Stationarity and unit roots in spatial autoregressive models

Andrés Ramírez Hassan

Universidad Nacional de Colombia
Facultad de Ciencias, Escuela de Estadística
Medellín, Colombia
2012

Stationarity and unit roots in spatial autoregressive models

Andrés Ramírez Hassan

This thesis is presented as partial requirement to be a:
Doctor en Ciencias-Estadística

Advisers:
Ph.D. Estadística, Juan Carlos Correa
Ph.D. Estadística, Juan Carlos Salazar

Universidad Nacional de Colombia
Facultad de Ciencias, Escuela de Estadística
Medellín, Colombia
2012

To my parents, Orlando and Nancy, who have given me their unconditional support. They have taught me that the primary aspect of the human being's evolution is humility. Unfortunately, I have not learned very well their lesson.

Acknowledgement

The author wishes to thank his advisers Juan Carlos Correa and Juan Carlos Salazar for helpful comments on earlier versions of this thesis. Also, I wish to thank EAFIT University for giving its financial support.

Abstract

Stationarity is a common assumption in statistical inference when data come from a random field, but this hypothesis has to be checked in order to avoid falling into nonsense regressions and inconsistent estimates. In this thesis, consequences on statistical inference associated with non-stationary random fields are shown, specifically due to a spatial unit root. A statistical test to check a spatial unit root for spatial autoregressive models is built in the frequency domain, and its asymptotic distribution found. Monte Carlo simulations are used to obtain the small sample properties of the proposed statistical test, and it is found that the size of the test is good, and the power of the test improves if the spatial autocorrelation coefficient decreases. Additionally, we find that the size of our test is better than other spatial unit root tests when the data generating process is not a spatial autoregressive model. In order to get better small sample properties of the test when there is a spatial unit root near to one, a Monte Carlo test is performed. Finally, two applications are done; first, the Mercer-Hall dataset, which is one of the most analysed regular lattice data in the literature, is studied. It is found that the data do not have a spatial unit root, although the dataset is highly autocorrelated. And second, data of electricity demand in the Department of Antioquia (Colombia) are studied. Statistical evidence based on different tests suggest that electricity consumption does not have a spatial unit root; therefore, parameter estimates are sensible. Specifically, it is found that the price elasticity of electricity demand is -1.150 while the income elasticity is 0.408.

Keywords: Stationarity, Random Fields, Spatial Unit Root Test, Spatial Autoregressive Models, Periodogram, Covariance, Monte Carlo Simulation.

Resumen

La hipótesis de estacionariedad es un supuesto común cuando los datos provienen de una realización de un campo aleatorio, pero esta hipótesis debe ser verificada para evitar caer en problemas de regresiones sin sentido o inconsistencia de los parámetros estimados. En esta tesis se muestran las consecuencias sobre la inferencia estadística asociadas a la no estacionariedad de los campos aleatorios, específicamente debido a la presencia de una raíz unitaria espacial. Se propone un estadístico de prueba en el dominio de las frecuencias para corroborar la presencia de una raíz unitaria espacial y se encuentra su distribución asintótica. Se utiliza simulación Monte Carlo para obtener las propiedades para muestras pequeñas del estadístico propuesto, y se observa que el tamaño es bueno, y que la potencia del estadístico mejora si la autocorrelación espacial disminuye. Adicionalmente, se encuentra que el tamaño de nuestra prueba supera al obtenido con otras pruebas para corroborar la presencia de una raíz unitaria espacial cuando el proceso generador de datos no es un proceso espacial autorregresivo. Dado el objetivo de mejorar la potencia del estadístico de prueba cuando se presenta una raíz espacial cercana a uno, se construye un estadístico fundamentado en simulación Monte Carlo. Finalmente se realizan dos aplicaciones, la primera consiste en el análisis de los datos Mercer-Hall, los cuales son una de la base de datos más citada en la literatura de datos en rejillas regulares, y se encuentra que las series en consideración no presentan raíz unitaria espacial, aunque están espacialmente autocorrelacionadas. Y en la segunda, se estudian los datos de la demanda de electricidad en el Departamento de Antioquia (Colombia). La evidencia estadística fundamentada en diferentes pruebas indica que el consumo de electricidad no tiene una raíz espacial unitaria; lo cual implica que los parámetros estimados tienen sentido. Específicamente, se encuentra que la elasticidad precio de la demanda de electricidad es -1.150, mientras que la elasticidad ingreso de la demanda es 0.408.

