République algérienne démocratique et populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université Mouloud MAMMERI - Tizi-Ouzou

Faculté des Sciences
Département de Mathématiques

Thèse de doctorat 3ème cycle en Mathématiques

Spécialité: Probabilités & Statistique

Estimation et Prédiction des Processus Spatiaux

Présentée par : Amel SAIDI

Jury de soutenance:

Président:	**Farida** ACHEMINE	**Maître de Conférences A, Univ. Tizi-Ouzou, Algérie**
Rapporteur :	**Abdelghani** HAMAZ	**Maître de Conférences A, Univ. Tizi-Ouzou, Algérie**
Co-Rapporteur :	**Fazia** BEDOUHENE	**Professeur, Univ. Tizi-Ouzou, Algérie**
Examinateurs :	**Abdelouahab** BIBI,	**Professeur, Univ. Oum El Bouaghi, Algérie**
	Nabil ZOUGAB	**Professeur, Univ. Bejaia, Algérie**
	Aicha BARECHE	**Professeur, Univ. Bejaia, Algérie**
Invité :	**Mohamed** IBAZIZEN,	**Maître de Conférences, Univ. Poitier, France**

Soutenue le : 21 décembre 2022

Acknowledgments

I am grateful to my advisor Dr Abdelghani Hamaz for his guidance and cooperation during my years as a post graduate student.

It has been a pleasure to get a glimpse of the field of specials processes by his side. This thesis would not have seen the light without his help.

A special thanks also goes to my co-advisor Pr Fazia Bedouhene who has supported me throughout this research project, her affection and caring behaviour was never limited to a guide.

I'm thankful to Dr Ouardia Arezki for her sincere involvement in this thesis, her advices and availability.

I would like to thank Dr Farida Achemine, Pr Aicha Bareche, Pr Abdelouahab Bibi, Pr Nabil Zougab and Dr Mohamed Ibazizen, who have agreed to be part of my thesis comity, it is a real pleasure to feel their interest in this work.

A special appreciation to every one who had assisted me one way or another.

I cannot forget to thank my family for all the unconditional support... all words cannot express.

Contents

List of Tables

List of Algorithms

List of Figures

List of publications

- **International Journals**

    1. **Saidi, A.**, Hamaz, A., & Arezki, O. (2022). Estimation in nonlinear random fields models of autoregressive type with random parameters. Communications in Statistics-Theory and Methods, 1-16.
    2. **Saidi, A.**, Hamaz, A., & M. Ibazizen (2022). Measure of predictability of a stationary two dimensionally indexed autoregressive moving-average models, Communications in Statistics - Simulation and Computation, 1-16.

- **National Conferences**

    1. **Saidi, A.**, Arezki & O. Hamaz, A. (2021). Estimation of two dimensionally indexed Random Coefficient Autoregressive models, Première Conférences des Mathématiques Pures et Appliqués en ligne (CNMPA).

Introduction

The analysis of spatial processes has received much attention in recent years and has been studied in disciplines such as geography, geology, biology, and agriculture [1], [2], [3], [4]. The theory of two-dimensional (2D) systems has been developed during the past decades [5–9]. The main advantages of 2D ARMA models are to offer the possibilities: firstly, the use of the numerous algorithms of estimation parameters of models whose implementation is relatively easy. Secondly, the ability to realize the recursive processes following a lexicographic path, which is associated with any quarter plane (QP). Indeed, certain applications of these models such as in engineering problems, image restoration, texture analysis, image encoding, identification, and systems modeling are given in [10] and references therein.

2D system identification and parametric representations of 2D stationary random fields based on parametric 2D AR and ARMA models have received great attention in a wide range of image and signal processing applications. These applications include image restoration, image compression, stochastic texture analysis and synthesis, and modeling of 2D data. Also, some problems such as image restoration, pattern recognition, and texture analysis can be converted to the parameter estimation of 2D series. It is also stated the characterization of 2D, real-valued, discrete, and homogeneous moving average (MA) random fields is a fundamental problem that has been encountered in a wide range of signal and image processing areas. For example, this issue is important, especially in the processing of natural texture images from the perspective of analysis/synthesis and segmentation purposes. It is verified thoroughly in [11] that a texture field is considered as a realization of a regular homogeneous random field with a mixed spectral distribution, and is decomposed into a sum of two mutually orthogonal components called a deterministic component and a purely indeterministic component, by using the 2D Wold-decomposition.

Remarkably little attention has been paid to methods for measuring predictability, whereas predictability provides a succinct measure of a key aspect of series dynamics and is therefore useful for summarizing the behavior of series, as well as for assessing agreement between models and data. So, it is natural and informative to judge forecasts by their accuracy. The extent of a series' predictability depends, obviously, on how much information the past conveys regarding future values of this series; as a result, some processes are inherently easy to forecast, and others are more difficult. Perhaps the most used measure of predictability is the mean square error of a "perfect" forecast model. This criterion was used to quantify the quality of predictors of stationary random fields when the third quadrant is chosen as past support; see [12, 13].

Typically, the mean square error increases with lead time and asymptotically approaches a finite value, called the saturation value. The saturation value is comparable to the mean square difference between two randomly chosen fields from the system. Consequently, all predictability is said to be lost when the errors are comparable to the saturation value since the forecast offers no better prediction than a randomly chosen field from the system. Despite its widespread use, the mean square error turns out to be a limited measure of predictability.

Some models formulated in real-life applications are nonlinear, their linear approximations can lose a good deal of information, so the forecasts may suffer from inaccuracy. Indeed, the variability of a process may well depend on the available information. In recent years, interest in nonlinear models for spatial processes has grown slowly [14], [15].

This reality has motivated extensive research to relax the assumption of a constant variance imposed by the traditional linear models. Among the proposed models, we focus on autoregressive models with coefficients assumed to be not constant but subjected to random perturbations, called random models of Autoregressive coefficients (RCAR). These models have special appeal among the non-linear models so far considered in the statistical literature. RCAR models are a natural generalization of constant autoregressive coefficient models AR, where the coefficients vary over time. Indeed, these models are defined by allowing random additive perturbations of the AR coefficients of ordinary AR models. They were introduced in the one-dimensional case (1D-RCAR) by several authors [16]; [17]; [18] and continue to gain interest [19]; [20]; [21]; [22]; [23].

More recently, attention has been oriented to 2D-RCAR [14]. However, research activities in this area are still weak and many important problems remain to be explored such as the estimation of 2D-RCAR model parameters. Indeed, very little is known about the theoretical properties of these procedures and the resulting estimates. An interesting characteristic of random coefficient autoregressive models is that their second-order properties are similar to those of conditional heteroscedastic models 2D-(G)ARCH. In fact, the 2D-(G)ARCH processes can be regarded as special cases of the 2D-RCA and these models play a very important role in spatial econometrics [24]. They can be successfully used with financial data, which are characterized by fat tails and volatility clustering. To explore the relationship between these two types of models [25]; [26].

Thesis Contributions

In this thesis, we present new original results on estimation and prediction in random field models.

1. The first contribution is derived in Chapter 2, we focus on two-dimensionally indexed random coefficients autoregressive models with order 2D-RCAR(p_1, p_2) for short. We first develop a maximum likelihood estimation procedure for estimating the unknown parameters of 2D-RCAR(p_1, p_2). Moreover, we prove that the estimates are strongly consistent. Finally, these results are then applied to construct efficient estimates in 2D-RCAR model of order $(0, 1)$. Then, a simulation part is given to illustrate the effectiveness and accuracy of the estimates

2. The second contribution is derived in Chapter 3, we consider a measure of predictability of a random field following a stationary two-dimensionally indexed autoregressive moving average (2D ARMA). We give a characterization of equal predictability processes by providing necessary and sufficient conditions, which highlight the role of the coefficients of the moving average (MA) and the equivalent autoregressive (EAR) representations.

Thesis Organization

This thesis is divided according to the objectives stated above and is therefore structured into three chapters.

- **In Chapter** 1, we have presented a certain number of preliminaries needed in the comprehension of the manuscript. In Section 1.2, spatial ARMA models are considered using the nonsymmetric half plane and quarter plane ordering on a lattice of data identifiability conditions and stationarity conditions based in [27]. Then, the correlation properties for the first-order ARMA model are discussed, and simple conditions for stationarity of the process are also borrowed from [28]. In Section 1.3 Definition of the $2D - RCAR$ processes and its $\mathbb{L}_2$-structure are given. To the best of our knowledge, the first study of 2D-RCAR models has been well established by (Bibi (2016) [14]. The basic definitions and results of the second-order properties presented in this section are borrowed from (Bibi (2016) [14].

  Chapter 2 is organized as follows: In section 2.2, we recall some estimation methods of spatial ARMA models: Yule-Walker Estimator for the First-Order Spatial Bilateral AR Model and Exact likelihood function and maximum likelihood estimation methods for the first-order spatial ARMA model. In section 2.3, we present new original theoretical results on estimation in nonlinear random field models. We focus on two-dimensionally indexed random coefficients autoregressive models with order $(p_1, p_2) \in \mathbb{N}^2, 2D - RCAR(p_1, p_2)$ for short. We first develop a maximum likelihood estimation procedure for estimating the unknown parameters of $2D - RCAR(p_1, p_2)$. Moreover, we prove that the estimates are strongly consistent. In Section 2.4, we focus on developing two parameter estimation procedures for first-order two-dimensionally indexed random coefficients autoregressive models (2D-RCAR) of the first order particularly. First, the model is explained in detail, and the definition and the condition of stationarity are given. Then, we provide a simulation study to show the accuracy of the maximum likelihood estimates and we provide a comparison with the least squares estimates.

- **Chapter 3** is organized as follows: In Section 3.2, we have collected some basic results with an emphasis on the condition for the existence of a moving average and equivalent autoregressive representations. In section 3.3, we introduce a measure of predictability of a quarter-plane (QP) two-dimensional autoregressive moving-average (2D ARMA) model excited by an unknown zero mean white noise. We clarify the link between the concept of parallelism and equal predictability at all horizons. A new result is proposed for determining the test of equal predictability. The test is effected using a Wald statistic, which follows asymptotically a central chi-squared distribution under the null hypothesis so that its application is very easy. To illustrate the applicability of our results, in section 3.4, we present a simulation study and an application on wheat yield data.

Chapter 1

Preliminaries on Spatial Processes

Contents

1.1 Introduction

The purpose of this chapter is to describe and present the main classical results on spatial statistics which thus facilitate the reading of the thesis. Some of these results and notions will be used in the remain of this manuscript.
In Section 1.2, spatial ARMA models are considered using the nonsymmetric half plane and quarter plane ordering on a lattice of data identifiability conditions and stationarity conditions based in [27]. Then, the correlation properties for the first-order ARMA model are discussed, and simple conditions for stationarity of the process are also borrowed from [28]. In Section 1.3, definition of the $2D - RCAR$ processes and its $\mathbb{L}_2$-structure are given. To the best of our knowledge, the first study of 2D-RCAR models has been well established by (Bibi (2016) [14]. The basic definitions and results of the second-order properties presented in this section are borrowed from (Bibi (2016) [14].

1.2 Definitions and Structural Properties of 2D ARMA Models

An d-dimensional random field is a collection of random variables $X(\mathbf{t})$ on a probability space $(\Omega, \mathcal{A}, P)$, where $\mathbf{t}$ varies in the d-dimensional Euclidean Space $\mathbb{R}^d$. We are interested in two-dimensional discrete random fields, that is, when $\mathbf{t} \in \mathbb{Z}^2, \mathbb{Z}$ being the set of all integers. In a sense, a spatial series may be considered as a generalization of a time series; however, its own intrinsic characteristics make its analysis considerably more difficult. The key difference is that a time series is unidirectional with a natural ordering from past to future; on the other hand, this ordering does not exist for a general spatial series. This difficulty and the huge amount of data to process motivate the consideration of unilateral models for spatial series. Two main types of unilateral representations have been studied in the literature: nonsymmetric half-plane (NSHP) representation and quarter-plane (QP) representation. The NSHP ordering on $\mathbb{Z}^2$ is defined by considering the set

$$S = \{t = (t_1, t_2) : 1 \le t_2, \text{ or } t_2 = 0 \text{ and } t_1 > 0\}$$

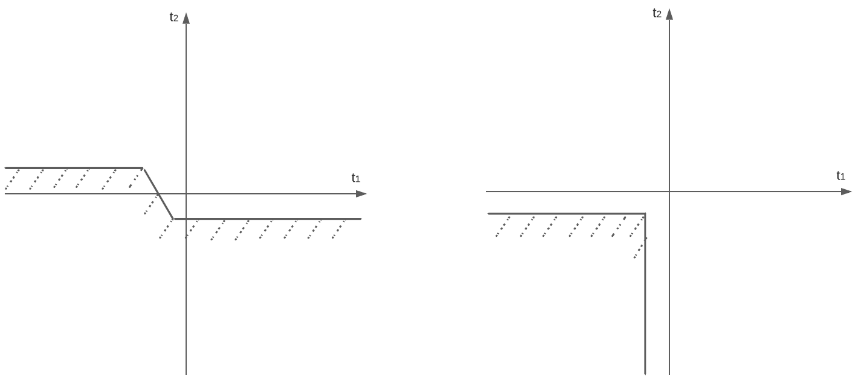

Figure 1.1: (a) $\{(t_1, t_2) < (0,0)\}$ in the NSHP ordering. (b) $\{(t_1, t_2) < (0,0)\}$ in the QP ordering.

(cf. Figure 1.1(a)). Then we say $x < y$ if $y - x \in S$. Thus, $\{(t_1, t_2) < (0,0)\} = -\mathbf{S}$. For the QP representation, a partial ordering on $\mathbb{Z}^2$ is defined as

$$(x_1, x_2) \le (y_1, y_2) \quad \text{if and only if } x_1 \le y_1, x_2 \le y_2 \tag{1.1}$$

(cf. Figure 1.1(b)). Whittle (1954) [29] first considered the NSHP representation. Helson & Lowdenslager (1958) [30] developed a mathematical theory for this type of representation in a general setting.

1.2.1 Stationary 2-D ARMA Models

Motivated by the definition of autoregressive-moving average (ARMA) time series models, a finite-order ARMA representation can be defined for spatial series. For example, in the QP case, an ARMA model can be defined as

$$\sum_{k_1=0}^{k_1=p_1} \sum_{k_2=0}^{k_2=p_2} \alpha_{k_1 k_2} X_{t_1-k_1, t_2-k_2} = \sum_{j_1=0}^{j_1=q_1} \sum_{j_2=0}^{j_2=q_2} \beta_{j_1 j_2} \epsilon_{t_1-j_1, t_2-j_2}, \quad \alpha_{00} = \beta_{00} = 1 \tag{1.2}$$

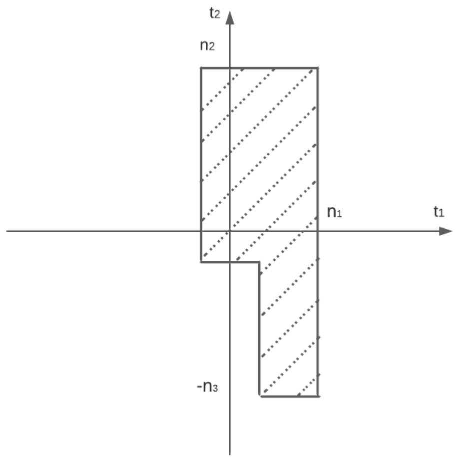

Figure 1.2: The support set $G(n)$

where $\{\epsilon_t\}$ is a spatial white noise process, p, q are finite and the ordering $k \leq p$ is as defined above. We concentrate on NSHP ARMA models. Let n_1, n_2, n_3 be nonnegative integers, $n = (n_1, n_2, -n_3)$ and let $G(n)$ denote the set $\{t = (t_1, t_2) : 1 \leq t_1 \leq n_1$ and $-n_3 \leq t_2 \leq n_2$, or $t_1 = 0$ and $0 \leq t_2 \leq n_2\}$.
In what follows we consider the spatial ARMA process

$$\sum\sum_{(k_1,k_2)\in G(\mathbf{p})} \alpha_{k_1 k_2} X_{t_1-k_1, t_2-k_2} = \sum\sum_{(j_1,j_2)\in G(\mathbf{p})} \beta_{j_1 j_2} \epsilon_{t_1-j_1, t_2-j_2}, \quad \alpha_{00} = \beta_{00} = 1 \tag{1.3}$$

Other spatial ARMA representations have been suggested in the literature; see, e.g. Kashyap (1984) [31] and the references therein for a review and applications of ARMA models.

We denote by D^2 the unit polydisc $\{(z_1, z_2) : |z_1| < 1, |z_2| < 1\}$, $\overline{D^2}$ its closure and T^2 its distinguished boundary $\{(z_1, z_2) : |z_1| = 1, |z_2| = 1\}$. We consider the spatial ARMA process defined by (1.3), the following conditions correspond to the stationarity condition for 2-D ARMA models [27]:

$$\sum\sum_{(k_1 k_2)\in G(\mathbf{p})} \alpha_{k_1 k_2} z_1^{k_1} z_2^{k_2} \neq 0 \, |z_1| \leq 1, |z_2| = 1 \tag{1.4}$$

$$\sum\sum_{(0,k_2)\in G(\mathbf{p})} \alpha_{0 k_2} z_2^{k_2} \neq 0, |z_2| \leq 1. \tag{1.5}$$

1.2.2 The Inverse 2-D ARMA Models

We consider the spatial ARMA process defined by (3.1), the following conditions correspond to the invertibility condition for 2-D ARMA models [27]:

$$\sum\sum_{(j_1,j_2)\in G(\mathbf{q})} \beta_{j_1 j_2} z_1^{j_1} z_2^{j_2} \neq 0 \, |z_1| \leq 1, |z_2| = 1 \tag{1.6}$$

$$\sum\sum_{(0,j_2)\in G(\mathbf{q})} \beta_{(0,j_2)} z_2^{j_2} \neq 0, |z_2| \leq 1. \tag{1.7}$$

Given two complex-valued functions $\sum\sum_{(k_1k_2)\in G(\mathbf{p})} \alpha_{k_1k_2} z_1^{k_1} z_2^{k_2}$ and $\sum\sum_{(j_1j_2)\in G(\mathbf{q})} \beta_{j_1j_2} z_1^{j_1} z_2^{j_2}$, if there exist two complex-valued functions

$$\sum\sum_{(k_1k_2)\in G(\mathbf{p}')} \alpha_{k_1k_2} z_1^{k_1} z_2^{k_2} \text{ and } \sum\sum_{(j_1j_2)\in G(\mathbf{q}')} \beta_{j_1j_2} z_1^{j_1} z_2^{j_2}$$

with $\sum\sum_{(j_1j_2)\in G(\mathbf{q}')} |b_{j_1j_2}|^2 > 0$ such that

$$\frac{\sum\sum_{(k_1k_2)\in G(\mathbf{p})} \alpha_{k_1k_2} z_1^{k_1} z_2^{k_2}}{\sum\sum_{(j_1j_2)\in G(\mathbf{q}')} \beta_{j_1j_2} z_1^{j_1} z_2^{j_2}} \sum\sum_{(j_1j_2)\in G(\mathbf{q}')} b_{j_1j_2} z_1^{j_1} z_2^{j_2} = \sum\sum_{(k_1k_2)\in G(\mathbf{p}')} \alpha_{k_1k_2} z_1^{k_1} z_2^{k_2}, \quad (z_1z_2) \in \mathbf{T}^2. \tag{1.8}$$

then

$$G(\mathbf{p}) \subset G(\mathbf{p}') \text{ and } G(\mathbf{q}) \subset G(\mathbf{q}'). \tag{1.9}$$

The two equations (1.8) and (1.10) are an identifiability condition for spatial ARMA [27].

Proposition 1.1: [27]

Under conditions (1.6) and (1.7), there exists $\{\phi_{t_1t_2} : t_1 \geq 0,\ t_2 \geq 0\}$ such that, for $(z_1z_2) \in \mathbf{T}^2$,

$$\frac{\sum\sum_{(k_1k_2)\in G(\mathbf{p})} \alpha_{k_1k_2} z_1^{k_1} z_2^{k_2}}{\sum\sum_{(j_1j_2)\in G(\mathbf{q})} \beta_{j_1j_2} z_1^{j_1} z_2^{j_2}} = \sum\sum_{(t_1t_2)\geq(0,0)} \phi_{t_1t_2} z_1^{t_1} z_2^{t_2}, \quad \phi_{00} = 1 \tag{1.10}$$

$$\sum\sum_{(j_1j_2)\in G(\mathbf{q})} \beta_{j_1j_2} \phi_{t_1-j_1,t_2-j_2} = \begin{cases} \alpha_{t_1t_2}, & (t_1t_2) \in G(\mathbf{p}) \\ 0, & (t_1t_2) \notin G(\mathbf{p}) \end{cases} \tag{1.11}$$

Proof: Define

$$F(z_1, z_2) = \frac{\sum\sum_{(k_1k_2)\in G(\mathbf{p})} \alpha_{k_1k_2} z_1^{k_1} z_2^{k_2}}{\sum\sum_{(j_1j_2)\in G(\mathbf{q})} \beta_{j_1j_2} z_1^{j_1} z_2^{j_2}} \quad |z_1| \leq 1, |z_2| \leq 1$$

Then on $z_2 = e^{i\lambda_2}, F(z_1, e^{i\lambda_2})$ can be written as

$$F(z_1, e^{i\lambda_2}) = \frac{\sum_{k_1=0}^{p_1} \alpha_{k_1}(e^{i\lambda_2}) z_1^{k_1}}{\sum_{j_1=0}^{q_1} \beta_{j_1}(e^{i\lambda_2}) z_1^{j_1}}. \tag{1.12}$$

Condition (1.6) implies that $F(z_1, e^{i\lambda_2})$ is analytic in $|z_1| < 1$, so it can be expanded as

$$F(z_1, e^{i\lambda_2}) = \sum_{t_1=0}^{\infty} \phi_{t_1}(\lambda_2) z_1^{t_1}. \tag{1.13}$$

Also, in view of condition (1.7), there exists a finite M such that

$$\frac{1}{2\pi}\int_{-\pi}^{x}\left|\frac{\sum_{k_1=0}^{p_1}\alpha_{k_1}\left(e^{i\lambda_2}\right)e^{ik_1\lambda_1}}{\sum_{j_1=0}^{q_1}\beta_{j_1}\left(e^{i\lambda_2}\right)e^{ij_1\lambda_1}}\right|^2 d\lambda_1 \le \frac{1}{2\pi}M^2\times 2\pi<\infty$$

i.e. $F\left(z_1,e^{i\lambda_2}\right)$ belongs to the Hardy space H^2. Parseval's theorem now implies that

$$\sum_{t_1=0}^{\infty}\left|\phi_{t_1}\left(\lambda_2\right)\right|^2<\infty$$

i.e. $\phi_{t_1}\left(\lambda_2\right)\in L^2\left(d\lambda_2\right)$. It then follows from Plancherel's Theorem that $\phi_{t_1}\left(\lambda_2\right)$ can be decomposed as

$$\phi_{t_1}\left(\lambda_2\right)=\sum_{t_2=-\infty}^{\infty}\phi_{t_1,t_2}e^{it_2\lambda_2}\qquad\left(t_1\ge 0\right)\tag{1.14}$$

Furthermore, we have from (1.12) and (1.13) that

$$\phi_{00}\left(\lambda_2\right)=\frac{\alpha_{00}\left(e^{i\lambda_2}\right)}{\beta_{00}\left(e^{i\lambda_2}\right)}.$$

Condition (1.7) now implies that $\alpha_{00}\left(z_2\right)/\beta_{00}\left(z_2\right)$ is analytic in $|z_2|<1$. Hence, it can be expanded as

$$\phi_{00}\left(e^{i\lambda_2}\right)=\frac{\alpha_{00}\left(e^{i\lambda_2}\right)}{\beta_{00}\left(e^{i\lambda_2}\right)}=\sum_{t_2=0}^{\infty}\phi_{0,t_2}e^{it_2\lambda_2}.\tag{1.15}$$

Putting (1.15) and (1.14) into (1.13), we get

$$\frac{\sum_{k_1=0}^{p_1}\alpha_{k_1}\left(e^{i\lambda_2}\right)e^{ik_1\lambda_1}}{\sum_{j_1=0}^{q_1}\beta_{j_1}\left(e^{i\lambda_2}\right)e^{ij_2\lambda_1}}=\sum\sum_{(t_1t_2)\ge(0,0)}\phi_{t_1t_2}e^{i(t_1\lambda_1+t_2\lambda_2)}$$

where $\sum\sum_{(t_1t_2)\ge(0,0)}=\sum_{t_2=0}^{\infty}+\sum_{t_1=0}^{\infty}\sum_{t_2=-\infty}^{\infty}$. Thus (1.10) is established. In view of this result, we get, for $z\in T^2$,

$$\begin{aligned}\sum_{k\in G(\mathbf{p})}a_k z^k&=\sum\sum_{(j_1j_2)\in G(\mathbf{q})}\beta_{j_1j_2}z_1^{j_1}z_2^{j_2}\sum_{t\ge 0}\phi_t z^t\\&=\sum_{t\ge 0}\sum_{j\in G(q)}\beta_j\phi_t z^{t+j}=\sum_{t\in\mathbb{Z}^2}\left(\sum_{j\in G(q)}\beta_j\phi_{t-j}\right)z^t\end{aligned}\tag{1.16}$$

using the change of variable $t'=t+j$ and defining $\phi_t=0$ for $t<0$. It follows from (1.16) that, on $z=e^{i\lambda}$,

$$e^{-is\lambda}\sum_{t\in\mathbb{Z}^2}\left(\sum_{j\in G(q)}\beta_j\phi_{t-j}\right)e^{it\lambda}=e^{-is\lambda}\sum_{k\in G(\mathbf{p})}\alpha_k e^{ik\lambda}$$

Consequently, setting $\alpha_t=0$ for $t\notin G(\mathbf{p})$,

$$\sum_{t\in\mathbb{Z}^2}\sum_{j\in G(q)}\beta_j\phi_{t-j}\int_{[-\pi,\pi]^2}e^{i(t-s)\lambda}d\lambda=\sum_{k\in G(\mathbf{p})}\alpha_k\int_{[-\pi,\pi]^2}e^{i(k-s)\lambda}d\lambda$$

$$= \sum_{t \in \mathbb{Z}^2} \alpha_t \int_{[-\pi,\pi]^2} e^{i(t-s)\lambda} d\lambda \tag{1.17}$$

Therefore,

$$\sum_{j \in G(q)} \beta_j \phi_{t-j} = \begin{cases} \alpha_t & t \in G(\mathbf{p}) \\ 0 & t \notin G(\mathbf{p}) \end{cases}$$

□

1.2.3 Correlations for 2-D ARMA models

We consider spatial processes defined on a regular rectangular grid in two dimensions with sites labeled (i, j), with an associated random variable Y_{ij} defined at each site. We first briefly review some different models used to analyze two-dimensional spatial data, and then concentrate on the first-order unilateral model of our interest.

Some models for spatial data

Three broad categories of spatial models that have been studied are the simultaneous autoregressive (SAR) models, Whittle (1954) [29], the conditional autoregressive (CAR) models (Bartlett (1971) [32] , Besag (1974) [33]), and the moving average (MA) models (Haining (1978b) [34]). The models considered in practice were mostly of first order, that is, only the immediate neighboring sites are used to model the value of a particular site. The simultaneous AR model (Whittle (1954) [29]) is defined as $\Phi(B_1, B_2) Y_{ij} = \varepsilon_{ij}$, where

$$\Phi(B_1, B_2) = \Sigma_k \Sigma_l \alpha_{kl} B_1^k B_2^l$$

with $B_1 Y_{ij} = Y_{i-1,j}$ and $B_2 Y_{ij} = Y_{i,j-1}$, and the ε_{ij} are a collection of independent random variables with $\boldsymbol{E}(\varepsilon_{ij}) = 0$ and $\operatorname{Var}(\varepsilon_{ij}) = \sigma^2$. Note that in the SAR model k and l can take both positive and negative values. Whittle (1954) [29] proved that given a set of autocorrelations of a SAR process $\{Y_{ij}\}$, there is a unique representation of the process in which Y_{ij} can be expressed as an autoregression on $Y_{il}(l < j)$ and $Y_{kl}(k < i, l$ unrestricted)

$$Y_{ij} = \sum_{k=1}^{\infty} \sum_{l=-\infty}^{\infty} \alpha_{kl} Y_{i-k,j-l} + \sum_{l=1}^{\infty} \alpha_{0l} Y_{i,j-l} + \varepsilon_{ij}. \tag{1.18}$$

The general form of models (1.18) are called unilateral AR models and special cases of these models will be considered in this paper. The correlation properties of the unilateral AR models have been studied by Besag (1972) [35] .

