, j_b)$ of ϕ_1 , ϕ_2 and ϕ_3 corresponding to the block $B_X(i_b, j_b)$ extended to $B'_X(i_b, j_b) = [X_{r,s}]_{(k-1)(i_b-1) \leq r \leq (k-1)i_b, (k-1)(j_b-1) \leq s \leq (k-1)j_b}$
- 4: Define \hat{X} on the block $B_X(i_b, j_b)$ by

$$\hat{X}_{r,s} = \hat{\phi}_1(i_b, j_b)X_{r-1,s} + \hat{\phi}_2(i_b, j_b)X_{r,s-1} + \hat{\phi}_3(i_b, j_b)X_{r-1,s-1}$$

where $(k-1)(i_b-1)+1 \leq r \leq (k-1)i_b$ and $(k-1)(j_b-1)+1 \leq s \leq (k-1)j_b$

- 5: Define \hat{Z} as $\hat{Z} = \hat{X} - \bar{Z}$
-

Algorithm 2 Segmentation

Require: Original image Z

Ensure: Segmentated image W

- 1: Generate an approximated image \hat{Z} of Z with the Algorithm 1.
 - 2: Compute the residual image W defined as $W = Z - \hat{Z}$.
-

In the first experiment, Algorithm 1 was applied to image representation. We locally adjusted an AR-2D process to the original image for different window sizes, and estimated the parameters of the model with the BMM estimator. Fig. 9, (a), (b), (c) and (d) exhibits the BMM reconstructed images obtained respectively using the window sizes 8×8 , 16×16 , 32×32 and 57×57 . For all window sizes, the BMM reconstructed images are visually good; although a quantitative analysis of the similarity between each BMM reconstructed image and the original image showed differences. We calculated the SSIM, CQ (1,1) and CQ_{max} index, between each reconstructed image and the original image. The three indexes revealed that the similarity decreases as the size of the window increases (Table 1); so the best fits were obtained with small window sizes. This result reflects the assumption that the two-dimensional autoregressive model is a local adjustment model. Next, we applied Algorithm 2 and generated four difference images (e), (f), (g) and (h) shown in Fig. 9. We observed that the difference image (h) highlights the edges more than the others do. This shows that when we performed the reconstruction with a 57×57 window size, (Fig. 9 (d)), a lot of information got lost and this is reflected by the difference image (Fig. 9 (h)).



Figure 9: Image (I) Lena original image. (Right) The first row has the reconstructions done for BMM estimator adjusting window sizes 8×8 , 16×16 , 32×32 and 57×57 respectively ((a) to (d)). The second row has the respective differences ((e) to (h)) with respect to the original image (I).

(I)



(a)



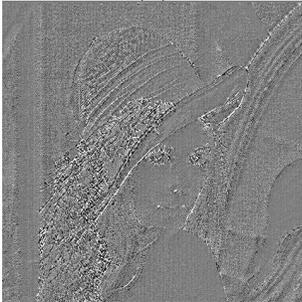
(b)



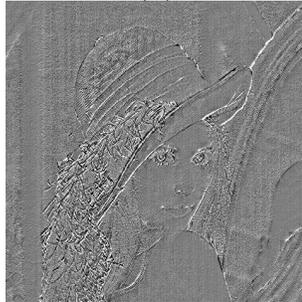
(c)



(d)



(e)



(f)



(g)



(h)

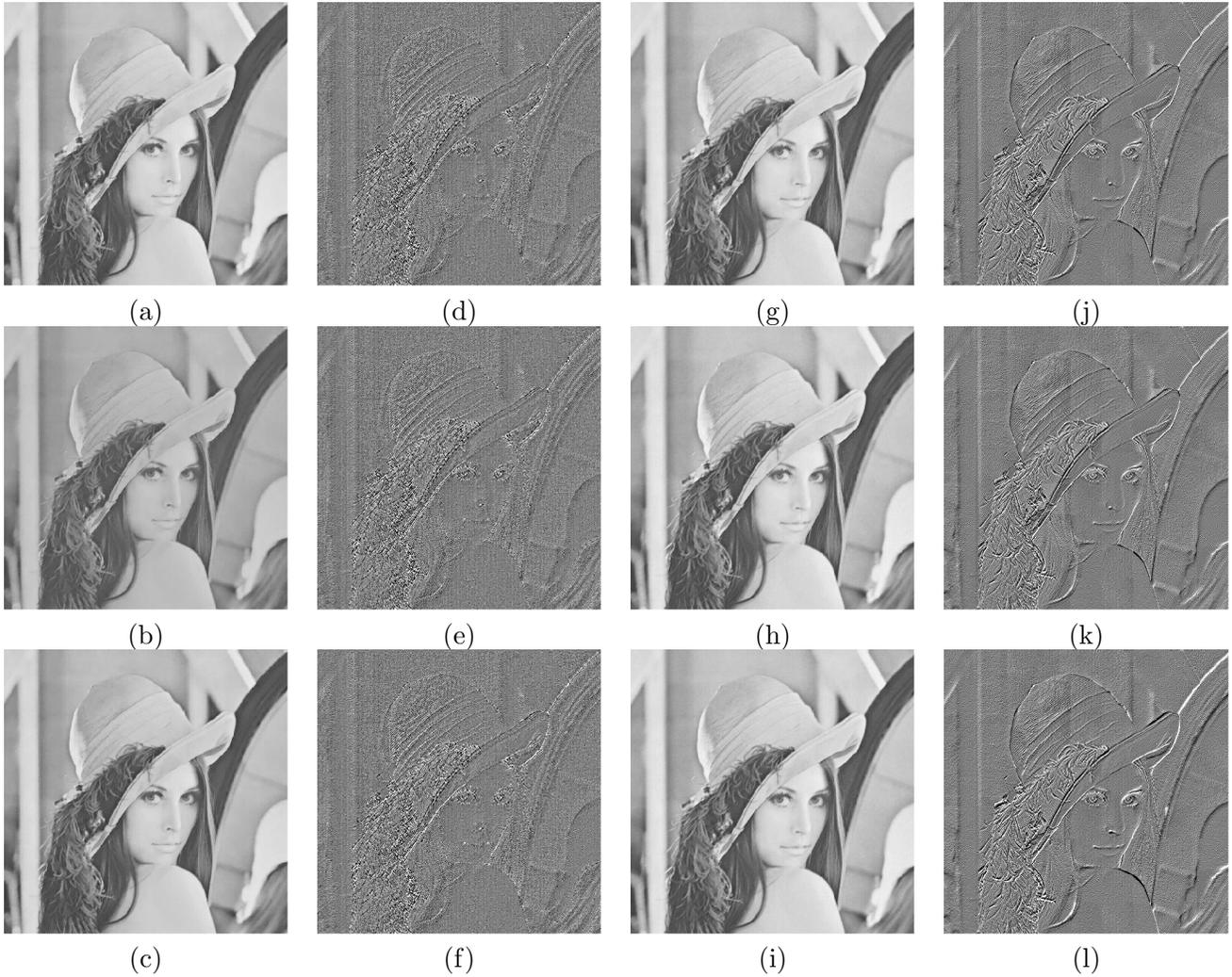
Window size	SSIM	CQ(1,1)	CQ _{max}
8×8	0.9948914	0.8582201	0.9706984
16×16	0.9827996	0.8309626	0.9544317
32×32	0.9779204	0.8151581	0.9462133
57×57	0.9762065	0.8073910	0.9423786