Palabras claves: Estacionariedad, Campos Aleatorios, Prueba de Raíz Unitaria Espacial, Modelo Espacial Autorregresivo, Periodograma, Covarianza, Simulación Monte Carlo.

Contents

1 Introduction

> "...Perhaps the foremost reason for studying spatial statistics is that we are often not only interested in answering the "how much" question, but the "how much is where" question."
>
> *Schabenberger and Gotway (2005, pp 1)*

In statistics, spatial analysis or spatial statistics include any of the formal techniques which study entities using their topological, geometric or geographic properties.

Most authors applying statistical methods for spatial data agree that one of the key features of this kind of data is the presence of spatial autocorrelation (Schabenberger and Gotway, 2005). This idea is summarized by a fundamental concept in geography which says that nearby entities often share more similarities than entities which are far apart. This idea is often labeled Tobler's first law of geography and may be summarized as

> "...everything is related to everything else, but near things are more related than distant things."
>
> *Tobler (1970, pp 237)*

Given this fact, it is necessary to establish a series of statistical tools in order to handle this kind of characteristic. Technically, the data generating process in space is seen as a random field (see Chapter 2). This means that a sample of size n in the space represents a single realization of a random experiment; a sample of size one from a n-dimensional distribution.

The classification in spatial statistics is characterized by the nature of the spatial domain (Cressie, 1993) where the spatial domain is a subset in $\mathbb{R}^d$. Typically, these data fall into three categories: geostatistical data, regional data (lattice data) and point pattern data. In geostatistical data and regional data the domain is fixed (the points in a subset of $\mathbb{R}^d$ are non-stochastic) but regional data are characterized by a discrete domain and geostatistical data by a continuous domain. On the other hand, the important feature of point pattern data is the random domain.

The principal concern in this dissertation is related to regional data; specifically, unit root processes in the spatial domain and the implications of assuming weak stationarity. In this context, where a simultaneous spatial autoregressive model is assumed with a row standardised contiguity matrix, a spatial unit root process is characterized by a spatial autocorrelation coefficient equal to one. This implies that the spatial impulse response function will not tend to zero as the distance between a pair of locations tends to infinity. This phenomenon means that the process is not stationary because its mean and variance depend on the absolute location of the spatial units (see Chapter 2).

This topic is important because omitting spatial effects in the estimation process causes inconsistency in the parameter estimates (Anselin, 1988), and regressions between non-stationary spatial series leads to nonsense outcomes. Specifically, Fingleton (1999) shows some analogies between time unit root processes and spatial unit root processes. In his paper is evidenced through simulation exercises that the variance of spatial autoregressive unit root processes depends on locations, and also, this tends to increase with an increment in the sample size. Additionally, the author shows how regressions between spatial autoregressive unit root processes generate nonsense outcomes. These outcomes are similarly found by Lauridsen and Kosfeld (2006, 2007), and equally, we achieve them in this dissertation. On the other hand, given a spatial autoregressive process with row standardised contiguity matrix, the spatial autocorrelation parameter should be between $1/\omega_{min}$ and 1, where ω_{min} is the smallest negative eigenvalue of the contiguity matrix (Ord, 1975; Anselin, 1982). This condition is strongly related to weak stationarity of a random field. Specifically, a spatial autocorrelation parameter in this range is necessary to ensure weak stationarity but is not sufficient due to edge and corner effects (Haining, 1990).

Although Mur and Trívez (2003) and Paelinck et al. (2004) warn about the application of the concept of unit root in the spatial context, there is a fact that is undeniable, a collection of data in the space is just one realisation of a random field. The implications are formidable: how does a researcher learn anything about the statistical properties of a random field if only a single realisation is available? Thus, it is necessary to develop statistical tests that contribute to check the stationarity hypothesis.