Tjostheim (1983) [36] examined in detail, including asymptotic properties of the corresponding Yule-Walker estimators, a special case of the unilateral AR models as defined by Whittle (1954) [29] where the value at the site (i, j) is a (finite) autoregression on the values at the sites which are in the lower quadrant of (i, j). In two dimensions this model becomes

$$Y_{ij} = \sum_{k=0}^{p_1} \sum_{l=0}^{p_2} \alpha_{kl} Y_{i-k,j-l} + \varepsilon_{ij}, \quad \alpha_{00} = 0. \tag{1.19}$$

In the following, we will examine unilateral first-order autoregressive moving average (ARMA) models of the quadrant type, given by

$$Y_{ij} = \alpha_1 Y_{i-1,j} + \alpha_2 Y_{i,j-1} + \alpha_3 Y_{i-1,j-1} + \varepsilon_{ij} + \theta_1 \varepsilon_{i-1,j} + \theta_2 \varepsilon_{i,j-1} + \theta_3 \varepsilon_{i-1,j-1} \tag{1.20}$$

Note that although these models may not possess the same intuitive appeal as the general SAR or CAR models because of their unilateral rather than quadrilateral structure, the models are capable of representing a wide range of general spatial correlation patterns. In particular, the special case of (1.20) where $\alpha_3 = -\alpha_1\alpha_2$ and $\theta_3 = \theta_1\theta_2$ leads to the linear-by-linear (or separable) form of spatial ARMA (1.18) model which has recently been studied extensively by Martin (1990) [37], who has argued that these models have wide uses in practical modeling.

General correlation structure

Defining $B_1Y_{ij} = Y_{i-1,j}$ and $B_2Y_{ij} = Y_{i,j-1}$, we can write model (1.20) as

$$Y_{ij} = (1 - \alpha_1 B_1 - \alpha_2 B_2 - \alpha_3 B_1 B_2)^{-1} (\varepsilon_{ij} + \theta_1 \varepsilon_{i-1,j} + \theta_2 \varepsilon_{i,j-1} + \theta_3 \varepsilon_{i-1,j-1}).$$

Note that if $1 - \alpha_1 z_1 - \alpha_2 z_2 - \alpha_3 z_1 z_2 \neq 0$ for all z_1, z_2 within the closed unit polydisc, $|z_1| \leqq 1, |z_2| \leqq 1$ ([38], Lemma 5.1), then use of a multinomial expansion for $(1 - \alpha_1 B_1 - \alpha_2 B_2 - \alpha_3 B_1 B_2)^{-1}$ implies the convergent representation

$$\begin{aligned} Y_{ij} = \sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\sum_{r=0}^{\infty} \frac{(k+l+r)!}{k!l!r!} \alpha_1^k \alpha_2^l \alpha_3^r \big(&\varepsilon_{i-k-r,j-l-r} + \theta_1 \varepsilon_{i-k-r-1,j-l-r} \\ &+\theta_2 \varepsilon_{i-k-r,j-l-r-1} + \theta_3 \varepsilon_{i-k-r-1,j-l-r-1}\big). \end{aligned} \quad (1.21)$$

The following gives a convenient set of conditions on the AR parameters $\alpha_1, \alpha_2, \alpha_3$ to ensure the existence of the stationary representation of the model (1.20)

Proposition 1.2: [28]

For the first-order spatial model (1.20) with $\Phi(z_1, z_2) = 1 - \alpha_1 z_1 - \alpha_2 z_2 - \alpha_3 z_1 z_2$, none of the roots of $\Phi(z_1, z_2) = 0$ lie within the closed unit polydisc $(|z_1| \leqq 1, |z_2| \leqq 1)$ if and only if

(i) $|\alpha_i| < 1$ for all i,

(ii) $\left(1 + \alpha_1^2 - \alpha_2^2 - \alpha_3^2\right)^2 - 4(\alpha_1 + \alpha_2\alpha_3)^2 > 0$,

(iii) $1 - \alpha_2^2 > |\alpha_1 + \alpha_2\alpha_3|$.

Proof: (a) Sufficiency. Assume the condition on the roots of $\Phi(z_1, z_2) = 0$ holds. We first prove condition (1) that $|\alpha_i| < 1$. Taking $z_2 = 0$ in $\Phi(z_1, z_2) = 0$ implies that $|z_1| = \frac{1}{|\alpha_1|} > 1$, by assumption, so that $|\alpha_1| < 1$.

Similarly, taking $z_2 = 0$ in $\Phi(z_1, z_2) = 0$, we prove that $|\alpha_1| < 1$.
To prove that $|\alpha_1| < 1$, note that the roots of $\Phi(z_1, z_2) = 0$ have $z_2 = \frac{1-\alpha_1 z_1}{|\alpha_1 + \alpha_3 z_1|}$. Taking z_1, equal to 1 and −1, from $\frac{1-\alpha_1}{|\alpha_1+\alpha_3|} > 1$. and $\frac{1-\alpha_1}{|\alpha_1-\alpha_3|} > 1$ we can readily obtain that $|\alpha_3| < 1$

To prove condition (2), suppose that $\Delta^2 = (1 + \alpha_1^2 - \alpha_2^2 - \alpha_2^2)^2 - 4(\alpha_1 + \alpha_2\alpha_3)^2 \leq 0$ is true and consider the values $(z_1^* z_2^*)$, where

$$\frac{1 + \alpha_1^2 - \alpha_2^2 - \alpha_2^2 + \Delta}{2(\alpha_1 + \alpha_2\alpha_3)}$$

and

$$\frac{1+\alpha_1^2-\alpha_2^2-\alpha_2^2+\Delta}{2(\alpha_2+\alpha_2\alpha_3)}$$

Note that $(z_1^* z_2^*)$ is a root of $\Phi(z_1, z_2) = 0$ and it can be shown that Δ^2 is also equivalent to

$$(1+\alpha_1^2-\alpha_2^2-\alpha_2^2)^2-4(\alpha_2+\alpha_1\alpha_3)^2$$

Hence $|z_1^*| = |z_2^*| = 1$ if $\Delta^2 \leq 0$, which is within closed unit polydisc and so we have contradiction. Thus condition (2) is established.
suppose that

$$\Delta^2 = \left(1+\alpha_1^2-\alpha_2^2-\alpha_3^2\right)^2 - 4\left(\alpha_1+\alpha_2\alpha_3\right)^2 \leqq 0$$

is true and consider the values (z_1^*, z_2^*), where

$$z_1^* = \frac{\left(1+\alpha_1^2-\alpha_2^2-\alpha_3^2+\Delta\right)}{2\left(\alpha_1+\alpha_2\alpha_3\right)}$$

and

$$z_2^* = \frac{\left(1-\alpha_1^2+\alpha_2^2-\alpha_3^2-\Delta\right)}{2\left(\alpha_2+\alpha_1\alpha_3\right)}$$

Note that (z_1^*, z_2^*) is a root of $\Phi\left(z_1, z_2\right) = 0$ and it can be shown that Δ^2 is also equivalent to

$$\left(1-\alpha_1^2+\alpha_2^2-\alpha_3^2\right)^2-4\left(\alpha_2+\alpha_1\alpha_3\right)^2$$

Hence $|z_1^*| = |z_2^*| = 1$ if $\Delta^2 \leqq 0$, which is within the closed unit polydisc and so we have a contradiction. Thus condition (ii) is established. Taking $z_2 = \alpha_2$ in $\Phi\left(z_1, z_2\right) = 0$, with $|z_2| < 1$ by (i), implies that

$$z_1 = \frac{\left(1-\alpha_2^2\right)}{\left(\alpha_1+\alpha_3\alpha_2\right)}$$

must be greater than 1 in absolute value which leads to condition (iii).

(b) Necessity. Assume conditions (i)-(iii) hold. Note that if (ω_1, ω_2) is a solution to $\Phi\left(z_1, z_2\right) = 0$, then

$$\omega_2 = \frac{\left(1-\alpha_1\omega_1\right)}{\left(\alpha_2+\alpha_3\omega_1\right)}$$

which implies

$$|\omega_2|^2 = \frac{1+\alpha_1^2\left|\omega_1\right|^2-2\alpha_1\operatorname{Re}\left(\omega_1\right)}{\alpha_2^2+\alpha_3^2\left|\omega_1\right|^2+2\alpha_2\alpha_3\operatorname{Re}\left(\omega_1\right)}.$$

Let $|\omega_1| \leqq 1$ and write $\omega_1 = r\exp(i\phi)$, where $0 \leqq r \leqq 1, 0 \leqq \phi \leqq 2\pi$. We need to establish that no roots of $\Phi\left(z_1, z_2\right) = 0$ lie within the closed unit polydisc, and hence we need to show that $|\omega_2| > 1$. It suffices to show that $1-\alpha_2^2 > M$ where

$$M = \sup_{0\leqq r\leqq 1} f(r) = \sup_{0\leqq r\leq 1}\left\{\left(\alpha_3^2-\alpha_1^2\right)r^2+2r\left|\alpha_1+\alpha_2\alpha_3\right|\right\}.$$

The function $f(r)$ may attain its maximum over $0 \leqq r \leqq 1$ at $r = 1$, in which case we need that

$$1-\alpha_2^2 > \alpha_3^2-\alpha_1^2+2\left|\alpha_1+\alpha_2\alpha_3\right|$$

which follows by condition (ii) and by using the fact that conditions (i) and (ii) also imply that $1 + \alpha_1^2 - \alpha_2^2 - \alpha_3^2 \geqq 0$. Otherwise (if $\alpha_3^2 - \alpha_1^2 < 0$), $f(r)$ may attain a maximum at an interior point

$$r_0 = \frac{|\alpha_1 + \alpha_2\alpha_3|}{\left(\alpha_1^2 - \alpha_3^2\right) < 1},$$

with maximum

$$M = f\left(r_0\right) = |\alpha_1 + \alpha_2\alpha_3|\, r_0 \leqq |\alpha_1 + \alpha_2\alpha_3|\,.$$

Hence, $1-\alpha_2^2 > M$ follows by condition (iii). Therefore, the proposition is established. □

It follows from (1.21) that

$$\boldsymbol{E}\left(Y_{kl}\varepsilon_{ij}\right) = 0 \text{ when } k < i \text{ or } l < j, \text{ and } \boldsymbol{E}\left(Y_{ij}\varepsilon_{ij}\right) = \sigma^2. \tag{1.22}$$

Let ρ_{st} denote the bivariate correlations at lags (s,t), which are defined as $\rho_{st} = \boldsymbol{E}\left(Y_{ij}Y_{i-s,j-t}\right)/\gamma$, where $\gamma = \mathrm{Var}\left(Y_{ij}\right)$. Note that

$$\rho_{-s,-t} = \boldsymbol{E}\left(Y_{ij}Y_{i+s,j+t}\right)/\gamma = \boldsymbol{E}\left(Y_{i-s,j-t}Y_{ij}\right)/\gamma = \rho_{st}$$

Hence we can concentrate on the half-plane $\{(s,t) \mid s \geqq 0,\ -\infty < t < \infty\}$ to obtain the bivariate correlations. Now from (1.20)

$$\begin{aligned}\gamma\rho_{st} =&\gamma\left(\alpha_1\rho_{s-1,t} + \alpha_2\rho_{s,t-1} + \alpha_3\rho_{s-1,t-1}\right) + \boldsymbol{E}\left(\varepsilon_{ij}Y_{i-s,j-t}\right)\\ &+ \theta_1\boldsymbol{E}\left(\varepsilon_{i-1,j}Y_{i-s,j-t}\right) + \theta_2\boldsymbol{E}\left(\varepsilon_{i,j-1}Y_{i-s,j-t}\right)\\ &+ \theta_3\boldsymbol{E}\left(\varepsilon_{i-1,j-1}Y_{i-s,j-t}\right).\end{aligned} \tag{1.23}$$

From (1.22), we see that $\boldsymbol{E}\left(\varepsilon_{i-1,j-1}Y_{i-s,j-t}\right) = 0$ when $s > 1$ or $t > 1$, so we have

$$\rho_{st} = \alpha_1\rho_{s-1,t} + \alpha_2\rho_{s,t-1} + \alpha_3\rho_{s-1,t-1}, \tag{1.24}$$

for $\{s > 1 \text{ and } -\infty < t < \infty\}$ and $\{s \leqq 1 \text{ and } t > 1\}$.
To solve this partial difference equation explicitly, we obtain the characteristic equations, for $s > 1$ and $t < -1$, from assuming a solution of the form of a constant times $\lambda^s\mu^{-t}$ for this region (see Mickens (1987) [39]. Noting that (1.23) also implies that $\rho_{st} = \alpha_1\rho_{s+1,t} + \alpha_2\rho_{s,t+1} + \alpha_3\rho_{s+1,t+1}$ for $s > 1, t < -1$, this leads to

$$\alpha_1\lambda + \alpha_2\mu^{-1} + \alpha_3\lambda\mu^{-1} = 1 = \alpha_1\lambda^{-1} + \alpha_2\mu + \alpha_3\lambda^{-1}\mu. \tag{1.25}$$

Eliminating μ from the two equations in (1.25) above leads to

$$\left(\alpha_1 + \alpha_2\alpha_3\right)\lambda^2 - \left\{1 + \alpha_1^2 - \left(\alpha_2^2 + \alpha_3^2\right)\right\}\lambda + \left(\alpha_1 + \alpha_2\alpha_3\right) = 0. \tag{1.26}$$

The two roots of (1.26) are reciprocals, λ and λ^{-1}, and hence a real value of λ with $|\lambda| < 1$ exists if $\left(1 + \alpha_1^2 - \alpha_2^2 - \alpha_3^2\right)^2 - 4\left(\alpha_1 + \alpha_2\alpha_3\right)^2 > 0$. Note that this, together with conditions (i) and (iii) of Proposition (2) are equivalent to the condition that no roots of $1 - \alpha_1 z_1 - \alpha_2 z_2 - \alpha_3 z_1 z_2 = 0$ lie within the closed unit polydisc.

In the special case where $\alpha_3 = 0$, the condition reduces to $|\alpha_1|+|\alpha_2| < 1$, as in Whittle (1954) [29] and Besag (1972) [35]. From (1.25) , the relations between the roots μ and λ are $\mu = (\alpha_2 + \alpha_3\lambda) / (1 - \alpha_1\lambda)$ and $\lambda = (\alpha_1 + \alpha_3\mu) / (1 - \alpha_2\mu)$.
To obtain an explicit representation for ρ_{st} for the remaining region of the quadrant $s \geqq 0$ and $t \leqq 0$, we divide this remaining region into two subregions and solve for ρ_{st}, using simultaneous partial difference equations, for these two subregions. First consider the relations for the region, $s > 1$ with $t = 0$ and $t = -1$. We observe from that $\rho_{s0} = \alpha_1\rho_{s-1,0} + \alpha_2\rho_{s,-1} + \alpha_3\rho_{s-1,-1}, s > 1$, or

$$(1 - \alpha_1 B_1)\,\rho_{s0} = (\alpha_2 + \alpha_3 B_1)\,\rho_{s,-1}, \quad s > 1.$$

Also, using $\boldsymbol{E}\left(\varepsilon_{ij}Y_{i+s,j}\right) = \sigma^2\left(\alpha_1 + \theta_1\right)\alpha_1^{s-1}$ for $s \geqq 1$ as obtained from (1.21), from (1.23) we have

$$\begin{aligned}\rho_{s,-1} = \rho_{-s,1} =&\alpha_1\rho_{s+1,-1} + \alpha_2\rho_{s0} + \alpha_3\rho_{s+1,0} \\ &+ \left(\sigma^2/\gamma\right)(\alpha_1 + \theta_1)(\theta_2 + \alpha_1\theta_3)\,\alpha_1^{s-1}, \quad s > 1.\end{aligned} \tag{1.27}$$

This implies that, for $s > 1$,

$$\left(\alpha_2 + \alpha_3 B_1^{-1}\right)\rho_{s0} + \left(1 - \alpha_1 B_1^{-1}\right)\rho_{s,-1} = \left(\sigma^2/\gamma\right)(\alpha_1 + \theta_1)(\theta_2 + \alpha_1\theta_3)\,\alpha_1^{s-1}. \tag{1.28}$$

Solving for $\rho_{s,-1}$ and ρ_{s0} from (1.27) and (1.28) simultaneously, we obtain

$$\rho_{s,-1} = \rho_{2,-1}\lambda^{s-2}, \quad \rho_{s0} = \left(\rho_{20} - \rho_{2,-1}\mu^{-1}\right)\alpha_1^{s-2} + \rho_{2,-1}\mu^{-1}\lambda^{s-2}, \quad s > 1.$$

Next consider the relations for the region $t > 1$ with $s = 0$ and $s = -1$, which is equivalent to consideration of the region $t < -1$ with $s = 0$ and $s = 1$. Following in the same line of reasoning as in the case of $s > 1$ with $t = 0$ and $t = -1$, we can show that for $t > 1$,

$$\rho_{-1,t} = \rho_{1,-t} = \rho_{1,-2}\mu^{t-2},$$
$$\rho_{0,-t} = \rho_{0t} = \left(\rho_{02} - \rho_{1,-2}\lambda^{-1}\right)\alpha_2^{t-2} + \rho_{1,-2}\lambda^{-1}\mu^{t-2}.$$

The correlation properties of the unilateral model (1.20)are summarized in the following.

Proposition 1.3: [28]

For the first-order spatial model (1.20), the spatial correlations ρ_{st} are given explicitly, over the quadrant $s \geqq 0, t \leqq 0$, by

$$\begin{aligned}
\rho_{s0} &= \lambda^{s-1}\rho_{10} + \frac{\sigma^2}{\gamma}\frac{(\theta_2 + \alpha_1\theta_3)(\alpha_1 + \theta_1)}{(\alpha_2 + \alpha_1\alpha_3)}\left(\lambda^{s-1} - \alpha_1^{s-1}\right), \quad s > 0, \\
\rho_{s,-1} &= \rho_{1,-1}\lambda^{s-1}, \quad s > 0, \\
\rho_{0t} &= \mu^{t-1}\rho_{01} + \frac{\sigma^2}{\gamma}\frac{(\theta_1 + \alpha_2\theta_3)(\alpha_2 + \theta_2)}{(\alpha_1 + \alpha_2\alpha_3)}\left(\mu^{t-1} - \alpha_2^{t-1}\right), \quad t > 0, \\
\rho_{1t} &= \rho_{1,-1}\mu^{-t-1}, \quad t < 0, \\
\rho_{st} &= \rho_{1,-1}\lambda^{s-1}\mu^{-t-1}, \quad s \geqq 1 \quad \text{and} \quad t \leqq -1,
\end{aligned} \tag{1.29}$$

with the initial values given by

$$\begin{aligned}
\rho_{1,-1} &= \lambda\rho_{01} + [\lambda/(\alpha_1+\alpha_2\alpha_3)]\, f_4,\\
\rho_{1,-1} &= \mu\rho_{10} + [\mu/(\alpha_2+\alpha_1\alpha_3)]\, f_4^*,\\
\rho_{10} &= \lambda + [\lambda/(\alpha_1+\alpha_2\alpha_3)]\,(\mu f_4^* + \alpha_3 f_2 + f_3),\\
\rho_{01} &= \mu + [\mu/(\alpha_2+\alpha_1\alpha_3)]\,(\lambda f_4 + f_2 + \alpha_3 f_3).
\end{aligned} \tag{1.30}$$

The forms of the f functions are

$$\begin{aligned}
f_1 &= \left(\sigma^2/\gamma\right)[1+(\alpha_1+\theta_1)(\theta_1+\alpha_2\theta_3)+(\alpha_2+\theta_2)(\theta_2+\alpha_1\theta_3)+\theta_3(\alpha_3+\theta_3)];\\
f_2 &= \left(\sigma^2/\gamma\right)[\theta_2+\theta_3(\alpha_1+\theta_1)];\\
f_3 &= \left(\sigma^2/\gamma\right)[\theta_1+\theta_3(\alpha_2+\theta_2)];\\
f_4 &= \left(\sigma^2/\gamma\right)(\alpha_2+\theta_2)(\theta_1+\alpha_2\theta_3);\\
f_4^* &= \left(\sigma^2/\gamma\right)(\alpha_1+\theta_1)(\theta_2+\alpha_1\theta_3).
\end{aligned}$$

The proof of this Proposition is detailed in [28, page 646].

1.3 Definitions and Structural Properties of 2D RCAR Models

Two-dimensional indexed random coefficient autoregressive model (2D-RCAR) is a natural generalization of 2D-AR. The analysis of 2D-RCAR models is quite tractable. Fixed coefficients for 2D-AR Model is not very realistic, the autoregressive models with coefficients assumed to be not constant, but subjected to random perturbations allow to handle the non-homogeneity of a process.
To the best of our knowledge, the first study of 2D-RCAR models has been established by (Bibi (2016) [14]. The basic definitions and results of the second-order properties presented in this section are borrowed from (Bibi (2016) [14].

1.3.1 Notations and Definitions

- Throughout, all random variables are defined on the same probability space $(\Omega, \Im, P)$,
- $\mathbb{L}_p = \mathbb{L}_p(\Omega, \Im, P), p \geq 1$ denotes the Banach space of random variables X such that $E\{|X|^p\} < +\infty$ ($\mathbb{L}_p$ is endowed by the norm $\|X\|_p = E\{|X|^p\}^{1/p}\big)$.
- $I_{(k)}$ denotes the $k \times k$ identity matrix and $O_{(k)}$ (respectively $\underline{O}_{(k)}$) denotes the $k \times k$ matrix (respectively k-vector) whose entries are zeros.
- If M is an $n \times n$ squared matrix, $\rho(M)$ refers to the spectral radius of a square matrix M, i.e., $\rho(M) := \max_i \{|\lambda_i(A)|\}$ where $\lambda_i(M)$ are the eigenvalues of M.
- All vectors are underlined except those of $\mathbb{Z}^2$ are written in bold letters, so $\mathbf{0} = (0,0), \mathbf{e}_1 = (1,0), \mathbf{e}_0 = (0,1)$, and $\mathbf{1} = (1,1)$. For any $s = (s_1, s_2), \mathbf{t} = (t_1, t_2) \in \mathbb{Z}^2$, and $\underline{\mathbf{z}} = (z_1, z_2) \in \mathbb{C}^2$, we write $\mathbf{s} < \mathbf{t}$ (resp. $\mathbf{s} \leq \mathbf{t}$) if and only if $[s_1 < t_1$ and $s_2 < t_2]$ (resp. $s_1 \leq t_1$ and $s_2 \leq t_2$) and $\underline{\mathbf{z}}^{\mathrm{s}} = z_1^{s_1} z_2^{s_2}\ X(\mathbf{t}-\mathbf{1})$.

- If $\mathbf{a} = (a_1, a_2), \mathbf{b} = (b_1, b_2) \in \mathbb{Z}^2$, $\frac{\mathbf{a}}{\mathbf{b}} = \left(\frac{a_1}{b_1}, \frac{a_2}{b_2}\right)$ whenever $b_1, b_2 \neq 0$, when $\mathbf{a} \leq \mathbf{b}$, the following indexing subset in $\mathbb{Z}^2$ will be considered $S[\mathbf{a}, \mathbf{b}] := S[a_1, b_1] \times S[a_2, b_2]$, where $S[a, b] = \{x \in \mathbb{Z}, a \leq x \leq b\}$ if $\mathbf{a} = \mathbf{0}$ (resp. $\mathrm{a} = -\infty$), $S[\mathbf{0}, \mathbf{b}]$ is noted $S[\mathbf{b}]$ (resp. $S_-[\mathbf{b}]$) and $S]\mathbf{b}] = S[\mathbf{b}] \backslash \{\mathbf{0}\}$.

In this section, we introduce the class of real-valued $2D - RCAR(\mathbf{p})$ models subject to the following definition

Definition 1.1: [14]

A $\mathbb{R}$-valued random field $\left(X(\mathbf{t}), \mathbf{t} \in \mathbb{Z}^2\right)$ defined on some probability space $(\Omega, \mathfrak{I}, P)$ is called two-dimensionally indexed random coefficient autoregressive model with order $\mathbf{p} = (p_1, p_2) \in \mathbb{N}^2$ denoted by $2D - RCAR(\mathbf{p})$ if it is a solution of the following stochastic difference equation

$$X(\mathbf{t}) = \sum_{\mathbf{j} \in S]\mathbf{p}]} \left(\beta_{\mathbf{j}} + \beta_{\mathbf{j}}(\mathbf{t})\right) X(\mathbf{t} - \mathbf{j}) + e(\mathbf{t}), \mathbf{t} \in \mathbb{Z}^2 \tag{1.31}$$

in which the innovation process $\left(e(\mathbf{t}), \mathbf{t} \in \mathbb{Z}^2\right)$ is an independent and identically distributed (i.i.d.) sequence of random fields with mean zero and variance $\sigma^2, \beta = \left(\beta_{\mathbf{j}}, \mathbf{j} \in S\right]\mathbf{p}]\right)'$ is a real vector of constants. If $\underline{\beta}(\mathbf{t}) = \left(\beta_{\mathbf{j}}(\mathbf{t}), \mathbf{j} \in S\right]\mathbf{p}], \mathbf{t} \in \mathbb{Z}^2)'$, then the perturbation process $\left(\beta(\mathbf{t}), \mathbf{t} \in \mathbb{Z}^2\right)$ is a strictly stationary of an independent vectors of random fields with zero mean, $\mathrm{Var}\{\beta(\mathbf{t})\} = \Pi$ and independent of $\left(e(\mathbf{t}), \mathbf{t} \in \mathbb{Z}^2\right)$.

1.3.2 Second-Order Properties $2D - RCAR$ models

The stochastic recurrence Equation (1.31) defines a nonlinear random field model that has received considerable attention in the context of random perturbation of dynamically multidimensional systems and signal processing.
Hence, and for the statistical purpose, it is often important in practice that the process $\left(X(\mathbf{t}), \mathbf{t} \in \mathbb{Z}^2\right)$ should be stable, ergodic and $\mathfrak{I}_{\mathbf{t}} := \mathfrak{I}_{\mathbf{t}}^{(e)} \vee \mathfrak{S}_{\mathbf{t}}^{(\beta)}-$ measurable where $\mathfrak{I}_{\mathbf{t}}^{(e)} := \sigma(e(\mathbf{s}), \mathbf{s} \in S_-[\mathbf{t}])$ and $\mathfrak{I}_{\mathbf{t}}^{(\beta)} := \sigma(\beta(\mathbf{t}), \mathbf{s} \in S_-[\mathbf{t}])$.
For this purpose, and according to $2D - AR$ model the equation (1.31) can be transformed into state-space form as $X(\mathbf{t}) = \underline{H'Y}(\mathbf{t})$ and

$$\underline{Y}(\mathbf{t}) = A_0(\mathbf{t})\underline{Y}(\mathbf{t} - \mathbf{e}_0) + A_1(\mathbf{t})\underline{Y}(\mathbf{t} - \mathbf{e}_1) + \underline{H}e(\mathbf{t}) \tag{1.32}$$

or equivalently $\Phi_{\mathbf{t}}(\mathbf{B})\underline{Y}(\mathbf{t}) = \underline{He}(\mathbf{t})$ where

$$\Phi_{\mathbf{t}}(\underline{z}) = I_{(n)} - A_0(\mathbf{t})\underline{z}^{\mathbf{e}_0} - A_1(\mathbf{t})\underline{z}^{\mathbf{e}_1}$$

for some appropriate vectors $\underline{Y}(\mathbf{t}), \underline{H}$, and random matrices $A_i(\mathbf{t}), i = 0, 1$ (not necessarily unique).
Since (1.32) is valid for all $\mathbf{t} \in \mathbb{Z}^2$, by successive substitution, we obtain a formal $\mathfrak{I}_{\mathbf{t}}-$ measurable solution $\left(\underline{Y}(\mathbf{t}), \mathbf{t} \in \mathbb{Z}^2\right)$ for (1.32) given by

$$\underline{Y}(\mathbf{t}) = e(\mathbf{t})\underline{H} + \sum_{k \geq 1} M_k(\mathbf{t})\underline{H} \tag{1.33}$$

where

$$M_k(\mathbf{t}) := \sum_{i_1,\dots,i_k=0}^{1} \left\{ \prod_{j=1}^{k} A_{i_j} \left(\mathbf{t} - \sum_{l=1}^{j-1} \mathbf{e}_{i_l} \right) \right\} e\left(\mathbf{t} - \sum_{l=1}^{k} \mathbf{e}_{i_l} \right), k \geq 1 \tag{1.34}$$

which satisfies the following recursive equation

$$M_k(\mathbf{t}) = A_0(\mathbf{t}) M_{k-1}(\mathbf{t} - \mathbf{e}_0) + A_1(\mathbf{t}) M_{k-1}(\mathbf{t} - \mathbf{e}_1), k \geq 1$$

In the problem of estimation and some related statistical questions, it is important to have conditions on $\left(A_i(\mathbf{t}), \mathbf{t} \in \mathbb{Z}^2\right), i = 0, 1$, and $\left(e(\mathbf{t}), \mathbf{t} \in \mathbb{Z}^2\right)$ which guarantee the convergence of infinite series in (1.33) in $\mathbb{L}_2$.
In such a case, we say that the Equation (1.32) (and hence Equation (1.31)) has a second-order stationary solution and (1.33) is actually its doubly moving-average representation. It is worth noting that when almost surely (a.s.) $\underline{A}(\mathbf{t}) = A_i, i = 0, 1$, for all $\mathbf{t} \in \mathbb{Z}^2$ (non stochastic), a necessary and sufficient condition for the existence of second-order strictly stationary and $\mathfrak{I}_t^{(e)}-$ measurable solution is that $\det\left(I_{(n)} - A_0 \underline{z}^{\mathbf{e}_0} - A_1 \underline{z}^{\mathbf{e}_1}\right) \neq 0$ for all $\underline{z} \in \mathbb{U}$ where

$$\mathbb{U} := \left\{ \underline{\mathbf{z}} = (z_1, z_2) \in \mathbb{C}^2 : |z_1| \leq 1, |z_2| \leq 1 \right\}$$

However, very little effort has been made toward the analysis of $2D - AR(\mathbf{p})$ whose coefficients are stochastic.