Table 1: SSIM, CQ and CQ_{max} index between the original image and each one of the BMM reconstructed images (a), (b), (c) and (d) in Figure 9.

In the second experiment, the original image was 10% additively contaminated (Fig. 10 (II)), and we used it as input in Algorithm 1. We obtained four reconstructed images using the LS, GM and the BMM estimators. Next, the Algorithm 2 was performed. The studies were carry out considering 8×8 , and 57×57 window sizes. In the first two columns in Figure 10, we can observe the results obtained considering 8×8 windows. Visually, there are not great differences between the different images reconstructed. When analyzing Table 2, we verify this because the measured indexes are comparable to each other. On the other hand, the third and fourth columns of Figure 10 show the results obtained by adjusting 57×57 windows. It is observed that the image (l), corresponding to the difference between the image restored with BMM (l) and the one contaminated with additive noise, highlights the edges slightly more.



Figure 10: Image (I) Lena original image; Image (II) image with 10% additive contamination and $\sigma^2 = 50$. In the first row, adjustments made with LS; in the second row, with GM; and in the third row, with BMM. Columns 1 and 2 correspond to adjustments with 8×8 windows, and columns 3 and 4 to 57×57 windows. Columns 1 and 3 are the reconstruction for Algorithm 1 and columns 2 and 4 are the segmented images for Algorithm 2.



Estimate	SSIM	CQ(1,1)	CQ _{max}
LS	0.9836079	0.8416351	0.9588826
GM	0.9390820	0.7821954	0.9103257
BMM	0.9846007	0.8328356	0.9577393

Table 2: Similarity between the original image and the reconstructions of Lena with additive contamination (Figure 9 - II).

6 Conclusions and discussions

The following comments give a brief summary of the results obtained in this paper.

A new estimator called BMM was proposed to estimate the parameters in first-order two-dimensional autoregressive models with three parameters. The new estimator is a two-dimensional extension of the BMM estimator proposed by Muler et al. [50] for autoregressive models of time series. We also extended the definition of replacement contamination, given for one-dimensional models ([46]), to the case of AR-2D models; this type of contamination includes additive-type contamination. The performance of the proposed estimator for AR-2D models with replacement contamination and without contamination was analyzed. Besides, the new estimator was compared with the classical least square estimator (LS) and robust estimators M, GM and RA. The comparative analysis was performed from six experiments, each of which involved several Monte Carlo studies, considering different replacement contamination patterns and varying the level of contamination. The LS estimator produced estimates that are very sensitive to the presence of atypical values, while the other estimators had better results. Using Monte Carlo simulation, we observed that the GM, RA and BMM estimators are highly superior than the M and LS estimators. However, the new estimator showed the best behavior, in both accuracy and precision, followed by the RA estimator in accuracy and by the GM in precision. An analysis of the computational cost showed that the RA estimator is the most expensive, followed by the BMM, GM, M and LS estimators, in that order. Finally, in an application to real data, we introduced a variant of the algorithm developed by [53], to perform image segmentation, using an AR-2D model with three parameters, and BMM estimators. In the light of the examples shown in Section 5, we conclude that the adapted algorithm is able to highlight the borders and contours in the images. The following proposals outline some directions for future work.

Under mild regularity conditions, in [52], the authors established the asymptotic normality and consistency of the robust RA estimators for the parameter ϕ of a two-dimensional autoregressive unilateral process. That work extended the asymptotic theory of the RA estimators, which was available only for time series (see Bustos et al. [16]). Although the estimators M, GM and BMM are reasonable to estimate parameter ϕ , their asymptotic behavior are still open problems. In this paper, the difference between a real image and an BMM approximated image was computed. The resulting image could be used to detect the borders and to classify the original image. It would be interesting to explore the limitations of a segmentation method based on the difference image between a real and a fitted image. It is also important to analyze the behavior of the BMM estimator in combination with image restoration techniques. The same relevance has the study of properties of BMM estimator, in particular, and robust estimators, in general, as alternatives to the least square estimators under non causal and semi causal AR-2D models.