Regional data is the nearest spatial category to time series; although most work has been done to test the non-stationarity hypothesis in time series (Priestley and Rao, 1969; Dickey and Fuller, 1979; Phillips and Perron, 1988; Kwiatkowski et al., 1992), which is a single realization of a random process in one dimension. There is not much literature about formal procedures to test stationarity in spatial stochastic processes (Fuentes, 2005). However, this is an old subject since Whittle (1954) put it in discussion.

Bhattacharyya et al. (1997) and Baran et al. (2004) develop asymptotic inference for near unit process in the spatial autoregressive model $z(s_{1i}, s_{2j}) = \alpha z(s_{1i-1}, s_{2j}) + \beta z(s_{1i}, s_{2j-1}) - \alpha\beta z(s_{1i-1}, s_{2j-1}) + \epsilon(s_{1i}, s_{2j})$ which can be considered as being very simple because this model can be reduced to two one-dimensional autoregressions, $(1-\alpha L_1)(1-\beta L_2)z(s_{1i},s_{2j}) = \epsilon(s_{1i},s_{2j})$ where $L_1 z(s_{1i},s_{2j}) = z(s_{1i-1},s_{2j})$ and $L_2 z(s_{1i},s_{2j}) = z(s_{1i},s_{2j-1})$. Paulauskas (2007) considers the unit root case in the autoregressive model $z(s_{1i},s_{2j}) = \alpha z(s_{1i-1},s_{2j}) + \beta z(s_{1i},s_{2j-1}) + \epsilon(s_{1i},s_{2j})$. This author shows that the growth of variance of the process depends on the dimension of the lattice. Baran and Pap (2011) analyse the spatial autoregressive model $z(s_{1i},s_{2j}) = \alpha z(s_{1i-1},s_{2j}) + \beta z(s_{1i},s_{2j-1}) + \gamma z(s_{1i-1},s_{2j-1}) + \epsilon(s_{1i},s_{2j})$ in the unit root case, and find the limiting distribution and the rate of convergence of the least square estimator. It can be seen from the previous formulations that these models do not incorporate all the possible spatial interactions that $z(s_{1i},s_{2j})$ could have. Fingleton (1999) is the first author that introduces the concept of a unit root in regional data where all possible spatial interactions are considered. This author illustrates its implications via Monte Carlo simulations. He finds that a unit root leads to spurious spatial regression, like in the well known case of time series, and demonstrates that Ordinary Least Square estimation of spatial error correction models is not consistent. This author postulates the Moran's test as a diagnostic indicator of the presence of a unit root in spatial context. Specifically, Fingleton (1999) performs a Monte Carlo experiment in which a couple of independent Spatial Autoregressive processes are generated (see Chapter 2). These two spatial processes are used to conduct bivariate regressions, and the t-ratios for the null hypothesis of no relation between the processes are recorded. This author shows that when there are spatial unit root processes, the empirical size of the statistical test is considerably bigger than the nominal size, and this problem worsens if the sample size increases. This phenomenon implies a spurious relation between independent variables. Therefore, the author proposes the Moran's test for regression residuals as a diagnostic indicator of the presence of a spatial unit root. But there is one open question in his article: the null hypothesis in the Moran's test is not spatial autocorrelation, so rejecting the null hypothesis means autocorrelation or non-stationarity.

Recently, there has been an increasing interest in analysing spurious spatial regression when there is a near unit root process ($\rho = 1 - 1/\psi, \psi \to \infty$). Specifically, Lee and Yu (2009) and Baltagi and Liu (2010) investigate spurious spatial regression where the regressant and regressors may be generated from possible non-stationary processes. The former find that with a row-normalized spatial weights matrix, the possible spurious regression phenomena in the spatial setting are weaker than those in the non-stationary time series case. The latter study the case where the weight matrix is normalized and has equal elements, it is shown that the spurious spatial regression does not occur in a spatially autoregressive model. Actually, the asymptotic distribution of the OLS estimate converges to its true value zero. Additionally, Lee and Yu (2007), Martellosio (2010) and Roknossadati and Zarepour (2011) have studied properties of different mechanisms to estimate spatial unit root processes when there is a spatial effect parameter near to unity.