The next proposition provides some information about wide sense stationary solutions processes of (1.32)

Proposition 1.4: [14]

Consider Model (1.31) with associated state space (1.32). Then, a necessary and sufficient condition for Model (1.31) to have a unique $\mathfrak{I}_t$-measurable second-order strictly stationary solution given by (1.33) is that

$$E\left\{ \int_{[-\pi,\pi]^2} \left| \sum_{k\geq 0} \underline{H}' M_k(\mathbf{t}, \underline{\lambda}) \underline{H} \right|^2 d\underline{\lambda} \right\} < +\infty \tag{1.35}$$

where

$$M_k(\mathbf{t}, \underline{\lambda}) := \sum_{i_1,\dots,i_k=0}^{1} \left\{ \prod_{j=1}^{k} A_{i_j} \left(\mathbf{t} - \sum_{l=1}^{j-1} \mathbf{e}_{i_l} \right) e^{-i\underline{\lambda}.e_{ij}} \right\}$$

with $\underline{\lambda} = (\lambda_1, \lambda_2), d\underline{\lambda} = d\lambda_1 d\lambda_2$, and $\underline{\lambda}$ denotes the usual inner product.

Proof: Using the spectral representation of $e(\mathbf{t})$, i.e.,

$$e(\mathbf{t}) = \int_{[-\pi,\pi]^2} \underline{H} e^{it.\underline{\lambda}} dZ(\underline{\lambda})$$

where $Z(.)$ is an orthogonal stochastic measure with $E\{dZ(\underline{\lambda})\} = 0$ and spectral measure $E\left\{|dZ(\underline{\lambda})|^2\right\} = \frac{\sigma^2}{(2\pi)^d} d\underline{\lambda}$ where $d\underline{\lambda}$ means the Lebesgue measure on $\mathbb{R}^2$, then we get

$$X(\mathbf{t}) = \sum_{k\geq 0} \int_{[-\pi,\pi]^2} \underline{H}' M_k(\mathbf{t}, \underline{\lambda}) \underline{H} e^{it.\underline{\lambda}} d\bar{Z}(\underline{\lambda})$$

and the process $X(\mathbf{t})$ is in $\mathbb{L}_2$, that is $E\left\{X^2(\mathbf{t})\right\} < +\infty$ if and only if

$$E\left\{\sum_{k\geq 0}\int_{[-\pi,\pi]^2}\underline{H}'M_k(\mathbf{t},\underline{\lambda})\underline{H}e^{it.\underline{\lambda}}dZ(\underline{\lambda})\right\}^2$$
$$= E\left\{\frac{\sigma^2}{4\pi^2}\int_{[-\pi,\pi]^2}\left|\sum_{k\geq 0}\underline{H}'M_k(\mathbf{t},\underline{\lambda})\underline{H}\right|^2 d(\underline{\lambda})\right\} < +\infty \tag{1.36}$$

and the results follow.
Based on (1.35), several equivalent conditions can be derived to guarantee the convergence of the Series (1.33) in $\mathbb{L}_2$. Indeed, to verify that the process (1.33) is well-defined in $\mathbb{L}_2$, it is sufficient to show that $M_k(\mathbf{t})\underline{H}$ converges to zero in $\mathbb{L}_2$ (endowed with any matrix norm) at an exponential rate, as $k\to\infty$, since

$$\|\underline{Y}(\mathbf{t})\|_{\mathbb{L}_2} \leq \sum_{k\geq 0}\|M_k(\mathbf{t})\underline{H}\|_{\mathbb{L}_2}$$

For this purpose, let $\Gamma_k(\mathbf{h}) = E\{M_k(\mathbf{t})\otimes M_k(\mathbf{t}+\mathbf{h})\}$, then from (1.34) we have

$$\Gamma_k(\mathbf{0}) = A_0^{(2)}\Gamma_{k-1}(\mathbf{0}) + A_1^{(2)}\Gamma_{k-1}(\mathbf{0}) + \widetilde{A_0\otimes A_1}\Gamma_{k-1}(\mathbf{e}_1-\mathbf{e}_0) + \widetilde{A_1\otimes A_0}\Gamma_{k-1}(\mathbf{e}_0-\mathbf{e}_1),$$

where

$$A_i^{(2)} = E\left\{A_i^{\otimes 2}(\mathbf{t})\right\}, i = 0,1$$

and

$$\widetilde{A_i\otimes A_j} = E\{A_i(\mathbf{t})\otimes A_j(\mathbf{t})\}, 0\leq i\neq j\leq 1$$

So

$$\Gamma_k(\mathbf{0}) = \left(A_0^{(2)} + A_1^{(2)} + \widetilde{A_0\otimes A_1}z + \widetilde{A_1\otimes A_0}z^{-1}\right)\Gamma_{k-1}(\mathbf{0}), z = \frac{z_1}{z_2}.$$

Hence, the above recursive equation admits a non-trivial solution if and only if

$$\det\left(I_{(n^2)} - \widetilde{A_1\otimes A_0}Z_1 - \widetilde{A_1\otimes A_0}Z_2 - \left(A_0^{(2)} + A_1^{(2)}\right)Z_1Z_2\right)$$
$$\neq 0 \text{ for all } Z_i\in\mathbb{C}: |Z_i|\leq 1, i = 1,2 \tag{1.37}$$

□

Using Condition (1.37), we can obtain the following theorem.

Theorem 1.1: [14]

Consider Model (1.31) with state-space representation (1.32). Then Model (1.32) has a second-order strictly stationary and $\mathfrak{I}_t$-measurable solution if one of the following conditions holds:

1. $\rho\left(\left|A_0^{(2)} + A_1^{(2)}\right|\right) < 1$ and

$$\sup\rho\left(I_{(n^2)} - \left(A_0^{(2)} + A_1^{(2)}\right)Z_3\right)^{-1}(\widetilde{A_0\otimes A_1}Z_1 + widetilde A_1\otimes A_0))Z_2 < 1$$

2. $\rho\left(|\widetilde{A_0\otimes A_1}|\right) < 1$ and

$$\sup\rho\left(\left(I_{(n^2)} - \widetilde{A_0\otimes A_1}Z_1\right)^{-1}\left(\left(A_0^{(2)} + A_1^{(2)}\right)Z_3 + \widetilde{A_1\otimes A_0}Z_2\right)\right) < 1$$

3. $\rho\left(\left|\widetilde{A_1 \otimes A_0}\right|\right) < 1$ and

$$\sup \rho((I_{(n^2)} - \widetilde{A_1 \otimes A_0} Z_2)^{-1}((A_0^{(2)} + A_1^{(2)})Z_3 + A_0 \otimes A_1 Z_1)) < 1$$

4. $\rho\left(\left|A_0^{(2)} + A_1^{(2)}\right| + \left|\widetilde{A_1 \otimes A_0}\right| + \left|\widetilde{A_1 \otimes A_0}\right|\right) < 1$

5. $\left\|A_0^{(2)} + A_1^{(2)}\right\| + \left\|\widetilde{A_1 \otimes A_0}\right\| + \left\|\widetilde{A_1 \otimes A_0}\right\| < 1.$

Here $Z_3 = Z_1 Z_2$ such that $|Z_1| \leq 1, |Z_2| \leq 1, |M|$ means the non negative matrix with elements $(|m_{ij}|)$ and $\|.\|$ denotes any induced matrix norm.

Proof: From the theory of non-negative matrices, for any squared matrices

$$A, B, \rho(A) \leq \rho(|A|) \quad \text{and} \rho(|A|) \leq \rho(|A| + |B|)$$

Since $|Z_1| \leq 1$ and $|Z_2| \leq 1$, the first condition of Theorem 1 implies that

$$\left(I_{(n^2)} - \left(A_0^{(2)} + A_1^{(2)}\right) Z_1 Z_2\right)^{-1}$$

exists and that

$$\begin{aligned} 0 \neq \det(I_{(n^2)} - \left(I_{(n^2)} - \left(A_0^{(2)} + A_1^{(2)}\right) Z_1 Z_2\right)^{-1} (\widetilde{A_0 \otimes A_1} Z_1 + \widetilde{A_1 \otimes A_0} Z_2)) \\ = \det\left(\left(I_{(n^2)} - \left(A_0^{(2)} + A_1^{(2)}\right) Z_1 Z_2\right)^{-1}\right) \\ \det\left(I_{(n^2)} - \widetilde{A_0 \otimes A_1} Z_1 - \widetilde{A_1 \otimes A_0} Z_2 - \left(A_0^{(2)} + A_1^{(2)}\right) Z_1 Z_2\right) \end{aligned}$$

so Condition (1.37) follows. By similar arguments, we can show that conditions 2 and 3 are sufficient for (1.37). Moreover, it can be shown that the conditions $1-4$ are equivalent and condition 5 implies that condition 4 which is due to any induced norm matrix is an upper bound for its spectral radius. □

Remark 1.1: [14]

The above discussion may be extended to find sufficient conditions ensuring the existence of strictly stationary and $\Im_t-$ measurable solution having moments up to m-order $(m \geq 3)$. More precisely, a sufficient condition for the convergence of the Series (1.33) in $\mathbb{L}_m$ $(m \geq 3)$ is that for any $k \geq 1$

$$\rho\left(E\left\{A_0(\mathbf{t}) M_{k-1}(\mathbf{t} - \mathbf{e}_0) + A_1(\mathbf{t}) M_{k-1}(\mathbf{t} - \mathbf{e}_1)\right\}^{\otimes m}\right) < 1$$

Whenever the process $\left((\beta_i(\mathbf{t}), e(\mathbf{t})), t \in \mathbb{Z}^2\right), i = 0, 1$ admits a moments up to m-order.

1.4 Conclusion

This Chapter gave an overview of some basic results of spatial models: definitions and Structural Properties of 2D-ARMA and 2D-RCAR Models. This chapter contains all the necessary tools for understanding and developing the contributions and the rest of the manuscript and allows situating the contributions of this work in relation to existing work in the literature.

Chapter

2

Estimation of Spatial Models

Contents

2.1 Introduction

In section 2.2, we recall some estimation methods of spatial ARMA models: Yule-Walker Estimator for the First-Order Spatial Bilateral AR Model and Exact likelihood function and maximum likelihood estimation methods for the first-order spatial ARMA model.

In section 2.3, we present new original theoretical results on estimation in nonlinear random field models. We focus on two-dimensionally indexed random coefficients autoregressive models with order $(p_1, p_2) \in \mathbb{N}^2, 2D{-}RCAR(p_1, p_2)$ for short. We first develop a maximum likelihood estimation procedure for estimating the unknown parameters of $2D-RCAR(p_1,p_2)$. Moreover, we prove that the estimates are strongly consistent.

Finally, in Section 2.4, we focus on developing two parameter estimation procedures for first-order two-dimensionally indexed random coefficients autoregressive models (2D-RCAR) of the first order particularly. First, the model is explained in detail, and the definition and the condition of stationarity are given. Then, we provide a simulation study to show the accuracy of the maximum likelihood estimates and we provide a comparison with the least squares estimates

2.2 Estimation of Spatial ARMA Models

2.2.1 Exact likelihood function and maximum likelihood estimation

A good feature of the first-order spatial ARMA model (1.20) is that the exact likelihood function can be obtained in a convenient computational form. For values of Y_{ij} observed on a $m \times n$ rectangular grid for $i = 1, \cdots, m$ and $j = 1, \cdots, n$, let

$$\boldsymbol{Y} = (Y_{11}, Y_{21}, \cdots, Y_{m1}, \cdots, Y_{mn})'$$

and

$$\boldsymbol{Y}_0 = (Y_{00}, Y_{10}, \cdots, Y_{m0}, Y_{01}, \cdots, Y_{0n})'$$

with ε and ε_0 defined similarly, and set $a = (\varepsilon_0', Y_0')'$, so that Y is the vector of observations and $\boldsymbol{a}$ is the vector of unobserved initial values of the process and the errors. Let the covariance matrix of $\boldsymbol{a}$ be denoted as Ω, whose components are $\mathrm{Cov}(\varepsilon_0) = \sigma^2 I$, $\mathrm{Cov}(\boldsymbol{Y}_0) = \Gamma_0$ which is easily determined through the results in (1.29), and $\mathrm{Cov}(\varepsilon_0, \boldsymbol{Y}_0) = \sigma^2 C$ which is determined directly using the expansion for Y_{ij} as in (1.21). Also define

$$\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \alpha_3)', \boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3)', \text{ and } \boldsymbol{\xi} = \left(\boldsymbol{\alpha}', \boldsymbol{\theta}', \sigma^2\right)'$$

The model (1.20) for $\boldsymbol{Y}$ can be written as $A\boldsymbol{Y} = B\boldsymbol{\varepsilon} + F\boldsymbol{a}$, where A, B, and F are appropriately defined matrices, with A and B lower triangular having ones on the diagonal. Note that from (1.22), ε and $\boldsymbol{a}$ are uncorrelated and hence the covariance matrix of $\boldsymbol{Y}$ is

$$\Gamma = \mathrm{Cov}(\boldsymbol{Y}) = A^{-1}\left[\sigma^2 BB' + F\Omega F'\right]A'^{-1}.$$

Assuming normality, the p.d.f. of $\boldsymbol{Y}$ is

$$f(\boldsymbol{Y};\boldsymbol{\xi}) = |\Gamma|^{-1/2}\exp\left\{-(1/2)\boldsymbol{Y}'\Gamma^{-1}\boldsymbol{Y}\right\}$$

By use of a standard matrix inversion formula ([40], p. 33) for Γ^{-1}, the quadratic form in the likelihood function can be expressed as

$$\boldsymbol{Y}'\Gamma^{-1}\boldsymbol{Y} = \left(1/\sigma^2\right)\left[\boldsymbol{Y}'A'\left(BB'\right)^{-1}A\boldsymbol{Y} - \hat{\boldsymbol{a}}'\left(\sigma^2\Gamma_{\boldsymbol{a}}^{-1} \mid \boldsymbol{Y}\right)\hat{\boldsymbol{a}}\right], \tag{2.1}$$

where

$$\Gamma_{a|Y} = \sigma^2\left[F'B'^{-1}B^{-1}F + \sigma^2\Omega^{-1}\right]^{-1}$$

and

$$\hat{\boldsymbol{a}} = \mu_{a|Y} = \left(1/\sigma^2\right)\Gamma_{a|Y}F'B'^{-1}B^{-1}A\boldsymbol{Y}. \tag{2.2}$$

Note that the expressions for $\mu_{a|Y}$ and $\Gamma_{a|Y}$ are the conditional mean and covariance matrix of $\boldsymbol{a}$, given $\boldsymbol{Y}$. Also, it can be shown that

$$|\Gamma| = \left(\sigma^2\right)^{mn} \left|\left(1/\sigma^2\right)\Omega\right| \left|\sigma^2\Gamma_a^{-1}\right|\mathbf{y}\,|$$

and in addition, the form of Ω implies that

$$\left|\left(1/\sigma^2\right)\Omega\right| = \left|\left(1/\sigma^2\right)\Gamma_0 - C'C\right|$$

It also follows that in (2.1), the 'exact sum of squares' function can be written as

$$S(\boldsymbol{\alpha},\boldsymbol{\theta}) = \boldsymbol{Y}'A'\left(BB'\right)^{-1}A\boldsymbol{Y} - \hat{\boldsymbol{a}}'\left(\sigma^2\Gamma_a^{-1} \mid \boldsymbol{\gamma}\right)\hat{\boldsymbol{a}} = \hat{\boldsymbol{\varepsilon}}'\hat{\boldsymbol{\varepsilon}} + \hat{\boldsymbol{a}}'\left(\sigma^2\Omega^{-1}\right)\hat{\boldsymbol{a}}, \tag{2.3}$$

where $\hat{\boldsymbol{\varepsilon}} = B^{-1}(A\boldsymbol{Y} - F\hat{\boldsymbol{a}})$. So from the relation $B\hat{\boldsymbol{\varepsilon}} = A\boldsymbol{Y} - F\hat{\boldsymbol{a}}$, the $\hat{\varepsilon}_{ij}$ are computed recursively directly from the model equation (1.20) as

$$\hat{\varepsilon}_{ij} = Y_{ij} - \alpha_1\hat{Y}_{i-1,j} - \alpha_2\hat{Y}_{i,j-1} - \alpha_3\hat{Y}_{i-1,j-1} - \theta_1\hat{\varepsilon}_{i-1,j} - \theta_2\hat{\varepsilon}_{i,j-1} - \theta_3\hat{\varepsilon}_{i-1,j-1}$$

for $i = 1,\cdots,m, j = 1,\cdots,n$, with initial values $\hat{\varepsilon}_{ij} = \hat{\varepsilon}^0_{ij}$ for $i = 0$ or $j = 0$ and $\hat{Y}_{ij} = \hat{Y}^0_{ij}$ for $i = 0$ or $j = 0$ as obtained from (2.2). In addition, the initial values $\hat{a}$ in (2.2) are obtained as

$$\hat{\boldsymbol{a}} = \left(1/\sigma^2\right)\Gamma_{a|Y}F'B'^{-1}\varepsilon^0 = \left(1/\sigma^2\right)\Gamma_{a|Y}F'\boldsymbol{U},$$

where $\boldsymbol{\varepsilon}^0 = B^{-1}A\boldsymbol{Y}$ has elements ε^0_{ij} which are first calculated from the same recursion $B\boldsymbol{\varepsilon}^0 = A\boldsymbol{Y}$ as for $\hat{\boldsymbol{\varepsilon}}$, but with zero initial values $(a = 0)$, and then the elements of $\boldsymbol{U} = B'^{-1}\varepsilon^0$ are easily computed through the backward recursion

$$B'\boldsymbol{U} = \boldsymbol{\varepsilon}^0 \text{ as } U_{ij} = \varepsilon^0_{ij} - \theta_1U_{i+1,j} - \theta_2U_{i,j+1} - \theta_3U_{i+1,j+1}$$

for $i = 1,\cdots,m$ and $j = 1,\cdots,n$, with $U_{ij} = 0$ for $i = m+1$ or $j = n+1$.
Thus the explicit form of the likelihood function is

$$L\left(\boldsymbol{\alpha},\boldsymbol{\theta},\sigma^2 \mid \boldsymbol{Y}\right) = \left(\sigma^2\right)^{-mn/2}\left|\sigma^2\Omega^{-1} + F'\left(BB'\right)^{-1}F\right|^{-\frac{1}{2}}\left|\left(1/\sigma^2\right)\Gamma_0 - C'C\right|^{-\frac{1}{2}} \times \exp\left\{-\frac{1}{2\sigma^2}\left[\hat{\boldsymbol{\varepsilon}}'\hat{\boldsymbol{\varepsilon}} + \hat{\boldsymbol{a}}'\left(\sigma^2\Omega^{-1}\right)\hat{\boldsymbol{a}}\right]\right\}, \tag{2.4}$$

which is relatively easy to evaluate. In particular, for the first-order spatial AR model (all $\theta_i = 0$), it can be shown that the likelihood simplifies to

$$L\left(\boldsymbol{\alpha},\sigma^2 \mid \boldsymbol{Y}\right) = \left(\sigma^2\right)^{-mn/2}\Delta^{(m+n-1)/2}\left(1-\lambda^2\right)^{-(m-1)/2}\left(1-\mu^2\right)^{-(n-1)/2} \times \exp\left\{-\frac{1}{2\sigma^2}S(\boldsymbol{\alpha})\right\}, \tag{2.5}$$

where

$$\Delta = 1 - \alpha_1\lambda - \alpha_2\mu - \alpha_3\left(\alpha_1\mu + \alpha_2\lambda + \alpha_3\right)$$

is such that $\gamma = \operatorname{Var}\left(Y_{ij}\right) = \sigma^2/\Delta, \lambda$ and μ are obtained from (1.25) with $|\lambda|,|\mu| < 1$, and

$$S(\boldsymbol{\alpha}) = \sum_{i=2}^{m}\sum_{j=2}^{n}\varepsilon^2_{ij}$$

$$+\frac{\sigma^2}{\gamma}\left[\frac{1}{1-\lambda^2}\sum_{i=2}^{m}\left(Y_{i1}-\lambda Y_{i-1,1}\right)^2+\frac{1}{1-\mu^2}\sum_{j=2}^{n}\left(Y_{1j}-\mu Y_{1,j-1}\right)^2+Y_{11}^2\right]. \quad (2.6)$$

If in addition, we have $\alpha_3 = -\alpha_1\alpha_2$, then $\lambda = \alpha_1, \mu = \alpha_2, \Delta = \left(1-\alpha_1^2\right)\left(1-\alpha_2^2\right)$, and the determinant factor in the likelihood (2.5) is reduced to

$$\left(\sigma^2\right)^{-mn/2}\left(1-\alpha_1^2\right)^{n/2}\times\left(1-\alpha_2^2\right)^{m/2}$$

as in Martin [41]. Maximum likelihood estimates of the spatial unilateral first-order ARMA model (1.20) can be obtained by gradient algorithms, such as Newton-Raphson or scoring methods. Some explicit details on the score vector and information matrix for the spatial AR model are given by Basu and Reinsel [42]. The asymptotic properties of consistency and asymptotic normality of the maximum likelihood estimator can also be established. In particular, for the spatial unilateral first-order AR model, it can be shown for the ML estimator $\hat{\boldsymbol{\alpha}}$ that $(mn)^{\frac{1}{2}}(\hat{\boldsymbol{\alpha}}-\boldsymbol{\alpha})$ converges in distribution to $\mathrm{N}\left(\mathbf{0}, I(\boldsymbol{\alpha})^{-1}\right)$ as $m,n \to \infty$ (with $m/n \to$ constant > 0), where

$$I(\boldsymbol{\alpha}) = \frac{1}{\Delta}\begin{bmatrix} 1 & \lambda\mu & \mu \\ \lambda\mu & 1 & \lambda \\ \mu & \lambda & 1 \end{bmatrix}.$$

In the special case of the multiplicative AR model $(\alpha_3 = -\alpha_1\alpha_2)$, a similar asymptotic distribution result holds for $\hat{\boldsymbol{\alpha}} = (\hat{\alpha}_1, \hat{\alpha}_2)'$ with

$$I(\boldsymbol{\alpha}) = \mathrm{Diag}\left[\left(1-\alpha_1^2\right)^{-1}, \left(1-\alpha_2^2\right)^{-1}\right]$$

2.2.2 Yule-Walker Estimator

A special case that is of interest is the first-order spatial bilateral AR model (1.19). This model has been studied by [28], [41], and [37] studied the special case of $\alpha_{11} = -\alpha_{10}\alpha_{01}$. In this case we have $p_1 = 1$ and $p_2 = 1$, and the model is given by

$$Y_{ij} = \alpha_{01}Y_{i,j-1} + \alpha_{10}Y_{i-1,j} + \alpha_{11}Y_{i-1,j-1} + \epsilon_{ij} \quad (2.7)$$

for $i = 1,\ldots,m$ and $j = 1,\ldots,n$. Recall that

$$\begin{aligned} i^* &= \max(1, k-s+1), \\ m^* &= \min(m, m-s+k), \\ j^* &= \max(1, l-t+1), \\ n^* &= \min(n, n-t+l) \end{aligned}$$

and (s,t) and (k,l) take on only the values $(0,1),(1,0)$ and $(1,1)$ for the spatial YW estimation in this model. We will obtain explicit expressions for all the relevant terms $B(s,t,k,l), R^*(s,t,k,l)$, and $I^*(s,t,k,l)$ for the special case of first-order spatial bilateral AR model. The expressions for $B(s,t,k,l)$ are as follows:

$$mnB(0,1,0,1) = \sum_{i=1}^{m}\sum_{j=1}^{n-1}Y_{ij}^2 = \sum_{i=1}^{m}\sum_{j=1}^{n}Y_{ij}^2 - \sum_{i=1}^{m}Y_{in}^2$$

$$mnB(0,1,1,0) = \sum_{i=1}^{m}\sum_{j=1}^{n-1}Y_{ij}Y_{i-1,j+1} = \sum_{i=2}^{m}\sum_{j=1}^{n-1}Y_{ij}Y_{i-1,j+1} + \sum_{j=1}^{n-1}Y_{1j}Y_{0,j+1},$$

$$mnB(0,1,1,1) = \sum_{i=1}^{m}\sum_{j=1}^{n-1} Y_{ij}Y_{i-1,j} = \sum_{i=2}^{m}\sum_{j=1}^{n} Y_{ij}Y_{i-1,j} - \left(\sum_{i=1}^{m} Y_{in}Y_{i-1,n} - \sum_{j=1}^{n} Y_{1j}Y_{0j}\right)$$

$$\mathrm{mn}\,B(1,0,0,1) = \sum_{i=1}^{m-1}\sum_{j=1}^{n} Y_{ij}Y_{i+1,j-1} = \sum_{i=1}^{m-1}\sum_{j=2}^{n} Y_{ij}Y_{i+1,j-1} + \sum_{i=1}^{m-1} Y_{i1}Y_{i+1,0}$$

$$mnB(1,0,1,0) = \sum_{i=1}^{m-1}\sum_{j=1}^{n} Y_{ij}^2 = \sum_{i=1}^{m}\sum_{j=1}^{n} Y_{ij}^2 - \sum_{j=1}^{n} Y_{mj}^2$$

$$\mathrm{mn}\,B(1,0,1,1) = \sum_{i=1}^{m-1}\sum_{j=1}^{n} Y_{ij}Y_{i,j-1} = \sum_{i=1}^{m}\sum_{j=2}^{n} Y_{ij}Y_{i,j-1} + \left(\sum_{i=1}^{m} Y_{i1}Y_{i0} - \sum_{j=1}^{n} Y_{mj}Y_{m,j-1}\right)$$

$$mnB(1,1,0,1) = \sum_{i=1}^{m-1}\sum_{j=1}^{n-1} Y_{ij}Y_{i+1,j} = \sum_{i=1}^{m-1}\sum_{j=1}^{n} Y_{ij}Y_{i+1,j} - \sum_{i=1}^{m-1} Y_{in}Y_{i+1,n}$$

$$\mathrm{mn}\,B(1,1,1,0) = \sum_{i=1}^{m-1}\sum_{j=1}^{n-1} Y_{ij}Y_{i,j+1} = \sum_{i=1}^{m}\sum_{j=1}^{n-1} Y_{ij}Y_{i,j+1} - \sum_{j=1}^{n-1} Y_{mj}Y_{m,j+1}$$

$$mnB(1,1,1,1) = \sum_{i=1}^{m-1}\sum_{j=1}^{n-1} Y_{ij}^2 = \sum_{i=1}^{m}\sum_{j=1}^{n} Y_{ij}^2 - \left(\sum_{i=1}^{m} Y_{in}^2 + \sum_{j=1}^{n} Y_{mj}^2 - Y_{mn}^2\right)$$

where for each of the expressions above, the first term on the right-hand side corresponds to $mnR(s-k,t-l)$, while the remaining terms equal $mnR^*(s,t,k,l)$. It is shown in Basu and Reinsel [43] that for model

$$\gamma(1,0) = \lambda\gamma, \gamma(0,1) = \mu\gamma, \gamma(1,-1) = \lambda\mu\gamma,$$

where λ and μ are solved using the relations

$$\alpha_{10}\lambda + \alpha_{01}\mu^{-1} + \alpha_{11}\lambda\mu^{-1} = 1 = \alpha_{10}\lambda^{-1} + \alpha_{01}\mu + \alpha_{11}\lambda^{-1}\mu v \tag{2.8}$$

with $|\lambda| < 1$ and $|\mu| < 1$. Supposing that $m/n \to c^2, 0 < c < \infty$, as $m, n(mn)^{-1/2} = (m/n)^{-1/2}n^{-1} = (n/m)^{-1/2}{}^{-1}_{m}$ we obtain from the above relation and $\mathbf{R} \overset{P}{\to} \gamma I(\alpha)$ where

$$\gamma I(\alpha) = \begin{bmatrix} \gamma & \gamma(1,-1) & \gamma(1,0) \\ \gamma(1,-1) & \gamma & \gamma(0,1) \\ \gamma(1,0) & \gamma(0,1) & \gamma \end{bmatrix} = \gamma \begin{bmatrix} 1 & \lambda\mu & \lambda \\ \lambda\mu & 1 & \mu \\ \lambda & \mu & 1 \end{bmatrix} \tag{2.9}$$

$$\gamma I^* = \gamma \begin{bmatrix} -c & \frac{1}{c}\lambda\mu & \left(\frac{1}{c} - c\right)\lambda \\ c\lambda\mu & -\frac{1}{c} & \left(c - \frac{1}{c}\right)\mu \\ -c\lambda & -\frac{1}{c}\mu & \left(c + \frac{1}{c}\right) \end{bmatrix}.$$

Using the relation $I(\alpha)^{-1}r^*\alpha = b$, we obtain

$$b = \begin{bmatrix} -c & \frac{2\lambda\mu}{c(1-\lambda^2)} & \frac{2\lambda}{c(1-\lambda^2)} \\ \frac{2c\lambda\mu}{1-\mu^2} & -\frac{1}{c} & \frac{2c\mu}{1-\mu^2} \\ -\frac{2c\lambda\mu^2}{1-\mu^2} & -\frac{2\lambda^2\mu}{c(1-\lambda^2)} & -\left(\frac{1+\lambda^2}{c(1-\lambda^2)} + \frac{c(1+\mu^2)}{1-\mu^2}\right) \end{bmatrix} \begin{bmatrix} \alpha_{01} \\ \alpha_{10} \\ \alpha_{11} \end{bmatrix}. \tag{2.10}$$

Now using the relations (2.8) and simplifying equation (2.10), we obtain the elements of the asymptotic bias vector $b = (b_{01}, b_{10}, b_{11})$ in (6) explicitly as

$$b_{01} = -c\alpha_{01} + \frac{2\lambda}{c(1-\lambda^2)}(\mu\alpha_{10} + \alpha_{11}) = -\left(c + \frac{2}{c(1-\lambda^2)}\right)\alpha_{01} + \frac{2\mu}{c(1-\lambda^2)},$$

$$b_{10} = -\frac{1}{c}\alpha_{10} + \frac{2c\mu}{(1-\mu^2)}(\lambda\alpha_{01} + \alpha_{11}) = -\left(\frac{1}{c} + \frac{2c}{(1-\mu^2)}\right)\alpha_{10} + \frac{2c\lambda}{1-\mu^2},$$

$$b_{11} = -\left(\frac{c\left(1+\mu^2\right)}{1-\mu^2} + \frac{1+\lambda^2}{c(1-\lambda^2)}\right)\alpha_{11} - \frac{2c\lambda\mu^2}{1-\mu^2}\alpha_{01} - \frac{2\lambda^2\mu}{c(1-\lambda^2)}\alpha_{10}.$$

Hence we find that the asymptotic bias of the spatial YW estimator for this model may be particularly large when the roots of the characteristic equations (2.8), λ or μ, have an absolute value close to one. As a special case of model (2.7), consider the multiplicative AR(1) model where $\alpha_{11} = -\alpha_{10}\alpha_{01}$. In this special case, (2.10) can be simplified and the asymptotic biases of the spatial YW estimator reduce to

$$b_{01} = -c\alpha_{01}, \quad b_{10} = -\frac{1}{c}\alpha_{10}, \quad b_{11} = -\left(c + \frac{1}{c}\right)\alpha_{01}\alpha_{10}$$

This simple result indicates that the biases can be quite substantial when the grid sizes m and n are small or moderate.