References

- [1] ALLENDE, H., GALBIATI, J. AND VALLEJOS, R.. *Digital Image Restoration Using Autoregressive Time Series Type Models*. Proceedings of the 2nd. Latino-American Seminar on Radar Remote Sensing held at Santos, Sao Pablo, Brazil. ESA SP-434 (1998).
- [2] ALLENDE, H., GALBIATI, J. AND VALLEJOS, R.. *Robust image modelling on image processing*. Pattern Recognition Letters 22 (2001), pp. 1219-1231.
- [3] H. ALLENDE AND J. GALBIATI. *A non-parametric filter for digital image restoration, using cluster analysis*. Pattern Recognition Letters 25 (2004), pp. 841-847.
- [4] ALLENDE, H. AND PIZARRO, L.. *Robust estimation of roughness parameter in SAR amplitude images*. In: Sanfeliu, A., Ruiz-Shulcloper, J. (Eds.), CIARP, Lecture Notes in Computer Science, vol. 2905, (2003). Springer, Berlin.
- [5] AYSAL, T.C. AND BARNER, K.E.. *Quadratic weighted median filters for edge enhancement of noisy images*. IEEE Transactions on Image Processing 15 (2006), pp. 3294-3310.

- [6] BARAN, S., PAP, G., VAN ZUIJLEN, M.C.A.. *Asymptotic inference for a nearly unstable sequence of stationary spatial AR models*. Statistics and Probability Letters 69 (2004), pp. 53-61.
- [7] BASU, S. AND REINSEL, G.C.. *Properties of the spatial unilateral first-order ARMA model*. Advances in Applied Probability 25 (1993), pp. 631-648.
- [8] BEN HAMZA, A. AND KRIM, H.. *Image denoising: a nonlinear robust statistical approach*. IEEE Transactions on Signal Processing 49 (2001), pp. 3045-3054.
- [9] BESAG, J.. *Spatial interaction and the statistical analysis of lattice systems (with discussion)*. Journal of the Royal Statistical Society Series B 55 (1974), pp. 192-236.
- [10] BLACK, M.J. AND RANGARAJAN, A.. *On the unification of line processes, outlier rejection, and robust statistics with applications in early vision*. International Journal of Computer Vision 19 (1996), pp. 57-91.
- [11] BLACK, M.J., SAPIRO, G., MARIMONT, D. AND HEEGER, D.. *Robust anisotropic diffusion: connections between robust statistics, line processing, and anisotropic diffusion*. In: Proceedings of the First International Conference on Scale-Space Theory in Computer Vision. Lecture Notes in Computer Science, vol. 1252 (1997). Springer, Berlin, pp. 323-326.
- [12] BOUZOUBA, K. AND RADOUANE, L.. *Image identification and estimation using the maximum entropy principle*. Pattern Recognition Letters 21 (2000), pp. 691-700.
- [13] BRANDLE, N., BISCHOF, H. AND LAPP, H.. *Robust DNA microarray image analysis*. Machine Vision and Applications 15 (2003), pp. 11-28.
- [14] BRUNELLI, R. AND MESSELODI, S.. *Robust estimation of correlation with applications to computer vision*. Pattern Recognition 28 (1995), pp. 833-841.
- [15] O. BUSTOS. *Robust Statistics in SAR image processing*. ESA-SP 407 (1997), pp. 81-89.
- [16] BUSTOS, O. AND FRAIMAN, R.. *Asymptotic behavior of the estimates based on residual autocovariances for ARMA models*. In: Robust and Nonlinear Time Series. Lectures Notes in Statistics, vol. 29 (1984). Springer, New York, pp. 26-49.
- [17] BUSTOS, O.H., LUCINI, M.M. AND FRERY, A.C.. *M-estimators of roughness and scale for GA0-modelled SAR imagery*. EURASIP Journal on Applied Signal Processing 1 (2002), pp. 105-114.
- [18] BUSTOS, O., OJEDA, S., RUIZ, M., VALLEJOS, R. AND FRERY, A.. *Asymptotic Behavior of RA-estimates in Autoregressive 2D Gaussian Models*. Journal of Statistical Planning and Inference 139 (2009), pp. 3649-3664.
- [19] BUSTOS, O., OJEDA, S. AND VALLEJOS, R.. *Spatial ARMA models and its applications to image filtering*. Brazilian Journal of Probability and Statistics 23 (2009), pp. 141-165.
- [20] BUSTOS, O. AND YOHAI, V.. *Robust estimates for ARMA models*. Journal of the American Statistical Society 81 (1986), pp. 155-168.
- [21] BUTLER, N.. *A frequency domain comparison of estimation methods for Gaussian fields*. Communications in Statistics Theory and Methods 27 (1998), pp. 2325-2342.
- [22] CHEN, J.-H., CHEN, C.-S. AND CHEN, Y.-S.. *Fast algorithm for robust template matching with m-estimators*. IEEE Transactions on Signal Processing 51 (2003), pp. 230-243.
- [23] CHOI, B.. *On the asymptotic distribution of mean, autocovariance, autocorrelation, crosscovariance and impulse response estimators of a stationary multidimensional random field*. Communications in Statistics. Theory and Methods 29 (2000), pp. 1703-1724.