With regard to statistical tests to check stationarity in random fields, we can do a taxonomical classification. A first approximation can be done by the nature of the spatial domain. Specifically, Guan (2008) develops a statistical test to check stationarity for spatial point processes. This test is based on the integrated squared deviations of observed counts of events from their means estimated under stationarity, and the convergence of their partial sum processes. On the other hand, Bose and Steinhardt (1996), Ephraty et al. (2001), Mateu and Juan (2004) and Fuentes (2005) develop statistics to test stationarity for geostatistical data. Bose and Steinhardt (1996) build a statistic based on the covariance structure of the process to test stationarity. In particular, the centrosymmetric property is used which is exhibited by the sample covariance matrix of spatially stationary fields sampled at a uniform linear array. The centrosymmetric property reflects the fact that the operations of reversing the indexing of the elements do not alter the correlation matrix as long as spatial stationarity holds. The authors use invariance principles to ensure that the tests have constant significance under the null hypothesis of a stationary Gaussian Random Field (see Chapter 2). Specifically, they use the maximum eigenvalue test, the determinant test and the trace test associated with the projection onto the row space of subvectors of the measured random field. Ephraty et al. (2001) develop a statistical test to check stationarity in spatio-temporal geostatistical data. This test is based on the spatial cumulant spectrum, its properties in the stationary case and the assumption of temporal ergodicity of the measured random field. Under the null hypothesis of a Gaussian Random Field, the asymptotic distribution of their statistical test is proportional to a Chi-square distribution with known degrees of freedom. Mateu and Juan (2004) develop a statistical test based on the spectral representation of a non-stationary random field. The practical implementation of the test, given a spatial process sampled at regularly spaced data, is the following: First, select a number of subregions with equal sizes. Then, for each subregion, estimate the tapered periodogram. Third, estimate the coefficients (intercept and slope) of the regression $log(I_n^i(\xi_1, \xi_2)) = \beta_{0i} + \beta_{1i} log(\|\boldsymbol{\xi}\|) + \epsilon_i$ where $I_n^i(\xi_1, \xi_2)$ is the periodogram at frequencies ξ_1 and ξ_2 for subregion i, $\|\boldsymbol{\xi}\|$ is the Euclidean norm, ϵ_i is a stochastic perturbation and β_{0i} and β_{1i} are coefficients to be estimated. Finally, calculate the statistical test to check any difference between the parameters estimates in each subregion. Any statistical difference implies that the random field is not stationary. Finally, Fuentes (2005) uses the concept of evolutionary spatial spectrum, which means that the spatial spectral density function varies in space. The proposed method consists in testing the homogeneity of a set of spatial spectra evaluated at different locations. In particular, the evolutionary spatial spectral density is estimated at n nodes that constitute a systematic sample of all locations in a regular grid, and given the asymptotic results in the paper, the test to check stationarity is reduced to a simple two-factor analysis of variance of spectral estimates at different locations. In the context of regional data, Bhattacharyya et al. (2000), Lauridsen and Kosfeld (2004, 2006, 2007) and Beenstock and Felsenstein (2008) have proposed statistical tests to check non-stationarity associated with a spatial unit root process. Bhattacharyya et al. (2000) propose two statistical tests to check stationarity in the process $z(s_{1i}, s_{2j}) = \alpha z(s_{1i-1}, s_{2j}) + \beta z(s_{1i}, s_{2j-1}) - \alpha\beta z(s_{1i-1}, s_{2j-1}) + \epsilon(s_{1i}, s_{2j})$. Specifically, given the testing problem $H_0 : \alpha = \beta = 1$ versus $H_1 : 0 < \alpha, \beta < 1$ a periodogram test $\phi_n = 16\pi^2 I(\xi_1, \xi_2)/(\sigma^2 n^4) \xrightarrow{d} 3\chi_2^2$ is postulated. Once a significance level is established, a decision can be taken . Additionally, a spatial domain test $\psi_n = n^{3/2}(\hat{\theta} - (1,1)') \xrightarrow{d} \mathcal{N}(0, \Gamma)$ where $\Gamma = diag(2,2)$ for testing $H_0 : \alpha = \beta = 1$ is also proposed. It is found that the asymptotic power of the spatial test is one and over performs the power of the periodogram test. This outcome is due to the parametric test (ψ) performing better under the assumptions of the parametric models. However, the advantage of the periodogram test is that the periodogram can always be computed and is less sensitive to model assumptions. Lauridsen and Kosfeld (2004) propose a Wald test based on the maximum likelihood estimation. Their proposal resembles the Dickey-Fuller approach applied to time series. However, it is known that $(1-\hat{\rho})/\widehat{s.e.(\hat{\rho})}$ does not converge to the standard normal or t distribution under the null hypothesis of a spatial unit root process, i.e. $\rho = 1$. Thus, a Monte Carlo simulation is used to deduce the distribution of the statistical test. It is seen from the simulation exercises that the critical limits of the Wald test under the null hypothesis are higher than for the χ_1^2 distribution. Additionally, Lauridsen and Kosfeld (2006) propose a two-step Lagrange Multiplier test to check spatial non-stationarity. In the first step, the LM error statistic developed by Anselin (1988) is used to test the null hypothesis of no spatial autocorrelation. Therefore, a large statistic value indicates either a spatial unit root process or a stationary spatial autocorrelated process. In the second step, it is proposed to make the regression using the spatial differenced process, $\Delta z(s_{1i}, s_{2j})$ (see Chapter 4). If there is a spatial unit root in the original process, the differenced process is a white noise process, so that the LM error test statistic for this spatially differenced model will be close to zero. On the other hand, if the null hypothesis of non-stationarity does not hold, the errors resulting from spatial overdifferencing are expected to go along with a positive differenced LM value. Lauridsen and Kosfeld (2007) generalize this procedure by incorporating control for unobserved heteroscedasticity through a Lagrange Multiplier test developed by Anselin (1988) which adjusted for unobserved heteroscedasticity. Recently, Beenstock and Felsenstein (2008) point out that if the residuals contain a spatial unit root, the regression coefficient estimates will be *nonsense* rather than *spurious*.[1] Beenstock and Felsenstein (2008) develop a spatial "Dickey-Fuller" test under the null hypothesis of a spatial unit root, and find its empirical distribution through Monte Carlo simulations. These authors establish in a spatial autoregressive model that a spatial unit root process is generated by an autoregressive coefficient equal to the reciprocal of the number of neighbours.[2] Therefore, these authors perform Monte Carlo exercises where spatial unit root processes are simulated. Then, they use these synthetic datasets to estimate spatial autoregressive models by maximum likelihood. After that, they build the empirical distribution of the spatial autoregressive coefficient estimates, but they have to truncate the