2.3 Estimation in 2D-RCAR(p_1, p_2) Models

We tackle the problems concerning the estimation of the unknown parameters of the nonlinear model defined by the equation (2.11) which is identical to the model considered by Bibi [14], the following stochastic equation

$$X_{s,t} = \sum_{\substack{i=0 \\ (i,j)\neq(0,0)}}^{p_1}\sum_{j=0}^{p_2} (\zeta(i,j) + \zeta_{s,t}(i,j))X_{s-i,t-j} + \varepsilon_{s,t}, \quad (s,t) \in \mathbb{Z}^2. \tag{2.11}$$

For the model (2.11) the following assumptions are made by Bibi [14]:

- On the set of pairs of subscripts $(s,t); s,t \in \mathbb{Z}$, we set an order assuming that $(p,q) < (s,t)$ if $q < t$ or $q = t$ but $p < s$. Let us denote by $\mathcal{F}_{st}$ the $\sigma-$algebra generated by $\{(\eta_{p,q}, \varepsilon_{p,q});\ (p,q) < (s,t)\}$;
- Stationarity is indispensable for the rest of the work. A necessary and sufficient condition on the parameters $\underline{\zeta}$ and Λ such that model (2.11) to have a unique measurable stationary solution with respect to $\mathcal{F}_{s,t}$ has been presented and proved in detail in Proposition 2.1, [14].

In what follows we use linear algebra, matrix theory and the vectorization operator. Indeed, for any matrix X, $vec(X)$ stacks the columns of X one on top of the other. Moreover, for any square matrix X the half vector, $vech(X)$ is formed by vectorizing only the lower triangular part of X. Let

$$W_{st} = K_{m^2-1}vec(Y_{s,t-1}Y_{s,t-1}^{\top}),$$

where

$$Y_{s,t-1} = (X_{k,l}; (1,1) \leq (k,l) \ < (m,m))^{\top}.$$

In order to be able to use nonlinear random field models in practice, it is necessary to be able to fit the models to the data and to estimate the unknown parameters.

The computational procedure used to determine the unknown parameters of the process (2.11) is based on the optimization of the likelihood function which, under certain conditions, is shown to give strongly consistent estimates of the true parameters. Model (2.11) may be rewritten in the form of autoregressive process as:

$$X_{s,t} = \underline{\zeta}^{\top} Y_{s,t-1} + u_{s,t}, \tag{2.12}$$

where $u_{s,t}$ is the residual function defined as

$$u_{s,t} = \underline{\zeta}_{s,t} Y_{s,t-1} + \varepsilon_{s,t}$$

with $E(u_{s,t}|\mathcal{F}_{s,t}) = Y_{s,t-1}E(\underline{\zeta}_{s,t}) + E(\varepsilon_{s,t}) = 0$. Since $\underline{\zeta}_{s,t}$ and $\varepsilon_{s,t}$ are independent of the $\sigma-$field $\mathcal{F}_{s,t}$ and $X_{s,t}$ is $\mathcal{F}_{st}$-measurable. From the structure of the model (2.11), we have

$$E(X_{s,t}|\mathcal{F}_{s,t}) = \underline{\zeta}^{\top} Y_{s,t-1} \text{ and } E\left(u^2_{s,t}|\mathcal{F}_{s,t}\right) = \sigma^2 + \Gamma^{\top} W_{s,t}, \quad \Gamma = vech(\Lambda). \tag{2.13}$$

2.3.1 Construction of the Maximum Likelihood Estimates

Given a sample $\left(X_{k,l}; k = 1, ..., m; \ l = 1, ..., m\right)$ of size m^2 from the random field $\{X_{s,t}\}$ generated by the model (2.11) which is strictly stationary $\mathcal{F}_{s,t}$-measurable satisfying conditions (i) and (iv). To avoid cumbersome notations we put $\lambda = m^2$. The purpose is to obtain the likelihood function conditional on the values

$$\{X_{s,t}; 1 - p_1 < s \leq 0, 1 - p_2 < t \leq 0\}$$

as if we were assuming the joint normality of $\varepsilon_{s,t}$ and $\zeta_{s,t}$.

So that the derivation of the maximum likelihood estimates to be clear, we use the following step-by-step estimation procedure.
Step 1: given $A_{s-q_1,t-q_2} \in \mathcal{F}_{s,t}$, let

$$f(X_{s,t}, ..., X_{s-q_1,t-q_2}|A_{s-q_1,t-q_2+1})$$

denotes the density of $X_{s,t}, ..., X_{s-q_1,t-q_2+1}$. Hence, it follows from (2.13) that

$$\begin{aligned} &f_{mm}\{X_{11}, ..., X_{m,m}|X_{0,1}, \dots, X_{1-p_1,1-p_2}\} \\ &= \prod_{s=1}^{m}\prod_{t=1}^{m} f(X_{s,t}|X_{0,1}, \dots, X_{1-p_1,1-p_2}) \\ &= (2\pi)^{-\frac{\lambda}{2}} \prod_{s=1}^{m}\prod_{t=1}^{m} (\sigma^2 + \Gamma W_{s,t})^{-\frac{1}{2}} \times \exp\left\{-\frac{1}{2}\frac{(X_{s,t} - \underline{\zeta}^{\top} Y_{s,t-1})^2}{\sigma^2 + \Gamma^{\top} W_{s,t}}\right\} \\ &= \tilde{L}_{mm}\left(\underline{\zeta}, \Gamma, \sigma^2\right), \end{aligned}$$

the resulting $\tilde{L}_{mm}\left(\underline{\zeta}, \Gamma, \sigma^2\right)$ is the likelihood function conditional on $\{X_{00}, \dots X_{1-p_1,1-p_2}\}$. However, instead of the maximization of this latter, it will be more practical and appropriate to think to the minimization of the monotone function of the likelihood $\tilde{\mathcal{L}}_{mm}\left(\underline{\zeta}, \Gamma, \sigma^2\right)$, namely

$$\begin{aligned}\tilde{\mathcal{L}}_{mm}\left(\underline{\zeta}, \Gamma, \sigma^2\right) &= -2\lambda^{-1}\ln\left(\tilde{L}_{mm}\left(\underline{\zeta}, \Gamma, \sigma^2\right)\right) - \ln(2\pi)\\ &= \lambda^{-1}\sum_{s=1}^{m}\sum_{t=1}^{m}\ln(\sigma^2 + \Gamma^\top W_{st}) + \lambda^{-1}\sum_{s=1}^{m}\sum_{t=1}^{m}\frac{(X_{s,t} - \underline{\zeta}^\top Y_{s,t-1})^2}{\sigma^2 + \Gamma^\top W_{s,t}}.\end{aligned}$$

It will be seen that the likelihood function to be optimized is nonlinear in the parameters (to be estimated) so it is necessary to use the following procedure to find the estimates.

Step 2: the function $\tilde{\mathcal{L}}_{mm}\left(\underline{\zeta}, \Gamma, \sigma^2\right)$ is nonlinear in σ^2 and Γ and there is no fully defined expression for the estimates $\hat{\underline{\zeta}}_{mm}$, $\hat{\Gamma}_{mm}$ and $\hat{\sigma}^2_{mm}$ of $\underline{\zeta}$, Γ and σ^2, respectively, which minimize $\tilde{\mathcal{L}}_{mm}\left(\underline{\zeta}, \Gamma, \sigma^2\right)$. Hence, by letting $\varrho = \frac{\Gamma}{\sigma^2}$, we may equivalently minimize a function of ϱ alone by concentrating out the parameters $\underline{\zeta}$ and σ^2.

$$\begin{aligned}\tilde{L}_{mm}\left(\underline{\zeta}, \varrho, \sigma^2\right) &= \ln(\sigma^2) + \lambda^{-1}\sum_{s=1}^{m}\sum_{t=1}^{m}\ln(1 + \varrho^\top W_{s,t})\\ &\quad + \sigma^{-2}\lambda^{-1}\sum_{s=1}^{m}\sum_{t=1}^{m}\frac{(X_{s,t} - \underline{\zeta}^\top Y_{s,t-1})^2}{1 + \varrho^\top W_{s,t}}.\end{aligned} \tag{2.14}$$

For fixed ϱ, the function $\tilde{L}_{mm}\left(\underline{\zeta}, \varrho, \sigma^2\right)$ reaches its minimum at $\underline{\zeta} = \underline{\zeta}_{mm}(\varrho)$ and $\sigma^2 = \sigma^2_{mm}(\varrho)$, where

$$\underline{\zeta}_{mm}(\varrho) = \left(\lambda^{-1}\sum_{s=1}^{m}\sum_{t=1}^{m}\frac{Y_{s,t-1}Y^\top{}_{s,t-1}}{1 + \varrho^\top W_{s,t}}\right)^{-1} \times \lambda^{-1}\sum_{s=1}^{m}\sum_{t=1}^{m}\frac{X_{s,t}Y_{s,t-1}}{1 + \varrho^\top W_{s,t}},$$

and

$$\sigma^2_{mm}(\varrho) = \lambda^{-1}\sum_{s=1}^{m}\sum_{t=1}^{m}\frac{(X_{s,t} - \underline{\zeta}^\top{}_{mm}(\varrho)Y_{s,t-1})^2}{1 + \varrho^\top W_{s,t}}.$$

Finally, the maximum likelihood estimates $\hat{\underline{\zeta}}_{mm}$, $\hat{\Gamma}_{mm}$ and $\hat{\sigma}^2_{mm}$ are obtained by calculating $\hat{\varrho}_{mm}$, where $\hat{\varrho}_{mm}$ minimizes the function

$$L^*_{mm}(\varrho) = \ln(\sigma^2_{mm}(\varrho)) + \lambda^{-1}\sum_{s=1}^{m}\sum_{t=1}^{m}\ln(1 + \hat{\varrho}_{mm}W_{s,t}).$$

The estimates $\hat{\underline{\zeta}}_{mm}$, $\hat{\Gamma}_{mm}$ and σ^2_{mm} are given by $\hat{\sigma}^2_{mm} = \sigma^2_{mm}(\hat{\varrho})$, $\hat{\underline{\zeta}}_{mm} = \underline{\zeta}_{mm}(\hat{\varrho})$ and $\hat{\Gamma}_{mm} = \hat{\sigma}^2_{mm}\hat{\varrho}_{mm}$. Note that for λ large enough, the matrix

$$\lambda^{-1}\sum_{s=1}^{m}\sum_{t=1}^{m}\frac{(Y_{s,t-1}Y^\top_{s,t-1})}{1 + \varrho^T_{mm}W_{s,t}}$$

is invertible almost everywhere since it is always positive definite. In the opposite case, if it was not positive definite, it would exist a non zero $(p_1 \times p_2 - 1)$-component vector κ such that $\kappa^\top Y_{s,t-1} = 0$ almost everywhere.

Remark 2.1

The procedure above is useful if we use an optimal algorithm which doesn't require the first and second derivatives of the function $L^*_{mm}(\varrho)$. For this reason, it is better to minimize $\tilde{L}_{mm}(\underline{\zeta}, \varrho)$, indeed we have

$$\begin{aligned} \tilde{L}_{mm}(\underline{\zeta}, \varrho) &= \inf_{\sigma^2} \tilde{L}_m(\underline{\zeta}, \varrho, \sigma^2) - 1 \\ &= \lambda^{-1} \sum_{s=1}^{m} \sum_{t=1}^{m} \ln(1 + \varrho^\top W_{s,t}) \\ &\quad + \lambda^{-1} \ln \left(\sum_{s=1}^{m} \sum_{t=1}^{m} \frac{(X_{s,t} - \underline{\zeta}^\top Y_{s,t-1})^2}{1 + \varrho^\top W_{s,t}} \right). \end{aligned} \tag{2.15}$$

Finally, the maximum likelihood estimates $\hat{\zeta}_{mm}$, $\hat{\varrho}_{mm}$ and $\hat{\sigma}^2_{mm}$ are obtained from the following equations:

$$\tilde{L}_{mm}(\hat{\underline{\zeta}}_{mm}, \hat{\varrho}_{mm}) = \inf_{(\underline{\zeta}^\top, \varrho^\top) \in \Theta} \tilde{L}_{mm}(\underline{\zeta}, \varrho);$$

$$\hat{\sigma}^2_{mm}(\hat{\varrho}) = \lambda^{-1} \sum_{s=1}^{m} \sum_{t=1}^{m} \frac{(X_{s,t} - \hat{\underline{\zeta}}_{mm} Y_{s,t-1})^2}{1 + \varrho^\top_{mm} W_{s,t}};$$

$$\hat{\Gamma}_{mm} = \hat{\sigma}^2_{mm} \hat{\varrho}_{mm},$$

where the set Θ is the set of all $\frac{\lambda^2+3\lambda}{2}$ dimensional vectors $(\underline{\zeta}^\top, \varrho^\top)^\top$, with $\underline{\zeta}$ and ϱ containing λ and $\frac{\lambda^2+\lambda}{2}$ elements, respectively. It is the set over which $\tilde{L}_{mm}(\underline{\zeta}_{mm}, \varrho_{mm})$ is to be minimized.

Note that the same estimates will be obtained by minimizing either function (2.14) or (2.15). Using function (2.14) is suitable only for optimization algorithms that don't require the first and second derivatives. The calculation of the derivatives of the function (2.14) seems to complicated and may lead to some loss of accuracy of the estimation procedure. Therefore it is better to minimize the function (2.15) of $\underline{\zeta}$ and ϱ.

2.3.2 Strong Consistency of the Maximum Likelihood Estimates

In this section we prove the strong consistency of the maximum likelihood estimates $\hat{\underline{\zeta}}_{mm}$, $\hat{\Gamma}_{mm}$ and $\hat{\sigma}^2_{mm}$. Suppose that the real value $\theta_{0,0} = (\underline{\zeta}^\top{}_{0,0}, \varrho^\top{}_{0,0})^\top$ and that $\{X_{s,t}\}$ strictly stationary, and $\mathcal{F}_{s,t}$−measurable process $X_{s,t}$ satisfying (2.11) with $\underline{\zeta} = \underline{\zeta}_{0,0}$, $\sigma^2 = \sigma^2_{0,0}$ and $\Gamma = \Gamma_{0,0}$. Assumption **(iv)** is necessary as the vector of parameters of the 2D-RCA model may lie on the natural boundary of the parameter space, which causes difficulties when obtaining consistency of estimates.

An Additional assumption is needed for the analytical development of the main results in this paper. We need the set Θ to be compact so that several results from real analysis can be used. In fact, we show that Θ is compact by showing that it is bounded and closed, which follows from the construction of the model (2.11) using Borel's theorem. Moreover, the assumptions on the parameters of model (2.11) are made in such a way to guarantee that set Θ contains the vector of true values $\theta_{0,0} = (\underline{\zeta}^{\top}{}_{0,0}, \varrho^{\top}{}_{0,0})^{\top}$ within its interior.

We introduce the following lemma which is necessary to address the problem of concern in this section.

Lemma 2.1

Under assumptions **(i)**, **(iii)** and **(iv)**, there is no vector κ and constant K with $\kappa^t Y_{s,t-1} = K$ almost surely.

Proof: The proof is technical and very long. It is based on the rewriting of the Model (2.11) in the form $X_{s,t} = \underline{\zeta}^{\top} Y_{s,t-1} + u_{s,t}$, under assumption **(i)**, where $u_{s,t}$ is the residual function defined as $u_{s,t} = \underline{\zeta}_{s,t} Y_{s,t-1} + \varepsilon_{s,t}$. □

Now, we have gathered all the necessary tools to prove the strong consistency of the estimates. The following theorem provides the main result required in the proof of the strong consistency of $\hat{\underline{\zeta}}_{mm}$, $\hat{\Gamma}_{mm}$ and $\hat{\sigma}^2_{mm}$.

Theorem 2.1

For a strictly stationary, and $\mathcal{F}_{s,t}$−measurable process $X_{s,t}$ satisfying (2.11) with $\underline{\zeta} = \underline{\zeta}_{0,0}$, $\sigma^2 = \sigma^2_{0,0}$ and $\Gamma = \Gamma_{0,0}$ under assumptions **(i)**, **(iv)**, and let $\theta_{0,0} = (\underline{\zeta}^{\top}{}_{0,0}, \varrho^{\top}{}_{0,0})^{\top}$, where $\varrho_{0,0} = \frac{\Gamma}{\sigma^2_{0,0}}$. Then $\lim\limits_{m\to\infty} \tilde{L}_{mm}(\underline{\zeta}, \varrho)$ exists almost surely for all $(\underline{\zeta}^{\top}, \varrho^{\top})^{\top} \in \Theta$ and the limit $\tilde{L}(\underline{\zeta}, \varrho)$ is uniquely minimized over Θ at $(\underline{\zeta}^{\top}, \varrho^{\top})^{\top} = \theta_{0,0}$ provided that $\theta_{0,0} \in int(\Theta)$.

Proof: we have $0 \le \ln(1 + \varrho^{\top} W_{s,t}) \le \varrho^{\top} W_{s,t}$ and $E(W_{s,t})$ exists. The ergodic theorem ensures that $\lambda^{-1} \sum_{s=1}^{m} \sum_{t=1}^{m} (1 + \varrho^{\top} W_{s,t})$ converges almost surely to $E(1 + \varrho^{\top} W_{s,t})$, furthermore, we have that

$$0 < \lambda^{-1} \sum_{s=1}^{m} \sum_{t=1}^{m} \frac{(X_{s,t} - \underline{\zeta}^{\top} Y_{s,t-1})^2}{1 + \varrho^{\top} W_{s,t}} \le \lambda^{-1} \sum_{s=1}^{m} \sum_{t=1}^{m} (X_{s,t} - \underline{\zeta}^{\top} Y_{s,t-1})^2,$$

We thus have that

$$\lambda^{-1} \sum_{s=1}^{m} \sum_{t=1}^{m} \frac{(X_{s,t} - \underline{\zeta}^{\top} Y_{s,t-1})^2}{1 + \varrho^{\top} W_{s,t}} \xrightarrow{a.s} E\left(\frac{(X_{s,t} - \underline{\zeta}^{\top} Y_{s,t-1})^2}{1 + \varrho^{\top} W_{s,t}} \right).$$

As the right-hand side of the above is strictly greater than zero, for otherwise we would have $X_{s,t} = \underline{\zeta}^{\top} Y_{s,t-1}$ almost surely. So the function $\tilde{L}_{mm}(\underline{\zeta}, \varrho)$ converges almost surely to

$$\tilde{L}(\underline{\zeta}, \varrho) = E\left(\ln(1 + \varrho^{\top} W_{s,t}) + \ln E\left(\frac{(X_{s,t} - \underline{\zeta}^{\top} W_{s,t})^2}{1 + \varrho^{\top} W_{s,t}} \right) \right).$$

On the other hand we have

$$\begin{aligned} E\left(\frac{(X_{s,t}-\underline{\zeta}^\top W_{s,t})^2}{1+\varrho^\top W_{s,t}}\right) &= E\left(\frac{\left(X_{s,t}-\underline{\zeta}^\top{}_{0,0}Y_{s,t-1}+(\underline{\zeta}_{0,0}-\underline{\zeta})^\top Y_{s,t-1}\right)^2}{1+\varrho^\top W_{s,t}}\right) \\ &\geq E\left(\frac{\sigma^2_{0,0}+\Gamma^\top{}_{0,0}W_{s,t}}{1+\varrho^\top w_{s,t}}\right)=\sigma^2_{0,0}E\left(\frac{1+\varrho^\top{}_{0,0}W_{s,t}}{1+\varrho^\top W_{s,t}}\right), \end{aligned}$$

since

$$E\left(X_{s,t}-\underline{\zeta}^\top{}_{0,0}Y_{s,t-1}|\mathcal{F}_{s,t}\right)=0$$

and

$$E\left(\left(\underline{\zeta}_{s,t}Y_{s,t-1}+\varepsilon_{s,t}\right)^2|\mathcal{F}_{s,t}\right)=\sigma^2_{0,0}+\Gamma^\top{}_{0,0}W_{s,t}$$

and $\Gamma_{0,0}=\sigma^2_{0,0}\varrho_{0,0}$. Moreover, the previous equality hold only when $(\underline{\zeta}_{0,0}-\underline{\zeta})^\top Y_{s,t-1}=0$ almost surely that is when $\underline{\zeta}=\underline{\zeta}_{0,0}$, and, as a result, we have that

$$\begin{aligned} \inf_{\underline{\zeta}}\tilde{L}(\underline{\zeta},\varrho) &= \tilde{L}(\underline{\zeta}_{0,0},\varrho) \\ &= \ln(\sigma^2_{0,0})+\ln\left(\frac{E(1+\varrho^\top{}_{0,0}W_{s,t})}{1+\varrho^\top W_{s,t}}\right)+E\left(\ln(1+\varrho^\top W_{s,t})\right), \end{aligned}$$

and $\tilde{L}(\underline{\zeta},\varrho)=\inf_{\underline{\zeta}}\tilde{L}(\underline{\zeta},\varrho)$ only at $\underline{\zeta}=\underline{\zeta}_{0,0}$. We have that

$$E\left(\ln\frac{(1+\varrho^\top_{0,0}W_{s,t})}{1+\varrho^\top W_{s,t}}\right)\geq-\ln E\left(\frac{(1+\varrho^\top{}_{0,0}W_{s,t})}{1+\varrho^\top W_{s,t}}\right)$$

with equality only if we have

$$(1+\varrho^\top{}_{0,0}W_{s,t})=E\left(\frac{1+\varrho^\top_{0,0}W_{s,t}}{1+\varrho^\top W_{s,t}}\right)(1+\varrho^\top W_{s,t})$$

almost surely, that is when

$$\left(\varrho_{0,0}-\varrho E\left(\frac{1+\varrho^\top_{0,0}W_{s,t}}{1+\varrho^\top W_{s,t}}\right)\right)W_{s,t}=E\left(\frac{1+\varrho^\top_{0,0}W_{s,t}}{1+\varrho^\top W_{s,t}}\right)-1$$

almost surely. By Lemma 1, the last claim occurs only when

$$\varrho_{0,0}=\varrho E\left(\frac{1+\varrho^\top_{0,0}W_{s,t}}{1+\varrho^\top W_{s,t}}\right)$$

and

$$E\left(\frac{1+\varrho^\top_{0,0}W_{s,t}}{1+\varrho^\top W_{s,t}}\right)=1$$

that is, when $\varrho=\varrho_{0,0}$. Hence, $\tilde{L}_{mm}(\underline{\zeta},\varrho)$ is uniquely minimized at $\underline{\zeta}=\underline{\zeta}_{0,0}$ and $\varrho=\varrho_{0,0}$. □

Corollary 2.1

$\lim_{m\to\infty} \tilde{L}_{mm}(\underline{\zeta}, \Gamma, \sigma^2)$ exists almost surely and is minimized uniquely at $\underline{\zeta} = \underline{\zeta}_{0,0}$, $\Gamma = \Gamma_{0,0}$, $\sigma = \sigma^2_{0,0}$.

Proof: From Theorem 1 and expression of $\tilde{L}_{mm}(\underline{\zeta}, \varrho)$ we have that $\lim_{m\to\infty} \tilde{L}_{mm}(\underline{\zeta}, \Gamma, \sigma^2)$ is seen to exist almost everywhere and to be uniquely minimized at $\underline{\zeta} = \underline{\zeta}_{0,0}$. Furthermore, we have that

$$\Gamma = \varrho_{0,0} E\left(\frac{\left(X_{s,t} - \underline{\zeta}^\top{}_{0,0} Y_{s,t-1}\right)^2}{1 + \varrho^\top{}_{0,0} W_{s,t}}\right),$$

and we also have

$$\begin{aligned} E\left(\frac{\left(X_{s,t} - \underline{\zeta}^\top{}_{0,0} Y_{s,t-1}\right)^2}{1 + \varrho^\top{}_{0,0} W_{s,t}}\right) &= E\left(\frac{E(\left(X_{s,t} - \underline{\zeta}^\top{}_{0,0} Y_{s,t-1}\right)^2 | \mathcal{F}_{s,t})}{1 + \varrho^\top{}_{0,0} W_{s,t}}\right) \\ &= E\left(\frac{\sigma^2_{0,0} + \Gamma^\top{}_{0,0} W_{s,t}}{1 + \varrho^\top{}_{0,0} W_{s,t}}\right) = \sigma^2_{0,0}, \end{aligned}$$

and so $\lim_{m\to\infty} \tilde{L}_{mm}(\underline{\zeta}, \Gamma, \sigma^2)$ is uniquely minimized at $\underline{\zeta} = \underline{\zeta}_{0,0}$; $\Gamma = \Gamma_{0,0}$ and $\sigma^2 = \sigma^2_{0,0}$. □

We are now in a position to prove the strong consistency of the procedure. The following theorem states that the estimation procedure by minimizing the function (2.15) over $(\underline{\zeta}\top, \varrho\top) \in int(\Theta)$ is strongly consistent when $\theta_{00} = (\underline{\zeta\top}_{00}, \varrho\top_{00}) \in int(\Theta)$.