- [24] CHU, C.K., GLAD, I.K., GODTLIEBSEN, F. AND MARRON, J.S.. *Edge-preserving smoothers for image processing*. Journal of the American Statistical Association 93 (1988), pp. 526-541.
- [25] COMPORT, A., MARCHAND, E. AND CHAUMETTE, F.. *Statistically robust 2-d visual serving*. IEEE Transactions on Robotics 22 (2006), pp. 415-420.
- [26] CULLIS, B.R. AND GLESSON, A.C.. *Spatial analysis of field experiments an extension to two dimensions*. Biometrics 47 (1991), pp. 1449-1460.
- [27] L. DENBY and D. R. MARTIN, *Robust Estimation on the First Order Autoregressive Parameter*, Journal of the American Statistic Association. Vol. 74 pp. 140-146. (1979)
- [28] DRYDEN, I., IPPOLITI, L. AND ROMAGNOLI, L.. *Adjusted maximum likelihood and pseudo-likelihood estimation for noisy Gaussian Markov random fields*. Journal of Computational and Graphical Statistics 11 (2002), pp. 370-388.
- [29] FOX, A.J.. *Outliers in Time Series*. Journal of the Royal Statistical Society, Ser. B, 34 (1972), pp. 350-363.
- [30] GLENDINNING, R.H.. *Robust shape classification*. Signal Processing 77 (1999), pp. 121-138.
- [31] GOTTARDO, R., RAFTERY, A.E., YEUNG, K.Y. AND BUMGARNER, R.E.. *Bayesian robust inference for differential gene expression in microarrays with multiple samples*. Biometrics 62 (2006), pp. 10-18.
- [32] GRONDONA, M.R., CROSSA, J., FOX, P.N. AND PFEIFFER, W.H.. *Analysis of variety yield trials using two-dimensional separable ARIMA processes*. Biometrics 52 (1996), pp. 763-770.
- [33] GUO, J. AND BILLARD, L.. *Some inference results for causal autoregressive processes on a plane*. Journal of Time Series Analysis 19 (1998), pp. 681-691.
- [34] GUYON, X.. *Champs Aléatoires Sur un Réseau Modélisations*. In: Statistique et Applications (1993). Masson, Paris.
- [35] GUYON, X.. *Random fields on a network*. In: Modeling, Statistics and Applications (1995). Springer, Berlin.
- [36] HAINING, R.P.. *The moving average model for spatial interaction*. Transactions and Papers, Institute of British Geographers (N.S.) 3 (1978), pp. 202-225.
- [37] HASLER, D., SBAIZ, L., SUSSTRUNK, S. AND VETTERLI, M.. *Outlier modeling in image matching*. IEEE Transactions on Pattern Analysis and Machine Intelligence 25 (2003), pp. 301-315.
- [38] HUANG, H.-C. AND LEE, T.. *Data adaptive median filters for signal and image denoising using a generalized sure criterion*. IEEE Signal Processing Letters 13 (2006), pp. 561-564.
- [39] HUBER, P.. *Robust estimation of a location parameter*. Ann. Math. Statist. 53 (1964). MR0161415. pp. 73-101.
- [40] JAIN, A.K., MURTY, M.N. AND FLYNN, P.J.. *Data clustering: a review*. ACM Computing Surveys 31 (1999), pp. 264-323.
- [41] KASHYAP, R. AND EOM, K.. *Robust images techniques with an image restoration application*. IEEE Trans. Acoust. Speech Signal Process, Vol. 36 (1988), pp. 1313-1325.
- [42] KIM, J.H. AND HAN, J.H.. *Outlier correction from uncalibrated image sequence using the triangulation method*. Pattern Recognition 39 (2006), pp. 394-404.
- [43] KOIVUNEN, V.. *A robust nonlinear filter for image restoration*. IEEE Transactions on Image Processing 4 (1995), pp. 569-578.

- [44] LI, S.Z., WANG, H. AND SOH, W.Y.C.. *Robust estimation of rotation angles from image sequences using the annealing M-estimator*. Journal of Mathematical Imaging and Vision 8 (1998), pp. 181-192.
- [45] MARTIN, R.D.. *Robust estimation of autoregressive models*. In: Brillinger, D.R., Tiao, G.C. (Eds.), Direction in Time Series. Institute of Mathematical Statistics, Hayward, CA. (1980).
- [46] MARONNA R., DOUGLAS MARTIN R. AND VÌCTOR YOHAI. *Robust Statistics: Theory and Methods*. John Wiley & Sons. ISBN 0-470-01092-4 (2006).
- [47] MARROQUIN, J.L., SERVIN, M. AND RODRIGUEZ-VERA, R.. *Robust filters for low-level vision*. Expert Systems with Applications 14 (1998), pp. 169-177.
- [48] MEER, P., MINTZ, D., ROSENFELD, A. AND KIM, D.Y.. *Robust regression methods for computer vision: a review*. International Journal of Computer Vision 6 (1991), pp. 59-70.
- [49] MIRZA, M. AND BOYER, K.. *Performance evaluation of a class of M-estimators for surface parameter estimation in noisy range data*. IEEE Transactions on Robotics and Automation 9 (1993), pp. 75-85.
- [50] MULER, N., PEÑA, D. AND YOHAI V.. *Robust Estimation for ARMA models*. The Annals of Statistics, Vol. 37, No. 2 (2009), pp. 816-840.
- [51] NIREL, R., MUGGLESTONE, M.A. AND BARNETT, V.. *Outlier-robust spectral estimation for spatial lattice processes*. Communications in Statistics Theory and Methods 27 (1998), pp. 3095-3111.
- [52] OJEDA, S.. *Robust RA estimators for bidimensional autoregressive models*. Ph.D. dissertation, Facultad de Matemáticas, Astronomía y Física, Universidad Nacional de Córdoba, Argentina (1999).
- [53] S. OJEDA, R. VALLEJOS AND O. BUSTOS. *A new image segmentation algorithm with applications to image inpainting*. Computational Statistics and Data Analysis, Vol. 54 (2010), pp. 2082-2093.
- [54] OJEDA, S., VALLEJOS, R. AND LAMBERTI, P.W.. *Measure of similarity between images based on the codispersion coefficient*. J. Electron. Imaging 21, 023019 (2012).
- [55] OJEDA, S.M., VALLEJOS, R.O. AND LUCINI, M.. *Performance of RA estimator for bidimensional autoregressive models*. Journal of Statistical Simulation and Computation 72 (2002), pp. 47-62.
- [56] PALENICHKA, R.M., ZINTERHOF, P., RYTSAR, Y.B. AND IVASENKO, I.B.. *Structure-adaptive image filtering using order statistics*. Journal of Electronic Imaging 7 (1998), pp. 339-349.
- [57] PALENICHKA, R.M., ZINTERHOF, P. AND IVASENKO, I.B.. *Adaptive image filtering and segmentation using robust estimation of intensity*. Advances in Pattern Recognition 1876 (2000), pp. 888-897.
- [58] PISTONESSI S., MARTINEZ J., OJEDA SILVIA M. AND VALLEJOS RONNY O.. *A Novel Quality Image Fusion Assessment Based on Maximum Codispersion*. Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications. 20th Iberoamerican Congress, CIARP 2015 Montevideo, Uruguay (2015).
- [59] PRASTAWA, M., BULLITT, E., HO, S. AND GERIG, G.. *A brain tumor segmentation framework based on outlier detection*. Medical Image Analysis 8 (2004), pp. 275-283.
- [60] RABIE, T.. *Robust estimation approach for blind denoising*. IEEE Transactions on Image Processing 14 (2005), pp. 1755-1765.
- [61] ROTH, V.. *Kernel Fisher discriminants for outlier detection*. Neural Computation 18 (2006), pp. 942-960.
- [62] SIM, K.S., CHENG, Z. AND CHUAH, H.T.. *Robust image signal-to-noise ratio estimation using mixed Lagrange time delay estimation autoregressive model*. Scanning 26 (2004), pp. 287-295.