[1] Spurious regression is induced by the fact that the mean of the series increases or decreases with the domain. Nonsense regression is induced by the fact that the variance increases with the domain.

[2] This outcome assumes that the contiguity matrix is based on a binary criteria. However, this argument leaves out edge and corner effects (Haining, 1990).

distribution because they obtain coefficient estimates greater or equal to the upper limit of stationarity. Finally, they report the critical values from the truncated distribution. Therefore, if an estimated coefficient in a spatial autoregressive model is greater than the critical value at some significance level, the null hypothesis of a spatial unit root cannot be rejected.

It can be helpful to do a taxonomical classification by domain. In this way, the statistical tests to check stationarity based on the spatial domain are: Bose and Steinhardt (1996), Lauridsen and Kosfeld (2004, 2006, 2007), Beenstock and Felsenstein (2008) and Guan (2008). On the other hand, the statistics based on the frequency domain are: Bhattacharyya et al. (2000), Ephraty et al. (2001), Mateu and Juan (2004) and Fuentes (2005). The most important advantage of the statistical tests in the spatial domain is that these statistics over perform the power of the periodogram tests under the assumption of the parametric model. However, the advantage of the periodogram tests is that the periodogram can always be computed and they are less sensitive to model assumptions. Unfortunately, conventional estimation of the spectral density function through the periodogram is based on a regular lattice. And, we have not found any application of stationarity tests on irregular lattices or regional data. So, one of the most important contributions of this dissertation is to propose a spatial unit root test based on the frequency domain and use it to check stationarity in regional data.