Theorem 2.2

Let $\tilde{L}_{mm}(\underline{\zeta}, \varrho)$ be minimized over Θ at $\underline{\zeta} = \hat{\underline{\zeta}}_{mm}$, $\varrho = \hat{\varrho}_{mm}$ and $\hat{\underline{\theta}}_{mm} = (\hat{\zeta}^\top{}_{mm}, \hat{\varrho}^\top{}_{mm})^\top$. Then, $\hat{\theta}_{mm}$ converge almost surely to $\theta_{00} = (\underline{\zeta}_{00}, \varrho_{00})$ provided that $\theta_{00} \in int(\Theta)$.

Proof: First, we show that $\tilde{L}_{mm}(\underline{\zeta}, \varrho)$ converges uniformly almost surely to $L(\underline{\zeta}, \varrho)$ on Θ. As Θ is compact, it suffices to show that $\tilde{L}_{mm}(\underline{\zeta}, \varrho)$ is equicontinuous almost surely, or letting $\theta = (\underline{\zeta}^\top, \varrho^\top)^\top$, that given $\varepsilon > 0$, there exists an integer m and a positive number δ, both depending on ε, such that $|\tilde{L}_{mm}(\theta_1) - \tilde{L}_{mm}(\theta_2)| < \varepsilon$ almost surely for $(m, m) > (m^*, m^*)$ whenever $\|\theta_1 - \theta_2\| < \delta$. Now, since $\tilde{L}_{mm}(\theta)$ is differentiable on Θ and by the mean value theorem, for each $\theta_1, \theta_2 \in \Theta$ we have that

$$|\tilde{L}_{mm}(\theta_1) - \tilde{L}_{mm}(\theta_2)| = (\theta_1 - \theta_2)^\top \frac{\partial}{\partial\theta} \tilde{L}_{mm}(\theta^*_{12}),$$

where $\theta^*_{12} = \lambda_1\theta_1 + (1 - \lambda_1)\theta_2$ for some $\lambda_1 \in [0, 1]$. Then, let $\Theta^* = f(\Theta, \Theta, [0, 1])$, where $f : \mathbb{R}^{(\lambda^2+3\lambda)/2} \times \mathbb{R}^{(\lambda^2+3\lambda)/2} \times \mathbb{R} \to \mathbb{R}^{(\lambda^2+3\lambda)/2}$ is a continuous function defined by $f(\theta_1, \theta_2, \lambda) = \lambda\theta_1 + (1-\lambda)\theta_2$. Then the set Θ^* is compact since $\Theta \times \Theta \times [0,1]$ is compact and as we also have

$$|(\theta_1 - \theta_2)^\top \frac{\partial}{\partial\theta}\tilde{L}_{mm}(\theta_{12}^*)|^2 \leq \|\theta_1 - \theta_2\| \; \| \frac{\partial}{\partial\theta}\tilde{L}_{mm}(\theta_{12}^*) \|,$$

it follows that $\tilde{L}_{mm}(\theta)$ is equicontinous if $\lim_{m\to\infty} \sup_{\theta\in\Theta^*} \|\frac{\partial}{\partial\theta}\tilde{L}_{mm}(\theta)\|$ is finite almost surely. The vector $\frac{\partial}{\partial\theta}\tilde{L}_{mm}(\theta)$ is obtained from the fact that

$$\frac{\partial}{\partial\underline{\zeta}}\tilde{L}_{mm}(\theta) = -2(\sigma_{mm}^2(\theta))^{-1}\lambda^{-1}\sum_{s=1}^{m}\sum_{t=1}^{m}\frac{\left(X_{s,t} - \underline{\zeta}^\top Y_{s,t-1}\right)Y_{s,t-1}}{1+\varrho^\top Ws,t},$$

and we also have

$$\frac{\partial}{\partial\varrho}\tilde{L}_{mm}(\theta) = \sum_{s=1}^{m}\sum_{t=1}^{m}\frac{W_{st}}{1+\varrho^\top Ws,t} - (\sigma_{mm}^2(\theta))^{-1}\lambda^{-1}\sum_{s=1}^{m}\sum_{t=1}^{m}\frac{\left(X_{s,t} - \underline{\zeta}^\top Y_{s,t-1}\right)Y_{s,t-1}}{1+\varrho^\top Ws,t},$$

given that

$$\sigma_{mm}^2(\theta) = \lambda^{-1}\sum_{s=1}^{m}\sum_{t=1}^{m}\frac{\left(X_{s,t} - \underline{\zeta}^\top Y_{s,t-1}\right)^2}{1+\varrho^\top Ws,t}.$$

On the other hand we have $\lim_{m\to\infty}\sup_{\theta\in\theta^*} \| \frac{\partial}{\partial\theta}\tilde{L}_{mm}(\theta) \|$ is finite almost surely can now be proved as follows

$$\inf_{\theta\in\Theta^*}\sigma_{mm}^2(\theta) \geq \lambda^{-1}\inf_{\theta\in\Theta^*}\sum_{s=1}^{m}\sum_{t=1}^{m}\frac{(X_{s,t} - \underline{\zeta}^\top Y_{s,t-1})^2}{1+\mathbf{k}(W_{s,t}^\top W_{s,t})^{\frac{1}{2}}},$$

where $\mathbf{k} = \sup_{\theta\in\Theta^*}(\varrho^\top\varrho)^{\frac{1}{2}}$, this last term exists because Θ^* is bounded. Hence,

$$\inf_{\theta\in\Theta^*}\sigma_{mm}^2(\theta) \geq \lambda^{-1}\sum_{s=1}^{m}\sum_{t=1}^{m}\frac{(X_{s,t} - \underline{\zeta}_{mm}^{*\top} Y_{s,t-1})^2}{1+\mathbf{k}(W_{s,t}^\top W_{s,t})^{\frac{1}{2}}},$$

where

$$\underline{\zeta}_{mm}^* = \left(\lambda^{-1}\sum_{s=1}^{m}\sum_{t=1}^{m}\frac{Y_{s,t-1}Y_{s,t-1}^\top}{1+K(W_{s,t}^\top W_{s,t})}\right)^{-1}\left(\lambda^{-1}\sum_{s=1}^{m}\sum_{t=1}^{m}\frac{Y_{s,t-1}X_{s,t}}{1+k(W_{s,t}^\top W_{s,t})^{\frac{1}{2}}}\right).$$

By using the ergodic theorem we can see that $\underline{\zeta}_{mm}^*$ converges almost surely to $\underline{\zeta}_{00}$ and so

$$\lambda^{-1}\sum_{s=1}^{m}\sum_{t=1}^{m}\frac{(X_{s,t} - \underline{\zeta}_{mm}^{*\top} Y_{s,t-1})^2}{1+\mathbf{k}(W_{s,t}^\top W_{s,t})^{\frac{1}{2}}} \xrightarrow{a.s} E\left(\frac{(X_{s,t} - \underline{\zeta}_{mm}^{*\top} Y_{s,t-1})^2}{1+\mathbf{k}(W_{s,t}^\top W_{s,t})^{\frac{1}{2}}}\right)$$

and therefore

$$\lim_{m\to\infty}\inf_{\theta\in\theta^*}\sigma_{mm}^2(\theta) > 0.$$

Since $\{\tilde{L}_{mm}(\theta)\}$ converges uniformly for any $\varepsilon > 0$: there exists an integer λ^* depending on ε such that

$$|\tilde{L}_{mm}(\hat{\theta}_{mm}) - \tilde{L}(\hat{\theta}_{mm})| < \frac{\varepsilon}{2} \quad \text{and} \quad |\tilde{L}_{mm}(\theta_{00}) - \tilde{L}(\theta_{00})| < \frac{\varepsilon}{2}$$

almost surely whenever $\lambda > \lambda^*$. Thus, since $\tilde{L}_{mm}(\hat{\theta}_{mm}) \leq \tilde{L}_{mm}(\theta_{00})$ and $L(\hat{\theta}_{mm}) \geq L(\theta_{00})$ it follows that

$$\begin{aligned} 0 \leq L(\hat{\theta}_{mm}) - L(\theta_{00}) &= \{L(\hat{\theta}_{mm}) - \tilde{L}_{mm}(\hat{\theta}_{mm})\} + \{\tilde{L}_{mm}(\hat{\theta}_{mm}) - \tilde{L}_{mm}(\theta_{00})\} \\ &+ \{\tilde{L}_{mm}(\hat{\theta}_{00}) - L(\theta_{00})\} \\ &\leq \frac{\varepsilon}{2} + 0 + \frac{\varepsilon}{2} \end{aligned}$$

almost surely when $\lambda > \lambda^*$ hence $\{L(\hat{\theta}_{mm})\}$ converges almost surely to $L(\theta_{00})$.

Now suppose that $\hat{\theta}_{mm}$ does not converge almost surely to $\tilde{L}_{mm}(\theta_{00})$. Then it is possible to find a positive δ and an infinite subsequence $\{\hat{\theta}_{m_i m_j}\}$ of $\{\hat{\theta}_{mm}\}$ for which $|\hat{\theta}_{m_i m_j} - \theta_{00}| < \delta$ for all i, j. Since Θ is compact, there exists an infinite subsequence $\{\hat{\theta}_{m'_i m'_j}\}$ of $\{\hat{\theta}_{m_i m_j}\}$ which converges to θ^*, where $|\theta^* - \theta_{00}| > \delta$. Thus, since $L(\theta)$ is continuous, it follows that

$$\lim_{\substack{m'_i \to \infty \\ m'_j \to \infty}} L(\hat{\theta}_{m'_i m'_j}) = L(\theta^*) \neq L(\theta_{00}),$$

since θ_{00} is the unique minimizer of $L(\theta)$ by Theorem 1. However,

$$\lim_{\substack{m'_i \to \infty \\ m'_j \to \infty}} L(\hat{\theta}_{m'_i m'_j}) = \lim_{m \to \infty} L(\hat{\theta}_{mm}) = L(\theta_{00})$$

since $L(\hat{\theta}_{mm})$ converges almost surely to $L(\theta_{00})$, this contradiction implies that $\hat{\theta}_{mm}$ converge almost surely to θ_{00}. □

Theorem 2.3

The estimates $\hat{\Gamma}_{mm}$ and $\hat{\sigma}^2_{mm}$ converge almost surely to $\Gamma_{0,0}$ and $\sigma^2_{0,0}$, respectively.

Proof: Theorem 2 specifies that $\hat{\varrho}_{mm}$ and $\hat{\underline{\zeta}}_{mm}$ converge almost surely to $\varrho_{0,0}$ and $\underline{\zeta}_{0,0}$, respectively. By Theorem 2, we have,

$$\sigma^2_{mm}(\theta) = \lambda^{-1} \sum_{s=1}^{m} \sum_{t=1}^{m} \frac{\left(X_{s,t} - \underline{\zeta}^\top Y_{s,t-1}\right)^2}{1 + \varrho^\top W s, t}$$

converges uniformly and almost surely on Θ to

$$\sigma^2(\theta) = E\left\{\lambda^{-1} \sum_{s=1}^{m} \sum_{t=1}^{m} \frac{\left(X_{s,t} - \underline{\zeta}^\top Y_{s,t-1}\right)^2}{1 + \varrho^\top W s, t}\right\}.$$

It is also clear that $\hat{\sigma}^2_{mm} \xrightarrow{a.s} \sigma^2(\theta_{0,0})$ which was seen in the proof of Theorem 2 to be equal to $\sigma^2_{0,0}$. Moreover, as we have $\hat{\Gamma}_{mm} = \hat{\varrho}_{mm}\hat{\sigma}^2_{mm}$ then $\hat{\Gamma}_{mm} \xrightarrow{a.s} \Gamma_{0,0} = \varrho_{0,0}\, \sigma^2_{0,0}$.
□

2.4 Application of the Estimation procedure on a 2D-RCAR$(0,1)$ Model

2.4.1 Model and Stationarity Condition

Herein, we are interested specifically in the two dimensional indexed random coefficient autoregressive model of order $(0,1)$ (i.e, $2D-RCAR(0,1)$) generating a random field $\{X_{s,t}\}$ which obeys the difference equation

$$X_{s,t} = \begin{cases} (a+\eta_{s,t})\mathrm{X}_{s,t-1}+\varepsilon_{s,t}, & \text{if } (s,t)\geq(0,1) \\ 0, & \text{otherwise,} \end{cases} \tag{2.16}$$

where a is the parameter of the autoregression model (the nonrandom component of the autoregression coefficient $a+\eta_{s,t}$). We consider throughout the paper the assumption:

Assumption 2.1

The pairs of random variables $\begin{pmatrix}\eta_{s,t}\\ \varepsilon_{s,t}\end{pmatrix}$, $(s,t)\geq(0,1)$ are independent and identically normal distributed on some probability space $(\Omega,\mathcal{F},P)$ with

$$E\begin{pmatrix}\eta_{s,t}\\ \varepsilon_{s,t}\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix} \text{ and } var\begin{pmatrix}\eta_{s,t}\\ \varepsilon_{s,t}\end{pmatrix}=\begin{pmatrix}\eta^2 & 0\\ 0 & \sigma^2\end{pmatrix}, \eta^2,\sigma^2>0.$$

Model (2.11) is solved by backward substitution yielding,

$$\mathrm{X}_{s,t}=\sum_{\substack{k=0\\(k,l)\notin\delta}}^{s}\sum_{l=0}^{t}\prod_{\substack{i=0\\(i,j)\neq(k,l)}}^{k}\prod_{j=0}^{l}(a+\eta_{s-i,t-j})\,\varepsilon_{s-k,t-l}+\varepsilon_{s,t},$$

where $\delta=\{(0,0),(s,t)\}$. On the set of pairs of subscripts $(s,t);s,t\in\mathbb{Z}$, we set an order assuming that $(p,q)<(s,t)$ if $q<t$ or $q=t$ but $p<s$. Let us denote by $\mathcal{F}_{st}$ the $\sigma-$algebra generated by $\{(\eta_{p,q},\varepsilon_{p,q}):(p,q)<(s,t)\}$. The condition which guarantees that there exists a $\mathcal{F}_{st}$-measurable strictly stationary ergodic solution of (2.11) is that the parameters a and η satisfy the following inequality

$$a^2+\eta^2<1. \tag{2.17}$$

This solution is expressed as

$$\mathrm{X}_{s,t}=\sum_{\substack{k=0\\(k,l)\neq(0,0)}}^{+\infty}\sum_{l=0}^{+\infty}D_{k,l}\,\varepsilon_{s-k,t-l}+\varepsilon_{s,t}, \tag{2.18}$$

where

$$D_{k,l}=\prod_{\substack{i=0\\(i,j)\neq(k,l)}}^{k}\prod_{j=0}^{l}(a+\eta_{s-i,t-j}).$$

The main task in the analysis of the model (2.11) lies in estimating the vector of parameters $(a,\,\eta^2,\,\sigma^2)$.

In the previous section, conditions were found for the existence of stationary solution of the model defined by (2.11). In practice, however, given that a stationary random field $(X_{s,t})$ satisfies such an equation, one must fit the models to data and estimate the unknown parameters in order to provide predictors based on the past values of the random field. This section will be concerned with the estimation, both least squares and maximum likelihood estimates of the vector of parameters $\left(a, \eta^2, \sigma^2\right)$, as well as a derivation of the asymptotic properties of the estimates.
We rewrite the model (2.11) as $X_{s,t} = aX_{s,t-1}+u_{s,t}$, where $u_{s,t}$ is the residual function defined as $u_{s,t} = \eta_{s,t}X_{s,t-1} + \varepsilon_{s,t}$.

$$E(u_{s,t}|\mathcal{F}_{s,t}) = X_{s,t-1}E(\eta_{s,t}) + E(\varepsilon_{s,t}) = 0$$

since η_{st} and $\varepsilon_{s,t}$ are independent of the $\sigma-$field $\mathcal{F}_{st}$ and $X_{s,t}$ is $\mathcal{F}_{st}$-measurable. It results that $X_{s,t}$ is normal with

$$E(X_{s,t}|\mathcal{F}_{s,t}) = aX_{s,t-1} \text{ and } E\left(u^2_{s,t}|\mathcal{F}_{s,t}\right) = \eta^2 X^2_{s,t-1} + \sigma^2. \tag{2.19}$$

2.4.2 Least Squares Estimates of the Parameters

Given a sample

$$\left(X_{k,l}; k = 0, ..., m;\ l = 0, ..., n;\ (k, l) \neq (0, 0)\right)$$

of size $(m+1)(n+1)-1$, from the random field $(X_{s,t})$ generated by the model (2.11). To avoid cumbersome notations we put $\lambda = (m+1)(n+1)-1$. It is well-suited firstly to calculate the least squares estimate $\hat{a}^{LS}_{mn}$ of the parameter a which is obtained by minimizing $\sum\limits_{\substack{k=0\\(k,l)\neq(0,0)}}^{m}\sum\limits_{l=0}^{n} u^2_{s,t}$ with respect to a, namely

$$L_{LS}(a) = \sum_{\substack{k=0\\(k,l)\neq(0,0)}}^{m}\sum_{l=0}^{n} (X_{s,t} - E(X_{s,t}|\mathcal{F}_{s,t}))^2.$$

It follows that

$$\hat{a}^{LS}_{mn} = \left(\sum_{\substack{k=0\\(k,l)\neq(0,0)}}^{m}\sum_{l=0}^{n} X^2_{s,t-1}\right)^{-1} \sum_{\substack{k=0\\(k,l)\neq(0,0)}}^{m}\sum_{l=0}^{n} X_{s,t}X_{s,t-1}, \tag{2.20}$$

this step is required as the residuals $\hat{u}_{s,t} = X_{s,t} - \hat{a}^{LS}_{mn}X_{s,t-1}$, $s = 0, ..., m$, $t = 0, ..., n$, $(s, t) \neq (0, 0)$ are used to derive the estimates $\hat{\sigma}^2_{mn}$ and $\hat{\mu}^2_{mn}$ of σ^2 and μ^2 respectively.

The second step in the estimation procedure begins by using (2.19) to form the residuals $\hat{u}_{s,t}$. Let $\theta_{s,t} = \hat{u}^2_{s,t} - \sigma^2 - \mu^2 X^2_{s,t-1}$, then the estimates $\hat{\mu}^2_{mn}$ and $\hat{\sigma}^2_{mn}$ of μ^2 and σ^2 respectively are obtained by minimizing $\sum\limits_{\substack{s=0\\(s,t)\neq(0,0)}}^{m}\sum\limits_{t=0}^{n} \theta^2_{s,t}$ with respect to μ^2 and σ^2. We have

$$\frac{\partial}{\partial\mu^2}\left(\sum_{\substack{s=0\\(s,t)\neq(0,0)}}^{m}\sum_{t=0}^{n} \theta^2_{s,t}\right) = -2\sum_{\substack{s=0\\(s,t)\neq(0,0)}}^{m}\sum_{t=0}^{n} X^2_{s,t-1}(\hat{u}^2_{s,t} - \sigma^2 - \mu^2 X^2_{s,t-1}),$$

which implies that

$$\hat{\mu}^2_{mn} = \left(\sum_{\substack{s=0 \\ (s,t)\neq(0,0)}}^{m} \sum_{t=0}^{n} \hat{u}^2_{s,t} X^2_{s,t-1} - \sigma^2 X^2_{s,t-1} \right) \left(\sum_{\substack{s=0 \\ (s,t)\neq(0,0)}}^{m} \sum_{t=0}^{n} X^4_{s,t-1} \right)^{-1},$$

and we have

$$\frac{\partial}{\partial \sigma^2} \left(\sum_{\substack{s=0 \\ (s,t)\neq(0,0)}}^{m} \sum_{t=0}^{n} \theta^2_{s,t} \right) = -2 \sum_{\substack{s=0 \\ (s,t)\neq(0,0)}}^{m} \sum_{t=0}^{n} (\hat{u}^2_{s,t} - \sigma^2 - \hat{\mu}^2_{mn} X^2_{s,t-1}).$$

Thus

$$\hat{\sigma}^2_{mn} = \lambda^{-1} \sum_{\substack{s=0 \\ (s,t)\neq(0,0)}}^{m} \sum_{t=0}^{n} \hat{u}^2_{s,t} - \hat{\mu}^2_{mn} \lambda^{-1} \sum_{\substack{s=0 \\ (s,t)\neq(0,0)}}^{m} \sum_{t=0}^{n} X^2_{s,t-1}.$$

Consistency Of The Least Squares Estimates

This part is concerned with the relative consitency of the least squares estimator. The fourth moment condition on $X_{s,t}$ is required in order to derive the strong consistency, we obtain in the following lemma a condition on the parameters a and η which ensures the existence of the fourth moment of the stochastic process generated by (2.11).

Lemma 2.2

The $\mathcal{F}_{s,t}$−measurable stationary solution (2.18) under Assumption 1 has a finite fourth moment if and only if $a^4 + 6a^2\eta^2 + 3\eta^4 < 1$.

Proof: In fact, if $E(X^4_{s,t}) < \infty$, then from (2.18)

$$\begin{aligned} E\left(X^4_{s,t}\right) &= E\left(\varepsilon^4_{s,t}\right) + 6E(\varepsilon^2_{s,t}) \sum_{\substack{k=0 \\ (k,l)\neq(0,0)}}^{+\infty} \sum_{l=0}^{+\infty} E\left(D^2_{k,l}\right) E(\varepsilon^2_{s-k,t-l}) \\ &+ E\left(\sum_{\substack{k=0 \\ (k,l)\neq(0,0)}}^{+\infty} \sum_{l=0}^{+\infty} D_{k,l}\, \varepsilon_{s-k,t-l} \right)^4 . \end{aligned}$$

The first two terms are finite since $\varepsilon_{s,t}$ is normal and $X_{s,t}$ has a finite second moment. However

$$\begin{aligned} E\left(\sum_{\substack{k=0 \\ (k,l)\neq(0,0)}}^{+\infty} \sum_{l=0}^{+\infty} D_{k,l}\, \varepsilon_{s-k,t-l} + \varepsilon_{s,t} \right) &\geq E\left(\varepsilon^4_{s,t}\right) \sum_{\substack{k=0 \\ (k,l)\neq(0,0)}}^{+\infty} \sum_{l=0}^{+\infty} \left(E\left(a + \eta_{s,t}\right)^4\right)^{k\times l-1} \\ &= E\left(\varepsilon^4_{s,t}\right) \sum_{\substack{k=0 \\ (k,l)\neq(0,0)}}^{+\infty} \sum_{l=0}^{+\infty} (a^4 + 6a^2\eta^2 + 3\eta^4)^{k\times l-1}. \end{aligned}$$

We show that the condition $a^4 + 6a^2\eta^2 + 3\eta^4 < 1$ is necessary. Furthermore, for $(k,l) < (p,q)$

$$E(D^2_{p,q}D^2_{k,l}) = E(D^4_{k,l}) + E\left(\prod_{i'=k}^{p}\prod_{\substack{j'=l\\(i',j')\neq(p,q)}}^{q}(a+\eta_{s-i',t-j'})^2\right).$$

Hence

$$\begin{aligned} E\left(\sum^{+\infty}\sum_{(k,l)<(p,q)}^{+\infty} D^2_{k,l}D^2_{p,q}\varepsilon^2_{s-k,t-l}\varepsilon^2_{s-p,t-q}\right) &= E(\epsilon^2_{s,t})\sum_{k=0}^{+\infty}\sum_{\substack{l=0\\(k,l)\neq(0,0)}}^{+\infty}(a^4+6a^2\eta^2+3\eta^4)^{k\times l-1} \\ &\times \sum_{k=0}^{+\infty}\sum_{\substack{l=0\\(k,l)\neq(0,0)}}^{+\infty}(a^2+\eta^2)^{q\times p-k\times l+1}, \end{aligned}$$

$$\sum_{k=0}^{+\infty}\sum_{\substack{l=0\\(k,l)\neq(0,0)}}^{+\infty}(a^2+\eta^2)^{q\times p-k\times l+1} = \frac{a^2+\eta^2}{1-a^2-\eta^2}$$

will converge also if $a^4+6a^2\eta^2+3\eta^4<1$. Thus, noting that

$$(a^2+\eta^2)^2 < a^4+6a^2\eta^2+3\eta^4,$$

the condition is seen to be both necessary and sufficient. □

The main result of this section is the following theorem, we shall see, using the ergodic theorem, that $\hat{a}^{LS}_{mn}$ is a strongly consistent estimate of a and asymptotically normal.

Theorem 2.4

For a strictly stationary, ergodic and $\mathcal{F}_{s,t}$−measurable process $X_{s,t}$ satisfying (2.11) and the estimator $\hat{a}^{LS}_{mn}$ given by (2.20), $\hat{a}^{LS}_{mn}$ converges almost surely to a. Furthermore, if $E(X^4_{s,t}) < \infty$ then the random variable $\sqrt{mn}(\hat{a}^{LS}_{mn} - a)$ is asymptotically Gaussian with zero expectation and variance given by

$$\frac{(1-a-\eta^2)^2(\sigma^2E(X^2_{s,t-1})+\eta^2E(X^4_{s,t-1}))}{\sigma^4}.$$

Proof: From (2.20)

$$\hat{a}^{LS}_{mn} - a = \left(\sum_{s=0}^{m}\sum_{\substack{t=0\\(s,t)\neq(0,0)}}^{n}(X_{s,t-1})^2\right)^{-1}\sum_{s=0}^{m}\sum_{\substack{t=0\\(s,t)\neq(0,0)}}^{n}X_{s,t-1}u_{s,t}. \tag{2.21}$$

Since $\{X_{s,t}\}$ is strictly stationary and ergodic so are $\{X^2_{s,t}\}$ and $\{X_{s,t-1}u_{s,t}\}$. Furthermore, $\nu = E\left(X^2_{s,t}\right)$ is finite and

$$E\left(X_{s,t-1}u_{s,t}\right) = E\left(E\left(X_{s,t-1}u_{s,t}|\mathcal{F}_{s,t}\right)\right) = E\left(X_{s,t-1}E\left(u_{s,t}|\mathcal{F}_{s,t}\right)\right) = 0.$$

Since $E\left(u_{s,t}|\mathcal{F}_{s,t}\right) = 0$, and $X_{s,t-1}$ is $\mathcal{F}_{s,t}-$measurable. Thus by the ergodic theorem

$$\lambda^{-1}\sum_{\substack{s=0\\(s,t)\neq(0,0)}}^{m}\sum_{t=0}^{n} X_{s,t-1}^2 \xrightarrow{a.s} E(X_{s,t-1}^2)$$

and

$$\lambda^{-1}\sum_{\substack{s=0\\(s,t)\neq(0,0)}}^{m}\sum_{t=0}^{n} X_{s,t-1}u_{s,t} \xrightarrow{a.s} 0,$$

showing that $(\hat{a}_{mn}^{LS} - a)$ converges almost surely to 0. Remembering that $X_{s,t-1}u_{s,t}$ is strictly stationary and ergodic, furthermore $E(X_{s,t-1}u_{s,t}|\mathcal{F}_{s,t}) = 0$ and in lemma 2 we have shown that $E(X_{s,t}^4) < \infty$, so we affirm that

$$E(X_{s,t-1}u_{s,t})^2 = \eta^2 E(X_{s,t-1}^4) + \sigma^2 E(X_{s,t-1}^2) < \infty.$$

Then $\lambda^{-1}\sum_{\substack{s=0\\(s,t)\neq(0,0)}}^{m}\sum_{t=0}^{n} X_{s,t-1}u_{s,t}$ converges in distribution to the normal distribution with mean zero and variance $\eta^2 E(X_{s,t-1}^4) + \sigma^2 E(X_{s,t-1}^2)$. Hence $\lambda^{-1}(\hat{a}_{mn}^{LS} - a)$ is asymptotically Gaussian with zero expectation and variance $\nu^{-2}(\eta^2 E(X_{s,t-1}^4) + \sigma^2 E(X_{s,t-1}^2))$, where $\nu = \frac{\sigma^2}{1-a^2-\eta^2}$.