- [63] SINGH, M., ARORA, H. AND AHUJA, N.. *A robust probabilistic estimation framework for parametric image models*. In: Proceedings of the European Conference on Computer Vision- ECCV, vol. I (2004). Springer, Berlin, pp. 508-522.
- [64] STELAND, A.. *On the distribution of the clipping median under a mixture model*. Statistics and Probability Letters 71 (2005), pp. 1-13.
- [65] STEWART, C.V.. *Robust parameter estimation in computer vision*. SIAM Review 41 (1999), pp. 513-537.
- [66] TAREL, J.P., IENG, S.-S. AND CHARBONNIER, P.. *Using robust estimation algorithms for tracking explicit curves*. In: Proceedings of the European Conference on Computer Vision (2002). Springer, Berlin, pp. 492-507.
- [67] TJOSTHEIM, D.. *Statistical spatial series modelling*. Advances in Applied Probability 10 (1978), pp. 130-154.
- [68] TOHKA, J., ZIJDENBOS, A. AND EVANS, A.. *Fast and robust parameter estimation for statistical partial volume models in brain MRI*. Neuroimage 23 (2004), pp. 84-97.
- [69] TORR, P.H.S. AND ZISSERMAN, A.. *MLEMAC: a new robust estimator with application to estimating image geometry*. Computer Vision and Image Understanding 78 (2000), pp. 138-156.
- [70] VALLEJOS, R.. *Estimaciòn Robusta en Procesos Autorregresivos Bidimensionales*, Departamento de Matemàticas. Universidad Tècnica Federico Santa Marìa (1998).
- [71] VALLEJOS, R. AND GARCIA-DONATO, G.. *Bayesian analysis of contaminated quarter plane moving average models*. Journal of Statistical Computation and Simulation 76 (2006), pp. 131-147.
- [72] R. VALLEJOS AND T. MARDESIC, *A recursive algorithm to restore images based on robust estimation of NSHP autoregressive models*, Journal of Computational and Graphical Statistics, Vol. 13 (2004), pp. 674-682.
- [73] VAN DE WEIJER, J. AND VAN DEN BOOMGAARD, R.. *Least squares and robust estimation of local image structure*. International Journal of Computer Vision 64 (2005), pp. 143-155.
- [74] VOLOSHYNOVSKIY, S.V., ALLEN, A.R. AND HRYTSKIV, Z.D.. *Robust edge-preserving image restoration in the presence of non-Gaussian noise*. Electronics Letters 36 (2000), pp. 2006-2007.
- [75] WANG, Z. AND BOVIK, C.. *A universal image quality index*. IEEE Signal Process. Lett. 9 (2002), 81-84.
- [76] WANG, H.Z. AND SUTER, D.. *MDPE: a very robust estimator for model fitting and range image segmentation*. International Journal of Computer Vision 59 (2004), pp. 139-166.
- [77] WANG, H.Z., SUTER, D.. *Robust adaptive-scale parametric model estimation for computer vision*. IEEE Transactions on Pattern Analysis and Machine Intelligence 26 (2004), pp. 1459-1474.
- [78] P. WHITTLE, *On stationary processes on the plane*, Biometrika, Vol. 41 (1954), pp. 434-449.
- [79] XU, P.L.. *Sign-constrained robust least squares, subjective breakdown point and the effect of weights of observations on robustness*. Journal of Geodesy 79 (2005), pp. 146-159.
- [80] V. J. YOHAI. *High breakdown point and high efficiency estimates for regression*, Ann. Statist., Vol. 15 (1987), pp. 642-665. MR0888431.
- [81] ZERVAKIS, M.E. AND KWON, T.M.. *Robust estimation techniques in regularized image-restoration*. Optical Engineering 31 (1992), pp. 2174-2190.
- [82] ZERVAKIS, M.E., KATSAGGELOS, A.K. AND KWON, T.M.. *A class of robust entropic functionals for image restoration*. IEEE Transactions on Image Processing 4 (1995), pp. 752-773.

Appendix Tables

Table 3: Estimate of $\phi_1 = 0.15$, $\phi_2 = 0.17$ and $\phi_3 = 0.2$ in an AR-2D model without contamination.