If we think about the disadvantages of the statistical tests to check stationarity of random fields, we can find some limitations associated with them and their possible application in regional data. Specifically, the asymptotic distribution of the statistical test proposed by Guan (2008) is based on the convergence of a partial sum of its integrated squared deviations of observed counts of events from its mean. This asymptotic outcome is based on a result given by Ivanoff (1982), and unfortunately, we have not found this theorem for regional data. Additionally, this test is based on the concept of a two-dimensional Brownian motion, so the test is sensitive to the determination of an initial point. The test proposed by Bose and Steinhardt (1996) is sensitive to the property of centrosymmetry which is exhibited by the sample covariance matrix of a spatially stationary field sampled at a uniform linear array. Thus, it can be used to test stationarity in regular lattices but not on regional data. On the other hand, Ephraty et al. (2001) argue that their test can be applicable to an arbitrary geometry; however, they do not show this extension in their paper. But, this is not the biggest restriction in this test; the biggest restriction is that it is necessary to have spatio-temporal data to build the test. The statistical tests proposed by Mateu and Juan (2004) and Fuentes (2005) are implemented in regular lattices and are based on the concept of an evolutionary spectrum; this concept can be complicated and implies a big computational burden. Although these tests are based on the periodogram, which is a nonparametric estimator, there are some assumptions about the family of the spectral density function. Specifically, these statistics assume a Matér spectral density function. The statistical methodology used by Bhattacharyya et al. (2000) to propose their statistic is too related to the methodology that we use to build our statistical test. However, there is a big difference: our model is a simultaneous spatial autoregressive model; while the model proposed in Bhattacharyya et al. (2000) does not include all the possible multilateral effects. The tests proposed by Lauridsen and Kosfeld (2004, 2006, 2007) and Beenstock and Felsenstein (2008) can be applied to regular and irregular data. However, the statistical test proposed by Lauridsen and Kosfeld (2006, 2007) is based on a two-step strategy; therefore, the second stage is conditioned on the conclusion from the first stage. This implies that the overall type I error of the test will grow. Finally, there is a problem with the procedure proposed by Lauridsen and Kosfeld (2004) and Beenstock and Felsenstein (2008); they use Maximum Likelihood to estimate the spatial autoregressive coefficients, but it is known that under the null hypothesis of a spatial unit root, the probability distribution function of a spatial autoregressive field degenerates to zero (Anselin, 1982). Specifically, the asymptotic ML estimates will only hold if the regularity conditions for the log-likelihood function are satisfied. This statement is not fulfilled under the null hypothesis.

The main objectives of this thesis are the following:

- Show the consequences on statistical inference associated with non-stationary random fields, specifically due to a spatial unit root.
- Propose a statistic to test the null hypothesis of a spatial unit root in spatial autoregressive models and find its asymptotic distribution.
- Use a Monte Carlo statistical test as a tool to improve the finite sample performance of the theoretical test.
- Perform Monte Carlo simulations to obtain the small sample properties of the statistical test.
- Apply the spatial statistical methodology which is proposed, and specifically, the statistic to test the null hypothesis of a spatial unit root in particular datasets. Specifically, the Mercer-Hall data and the electricity demand in the Department of Antioquia (Colombia) are used.

The main achievements in this dissertation is proposing a statistic to test the null hypothesis of a spatial unit root random field, and showing its asymptotic distribution under the null hypothesis (Theorem 1). Additionally, we extend the test to cover a generalised spatial autoregressive process (Theorem 2). This theoretical development contributes to spatial statistics, and specially regional data analysis, because stationarity of the random field which generates a specific realization of lattice data is an implicit assumption that needs to be checked.

With regard to the test, it is found that the size of the test is closer to the nominal size as the sample increases, and also, the power of the test improves if the spatial autocorrelation coefficient decreases. Although other spatial unit root tests over perform our test in some circumstances when the data generating process is a spatial autoregressive process, our test gets the best size when the data generating process is not a spatial autoregressive process.