□

2.4.3 Maximum likelihood estimates of the parameters

Given a sample

$$\left(X_{k,l}; k = 0, ..., m,\ l = 0, ..., n,\ (k,l) \neq (0,0)\right)$$

of size λ, from the random field $(X_{s,t})$ generated by the model (2.11), we obtain the likelihood function conditional on the values $\{X_{s,t}; -p_1 < s \leq 0, -p_2 < t \leq 0\}$. Given $A_{s-q_1,t-q_2} \in \mathcal{F}_{s,t}$, let

$$f(X_{s,t}, ..., X_{s-q_1,t-q_2}|A_{s-q_1,t-q_2})$$

denote the density of $X_{s,t}, ..., X_{s-q_1,t-q_2}$. Furthermore, it follows from equation (2.19) that

$$\begin{aligned}
f_{mn}\{X_{01}, ..., X_{m,n}|X_{00}\} &= \prod_{\substack{s=0\\(s,t)\neq(0,0)}}^{m}\prod_{t=0}^{n} f(X_{s,t}|X_{00}) \\
&= (2\pi)^{-\frac{\lambda}{2}}\prod_{\substack{s=0\\(s,t)\neq(0,0)}}^{m}\prod_{t=0}^{n} (\sigma^2 + \eta^2 X_{s,t-1}^2)^{-\frac{1}{2}} \times \exp\left\{-\frac{1}{2}\frac{(X_{s,t} - aX_{s,t-1})^2}{\sigma^2 + \eta^2 X_{s,t-1}^2}\right\} \\
&= L_{mn}\left(a, \eta^2, \sigma^2\right),
\end{aligned}$$

which is the Likelihood function conditional on X_{00}. Instead of the maximization of this latter, it will be more practical and appropriate to think to the minimization of the monotone function of the likelihood $\tilde{\mathcal{L}}_{mn}\left(a, \eta^2, \sigma^2\right)$

$$\tilde{\mathcal{L}}_{mn}\left(a, \eta^2, \sigma^2\right) = -2\lambda^{-1}\ln\left(L_{mn}\left(a, \eta^2, \sigma^2\right)\right) - \ln(2\pi)$$

$$= \lambda^{-1}\sum_{\substack{s=0\\(s,t)\neq(0,0)}}^{m}\sum_{t=0}^{n} \ln(\sigma^2+\eta^2 X_{s,t-1}^2) + \lambda^{-1}\sum_{\substack{s=0\\(s,t)\neq(0,0)}}^{m}\sum_{t=0}^{n} \frac{(X_{s,t}-aX_{s,t-1})^2}{\sigma^2+\eta^2X_{s,t-1}^2}.$$

The function $\tilde{\mathcal{L}}_{mn}\left(a,\eta^2,\sigma^2\right)$ is nonlinear in σ^2 and η^2 and there is no fully defined expression for the estimates $\hat{a}_{mn}$, $\hat{\eta}^2_{mn}$ and $\hat{\sigma}^2_{mn}$ of a, η^2 and σ^2, respectively, which minimize $\tilde{\mathcal{L}}_{mn}\left(a,\eta^2,\sigma^2\right)$. However,

$$\tilde{L}_{mn}\left(a,\tau,\sigma^2\right) = \ln(\sigma^2) + \lambda^{-1}\sum_{\substack{s=0\\(s,t)\neq(0,0)}}^{m}\sum_{t=0}^{n} \ln(1+\tau X_{s,t-1}^2) + \sigma^{-2}\lambda^{-1}\sum_{\substack{s=0\\(s,t)\neq(0,0)}}^{m}\sum_{t=0}^{n} \frac{(X_{s,t}-aX_{s,t-1})^2}{1+\tau X_{s,t-1}^2}.$$

For fixed τ, the function $\tilde{L}_{mn}\left(a,\tau,\sigma^2\right)$ reaches its minimum when $a = a_{mn}(\tau)$ and $\sigma^2 = \sigma^2_{mn}(\tau)$, where

$$a_{mn}(\tau) = \left(\lambda^{-1}\sum_{\substack{s=0\\(s,t)\neq(0,0)}}^{m}\sum_{t=0}^{n} \frac{X_{s,t-1}^2}{1+\tau X_{s,t-1}^2}\right)^{-1} \times \lambda^{-1}\sum_{\substack{s=0\\(s,t)\neq(0,0)}}^{m}\sum_{t=0}^{n} \frac{X_{s,t}X_{s,t-1}}{1+\tau X_{s,t-1}^2},$$

and

$$\sigma^2_{mn}(\tau) = \lambda^{-1}\sum_{\substack{s=0\\(s,t)\neq(0,0)}}^{m}\sum_{t=0}^{n} \frac{(X_{s,t}-aX_{s,t-1})^2}{1+\tau X_{s,t-1}^2}.$$

An interesting fact is that the estimates of a and σ^2 are found in terms of τ. Therefore, $\tilde{L}_{mn}$ will depend only on τ. Thus, minimizing $L_{mn}\left(a,\eta^2,\sigma^2\right)$ using this concentrated maximum likelihood approach amounts to minimizing a function of τ. If $L_{mn}\left(a,\eta^2,\sigma^2\right)$ is minimized at $\tau = \hat{\tau}_{mn}$, the maximum likelihood estimates of a, σ^2 and η^2 are given by $\hat{a}_{mn} = a_{mn}(\hat{\tau}_{mn})$, $\hat{\sigma}^2_{mn}(\hat{\tau}_{mn})$ and $\hat{\eta}^2_{mn} = \hat{\sigma}_{mn}\hat{\tau}_{mn}$, respectively.

This procedure can only be used in the case where the optimization algorithm does not require the second derivative of the function. Otherwise, if the second derivative is necessary and in the case of $\tilde{\mathcal{L}}_{mn}$ they are rather complicated. For this reason, it is better to minimize $\tilde{L}_{mn}(a,\tau)$.

$$\begin{aligned}\tilde{L}_{mn}(a,\tau) &= \inf_{\sigma^2}\tilde{L}_{mn}(a,\tau,\sigma^2) - 1\\ &= \lambda^{-1}\sum_{\substack{s=0\\(s,t)\neq(0,0)}}^{m}\sum_{t=0}^{n} \ln(1+\tau X_{s,t-1}^2) + \lambda^{-1}\ln\left(\sum_{\substack{s=0\\(s,t)\neq(0,0)}}^{m}\sum_{t=0}^{n} \frac{(X_{s,t}-aX_{s,t-1})^2}{1+\tau X_{s,t-1}^2}\right).\end{aligned}$$

The partial derivates of $\tilde{L}_{mn}(a,\tau)$ with respect to a and τ are easily obtained, the maximum likelihood estimates $\hat{a}_{mn}$, $\hat{\eta}^2_{mn}$ and $\hat{\sigma}^2_{mn}$ is obtained from the following equations.

$$\tilde{L}_{mn}(\hat{a}_{mn},\hat{\tau}_{mn}) = \inf_{(a,\tau)\in\Theta}\tilde{L}_{mn}(a,\tau) \tag{2.22}$$

.

$$\hat{\sigma}^2_{mn}(\hat{\tau}) = \lambda^{-1} \sum_{s=0}^{m} \sum_{\substack{t=0 \\ (s,t)\neq(0,0)}}^{n} \frac{(X_{s,t} - \hat{a}_{mn} X_{s,t-1})^2}{1 + \hat{\tau}_{mn} X^2_{s,t-1}}.$$

$$\hat{\mu}^2_{mn} = \hat{\sigma}^2_{mn} \hat{\tau}_{mn}.$$

Consistency Of The Maximum Likelihood Estimates

To show the strong consistency of the maximum likelihood estimators, it will be required that Θ be compact. Accordingly, any continuous function in Θ reaches its limits in Θ. Also, equicontinuity and uniform convergence are equivalent in Θ.
The following lemma provides one of the main results used to prove strong consistency of $\hat{a}_{mn}$, $\hat{\eta}^2_{mn}$ and $\hat{\sigma}^2_{mn}$.

Lemma 2.3

For a strictly stationary, ergodic and $\mathcal{F}_{s,t}$−measurable process $X_{s,t}$ satisfying (2.11) with $a = a_{0,0}$, $\sigma^2 = \sigma^2_{0,0}$ and $\mu^2 = \mu^2_{0,0}$ under assumption (1), and $\theta_{0,0} = (a_{0,0}, \tau_{0,0})'$, where $\tau_{0,0} = \frac{\mu^2_{0,0}}{\sigma^2_{0,0}}$.

Then $\lim_{\substack{m\to+\infty \\ n\to+\infty}} \tilde{L}_{mn}(a,\tau)$ exists almost surely for all $(a,\tau)' \in \Theta$ and the limit $\tilde{L}_{mn}(a,\tau)$ is uniquely minimized over θ at $(a,\tau)' = \theta_{0,0}$ provided that $\theta_{0,0} \in int(\Theta)$.

Proof: Since $0 \le \ln(1 + \tau X^2_{s,t-1}) \le \tau X_{s,t-1}$, and $E(X^2_{s,t-1})$ exists,

$$\lambda^{-1} \sum_{s=0}^{m} \sum_{\substack{t=0 \\ (s,t)\neq(0,0)}}^{n} \ln(1 + \tau X^2_{s,t-1})$$

converges almost surely to $E(\ln(1 + \tau X^2_{s,t-1}))$ by the ergodic theorem. Also, since

$$0 < \lambda^{-1} \ln\left(\sum_{s=0}^{m} \sum_{\substack{t=0 \\ (s,t)\neq(0,0)}}^{n} \frac{(X_{s,t} - aX_{s,t-1})^2}{1 + \tau X^2_{s,t-1}} \right) \le \lambda^{-1} \ln\left(\sum_{s=0}^{m} \sum_{\substack{t=0 \\ (s,t)\neq(0,0)}}^{n} (X_{s,t} - aX_{s,t-1})^2 \right)$$

and this later term has a finite expectation and converges by the ergodic theorem, we have

$$0 < \lambda^{-1} \ln\left(\sum_{s=0}^{m} \sum_{\substack{t=0 \\ (s,t)\neq(0,0)}}^{n} \frac{(X_{s,t} - aX_{s,t-1})^2}{1 + \tau X^2_{s,t-1}} \right) \xrightarrow{a.s} E\left(\frac{(X_{s,t} - aX_{s,t-1})^2}{1 + \tau X^2_{s,t-1}} \right).$$

Hence, $\tilde{L}_{mn}(a,\tau)$ converges almost surely to

$$\tilde{L}(a,\tau) = E\left(\ln(1 + \tau X^2_{s,t-1}) + \ln\left(\frac{(X_{s,t} - aX_{s,t-1})^2}{1 + \tau X^2_{s,t-1}} \right) \right).$$

$$\begin{aligned}
E\left(\frac{(X_{s,t}-aX_{s,t-1})^2}{1+\tau X_{s,t-1}^2}\right) &= E\left(\frac{(X_{s,t}-a_{0,0}X_{s,t-1}+(a_{0,0}-a)X_{s,t-1})^2}{1+\tau X_{s,t-1}^2}\right) \\
&= E\left(\frac{E\left((\eta_{s,t}X_{s,t-1}+\varepsilon_{s,t})^2|\mathcal{F}_{s,t-1}\right)}{1+\tau X_{s,t-1}^2}\right) \\
&\quad +E\left(\frac{(a_{0,0}-a)^2X_{s,t-1}^2}{1+\tau X_{s,t-1}^2}\right) \\
&\geq E\left(\frac{\sigma_{0,0}^2+\eta_{0,0}^2X_{s,t-1}^2}{1+\tau X_{s,t-1}^2}\right) \\
&= \sigma_{0,0}^2E\left(\frac{1+\tau_{0,0}X_{s,t-1}^2}{1+\tau X_{s,t-1}^2}\right),
\end{aligned}$$

since

$$E\left(X_{s,t}-a_{0,0}X_{s,t-1}|\mathcal{F}_{s,t}\right)=0,$$

and

$$E\left((\eta_{s,t}X_{s,t-1}+\varepsilon_{s,t})^2|\mathcal{F}_{s,t}\right)=\sigma_{0,0}^2+\eta_{0,0}^2X_{s,t-1}^2.$$

and $\eta_{0,0}^2=\sigma_{0,0}^2\tau_{0,0}^2$. Moreover, equality will hold in the above only when $(a_{0,0}-a)X_{s,t-1}$ almost surely, that is, when $a=a_{0,0}$. Thus,

$$\begin{aligned}
\inf_a \tilde{L}_{mn}(a,\tau) &= \tilde{L}_{mn}(a_{0,0},\tau) \\
&= \ln(\sigma_{0,0}^2)+\ln\left(\frac{E(1+\tau_{0,0}X_{s,t-1}^2)}{1+\tau X_{s,t-1}^2}\right)+E\left(\ln(1+\tau X_{s,t-1}^2)\right),
\end{aligned}$$

and $\tilde{L}_{mn}(a,\tau)=\inf_a \tilde{L}(a,\tau)$ only at $a=a_{0,0}$. Now, if X is any positive random variable with an expectation 1, then

$$E(\ln(X))\leq\ln(E(X))=0,$$

by Jensen's inequality, with equality only when $X=1$ almost surely.
Letting

$$X=c^{-1}\frac{1+\tau_{0,0}X_{s,t-1}^2}{1+\tau X_{s,t-1}}$$

, where $c=E\left(\frac{1+\tau_{0,0}X_{s,t-1}^2}{1+\tau X_{s,t-1}^2}\right)$, it is seen that

$$E\left(\ln\frac{(1+\tau X_{s,t-1}^2)}{1+\tau_{0,0}X_{s,t-1}^2}\right)\geq-\ln E\left(\frac{(1+\tau X_{s,t-1}^2)}{1+\tau_{0,0}X_{s,t-1}^2}\right)$$

with equality only when

$$(1+\tau_{0,0}X_{s,t-1}^2)=c(1+\tau X_{s,t-1}^2).$$

almost surely, that is when $(\tau_{0,0}-c\tau)X_{s,t-1}^2=c-1$ almost surely. This occurs only when $\tau_{0,0}=c\tau$ and $c=1$, that is, when $\tau=\tau_{0,0}$. Hence, $\tilde{L}_{mn}(a,\tau)$ is uniquely minimized at $a=a_{0,0}$ and $\tau=\tau_{0,0}$. □

Lemma 2.4

$\lim_{\substack{n\to\infty \\ m\to\infty}} L_{mn}(a,\mu^2,\sigma^2)$ exists almost surely and minimized uniquely at $a = a_{0,0}$, $\mu^2 = \mu^2_{0,0}$, $\sigma = \sigma^2_{0,0}$.

Proof: From Lemma 3 and the convergence of $\tilde{L}_{mn}(a,\tau)$, $\lim_{\substack{n\to\infty \\ m\to\infty}} L_{mn}(a,\mu^2,\sigma^2)$ is seen to exist almost everywhere and to be uniquely minimized at $a = a_{0,0}$,

$$\sigma^2 = \sigma^{2*} = E\left(\frac{(X_{s,t} - a_{0,0}X_{s,t-1})^2}{1+\tau_{0,0}X^2_{s,t-1}}\right),$$

and $\mu^2 = \tau_{0,0}\sigma^{2*}$. But

$$\begin{aligned}\sigma^{2*} &= E\left(E\left(\frac{(X_{s,t} - a_{0,0}X_{s,t-1})^2}{1+\tau_{0,0}X^2_{s,t-1}}\right)\right) \\ &= E\left(\frac{\sigma^2_{0,0} + \mu^2_{0,0}X^2_{s,t-1}}{1+\tau_{0,0}X^2_{s,t-1}}\right) = \sigma^2_{0,0},\end{aligned}$$

and so $\lim_{\substack{n\to\infty \\ m\to\infty}} L_{mn}(a,\mu^2,\sigma^2)$ is uniquely minimized at $a = a_{0,0}$, $\mu^2 = \mu^2_{0,0}$, $\sigma = \sigma^2_{0,0}$. We are now in position to prove the strong consistency of the procedure. □

Theorem 2.5

Let $\tilde{L}_{mn}(a,\tau)$ be minimized over Θ at $a = \hat{a}_{mn}$, $\tau = \hat{\tau}_{mn}$, and $\hat{\theta}_{mn} = (\hat{a}_{mn},\hat{\tau}_{mn})$. Then $\hat{\theta}_{mn}$ converge almost surely to $\theta_{00} = (a_{00},\tau_{00})$ provided that $\theta_{00} \in int(\Theta)$.

Corollary 2.2

$\hat{\eta}^2_{mn}$ and $\hat{\sigma}^2_{mn}$ converge almost surely to $\eta^2_{0,0}$ and $\sigma^2_{0,0}$, respectively.

Proof: Theorem 5 specifies that $\hat{\tau}_{mn}$ and $\hat{a}_{mn}$ converge almost surely to $\tau_{0,0}$ and $a_{0,0}$, respectively. Now,

$$\hat{\sigma}^2_{mn} = \lambda^{-1}\sum_{\substack{k=0 \\ (k,l)\neq(0,0)}}^{m}\sum_{l=0}^{n}\frac{(X_{s,t} - \hat{a}_{mn})^2}{1+\hat{\tau}_{mn}X^2_{s,t-1}}.$$

By theorem 5, the sequence $\left\{\sigma^2_{mn}(\theta)\right\}$, where

$$\sigma^2_{mn} = \lambda^{-1}\sum_{\substack{k=0 \\ (k,l)\neq(0,0)}}^{m}\sum_{l=0}^{n}\frac{(X_{s,t} - a_{mn})^2}{1+\tau_{mn}X^2_{s,t-1}}$$

converges uniformly and almost surely on Θ to

$$\sigma^2(\theta) = E\left\{\frac{(X_{s,t} - a_{mn})^2}{1 + \tau_{mn} X_{s,t-1}^2}\right\}.$$

It is also evident that $\hat{\sigma}^2_{mn} \xrightarrow{a.s} \sigma(\theta^2_{0,0})$ which was seen in the proof of lemma 3 to equal $\sigma^2_{0,0}$, since $\hat{\eta}^2_{mn} = \hat{\tau}_{mn}\hat{\sigma}^2_{mn}$, it follows that $\hat{\eta}^2_{mn} \xrightarrow{a.s} \eta^2_{0,0} = \tau_{0,0}\sigma^2_{0,0}$. □

In this section, we show the effectiveness and accuracy of the estimation method developed in the previous sections. Moreover, we establish a comparative study between the maximum likelihood estimates and the least square estimates.

2.4.4 Comparative study

In order to compare maximum likelihood and the least square estimates, the sample square of the estimates is retained as the criterion of comparison. The calculation steps that allow comparison are summarized in Algorithm 1.

Algorithm 1 Prediction Algorithm

for $(n, m) \in \{(100, 150)(150, 200)(250, 250)\}$ and for several values of (a, η^2, σ^2) **do**

  for $j \in \{1 : \text{rep} = 1000\}$ **do**

- Generate data as a (n, m) rectangular grid from spatial models of the form (2.16) where $\epsilon(s, t)$ is Gaussian white noise process with mean 0 and variance $\sigma^2 = 1$ and $\eta(s, t)$ is supposed Uniform on the $[0, b]$, with b is chosen so that the condition $a^2 + \eta^2 < 1$ holds.
- Estimate the vector of the parameters $(a, \eta^2, \sigma^2)'$ by the two methods cited in the subsections and .
- Calculate the sample square error of each estimate.

  end for

  Calculate the sample mean square error of each estimate.

end for

In Table 2.1, row (a) gives the true parameter values, row (b) gives the mean square errors of the estimates obtained by the least square method, and row (c) is dedicated to the mean square error of the estimates using the maximum likelihood procedure. Thus our MSE is defined as the expected value or the mean of values over the set of 1000 repetitions. This definition of MSE is convenient for computing. One reason for choosing as many as 1000 replications is to reduce this standard error and hence obtain a relatively reliable ranking of different values.

The results in Table 2.1 indicate that, with the increase in the values of n and m, the values of the estimated MSE of the two estimators decrease. Also, for fixed values of n and m and for the various values of the parameters considered, it can be seen that the MSE increases when approaching the zone of non-stationarity. Particularly as the sample size increases, the mean square errors of the maximum likelihood are less than those of estimates obtained by the least squares method. This is what one would expect of course since the maximum likelihood estimates are asymptotically unbiased and efficient. Also, the results recorded in Table 2.1 indicate that almost all of the true values of the parameters are within two sample standard deviations of their respective sample mean.

(n,m)		a	η^2	σ^2
(150,150)	(a)	0.40	0.25	0.25
	(b)	0.113	1.002	0.105
	(c)	0.112	0.099	0.098
(250,250)	(a)	0.40	0.25	0.25
	(b)	0.112	0.098	0.089
	(c)	0.09	0.091	0.810
(500,500)	(a)	0.40	0.25	0.25
	(b)	0.031	0.023	0.012
	(c)	0.002	0.009	0.006
(150,150)	(a)	0.60	0.36	0.25
	(b)	0.116	0.108	0.115
	(c)	0.112	0.099	0.108
(250,250)	(a)	0.60	0.36	0.49
	(b)	0.007	0.079	0.097
	(c)	0.005	0.069	0.053
(500,500)	(a)	0.60	0.36	0.49
	(b)	0.05	0.062	0.010
	(c)	0.01	0.004	0.006
(150,150)	(a)	0.70	0.36	1.00
	(b)	0.180	0.121	0.121
	(c)	0.117	0.160	0.103
(250,250)	(a)	0.70	0.36	1.00
	(b)	0.097	0.097	0.116
	(c)	0.008	0.009	0.108
(500,500)	(a)	0.70	0.36	1.00
	(b)	0.006	0.09	0.006
	(c)	0.001	0.003	0.002
(150,150)	(a)	0.70	0.49	1.00
	(b)	0.114	0.210	0.166
	(c)	0.091	0.115	0.098
(250,250)	(a)	0.70	0.49	1.00
	(b)	0.111	0.102	0.0105
	(c)	0.12	0.008	0.006
(500,500)	(a)	0.70	0.49	1.00
	(b)	0.023	0.059	0.007
	(c)	0.001	0.005	0.002

Table 2.1: Values of the mean square errors of the estimates.

2.5 Conclusion

In this chapter, we presented new results on the estimation problem, more precisely, we have developed an estimation procedure for estimating unknown parameters of two-dimensional random coefficient autoregressive models.

Chapter

3

Measure of predictability of 2D-ARMA models

Contents

3.1 Introduction

In this chapter, we consider a measure of predictability of a random field following a stationary two-dimensionally indexed autoregressive moving average (2D-ARMA). The purpose of this chapter is two-fold. First, we give a characterization of equal predictability processes by providing necessary and sufficient conditions, which highlight the role of the coefficients of the moving average (MA) and the equivalent autoregressive (EAR) representations. Second, we implement a procedure allowing us to carry out a test of equal predictability. A simulation study is presented and an application to real data of wheat yield is reported.

We deal with the equal predictability of a quarter-plane (QP) two-dimensional autoregressive moving-average (2D ARMA) model excited by an unknown zero-mean white noise. Our predictability measure is closely related to [44] used in time series;

they proposed a natural measure of one-step-ahead predictability for covariance stationary time series. Following this same order of ideas, we propose a measure of predictability of a quarter-plane (QP) two-dimensional autoregressive moving-average (2D ARMA) model excited by an unknown zero-mean white noise. The relations between the parameters of a 2D ARMA model and their 2D equivalent autoregressive and moving average representations are used in order to obtain necessary and sufficient conditions of equal predictability and parallelism. Based on certain intermediate results in the form of propositions and lemmas, we have set up a statistical test allowing us to affirm or not the equal predictability between two series of data generated by 2D ARMA models.

3.2 Model and Assumptions

In this section, we describe the problem to be treated in this chapter and introduce the necessary notations. In time series, there is a natural neighbor structure induced by the existing total order of $\mathbb{Z}$ (the set of all past values of $t \in \mathbb{Z}$ is the set of all integers that are less than t). However, for points on the plane, for instance, $(m,n) \in \mathbb{Z}^2$, there are several different notions of the neighborhood. In general, definitions of the neighborhood of a point (m,n) on the plane are motivated by the physical acquisition system of the data. This is the case with images that have been captured by satellites. In what follows we use the most frequent neighborhood found in the literature; [45], the third quadrant $\mathcal{Q} = \{(i,j); i \leq 0, j \leq 0, (i,j) \neq (0,0)\}$, which is called strongly causal region.

We recall that a random $2-D$ field process is a quarter plane $2-D$ causal ARMA (resp. AR, MA) of finite order if and only if the innovation of a horizontal path coincides with the innovation of a vertical path, characterization of such process is quoted in [46]. A random field $\left\{X(s,t),(s,t) \in \mathbb{Z}^2\right\}$, which is stationary to the second order, will have a representation called a quarter plane 2 – D causal ARMA process of finite order (p_1,p_2,q_1,q_2) if it obeys a differential equation :

$$X(s,t) + \sum_{\substack{i=0\\(i,j)\neq(0,0)}}^{p_1} \sum_{j=0}^{p_2} a_{i,j} X(s-i,t-j) = \epsilon(s,t) + \sum_{\substack{k=0\\(k,l)\neq(0,0)}}^{q_1} \sum_{l=0}^{q_2} b_{k,l}\epsilon(s-k,t-l) \quad (3.1)$$

and $\{\epsilon(s,t)\}$ is an unknown random excitation which is assumed to be the white noise with zero mean and variance σ_ϵ^2. Let the transfer function $H(z_1,z_2)$ associated with the differential equation of system (3.1), expressed in the form of a rational fraction of 2 – D polynomials in z_1 and z_2,

$$H(z_1,z_2) = \frac{\phi(z_1,z_2)}{\theta(z_1,z_2)}, \quad (3.2)$$

with

$$\phi(z_1,z_2) = 1 + \sum_{\substack{i=0\\(i,j)\neq(0,0)}}^{p_1} \sum_{j=0}^{p_2} a_{i,j} z_1^i z_2^j,$$

and

$$\theta(z_1,z_2) = 1 + \sum_{\substack{k=0\\(k,l)\neq(0,0)}}^{q_1} \sum_{l=0}^{q_2} b_{k,l} z_1^k z_2^l.$$

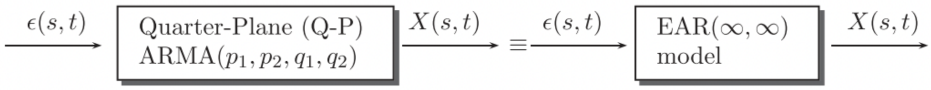

Figure 3.1: Relationship between ARMA and EAR models

Stationarity condition of system (3.1) (see [47])

$$\begin{cases} \sum_{i=0}^{p_1}\sum_{j=0}^{p_2} a_{ij} z_1^i z_2^j \neq 0 & \text{if } |z_1| \leq 1, |z_2| = 1 \\ \sum_{j=0}^{p_2} a_{0j} z_2^j \neq 0 & \text{if } |z_2| \leq 1 \end{cases} \tag{3.3}$$

with $a_{00} = 1$.
The conditions for the transfer function $H(z_1, z_2)$ to be invertible are formulated in [47] as

$$\begin{cases} \sum_{k=0}^{q_1}\sum_{l=0}^{q_2} b_{k,l} z_1^k z_2^l \neq 0 & \text{if } |z_1| \leq 1, |z_2| = 1 \\ \sum_{l=0}^{q_2} b_{0,l} z_2^l \neq 0 & \text{if } |z_2| \leq 1 \end{cases} \tag{3.4}$$

with $b_{00} = 1$
In Figure 3.1, the Relationship between ARMA and EAR models. Under condition (3.4), there exist a sequence $\{\pi_{ij}, i \geq 0, j \geq 0\}$ such that

$$\frac{1}{H(z_1, z_2)} = \frac{\sum_{i=0}^{p_1}\sum_{j=0}^{p_2} a_{i,j} z_1^i z_2^j}{\sum_{k=0}^{q_1}\sum_{l=0}^{q_2} b_{k,l} z_1^k z_2^l} = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} \pi_{ij} z_1^i z_2^j, \quad \pi_{00} = 1 \tag{3.5}$$

In this case, the Equivalent AutoRegressive $(\text{EAR}(\infty,\infty))$ process to the model (3.1) is

$$X(s,t) = \epsilon(s,t) + \sum_{\substack{i=0 \\ (i,j)\neq(0,0)}}^{\infty} \sum_{j=0}^{\infty} \pi_{i,j} X(s-i, t-j) \tag{3.6}$$

In Figure 3.1 , the relationship between ARMA (p_1, p_2, q_1, q_2) and EAR(∞,∞) models can also be constructed by 2D-EAR model. Indeed, by using (3.2) and (3.5) we can write

$$\frac{\sum_{i=0}^{q_1}\sum_{j=0}^{q_2} b_{i,j} z_1^{-i} z_2^{-j}}{1 + \sum_{\substack{i=0 \\ (i,j)\neq(0,0)}}^{p_1}\sum_{j=0}^{p_2} a_{i,j} z_1^{-i} z_2^{-j}} \approx \frac{1}{\sum_{i=0}^{\infty}\sum_{j=0}^{\infty} \pi_{i,j} z_1^{-i} z_2^{-j}} \tag{3.7}$$