N		ϕ_1						ϕ_2						ϕ_3					
		LS	M	GM	RA	BMM-2D		LS	M	GM	RA	BMM-2D		LS	M	GM	RA	BMM-2D	
8	$\hat{\phi}_i$	0.14642	0.14638	0.14758	0.14797	0.15260		0.16328	0.16474	0.16422	0.16959	0.16968		0.18854	0.18636	0.18907	0.19012	0.19083	
	mse_{ϕ_i}	0.01832	0.01955	0.02082	0.02509	0.02789		0.01658	0.01871	0.01914	0.02366	0.02766		0.02046	0.02252	0.02255	0.03273	0.03233	
	$\hat{\sigma}_i^2$	0.01835	0.01957	0.02085	0.02513	0.02794		0.01657	0.01872	0.01915	0.02370	0.02771		0.02037	0.02238	0.02247	0.03270	0.03231	
16	$\hat{\phi}_i$	0.14998	0.15056	0.15081	0.15153	0.15390		0.16794	0.16598	0.16706	0.16771	0.16953		0.19751	0.19826	0.19811	0.19560	0.20218	
	mse_{ϕ_i}	0.00430	0.00502	0.00498	0.00525	0.00586		0.00408	0.00458	0.00489	0.00497	0.00512		0.00453	0.00485	0.00520	0.00589	0.00537	
	$\hat{\sigma}_i^2$	0.00430	0.00503	0.00499	0.00526	0.00586		0.00408	0.00457	0.00490	0.00497	0.00513		0.00454	0.00485	0.00521	0.00589	0.00537	
32	$\hat{\phi}_i$	0.15014	0.15025	0.14993	0.15053	0.15127		0.16881	0.16884	0.16906	0.16930	0.17019		0.19897	0.19842	0.19851	0.19828	0.19945	
	mse_{ϕ_i}	0.00099	0.00108	0.00110	0.00113	0.00115		0.00099	0.00106	0.00114	0.00111	0.00108		0.00110	0.00120	0.00123	0.00128	0.00124	
	$\hat{\sigma}_i^2$	0.00099	0.00108	0.00110	0.00113	0.00115		0.00099	0.00106	0.00114	0.00111	0.00109		0.00110	0.00120	0.00123	0.00128	0.00124	
57	$\hat{\phi}_i$	0.14937	0.14937	0.14951	0.14942	0.14956		0.16864	0.16840	0.16845	0.16851	0.16858		0.19934	0.19971	0.19962	0.19944	0.19944	
	mse_{ϕ_i}	0.00031	0.00032	0.00033	0.00033	0.00033		0.00030	0.00031	0.00032	0.00033	0.00031		0.00031	0.00033	0.00034	0.00035	0.00033	
	$\hat{\sigma}_i^2$	0.00031	0.00032	0.00033	0.00033	0.00033		0.00030	0.00031	0.00032	0.00033	0.00031		0.00031	0.00033	0.00034	0.00035	0.00033	

Table 4: Estimate of $\phi_1 = 0.15$, $\phi_2 = 0.17$ and $\phi_3 = 0.2$ in an AR-2D model with additive contamination of $\mu = 0$ and $\sigma^2 = 50$.

N		ϕ_1						ϕ_2						ϕ_3					
		LS	M	GM	RA	BMM-2D		LS	M	GM	RA	BMM-2D		LS	M	GM	RA	BMM-2D	
8	$\hat{\phi}_i$	0.04048	0.05018	0.11614	0.12502	0.10983		0.04915	0.04851	0.12014	0.12216	0.11082		0.06266	0.06973	0.15827	0.12456	0.11873	
	mse_{ϕ_i}	0.03170	0.02101	0.02208	0.06115	0.02689		0.03560	0.02513	0.02582	0.06229	0.02848		0.04177	0.02690	0.02690	0.05767	0.02989	
	$\hat{\sigma}_i^2$	0.01975	0.01107	0.02098	0.06065	0.02533		0.02104	0.01039	0.02338	0.06012	0.02503		0.02296	0.00995	0.02521	0.05208	0.02333	
16	$\hat{\phi}_i$	0.03640	0.03670	0.10622	0.17498	0.12390		0.03880	0.03902	0.11819	0.19017	0.13301		0.05228	0.05043	0.15777	0.16202	0.13817	
	mse_{ϕ_i}	0.01750	0.01434	0.00669	0.01389	0.00729		0.02094	0.01872	0.00731	0.01249	0.00884		0.02581	0.02413	0.00677	0.01153	0.01208	
	$\hat{\sigma}_i^2$	0.00460	0.00151	0.00478	0.01329	0.00662		0.00374	0.00157	0.00463	0.01211	0.00749		0.00399	0.00176	0.00499	0.01010	0.00827	
32	$\hat{\phi}_i$	0.03426	0.03537	0.11038	0.19341	0.13797		0.03831	0.04006	0.12820	0.21266	0.16113		0.04690	0.04852	0.15217	0.15033	0.16094	
	mse_{ϕ_i}	0.01446	0.01344	0.00259	0.00522	0.00180		0.01843	0.01721	0.00291	0.00588	0.00180		0.02460	0.02329	0.00352	0.00475	0.00354	
	$\hat{\sigma}_i^2$	0.00107	0.00030	0.00103	0.00335	0.00166		0.00110	0.00033	0.00116	0.00407	0.00173		0.00117	0.00034	0.00124	0.00229	0.00202	
57	$\hat{\phi}_i$	0.03397	0.03567	0.11288	0.20015	0.14520		0.03860	0.03888	0.12653	0.21397	0.16261		0.04609	0.04812	0.15209	0.14793	0.17444	
	mse_{ϕ_i}	0.01378	0.01316	0.00172	0.00386	0.00049		0.01759	0.01729	0.00222	0.00332	0.00057		0.02401	0.02317	0.00266	0.00357	0.00125	
	$\hat{\sigma}_i^2$	0.00032	0.00009	0.00034	0.00135	0.00047		0.00032	0.00010	0.00033	0.00139	0.00052		0.00032	0.00010	0.00037	0.00086	0.00060	

Table 5: Estimate of $\phi_1 = 0.15$, $\phi_2 = 0.17$ and $\phi_3 = 0.2$ in an AR-2D model with additive contamination, window size 32×32 .