Assuming that the stationarity condition (3.3) is satisfied, the model (3.3) admits a finite order Moving Average $(\text{MA}(\infty,\infty))$ process

$$X(s,t) = \sum_{k=0}^{\infty}\sum_{l=0}^{\infty} \psi_{k,l} \epsilon(s-k, t-l),$$

where the sequence $\{\psi_{k,l}; (k,l) \in \mathbb{Z}^2\}$ is called its MA parameters, with $\psi_{0,0} = 1, \psi_{k,l} = 0$ when $k < 0$ or $l < 0$ and $\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}|\psi_{k,l}|^2 < \infty$. The sequence $\{\psi_{k,l}\}$ is determined by the relation

$$1 + \sum_{\substack{k=0 \\ (k,l)\neq(0,0)}}^{\infty} \sum_{l=0}^{\infty} \psi_{k,l} L^{k,l} = \frac{\theta(L)}{\phi(L)},$$

and it is well known (see [12]), that the MA(∞,∞) coefficients $\psi_{k,l}$ and EAR$(\infty,\infty)\pi_{k,l}$ are related via the following recursions:

$$\psi_{0,0} = \pi_{0,0} = 1 \tag{3.8}$$

$$\psi_{i,j} = \sum_{\substack{k=0 \\ (k,l)\neq(0,0)}}^{i} \sum_{l=0}^{j} \pi_{k,l}\pi_{i-k,j-l}. \tag{3.9}$$

3.3 Measure of Predictability

We now consider the problem of prediction. For convenience, we keep the notations used in [13]. More precisely, we use the third quadrant $\mathcal{Q} = \{(i,j); i \leq 0, j \leq 0; (i,j) \neq (0,0)\}$ as the past. On this past support, it is possible to define the lexicographic order, so the first future value is $X(0,0)$. Note $\hat{X}(0,0)$ the best linear predictor of $X(0,0)$ which is obtained by the orthogonal projection onto the closed space $\mathcal{H} = \{X(i,j); (i,j) \in \mathcal{Q}\}$. Also, the innovation process is white noise.
The standard tool used in the time series domain as a measure of predictability of $\{X_t; t \in \mathbb{Z}\}$, based on the past, is the index

$$R^2 = 1 - \frac{\text{var}(X(t+1) - \hat{X}(t+1))}{\text{var}\,(X_t)},$$

see for instance [44]. In the same vein, we define as a measure of predictability of $\{X(s,t); (s,t) \in \mathbb{Z}^2\}$, with finite variance, using the third quadrant as its past history, the index

$$R_X^2 = 1 - \frac{\text{var}(e(0,0))}{\text{var}(X(s,t))},$$

where $e(0,0) = X(0,0) - \hat{X}(0,0)$.
R_X^2 will bind between zero and unity. For a series that cannot be forecast, such as white noise, this ratio is 0. If the series can be predicted without error, R^2 will be equal to 1. Let $\{X(s,t); (s,t) \in \mathbb{Z}^2\}$ be a stationary ARMA (p_1,p_2,q_1,q_2) defined by (3.3). Suppose that the stationarity and invertibility conditions given by (3.3) and (3.4) are satisfied. As mentioned previously, by applying the 2D-EAR transformation, the model can be written as

$$X(s,t) = \sum_{\substack{i=0 \\ (i,j)\neq(0,0)}}^{\infty} \sum_{j=0}^{\infty} \pi_{i,j} X(s-i,t-j) + \epsilon(s,t),$$

where the sequence π_{ij} is determined by the relation

$$1 - \sum_{\substack{i=0 \\ (i,j)\neq(0,0)}}^{\infty} \sum_{j=0}^{\infty} \pi_{i,j} L^{i,j} = \frac{\phi(L)}{\theta(L)},$$

where $L^{i,j}X(s,t) = X(s-i,t-j)$.
Thus, it shows that the optimal forecast of $X(0,0)$ given its past history is

$$\mathbf{E}(X(0,0)\backslash X(i,j);(i,j)\in\mathcal{Q}) = \sum_{\substack{i=0\\(i,j)\neq(0,0)}}^{\infty}\sum_{j=0}^{\infty}\pi_{i,j}X(-i,-j)$$

and the innovation $\epsilon(0,0)$ is the corresponding one-step forecast error. Thus we have

$$R_X^2 = 1 - \frac{\sigma_\epsilon^2}{\operatorname{var}(X(s,t))}.$$

Also, since $\{X(s,t);(s,t)\in\mathbb{Z}^2\}$ is stationary, it can be written as a finite order moving average process

$$X(s,t) = \epsilon(s,t) + \sum_{\substack{k=0\\(k,l)\neq(0,0)}}^{\infty}\sum_{l=0}^{\infty}\psi_{k,l}\epsilon(s-k,t-l),$$

where the sequence $\{\psi_{k,l}\}$ is determined by the relation

$$1 + \sum_{\substack{k=0\\(k,l)\neq(0,0)}}^{\infty}\sum_{l=0}^{\infty}\psi_{k,l}L^{k,l} = \frac{\theta(L)}{\phi(L)}$$

and hence

$$R_X^2 = 1 - \frac{1}{\left(1 + \sum\limits_{\substack{k=0\\(k,l)\neq(0,0)}}^{\infty}\sum\limits_{l=0}^{\infty}\psi_{k,l}^2\right)} = \frac{\sum\limits_{\substack{k=0\\(k,l)\neq(0,0)}}^{\infty}\sum\limits_{l=0}^{\infty}\psi_{k,l}^2}{\left(1 + \sum\limits_{\substack{k=0\\(k,l)\neq(0,0)}}^{\infty}\sum\limits_{l=0}^{\infty}\psi_{k,l}^2\right)} \tag{3.10}$$

We call the sequence $\{\psi_{k,l}\}$ ψ-weights; it represents the dependence structure on the series. A series with small ψ-weights (with a "few structures") will be predictable than one with large ψ-weights (a series with "a more structure"). Thus, this predictability measure provides a synthetic evaluation of the dependence structure of stationary random fields.
Given two stationary and invertible 2D ARMA processes, $\{X(s,t);(s,t)\in\mathbb{Z}^2\}$ and $\{Y(s,t);(s,t)\in\mathbb{Z}^2\}$, the condition of equal predictability for the two processes can be formulated as follows:

$$R_X^2 - R_Y^2 = 0$$

It seems clear that this condition of equal predictability is limited to one-step prediction in the lexicographic order. However, it is possible to extend and define a measure of predictability relative to a $h = (h_1,h_2)$-steps, with $(h_1,h_2)\in\mathcal{F} = \{(i,j); i\geq 0, j\geq 0; (i,j)\neq(0,0)\}$, by

$$R_X^2(h) = 1 - \frac{\operatorname{var}(e(h))}{\operatorname{var}(X(s,t))} = 1 - \frac{1+\sum\limits_{\substack{k=0\\(k,l)\neq(0,0)}}^{h_1-1}\sum\limits_{l=0}^{h_2-1}\psi_{k,l}^2}{1+\sum\limits_{\substack{k=0\\(k,l)\neq(0,0)}}^{\infty}\sum\limits_{l=0}^{\infty}\psi_{k,k}^2}$$

$$= \frac{\sum_{k=h_1}^{\infty} \sum_{l=h_2}^{\infty} \psi_{k,l}^2}{1 + \sum_{\substack{k=0 \\ (k,l)\neq(0,0)}}^{\infty} \sum_{l=0}^{\infty} \psi_{k,l}^2}, \quad (h_1, h_2) \in \mathcal{F}$$

where $e(h_1, h_2) = X(h_1, h_2) - \mathbf{E}[X(h_1, h_2) \backslash X(i,j); (i,j) \in \mathcal{Q}]$. Given two stationary and invertible 2D ARMA processes, $\{X(s,t); (s,t) \in \mathbb{Z}^2\}$ and $\{Y(s,t); (s,t) \in \mathbb{Z}^2\}$, the condition of equal predictability at all horizon for the two processes can be formulated as follow:

$$R_X^2(h_1, h_2) - R_Y^2(h_1, h_2) = 0, \quad (h_1, h_2) \in \mathcal{F}.$$

Now, by giving a horizon of any forecast, we introduce specification predictability equal to all horizons between two stationary ARMA processes.
Let $\{X(s,t); (s,t) \in \mathbb{Z}^2\}$ and $\{Y(s,t); (s,t) \in \mathbb{Z}^2\}$ be two stationary 2D ARMA processes

$$\phi_x(B)X(s,t) = \theta_x(B)\epsilon_x(s,t) \quad \epsilon_x(s,t) \sim WN\left(0, \sigma_x^2\right)$$
$$\phi_y(B)Y(s,t) = \theta_y(B)\epsilon_y(s,t) \quad \epsilon_y(s,t) \sim WN\left(0, \sigma_y^2\right)$$

where

$$\phi_x(z_1, z_2) = \sum_{\substack{i=0 \\ (i,j)\neq(0,0)}}^{p_1} \sum_{j=0}^{p_2} a_{i,j}^x z_1^i z_2^j, \quad \theta_x(z_1, z_2) = \sum_{\substack{i=0 \\ (i,j)\neq(0,0)}}^{q_1} \sum_{j=0}^{q_2} b_{i,j}^x z_1^i z_2^j.$$

$$\phi_y(z_1, z_2) = \sum_{\substack{i=0 \\ (i,j)\neq(0,0)}}^{p_1'} \sum_{j=0}^{p_2'} a_{i,j}^y z_1^i z_2^j, \quad \theta_y(z_1, z_2) = \sum_{\substack{i=0 \\ (i,j)\neq(0,0)}}^{q_1'} \sum_{j=0}^{q_2'} b_{i,j}^y z_1^i z_2^j. \tag{3.11}$$

Since $\{X(s,t); (s,t) \in \mathbb{Z}^2\}$ and $\{Y(s,t); (s,t) \in \mathbb{Z}^2\}$ are stationary they can be represented in the following form:

$$X(s,t) = \sum_{\substack{i=0 \\ (i,j)\neq(0,0)}}^{\infty} \sum_{j=0}^{\infty} \psi_{i,j}\epsilon(s-i, t-j) + \epsilon(s,t) \tag{3.12}$$

$$Y(s,t) = \sum_{\substack{i=0 \\ (i,j)\neq(0,0)}}^{\infty} \sum_{j=0}^{\infty} \delta_{i,j}\eta(s-i, t-j) + \eta(s,t), \tag{3.13}$$

where the sequences $\{\psi_{i,j}\}$ and $\{\delta_{i,j}\}$ are determined by the relations

$$1 - \sum_{\substack{i=0 \\ (i,j)\neq(0,0)}}^{\infty} \sum_{j=0}^{\infty} \psi_{i,j} L^{i,j} = \frac{\theta_x(L)}{\phi_x(L)}$$

$$1 - \sum_{\substack{i=0 \\ (i,j)\neq(0,0)}}^{\infty} \sum_{j=0}^{\infty} \delta_{i,j} L^{i,j} = \frac{\theta_y(L)}{\phi_y(L)}.$$

Proposition 3.1: [48]

Let $\{X(s,t);(s,t)\in\mathbb{Z}^2\}$ and $\{Y(s,t);(s,t)\in\mathbb{Z}^2\}$ be two stationary 2D ARMA processes. The processes $\{X(s,t);(s,t)\in\mathbb{Z}^2\}$ and $\{Y(s,t);(s,t)\in\mathbb{Z}^2\}$ are equally predictable at all horizon if and only if

$$\psi^2_{h_1,h_2}=\delta^2_{h_1,h_2},\quad (h_1,h_2)\in\mathcal{F}$$

Proof: ($\Rightarrow$) If $\{X(s,t);(s,t)\in\mathbb{Z}^2\}$ and $\{Y(s,t);(s,t)\in\mathbb{Z}^2\}$ are equally predictable at all horizons, then

$$R^2_X(h_1,h_2)-R^2_Y(h_1,h_2)=0,\quad \forall (h_1,h_2)\in\mathcal{F}$$

In particular for $h=(1,0)$, we have

$$R^2_X(1,0)-R^2_Y(1,0)=\frac{1}{1+\sum\limits_{\substack{i=0\\(i,j)\neq(0,0)}}^{\infty}\sum\limits_{j=0}^{\infty}\psi^2_{i,j}}-\frac{1}{1+\sum\limits_{\substack{i=0\\(i,j)\neq(0,0)}}^{\infty}\sum\limits_{j=0}^{\infty}\delta^2_{i,j}}=0$$

It follows that

$$\sum_{\substack{i=0\\(i,j)\neq(0,0)}}^{\infty}\sum_{j=0}^{\infty}\psi^2_{i,j}=\sum_{\substack{i=0\\(i,j)\neq(0,0)}}^{\infty}\sum_{j=0}^{\infty}\delta^2_{i,j}.$$

By combining this condition with equal predictability at any horizon (h_1,h_2)

$$\frac{\sum\limits_{i=h_1}^{\infty}\sum\limits_{j=h_2}^{\infty}\psi^2_{i,j}}{1+\sum\limits_{\substack{i=0\\(i,j)\neq(0,0)}}^{\infty}\sum\limits_{j=0}^{\infty}\psi^2_{i,j}}=\frac{\sum\limits_{i=h_1}^{\infty}\sum\limits_{j=h_2}^{\infty}\delta^2_{i,j}}{1+\sum\limits_{\substack{i=0\\(i,j)\neq(0,0)}}^{\infty}\sum\limits_{j=0}^{\infty}\delta^2_{i,j}},\quad (h_1,h_2)\in\mathcal{F}$$

implies that

$$\sum_{i=h_1}^{\infty}\sum_{j=h_2}^{\infty}\psi^2_{i,j}=\sum_{i=h_1}^{\infty}\sum_{j=h_2}^{\infty}\delta^2_{i,j},\quad (h_1,h_2)\in\mathcal{F}.$$

and so

$$\begin{aligned}\psi^2_{h_1,h_2}&=\sum_{i=h_1}^{\infty}\sum_{j=h_2}^{\infty}\psi^2_{i,j}-\sum_{\substack{i=h_1\\(i,j)\neq(h_1,h_2)}}^{\infty}\sum_{j=h_2}^{\infty}\psi^2_{i,j}\\&=\sum_{i=h_1}^{\infty}\sum_{j=h_2}^{\infty}\delta^2_{i,j}-\sum_{\substack{i=h_1\\(i,j)\neq(h_1,h_2)}}\sum_{j=h_2}^{\infty}\delta^2_{i,j}\\&=\delta^2_{h_1,h_2},\quad (h_1,h_2)\in\mathcal{F}.\end{aligned}$$

If

$$R^2_X(h_1,h_2)=R^2_Y(h_1,h_2),\quad (h_1,h_2)\in\mathcal{F}$$

it follows that

$$R^2_X(h_1,h_2)=R^2_Y(h_1,h_2)=0,\quad (h_1,h_2)\in\mathcal{F},$$

that is $\{X(s,t);(s,t)\in\mathbb{Z}^2\}$ and $\{Y(s,t);(s,t)\in\mathbb{Z}^2\}$ are equally predictable at all horizons, which completes the proof. □

3.3.1 Equal predictability and parallelism

Let $\{X(s,t);(s,t)\in\mathbb{Z}^2\}$ and $\{Y(s,t);(s,t)\in\mathbb{Z}^2\}$ be two stationary 2D ARMA processes

$$\begin{aligned}\phi_x(B)X(s,t)&=\theta_x(B)\epsilon(s,t), &\epsilon(s,t)\sim WN\left(0,\sigma_\epsilon^2\right)\\ \phi_y(B)Y(s,t)&=\theta_y(B)\eta(s,t), &\eta(s,t)\sim WN\left(0,\sigma_\eta^2\right)\end{aligned}$$

where $\phi_x(z_1,z_2),\theta_x(z_1,z_2)\,\phi_y(z_1,z_2)$ and $\theta_y(z_1,z_2)$ are finite polynomials in (z_1,z_2) given by. We say that $\{X(s,t);(s,t)\in\mathbb{Z}^2\}$ and $\{Y(s,t);(s,t)\in\mathbb{Z}^2\}$ are parallel if

$$\phi_x(z_1,z_2)=\phi_y(z_1,z_2)\quad\text{and}\quad\theta_x(z_1,z_2)=\theta_y(z_1,z_2).$$

In the domain of time series, this notion has been introduced in [44]. Such is the case in the field of time series, two 2D-ARMA are in parallel if they have the same structure. The following lemma highlights the link between the concept of parallelism and equal predictability at all horizons.

Lemma 3.1: [48]

Let $\{X(s,t);(s,t)\in\mathbb{Z}^2\}$ and $\{Y(s,t);(s,t)\in\mathbb{Z}^2\}$ be two stationary ARMA processes. If $\{X(s,t);(s,t)\in\mathbb{Z}^2\}$ and $\{Y(s,t);(s,t)\in\mathbb{Z}^2\}$ are parallel, then they are equally predictable at all horizons.

Proof: If $\{X(s,t);(s,t)\in\mathbb{Z}^2\}$ and $\{Y(s,t);(s,t)\in\mathbb{Z}^2\}$ are parallel, then

$$\phi_x(z_1,z_2)=\phi_y(z_1,z_2)\quad\text{and}\quad\theta_x(z_1,z_2)=\theta_y(z_1,z_2).$$

Thus

$$\frac{\theta_x(z_1,z_2)}{\phi_x(z_1,z_2)}=\frac{\theta_y(z_1,z_2)}{\phi_y(z_1,z_2)}$$

and so $X(s,t)$ and $Y(s,t)$ have the same moving average representation, that is

$$\psi_{i,j}=\delta_{i,j},\quad\text{for }(i,j)\in\mathcal{F}.$$

It follows that

$$R_X^2(h_1,h_2)=R_Y^2(h_1,h_2),\quad\text{for }(h_1,h_2)\in\mathcal{F}.$$

This completes the proof. □

Through this lemma, it is clear that the condition of parallelism is sufficient to obtain equal predictability at all horizons between two 2D-ARMA processes.
Considering the case of a stationary 2D-ARMA process with equal weightings compared to the same delays, we give the following characterization of equal predictability in any horizon of prediction based on the notion of parallelism.

Proposition 3.2: [48]

Let $\{X(s,t);(s,t)\in\mathbb{Z}^2\}$ and $\{Y(s,t);(s,t)\in\mathbb{Z}^2\}$ be two stationary $2D-ARMA$ processes. Let us suppose that $\psi(i,j),\delta(i,j)\geq 0$ for all (i,j). $\{X(s,t);(s,t)\in\mathbb{Z}^2\}$ and $\{Y(s,t);(s,t)\in\mathbb{Z}^2\}$ are equally predictable at all horizons if and only if $\{X(s,t);(s,t)\in\mathbb{Z}^2\}$ and $\{Y(s,t);(s,t)\in\mathbb{Z}^2\}$ are parallel.

Proof: ($\Rightarrow$) If $\left\{X(s,t);(s,t)\in\mathbb{Z}^2\right\}$ and $\left\{Y(s,t);(s,t)\in\mathbb{Z}^2\right\}$ are equally predictable horizons, then, by Proposition 1, we have

$$\psi^2(h_1,h_2)=\delta^2(h_1,h_2)\quad (h_1,h_2)\in\mathcal{F}$$

Since, $\psi(h_1,h_2)$ and $\delta(h_1,h_2)$ have the same sign for all (h_1,h_2), it follows that $\psi(h_1,h_2)=\delta(h_1,h_2),(h_1,h_2)\in\mathcal{F}$, and hence

$$\psi(L_1,L_2)=\sum_{\substack{i=0\\(i,j)\neq(0,0)}}^{\infty}\sum_{j=0}^{\infty}\psi_{i,j}L^{i,j}=\sum_{\substack{i=0\\(i,j)\neq(0,0)}}^{\infty}\sum_{j=0}^{\infty}\delta_{i,j}L^{i,j}=\delta(L).$$

Hence,

$$\frac{\theta_x(z_1,z_2)}{\phi_x(z_1,z_2)}=\frac{\theta_y(z_1,z_2)}{\phi_y(z_1,z_2)}\tag{3.14}$$

Thus

$$\theta_x(z_1,z_2)=\frac{\theta_y(z_1,z_2)}{\phi_y(z_1,z_2)}\phi_x(z_1,z_2)\quad\text{and}\quad\theta_y(z_1,z_2)=\frac{\theta_x(z_1,z_2)}{\phi_x(z_1,z_2)}\phi_y(z_1,z_2)$$

If $\phi_x(z_1,z_2)\neq\phi_y(z_1,z_2)$, with $p_y\geq p_x$, and if the polynomials $\phi_x(z_1,z_2)$ and $\phi_y(z_1,z_2)$ have $0\leq k\leq p_x$ common roots, it is possible to express $\theta(z_1,z_2)$ in the form of factorization

$$\theta_x(z)=\frac{\theta_y(z)(1-\lambda_{x,1}z)(1-\lambda_{x,2}z)\cdots\left(1-\lambda_{x,(p_x-k)}z\right)}{(1-\lambda_{y,1}z)(1-\lambda_{y,2}z)\cdots\left(1-\lambda_{y,(p_y-k)}z\right)},$$

where $z=(z_1,z_2)'$, $\lambda_{x,i}=\left(\lambda^1_{x,i},\lambda^2_{x,i}\right)$ and $\lambda_{y,i}=\left(\lambda^1_{y,i},\lambda^2_{y,i}\right)$ are the non-common roots of the AR polynomials.
Thus $\theta_x(z)$ is a polynomial of infinite order, but this is absurd since, by hypothesis,

On the other hand, if $\phi_x(z)\neq\phi_y(z)$, with $p_x<p_y$, and if the polynomials $\phi_x(z)$ and $\phi_y(z)$ have $0\leq k\leq p_y$ common roots, factoring the AR operators we have that

$$\theta_y(z)=\frac{\theta_x(z)(1-\delta_{y,1}z)(1-\delta_{y,2}z)\cdots\left(1-\delta_{y,(p_y-k)}z\right)}{(1-\delta_{x,1}z)(1-\delta_{x,2}z)\cdots\left(1-\delta_{x,(p_x-k)}z\right)},$$

where $\delta_{x,i}=\left(\delta^1_{x,i},\delta^2_{x,i}\right)$ and $\delta_{y,i}=\left(\delta^1_{y,j},\lambda^2_{y,i}\right)$ are the non-common roots of the AR polynomials are the non-common roots of the AR polynomials. Thus, $\theta_y(z)$ is a polynomial of infinite order, but this is absurd since, by hypothesis, $q_y<\infty$. It follows that

$$\phi_x(z_1,z_2)=\phi_y(z_1,z_2)\tag{3.15}$$

Equations (3.14) and (3.15) imply that $\theta_x(z_1,z_2)=\theta_y(z_1,z_2)$. Thus, we can conclude that $\{X(s,t)\}$ and $\{Y(s,t)\}$ are parallel.
($\Leftarrow$) If $\{X_t\}$ and $\{Y_t\}$ are parallels, then, by Lemma 1, it follows that $\{X_t\}$ and $\{Y_t\}$ are equal predictable at all horizons. This completes the proof. □

3.3.2 Comparison of predictability in spatial GARCH models

The comparison of the predictability of squared 2DGARCH processes is a direct consequence of the various previous results.
The spatial GARCH model has been introduced by [49] for clutter modeling and anomaly detection. The 2DGARCH (p_1, p_2, q_1, q_2) process is defined as:

$$\begin{cases} \epsilon(s,t) = \sqrt{h(s,t)}\eta(s,t) \\ h_t = \alpha_{0,0} + \sum\limits_{\substack{i=0\\(i,j)\neq(0,0)}}^{q_1} \sum\limits_{l=0}^{q_2} \alpha_{i,j}\epsilon^2_{s-i,t-j} + \sum\limits_{\substack{k=0\\(k,l)\neq(0,0)}}^{p_1} \sum\limits_{l=0}^{p_2} \beta_{k,l} h_{t-k,s-l}, \end{cases} \tag{3.16}$$

where $\alpha_{0,0} > 0, \alpha_{i,j} \geq 0$, and $\beta_{k,l} \geq 0, h_t$ is the conditional variance of ϵ_t (i.e., $\epsilon_t \backslash \psi_t \sim \mathcal{N}(0, h_t)$) and ψ_t denotes all the informations

$$\psi_t = \left\{ \left(\epsilon_{s-k,t-l}\right), 0 \leq k \leq p_1, 0 \leq l \leq p_2, \left(h_{s-i,t-j}\right), 0 \leq i \leq p_1, 0 \leq l \leq p_2, 0 \leq j \leq p_2 \right\}$$

$\eta(s,t) \sim \mathcal{N}(0,1)$ is a stochastic 2D process independent of $h(k,l), \forall (k,l) \leq (s,t)$.
As we need of the second order stationary of the 2D GARCH model like (3.16), we suppose that

$$\sum_{\substack{i=0\\(i,j)\neq(0,0)}}^{q_1} \sum_{j=0}^{q_2} \alpha_{i,j} + \sum_{\substack{k=0\\(k,l)\neq(0,0)}}^{p_1} \sum_{l=0}^{p_2} \beta_{k,l} < 1.$$

It is well known that $\{\epsilon(s,t)\}$ admits the following 2DARMA (q_1^*, q_2^*, p_1, p_2) representation:

$$\epsilon(s,t) = \alpha_{0,0} + \sum_{\substack{i=0\\(i,j)\neq(0,0)}}^{q_1^*} \sum_{j=0}^{q_2^*} (\alpha_{i,j} + \beta_{i,j})\epsilon(s-i,t-j)^2 - \sum_{\substack{k=0\\(k,l)\neq(0,0)}}^{p_1} \sum_{l=0}^{p_2} \beta_{k,l}\eta(s-k,t-l) + \eta(s,t)$$

with $\eta(s,t) = \epsilon(s,t)^2 - h(s,t), q_i^* = \max(p_i, q_i), \alpha_{i,j} = 0 \quad \text{if} \quad (i,j) > (q_1, q_2), \beta_{i,j} = 0 \quad \text{if} \quad (i,j) > (p_1, p_2) . \eta(s,t)$ is, by construction, a martingale-difference sequence relative to the σ-field generated by $\{\epsilon(k,l), (k,l) \leq (s,t)\}$. We denote that the infinite moving average representation of $\epsilon(s,t)^2$ is

$$\epsilon(s,t) = \delta + \psi(L)\eta(s,t),$$

where $\psi(L) = \sum\limits_{i=0}^{\infty}\sum\limits_{j=0}^{\infty} \psi_{i,j} L^{i,j}$. The values of $\{\psi_{i,j}\}$ are obtained in the same way by the relations given previously.

Proposition 3.3: [48]

Let $\{\epsilon(s,t)\}$ and $\zeta(s,t)$ be two covariance stationary GARCH processes, $\left\{\epsilon(s,t)^2\right\}$ and $\left\{\zeta(s,t)^2\right\}$ are equally predictable if and only if $\left\{\epsilon(s,t)^2\right\}$ and $\left\{\zeta(s,t)^2\right\}$ are parallel.

3.3.3 Testing for equal predictability

Through the previous results, it is possible to implement an equal predictability test of two 2D ARMA processes. This test can extend naturally to 2D ARMA process. As future predictions depend only on their past, we assume that these models are estimated separately, it will spread the difficulties of calculations and those related to the identification. It is possible to use Proposition 1 to test the null hypothesis of equal predictability of 2D ARMA processes. Indeed, the null hypothesis can be characterized by:

$$H_0 : \psi_{i,j}^{(1)} = \psi_{i,j}^{(2)} \quad \forall (i,j) \tag{3.17}$$

In what follows, any vector with indices in $\mathbb{Z}^2$ is arranged in the lexicographic order that is

$$(i,j) < (k,l) \text{ if } i < j \text{ or } (i = j \text{ and } j \leq l).$$

Being $\psi_{k,l}^{(h)} = f_h\left(\gamma^{(h)}\right)$, where $\gamma^{(h)} = \left(\phi_{0,1}^{(h}, \phi_{0,2}^{(h)}, \cdots, \theta_{p_1,p_2}^{(h)}, \cdots, \theta_{q_1,q_2}^{(h)}\right)$, the equation which represents the null hypothesis of the test is a set on nonlinear restrictions on the AR and MA coefficients. Let $\hat{\gamma}^{(h)} = \left(\phi_{0,1}^{(h}, \phi_{0,2}^{(h)}, \cdots, \theta_{p_1,p_2}^{(h)}, \cdots, \theta_{q_1,q_2}^{(h)}\right)$, the maximum likelihood estimator of $\gamma^{(h)}$; the maximum likelihood estimator of $\left(\psi_{i,j}^{(h)}\right)^2$ will be $\left(\hat{\psi}_{k,l}^{(h)}\right)^2 = \left(f_h\left(\hat{\gamma}^{(h)}\right)\right)$. The random vector

$$\left(\hat{\psi}^{(h)}\right)^2 = \left(\left(\hat{\psi}_{0,1}^{(1)}\right)^2, \left(\hat{\psi}_{0,2}^{(1)}\right)^2, \cdots, \left(\hat{\psi}_{n_1,n_2}^{(1)}\right)^2, \left(\hat{\psi}_{0,1}^{(2)}\right)^2, \left(\hat{\psi}_{0,2}^{(2)}\right)^2, \cdots, \left(\hat{\psi}_{m_1,m_2}^{(2)}\right)^2\right)'$$

has asymptotic Normal distribution with mean

$$\hat{\psi}^2 = \left(\left(\psi_{0,1}^{(1)}\right)^2, \left(\psi_{0,2}^{(1)}\right)^2, \cdots, \left(\psi_{n_1,n_2}^{(1)}\right)^2, \left(\psi_{0,1}^{(2)}\right)^2, \left(\psi_{0,2}^{(2)}\right)^2, \cdots, \left(\psi_{m_1,m_2}^{(2)}\right)^2\right)'$$

and covariance matrix $G\hat{\wedge}G'$, where G is the $[(n_1m_1 + n_2 - m_2) \times (n_1 + n_2)(p_1 + p_2)(q_1 + q_2)]$ matrix whose the (k,l) the function in ψ^2 with respect to $\theta_{0,1}, \cdots, \theta_{n_1,n_2}$, and $\hat{\wedge}$ is the maximum likelihood estimator of the covariance matrix of $\hat{\theta}_{0,1}, \cdots, \hat{\theta}_{n_1,n_2}$. Being the two 2D ARMA processes independent, $\hat{\wedge}$ is a block diagonal matrix. Having the asymptotic distribution of $\hat{\psi}^2$, we can test the equation of the null hypothesis, rewriting the null hypothesis as a set of restrictions:

$$\mathbf{A}\psi^2 = 0$$

where $\mathbf{A}$ is a $(n_1 + n_2 - 1)(m_1 + m_2) \times (n_1 + n_2)(m_1 + m_2)$ matrix of the form

$$\mathbf{A} = \begin{bmatrix} I_{m_1+m_2} & -I_{m_1+m_2} & 0_{m1+m_2} & \cdots & 0_{m1+m_2} & 0_{m1+m_2} \\ 0_{m_1+m_2} & I_{m_1+m_2} & -I_{m_1+m_2} & \cdots & 0_{m_1+m_2} & 0_{m_1+m_2} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0_{m_1+m_2} & 0_{m_1+m_2} & 0_{m_1+m_2} & \cdots & I_{m_1+m_2} & -I_{m_1+m_2} \end{bmatrix}$$

with $I_{m_1+m_2}$ representing an $(m_1 + m_2) \times (m_1 + m_2)$ identity matrix and $0_{m_1+m_2}$ an $(m_1 + m_2) \times (m_1 + m_2)$ matrix with all the elements equal to zero. The Wald test statistic for the null hypothesis is given by:

$$W = \left(A\hat{\psi}^2\right)' \left(AG\hat{\wedge}\boldsymbol{G}'A'\right)^{-1} \left(\boldsymbol{A}\hat{\psi}^2\right)$$

W is asymptotically distributed as a central chi-square random variable with $(n_1 + n_2 - 1)(m_1 + m_2)$ degrees of freedom. Proposition 2 established the equivalence of the equal predictability and the parallelism between stationary 2D ARMA processes with ψ-weights with the same sign; this result can be extended to the case of equal predictability and parallelism among n 2D ARMA and allows us to verify simply the equality of AR and MA coefficients of n models. The null hypothesis of equal predictability is given by:

$$\begin{array}{lll} \phi_{i,j}^{(1)} = \phi_{i,j}^{(2)} = \cdots = \phi_{i,j}^{(n)} & \text{for each} & 0 \le i \le p_1; 0 \le j \le p_2 \\ \theta_{k,l}^{(1)} = \theta_{k,l}^{(2)} = \cdots = \theta_{k,l}^{(n)} & \text{for each} & 0 \le k \le p_1; 0 \le l \le p_2 \end{array} \tag{3.18}$$

The null hypothesis of equal predictability (Eq. (3.18)) can be expressed as a set of linear restrictions:

$$\boldsymbol{A\theta} = \mathbf{0} \tag{3.19}$$

where $\boldsymbol{\theta} = \left(\phi_{1,0}^{(1)}, \ldots, \phi_{p_1,p_2}^{(1)}, \theta_{1,1}^{(1)}, \ldots, \theta_{q_1,q_2}^{(1)}, \ldots, \phi_{1,0}^{(n)}, \ldots, \phi_{p_1,p_2}^{(n)}, \theta_{1,0}^{(n)}, \ldots, \theta_{q_1,q_2}^{(n)}\right)'$ and $\hat{\wedge}$ is the same matrix of Eq. (16). Of course W is asymptotically distributed as a central chi-square random variable with $(n_1 + n_2 - 1)(m_1 + m_2)$ degrees of freedom. Also, in this case, the matrix $\mathbf{A}$ has to be properly arranged if the models compared have different 2D ARMA orders.

3.4 Numerical results

3.4.1 Simulation study

We present here a simulation study carried out using the statistical software R4.0.0 in order to highlight the validity of the results of the previous sections. The steps involved in the computations are summarized as follows. Consider the stationary first order multiplicative spatial autoregressive model (MSAR(1)) defined by

$$X(s,t) = \phi_{1,0} X(s-1,t) + \phi_{0,1} X(s,t-1) + \phi_{1,1} X(s-1,t-1) \epsilon(s,t) \tag{3.20}$$

where $\left\{\epsilon(s,t); (s,t) \in \mathbb{Z}^2\right\}$ are independent random variables with

$$\mathbf{E}(\epsilon(s,t)) = 0, \mathrm{Var}(\epsilon(s,t)) = \sigma^2, \phi_{1,1} = \phi_{1,0} \cdot \phi_{0,1}$$

and $|\phi_{1,0}| < 1$ and $|\phi_{0,1}| < 1$.

Note that this model is a particular case of the ARMA (p_1, p_2, q_1, q_2) model, it has been reported extensively in the literature (see [49] and references therein).

The comparison of predictability will be made between the simulated series (S), and certain series obtained by transformation. These series are the square root of simulation (SS), its $\log_e$ (LS), and its reciprocal (RS).

Algorithm 2 Equal predictability Algorithm

Step 1. Simulations of DATA

- For several values of ϕ_{ij}
- Generate a date as a $n \times m$ rectangular grid from spatial models of the form (3.20), where $\{\epsilon(s,t)\}$ is a Gaussian White Noise process with mean 0 and variance $\sigma^2 = 25$. Different samples
- **for** the simulated data **do**
- simulate three other types of series,
  - the logarithm of the simulated data;
  - the square root of the simulated data;
  - the reciprocal of the simulated data
- **end for**

Step 2. Estimation of MODELS

- **for** each of the four series do **do**
  - Fit the model $\text{MSAR}_k(1), k = 1, \dots, 4$.
  - Estimate the parameters $\phi^k_{i,j}$ of each model $\text{MSAR}_k(1)$.

Step 3. Test of EQUAL PREDICTABILITY

- Perform the test for parallelism and equal predictability, i.e.
- **for** $k = 1$ to 4 **do**

  Test the hypothesis $H_0 : \phi^k_{i,j} = \phi^l_{i,j}, k \neq l$ against the hypothesis $H_1 : \phi^k_{i,j} \neq \phi^l_{i,j}$, by calculating
  - the covariance matrix $\hat{\wedge}^{(k)}$ for each model $\text{MASR}_k(1)$,
  - the statistic test given by (3.17),
  - the p-values in parenthesis
- **end for**

In Table 3.1, we report the values of the test statistic and p-values in parenthesis.

ϕ		S	RS	SS
$(0.1, 0.1)$	RS	2.3280 (0.2758)		
		18.114 (0.0020)		
	SS	0.551 (0.817)	3.698 (0.130)	
		0.687 (0.670)	8.901 (0.015)	
	LS	0.251 (0.859)	3.7891 (0.131)	0.0075 (0.991)
		0.582 (0.635)	8.880 (0.009)	0.020 (0.989)
$(0.3, 0.3)$	RS	2.2450 (0.3102)		
		18.215 (0.0041)		
	SS	0.462 (0.813)	3.699 (0.119)	
		0.701 (0.702)	8.912 (0.010)	
	LS	0.279 (0.858)	3.910 (0.129)	0.0091 (0.992)
		0.609 (0.704)	8.262 (0.014)	0.0348 (0.995)
$(0.5, 0.5)$	RS	2.2140 (0.2870)		
		18.108 (0.0030)		
	SS	0.456 (0.802)	3.712 (0.125)	
		0.692 (0.688)	8.892 (0.014)	
	LS	0.281 (0.860)	3.890 (0.125)	0.0089 (0.996)
		0.615 (0.698)	8.260 (0.015)	0.0350 (0.997)
$(0.7, 0.7)$	RS	2.2220 (0.2910)		
		17.892 (0.0080)		
	SS	0.551 (0.821)	4.001 (0.178)	
		0.688 (0.696)	8.622 (0.009)	
	LS	0.297 (0.912)	3.580 (0.125)	0.0089 (0.996)
		0.586 (0.499)	8.488 (0.020)	0.0250 (0.993)
$(0.9, 0.9)$	RS	2.980 (0.4410)		
		21.356 (0.0050)		
	SS	0.552 (0.768)	4.102 (0.135)	
		0.683 (0.713)	9.202 (0.018)	
	LS	0.329 (0.659)	4.152 (0.205)	0.0091 (0.920)
		0.705 (0.508)	7.689 (0.030)	0.040 (0.995)

Table 3.1: Values of the test statistic and p-values in parenthesis for the Simulated data. Where S represents the simulation, SS is the square root of simulation, LS is its $\log_e$, and RS is its reciprocal.

It emerges from Table 3.1 that the reciprocal of the simulated data is not parallel to the originally simulated data, the square root transform, and the $\log_e$, but they all have equal predictive ability. it should be noted that for the purpose of modeling and inference it will be appropriate to use either the square root or $\log_e$ transforms of the data since they possess a similar structure to the actual simulated data.

3.4.2 A real data application

To illustrate the application of the proposed results in a real data set, we use the classical Mercer & Hall wheat yield data (WY). This data set is available in the R library (spdep) using simple command data (wheat).

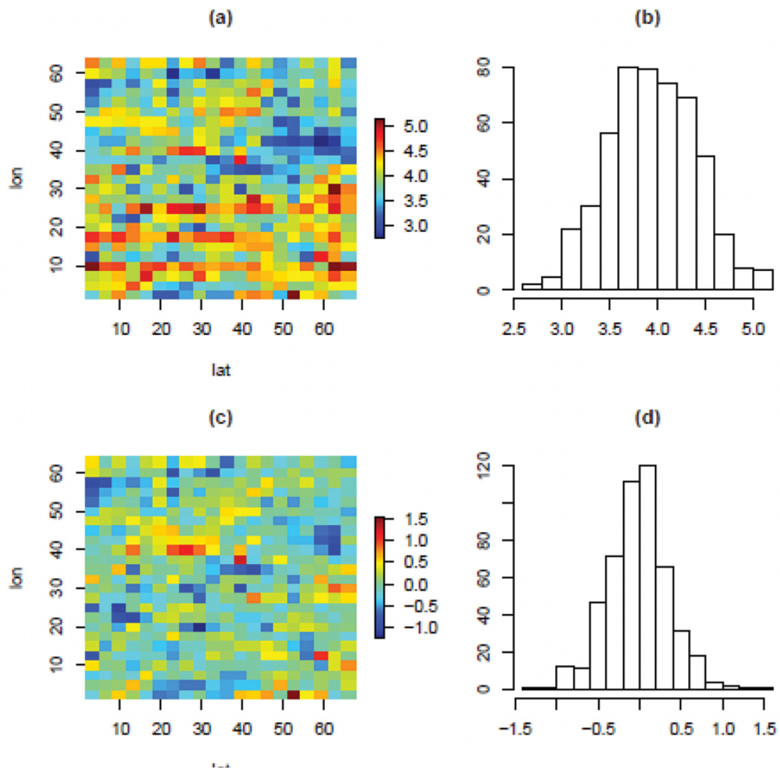

Figure 3.2: Wheat yields data: (a) spatial plot and (b) histogram. Median polish residuals: (c) spatial plot and (d) histogram. is close to normal.

A wheat yield experiment was conducted at Rothamsted Experimental Station in Great Britain. The experiment, a uniformity trial, consisted of giving a 20 by 25 lattice of plots the same treatment with approximately 1 acre area under each plot. The unit of measure of grain productivity is the pound. On the 20 by 25 layout each of the 20 rows runs in the east-west direction and each of the 25 columns runs in the north-south direction. The exact dimensions of the plots that formed the basis of this study are unknown, although some researchers used 3.30 meters from east to west and 2.51 meters from north to south. Previous analysis of this series was done by [50]. The graphical representation of the data is given in Figure 3.2 and its histogram in Figure 3.2(b) shows the familiar bell-shaped curve, indicating the nearly normal distribution for the 500 wheat yield measurements. The mean

(3.95), median (3.94), and mode (3.97) are nearly equal, and the skewness (.036) and kurtosis $(-.254)$ are close to zero, all providing evidence that the distribution Figure 3.2 [51, pp. 284-259] conducted an exploratory data analysis of this data in which he confirmed the presence of an irregular east-west trend in the data. The trend effect is removed by applying median polishing, an exploratory data analysis technique proposed by [52], to the data. The residuals of the median polished data are presented in Figure 2(c) along with its histogram in (d). Here, we use the generalized form of the corrected Akaike's information criteria (AICC) proposed by [5]for model selection in the context of time series models. The AICC is defined as:

$$AICC = |A|\log\left(\frac{RSS_{p_1,p_2,q_1,q_2}}{|A|}\right) + \frac{2R|A|}{|A|-R-1},$$

where RSS_{p_1,p_2,q_1,q_2} is the residual sum of squares, R is the number of parameters and $|A|$ is the cardinality of set A. Using AICC, from all possible pairs (p_1, p_2, q_1, q_2) the pair which minimizes this criterion is selected.
It should be emphasized that the coefficients of MA and AR representation and the autocovariance function of weakly stationary random fields decay at an exponential rate, see, which implies for N and M sufficiently large, only the first coefficients $a_{k,l}$ and $b_{k,l}$ are significant. Therefore, the double infinite sums are truncated at (N, M) for computation.

For this data set, we choose, among all possible pairs (p_1, p_2, q_1, q_2) for which $0 \leq p_1, p_2, q_1, q_2 \leq 10$, the one for which the value of AICC is minimum. We have retained $(p_1, p_2, q_1, q_2) = (2, 1, 3, 1)$. By following the same steps as those stated in Algorithm 1 of Section , we obtain the following results: Table 3.2 Values of the test statistic and p-values in parenthesis for classical Mercer & Hall wheat yield data.

	WY	RWY	SWY
RWY	19.981(0.0001)		
	20.831(0.0001)		
SWY	0.0042(0.991)	22.04(0.0001)	
	0.0137(0.997)	21.37(0.0001)	
LWY	0.0130(0.9996)	22.012(0.0001)	0.0031(0.998)
	0.0501(0.9982)	21.870(0.0001)	0.0141(0.9994)

Table 3.2: WY represents the classical Mercer & Hall wheat yield data, SWY is the square root of WY, LWY is its $\log_e$, and RWY is its reciprocal.

The results of Table 3.2 allow us to confirm that the reciprocal of the wheat yield data is not parallel to the wheat yield data, the square root transform, and the $\log_e$, but they all have equal predictability. But for the purpose of modeling and inference, it will be appropriate to use either the square root or $\log_e$ transforms of the data since they possess a similar structure to the actual data.

3.5 Conclusion and discussion

We recommend that when data need transforming, it will be necessary to also consider whether transformations are parallel to the original data this is because, in building models, the statisticians have devised an iterative strategy of model identification, estimation, and diagnostic checking. The identification stage of their model building cycle relies on the recognition of typical patterns of behavior in the sample autocorrelation function (its transform and partial autocorrelation). If the series to be fitted is subjected to transformation, it can be the case that the autocorrelation function of the transformed series exhibits markedly different behavior patterns from that of the original series which can be easily spotted by checking if there are parallel.

When different data sets (rather than transformations) which may be affected by similar innovations or economic variables are of interest, it will be much easier to investigate if the data sets are parallel, and if they are, it will lessen the task of analyzing and forecasting for each data set. This can be done by collecting the data into different parallel groups and since they are parallel, we select one of them and use it to make inferences about the others.

Conclusion and future prospects

We presented new results on the estimation and prediction problem, more precisely, we have developed an estimation procedure for estimating unknown parameters of 2D-RCAR models. Furthermore, for 2-D ARMA models, we gave a characterization of equal predictability processes by providing necessary and sufficient conditions, which highlight the role of the coefficients of the moving average (MA) and the equivalent autoregressive (EAR) representations.

Research prospects

Many questions are still open in this area. As future perspectives on the results presented in this manuscript.

1. we aim as future work, to develop a more general approach and discuss a general framework for the estimation of nonlinear random fields. This allows us to examine the results obtained in the chapter with new proofs as well as to obtain a number of new results on the asymptotic theory of maximum likelihood. The new approach will be based on Taylor's expansion of a general penalty function. It should be noted that by using a general penalty function, it will be possible to relax the moment conditions on $X_{s,t}$.

2. Furthermore, we have seen in this thesis that the proof of strong consistency of the maximum likelihood estimates requires Θ to be compact which is a quite restrictive condition. Generalizing to nonlinear random fields will imply the relaxation of this condition.

Bibliography

Bibliography

[1] A.D. Cliff and J.K. Ord. Spatial processes: Models and applications (pion limited), 1981.

[2] M. S. Bartlett. *The statistical analysis of spatial pattern*, volume 15. Springer Science & Business Media, 2013.

[3] B.D. Ripley. Spatial statistics john wiley & sons. *New York, New York*, 1981.

[4] G. Upton and B. Fingleton. *Spatial data analysis by example. Volume 1: Point pattern and quantitative data.* John Wiley & Sons Ltd., 1985.

[5] A. Kızılkaya and A. H. Kayran. Estimation of 2-D ARMA model parameters by using equivalent ar approach. *Journal of the Franklin Institute*, 342(1):39–67, 2005.

[6] A. Kayran. 2d arma lattice modelling using two-channel ar lattice approach. *Electronics Letters*, 32(16), 1996.

[7] Y. Xia, Z. Deng, and X. Zheng, Wei. Analysis and application of a novel fast algorithm for 2-D ARMA model parameter estimation. *Automatica*, 49(10):3056–3064, 2013.

[8] M. S. Sadabadi, M. Shafiee, and M. Karrari. Two-dimensional arma model order determination. *ISA transactions*, 48(3):247–253, 2009.

[9] D. Zhipo and X. Youshen. A neural network algorithm for fast blind image restoration using a novel 2D-ARMA parameter estimation. In *2010 International Conference on Audio, Language and Image Processing*, pages 1744–1748. IEEE, 2010.

[10] M. Zeinali and M. Shafiee. A new method for 2-d moving average model parameter estimation. *IETE Journal of Research*, 60(5):364–372, 2014.

[11] D.G. Manolakis, V.K. Ingle, and S.M. Kogon. Statistical and adaptive signal processing, series in electrical and computer engineering, 2000.

[12] P. Loubaton. *Champs stationnaires au sens large sur Z2: Propriétés structurelles et modeles paramétriques.* Ecole nationale supérieure des télécommunications, 1989.

[13] A. Hamaz, O. Arezki, and F. Achemine. Impact of missing data on the prediction of random fields. *Journal of Applied Statistics*, 47(1):132–149, 2020.

[14] A. Bibi. Two-stage generalized moment method approach for bidimensional random coefficient autoregressive models. *Communications in Statistics-Theory and Methods*, 45(14):4268–4284, 2016.

[15] A. Saidi, A. Hamaz, and O. Arezki. Estimation in nonlinear random fields models of autoregressive type with random parameters. *Communications in Statistics-Theory and Methods*, pages 1–16, 2022.

[16] J. Anděl. Autoregressive series with random parameters. *Mathematische Operationsforschung und Statistik*, 7(5):735–741, 1976.

[17] A. Brandt. The stochastic equation $Y_{n+1} = A_n Y_n + B_n$ with stationary coefficients. *Advances in Applied Probability*, 18(1):211–220, 1986.

[18] P. Bougerol and N. Picard. Strict stationarity of generalized autoregressive processes. *The Annals of Probability*, 20(4):1714–1730, 1992.

[19] A. Aue, L. Horváth, and J. Steinebach. Estimation in random coefficient autoregressive models. *Journal of Time Series Analysis*, 27(1):61–76, 2006.

[20] M. Shitan, A. Thavaneswaran, and T. Vazifedan. First-order random coefficient autoregressive (RCA (1)) model: Joint whittle estimation and information. *Acta et Commentationes Universitatis Tartuensis de Mathematica*, 19(1):3–10, 2015.

[21] V. B. Goryainov. Least-modules estimates for spatial autoregression coefficients. *Journal of Computer and Systems Sciences International*, 50(4):565, 2011.

[22] A. V. Goryainov and E.R Goryainova. Comparison of efficiency of estimates by the methods of least absolute deviations and least squares in the autoregression model with random coefficient. *Automation and Remote Control*, 77(9):1579–1588, 2016.

[23] M. Regis, P. Serra, and E. R. van den Heuvel. Random autoregressive models: A structured overview. *Econometric Reviews*, 41(2):207–230, 2022.

[24] S. Mousazadeh and I. Cohen. Two dimensional noncausal ar-arch model: Stationary conditions, parameter estimation and its application to anomaly detection. *Signal processing*, 98:322–336, 2014.

[25] R. S. Tsay. Conditional heteroscedastic time series models. *Journal of the American Statistical association*, 82(398):590–604, 1987.

[26] S. J Leybourne, B. PM McCabe, and A. R. Tremayne. Can economic time series be differenced to stationarity? *Journal of Business & Economic Statistics*, 14(4):435–446, 1996.

[27] D. Huang and V.V. Anh. Estimation of spatial arma models. *Australian Journal of Statistics*, 34(3):513–530, 1992.

[28] S. Basu and G. C. Reinsel. Properties of the spatial unilateral first-order ARMA model. *Advances in applied Probability*, 25(3):631–648, 1993.

[29] P. Whittle. On stationary processes in the plane. *Biometrika*, pages 434–449, 1954.

[30] H. Helson and D. Lowdenslager. Prediction theory and fourier series in several variables. *Acta mathematica*, 99(1):165–202, 1958.

[31] R.L. Kashyap and P.M. Lapsa. Orientation of anisotropic random fields and images. *Pattern Recognition Letters*, 2(4):249–256, 1984.

[32] M. S. Bartlett. Physical nearest-neighbour models and non-linear time-series. *Journal of Applied Probability*, 8(2):222–232, 1971.

[33] J. Besag. Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society: Series B (Methodological)*, 36(2):192–225, 1974.

[34] R. P. Haining. The moving average model for spatial interaction. *Transactions of the Institute of British Geographers*, pages 202–225, 1978.

[35] J. Besag. On the correlation structure of some two-dimensional stationary processes. *Biometrika*, 59(1):43–48, 1972.

[36] T. Tjøstheim. Statistical spatial series modelling ii: some further results on unilateral lattice processes. *Advances in Applied Probability*, 15(3):562–584, 1983.

[37] R. J. Martin. The use of time-series models and methods in the analysis of agricultural field trials. *Communications in Statistics-Theory and Methods*, 19(1):55–81, 1990.

[38] D. Tjøstheim. Statistical spatial series modelling. *Advances in Applied Probability*, 10(1):130–154, 1978.

[39] R. E. Mickens. *Difference equations*, volume 13. Van Nostrand Reinhold New York, 1987.

[40] C. R. Rao. Linear statistical inference and its applications 2nd ed wiley. *New york*, 1:973, 1973.

[41] R. J. Martin. A subclass of lattice processes applied to a problem in planar sampling. *Biometrika*, 66(2):209–217, 1979.

[42] G. C. Reinsel, S. Basu, and S. F. Yap. Maximum likelihood estimators in the multivariate autoregressive moving-average model from a generalized least squares viewpoint. *Journal of Time Series Analysis*, 13(2):133–145, 1992.

[43] S. Basu and G. C. Reinsel. Properties of the first-order spatial bilateral arma model, 1990.

[44] E. Otrano and U. Triacca. Testing for equal predictability of stationary arma processes. *Journal of Applied Statistics*, 34(9):1091–1108, 2007.

[45] O. Bustos, S. Ojeda, and R. Vallejos. Spatial arma models and its applications to image filtering. *Brazilian Journal of Probability and Statistics*, 23(2):141–165, 2009.

[46] J. P. Neary. Putting the "new" into new trade theory: Paul krugman's nobel memorial prize in economics. *Scandinavian Journal of Economics*, 111(2):217–250, 2009.

[47] J. M. Francos, A. Narasimhan, and J. W. Woods. Maximum-likelihood parameter estimation of discrete homogeneous random fields with mixed spectral distributions. *IEEE transactions on signal processing*, 44(5):1242–1255, 1996.

[48] A. Saidi, A. Hamaz, and M. Ibazizen. Measure of predictability of a stationary two dimensionally indexed autoregressive moving-average models. *Communications in Statistics-Simulation and Computation*, pages 1–16, 2022.

[49] A. Noiboar and I. Cohen. Two-dimensional garch model with application to anomaly detection. In *2005 13th European Signal Processing Conference*, pages 1–4. IEEE, 2005.

[50] N. Cressie. Statistics for spatial data: Wiley series in probability and statistics wiley-interscience. *NY*, 15:105–209, 1993.

[51] N. Cressie, C. A. Gotway, and M. O. Grondona. Spatial prediction from networks. *Chemometrics and Intelligent Laboratory Systems*, 7(3):251–271, 1990.

[52] J. W Tukey et al. *Exploratory data analysis*, volume 2. Reading, MA, 1977.

Abstract

This thesis aims at providing a new result on estimation an prediction of 2D- random field. First, we are interested in the study of a class of random coefficient autoregressive spatial models under different assumptions. We begin by studying the probabilistic properties such as stationary. Then, we present new theoretical results on estimation in nonlinear random field models. We focus on two-dimensionally indexed random coefficients autoregressive model with order $(p_1, p_2) \in \mathbb{N}^2$, $2D - RCAR(p_1, p_2)$ for short. We develop a Maximum Likelihood Estimation (MLE) procedure for estimating the unknown parameters of $2D - RCAR(p_1, p_2)$ and we prove that the estimates are strongly consistent. These results are then applied to construct efficient estimates in 2D-RCAR model of order $(0, 1)$ and throught a simulation part, we illustrate the effectiveness and accuracy of the estimates.

Secondly, we consider a measure of predictability of a random field following a stationary two-dimensionally indexed autoregressive moving average (2D ARMA). The purpose of this is two-fold. First, we give a characterization of equal predictability processes by providing necessary and sufficient conditions, which highlight the role of the coefficients of the moving average (MA) and the equivalent autoregressive representations <EAR>. Then, we implement a procedure allowing us to carry out a test of equal predictability. A simulation study is presented and an application to real data of wheat yield is reported.

Keywords: 2D-RCAR models, Nonlinear random fields, Maximum likelihood, Strong consistency, Stationary random fields; 2-D ARMA model; Moving average representation; Autoregressive representation; Predictability

Résumé

Cette thèse vise à fournir un nouveau résultat sur l'estimation et la prédiction des champs aléatoires 2D. Tout d'abord, nous nous intéressons à l'étude d'une classe de modèles spatiaux autorégressifs à coefficient aléatoire sous différentes hypothèses. Nous commençons par étudier les propriétés probabilistes telles que la stationnarité. Ensuite, nous présentons de nouveaux résultats théoriques sur l'estimation dans les modèles de champs aléatoires non linéaires. Nous nous concentrons sur le modèle autorégressif à coefficients aléatoires à indexation bidimensionnelle d'ordre $(p_1, p_2) \in \mathbb{N}^2$, $2D - RCAR(p_1, p_2)$ en abrégé. Nous développons une procédure d'estimation par maximum de vraisemblance (MLE) afin d'estimer les paramètres inconnus de $2D - RCAR(p_1, p_2)$ et nous prouvons que les estimateurs sont fortement consistants. Ces résultats sont ensuite appliqués pour construire des estimateurs efficaces dans le modèle $2D - RCAR$ d'ordre $(0, 1)$ et à travers une partie de simulation, nous illustrons l'efficacité et la précision des estimations.

Deuxièmement, nous considérons une mesure de la prédictabilité d'un champ aléatoire suivant une moyenne mobile autorégressive stationnaire à indexation bidimensionnelle (2D ARMA). L'objectif est double. D'abord, donner une caractérisation des processus de prédictabilité égale en fournissant des conditions nécessaires et suffisantes, qui mettent en évidence le rôle des coefficients de la moyenne mobile (MA) et des représentations autorégressives équivalentes (EAR). Ensuite, mettre en œuvre une procédure permettant d'effectuer un test d'égale prédictibilité. Une étude de simulation est présentée et une application sur des données réelles de rendement de blé est rapportée.

Mots-clés: Modèles 2D-RCAR, Champs aléatoires non linéaires, Maximum de vraisemblance, Cohérence forte, Champs aléatoires stationnaires ; Modèle ARMA 2-D ; Représentation moyenne mobile ; Représentation autorégressive ; Prédictibilité