%		ϕ_1				ϕ_2				ϕ_3						
		LS	M	GM	RA	BMM-2D	LS	M	GM	RA	BMM-2D	LS	M	GM	RA	BMM-2D
5%	$\hat{\phi}_i$	0.05354	0.05835	0.12624	0.17486	0.14533	0.06474	0.06533	0.14652	0.19550	0.16565	0.07732	0.08454	0.17818	0.17824	0.18322
	mse_{ϕ_i}	0.01035	0.00893	0.00164	0.00235	0.00139	0.012162	0.01153	0.00173	0.00245	0.00155	0.01613	0.01408	0.00164	0.00209	0.00188
	$\hat{\sigma}_i^2$	0.00104	0.00053	0.00108	0.00174	0.00137	0.00108	0.00058	0.00119	0.00181	0.00153	0.00108	0.00075	0.00116	0.00162	0.00160
10%	$\hat{\phi}_i$	0.03426	0.03537	0.11038	0.19341	0.13797	0.03831	0.04006	0.12820	0.21266	0.16113	0.04690	0.04852	0.15217	0.15033	0.16094
	mse_{ϕ_i}	0.01446	0.01344	0.00259	0.00522	0.00180	0.01843	0.01721	0.00291	0.00588	0.00180	0.02460	0.02329	0.00352	0.00475	0.00354
	$\hat{\sigma}_i^2$	0.00107	0.00030	0.00103	0.00335	0.00166	0.00110	0.00033	0.00116	0.00407	0.00173	0.00117	0.00034	0.00124	0.00229	0.00202
15%	$\hat{\phi}_i$	0.02257	0.02493	0.09578	0.19227	0.13160	0.02793	0.02815	0.10832	0.20706	0.14783	0.03318	0.03389	0.13131	0.12771	0.14057
	mse_{ϕ_i}	0.01729	0.01585	0.00397	0.00606	0.00232	0.02115	0.02035	0.00484	0.00628	0.00249	0.02881	0.02780	0.00582	0.00782	0.00553
	$\hat{\sigma}_i^2$	0.00106	0.00021	0.00103	0.00428	0.00198	0.00096	0.00023	0.00103	0.00492	0.00200	0.00099	0.00021	0.00110	0.00259	0.00200
20%	$\hat{\phi}_i$	0.01766	0.01886	0.08255	0.16534	0.12508	0.02144	0.02021	0.09349	0.17371	0.13534	0.02511	0.02647	0.11064	0.11122	0.12291
	mse_{ϕ_i}	0.01852	0.01737	0.00580	0.00612	0.00273	0.02306	0.02259	0.00704	0.00566	0.00336	0.03155	0.03031	0.00919	0.01082	0.00820
	$\hat{\sigma}_i^2$	0.00101	0.00017	0.00125	0.00589	0.00211	0.00100	0.00016	0.00118	0.00566	0.00216	0.00096	0.00020	0.00121	0.00295	0.00226

Table 6: Estimate of $\phi_1 = 0.15$, $\phi_2 = 0.17$ and $\phi_3 = 0.2$ in an AR-2D model with t-student contamination with 2.3 d.f., window size 57.

%		ϕ_1				ϕ_2				ϕ_3						
		LS	M	GM	RA	BMM-2D	LS	M	GM	RA	BMM-2D	LS	M	GM	RA	BMM-2D
5%	$\hat{\phi}_i$	0.12843	0.13359	0.14418	0.14803	0.14869	0.14606	0.15290	0.16332	0.16762	0.16841	0.17005	0.18190	0.19197	0.19234	0.19414
	mse_{ϕ_i}	0.00099	0.00063	0.00038	0.00035	0.00044	0.00113	0.00064	0.00036	0.00033	0.00044	0.00157	0.00070	0.00042	0.00044	0.00047
	$\hat{\sigma}_i^2$	0.00053	0.00036	0.00035	0.00035	0.00043	0.00055	0.00035	0.00032	0.00033	0.00043	0.00067	0.00037	0.00036	0.00038	0.00044
10%	$\hat{\phi}_i$	0.11243	0.11936	0.13663	0.14450	0.14653	0.12495	0.13453	0.15399	0.16093	0.16367	0.15040	0.16673	0.18506	0.18625	0.19089
	mse_{ϕ_i}	0.00193	0.00133	0.00056	0.00042	0.00048	0.00271	0.00169	0.00060	0.00042	0.00049	0.00322	0.00159	0.00059	0.00057	0.00056
	$\hat{\sigma}_i^2$	0.00052	0.00039	0.00038	0.00039	0.00047	0.00068	0.00043	0.00035	0.00034	0.00045	0.00076	0.00048	0.00037	0.00038	0.00047
15%	$\hat{\phi}_i$	0.09909	0.10950	0.13231	0.14231	0.14422	0.11271	0.12404	0.14860	0.15876	0.16274	0.13272	0.14937	0.17524	0.17647	0.18357
	mse_{ϕ_i}	0.00318	0.00207	0.00064	0.00041	0.00050	0.00396	0.00257	0.00081	0.00049	0.00058	0.00530	0.00300	0.00096	0.00092	0.00075
	$\hat{\sigma}_i^2$	0.00059	0.00043	0.00033	0.00035	0.00047	0.00068	0.00046	0.00035	0.00037	0.00053	0.00078	0.00044	0.00035	0.00037	0.00049
20%	$\hat{\phi}_i$	0.08895	0.09830	0.12530	0.13771	0.14167	0.09962	0.11246	0.14363	0.15533	0.16191	0.11802	0.13740	0.16826	0.16956	0.18116
	mse_{ϕ_i}	0.00426	0.00308	0.00095	0.00052	0.00052	0.00556	0.00370	0.00099	0.00053	0.00042	0.00753	0.00434	0.00137	0.00132	0.00087
	$\hat{\sigma}_i^2$	0.00053	0.00040	0.00034	0.00037	0.00045	0.00061	0.00039	0.00029	0.00031	0.00036	0.00081	0.00042	0.00037	0.00040	0.00052

Table 7: Estimate of $\phi_1 = 0.15$, $\phi_2 = 0.17$ and $\phi_3 = 0.2$ in an AR-2D model with replacement contamination by another AR process of parameters $\tilde{\phi}_1 = 0.1$, $\tilde{\phi}_2 = 0.2$ and $\tilde{\phi}_3 = 0.3$, window size 32.

%		ϕ_1						ϕ_2						ϕ_3					
		LS	M	GM	RA	BMM-2D	LS	M	GM	RA	BMM-2D	LS	M	GM	RA	BMM-2D			
5%	$\hat{\phi}_i$	0.13738	0.13788	0.13763	0.13795	0.13951	0.15559	0.15667	0.15660	0.15707	0.15873	0.18439	0.18481	0.18530	0.18601	0.18667			
	mse_{ϕ_i}	0.00115	0.00119	0.00120	0.00125	0.00116	0.00123	0.00125	0.00133	0.00128	0.00127	0.00126	0.00131	0.00140	0.00138	0.00128			
	$\hat{\sigma}_i^2$	0.00099	0.00104	0.00105	0.00111	0.00105	0.00102	0.00108	0.00115	0.00112	0.00114	0.00102	0.00108	0.00119	0.00118	0.00111			
	$\hat{\phi}_i$	0.12977	0.13134	0.13044	0.13186	0.13395	0.14436	0.14571	0.14477	0.14603	0.14756	0.17055	0.17244	0.17132	0.17310	0.17487			
10%	$\hat{\phi}_i$	0.00136	0.00140	0.00145	0.00139	0.00136	0.00166	0.00165	0.00173	0.00170	0.00160	0.00192	0.00191	0.00199	0.00192	0.00187			
	mse_{ϕ_i}	0.00096	0.00105	0.00107	0.00106	0.00111	0.00100	0.00107	0.00109	0.00113	0.00110	0.00106	0.00116	0.00120	0.00124				
	$\hat{\phi}_i$	0.11879	0.12068	0.11975	0.12008	0.12333	0.13053	0.13242	0.13282	0.13331	0.13515	0.15953	0.16200	0.16142	0.16245				
	$\hat{\phi}_i$	0.00197	0.00194	0.00203	0.00200	0.00188	0.00256	0.00249	0.00251	0.00247	0.00234	0.00255	0.00245	0.00249	0.00250	0.00236			
15%	$\hat{\phi}_i$	0.00100	0.00108	0.00112	0.00111	0.00117	0.00100	0.00108	0.00113	0.00113	0.00113	0.00091	0.00101	0.00100	0.00109	0.00110			
	$\hat{\phi}_i$	0.11002	0.11141	0.11104	0.11190	0.11384	0.11989	0.12046	0.11980	0.12060	0.12338	0.15032	0.15283	0.15146	0.15289	0.15556			
	mse_{ϕ_i}	0.00257	0.00251	0.00253	0.00250	0.00238	0.00356	0.00362	0.00369	0.00364	0.00344	0.00344	0.00337	0.00348	0.00344	0.00317			
	$\hat{\phi}_i$	0.00097	0.00102	0.00101	0.00105	0.00107	0.00105	0.00117	0.00117	0.00120	0.00127	0.00097	0.00115	0.00113	0.00123	0.00120			

Table 8: Estimate of $\phi_1 = 0.15$, $\phi_2 = 0.17$ and $\phi_3 = 0.2$ in an AR-2D model with replacement contamination by a white noise process of variance 50, window size 32.

%		ϕ_1						ϕ_2						ϕ_3					
		LS	M	GM	RA	BMM-2D	LS	M	GM	RA	BMM-2D	LS	M	GM	RA	BMM-2D			
5%	$\hat{\phi}_i$	0.05141	0.05608	0.12961	0.18022	0.14770	0.05960	0.06437	0.14453	0.19548	0.16614	0.06973	0.07972	0.17067	0.17307	0.17888			
	mse_{ϕ_i}	0.01085	0.00936	0.00152	0.00282	0.00141	0.01315	0.01174	0.00167	0.00230	0.00131	0.01828	0.01528	0.00203	0.00249	0.00205			
	$\hat{\sigma}_i^2$	0.00113	0.00054	0.00110	0.00191	0.00141	0.00097	0.00058	0.00103	0.00166	0.00130	0.00132	0.00082	0.00117	0.00177	0.00161			
	$\hat{\phi}_i$	0.02990	0.03274	0.10976	0.19717	0.13843	0.03144	0.03431	0.12180	0.20953	0.15686	0.03895	0.04424	0.14775	0.14994	0.16107			
10%	$\hat{\phi}_i$	0.01550	0.01409	0.00274	0.00560	0.00178	0.02037	0.01871	0.00349	0.00467	0.00182	0.02701	0.02462	0.00392	0.00485	0.00347			
	mse_{ϕ_i}	0.00108	0.00034	0.00112	0.00338	0.00165	0.00117	0.00030	0.00116	0.00311	0.00165	0.00107	0.00036	0.00119	0.00235	0.00196			
	$\hat{\phi}_i$	0.01994	0.02190	0.09124	0.19055	0.12807	0.01905	0.02397	0.10452	0.20163	0.14848	0.02439	0.02998	0.12630	0.12875	0.14057			
	$\hat{\phi}_i$	0.01796	0.01666	0.00453	0.00617	0.00237	0.02382	0.02154	0.00540	0.00536	0.00248	0.03185	0.02916	0.00673	0.00769	0.00576			
15%	$\hat{\phi}_i$	0.00104	0.00025	0.00108	0.00454	0.00190	0.00103	0.00022	0.00112	0.00436	0.00202	0.00101	0.00025	0.00130	0.00262	0.00224			
	$\hat{\phi}_i$	0.01369	0.01590	0.07760	0.16684	0.12467	0.01536	0.01802	0.08600	0.17541	0.13500	0.01779	0.02225	0.10617	0.10972	0.12265			
	mse_{ϕ_i}	0.01970	0.01816	0.00627	0.00464	0.00308	0.02498	0.02326	0.00824	0.00463	0.00345	0.03430	0.03180	0.00991	0.01091	0.00807			
	$\hat{\phi}_i$	0.00112	0.00018	0.00103	0.00436	0.00245	0.00106	0.00016	0.00119	0.00461	0.00223	0.00110	0.00021	0.00111	0.00277	0.00209			