In applications, we find that the Mercer-Hall data do not have a spatial unit root, although, this dataset is highly autocorrelated. With regard to the electricity demand in the Department of Antioquia (Colombia), all tests indicate that there is no spatial unit root in the electricity consumption. Additionally, we find that our test is robust to different schemes of discretisation of regional data. Therefore, the elasticity estimates are sensible, and imply that an increment of 1% in electricity price means a reduction of 1.150% in electricity consumption, while the income elasticity is 0.408.

In the next chapter, some elements of discrete random fields are given. Specifically, the representation of a bi-dimensional random field in the spatial and frequency domains is shown, and the relation of a pure spatial autoregressive process and its representation in the frequency domain is highlighted. Chapter 3 examines some consequences in the statistical inference because of a non-stationary spatial unit root process. In Chapter 4, a statistic is proposed to test the null hypothesis of a spatial unit root random field, its asymptotic distribution is found, and some Monte Carlo simulations are performed to analyse the small sample properties. In order to get better small sample properties, a Monte Carlo statistical test is proposed. Chapter 5 shows some applications, and the final chapter summarises key research issues and postulates some future research.

2 Some theory: an introduction to discrete random fields

This section is strongly based on Cressie (1993), Billingsley (1995), Ripley (2004) and Schabenberger and Gotway (2005).

Given a complete probability space $(\Omega, \mathcal{F}, \mathcal{P})$ and $D \subseteq \mathbb{R}^d$ a topological space, then a measurable mapping $\mathbf{z}(\mathbf{s}, \omega) : D \times \Omega \to \mathbb{R}^n$ is called a *vector valued random field* (Adler and Taylor, 2007). This means that each component $\mathbf{z}(\mathbf{s}, \omega)$, which is located in $\mathbf{s}$, is an outcome of a random experiment $\omega \in \Omega$ where a particular realization produces a surface $\mathbf{z}(.\,, \omega)$. As a consequence, the collection of n indexed observations that make up the dataset do not represent a sample of size n. They represent a single realization of a random experiment; a sample of size one from a n-dimensional distribution. In this context, the concept of stationarity is very important. Specifically, given $\omega \in \Omega$, a random field $\mathbf{z}(.\,, \omega)$ is called a *strong stationary field* if the spatial distribution is invariant under translation of the index, i.e.,

$$Pr(\mathbf{z}(\mathbf{s_1}) < z_1, \mathbf{z}(\mathbf{s_2}) < z_2, ..., \mathbf{z}(\mathbf{s_n}) < z_n) = Pr(\mathbf{z}(\mathbf{s_1} + \mathbf{h}) < z_1, \mathbf{z}(\mathbf{s_2} + \mathbf{h}) < z_2, ..., \mathbf{z}(\mathbf{s_n} + \mathbf{h}) < z_n) \tag{2.1}$$

for all n and $\mathbf{h}$.

A strong stationary random field repeats itself throughout the domain. But this is a stringent condition; most statistical methods for spatial analysis are satisfied with stationary conditions based on the moments of the spatial distribution.

Weak stationarity of a random field implies that $Cov_\Omega[\mathbf{z}(\mathbf{s}, \omega), \mathbf{z}(\mathbf{s} + \mathbf{h}, \omega)] = C(\mathbf{h})$ and $E_\Omega[\mathbf{z}(\mathbf{s}, \omega)] = \mu$. The mean of a weak stationary random field is constant and the covariance between attributes at different locations is only a function of their separation $\mathbf{h}$, this implies that the variance of a weak stationary process is constant. Weak stationarity reflects the lack of importance of absolute position. Strong stationarity implies weak stationarity but the converse is not true.

If $\mathbf{z}(.\,, \omega)$ is not weak stationary, the increments $\Delta(\mathbf{z}(.\,, \omega))$ might be, where

$$\Delta(\mathbf{z}(.\,, \omega)) = \{\mathbf{z}(\mathbf{s}, \omega) - \mathbf{z}(\mathbf{s} + \mathbf{h}, \omega) : \mathbf{s}, \mathbf{h} \in D \subset \mathbb{R}^d, \omega \in \Omega\} \tag{2.2}$$

A process that has this characteristic is said to have *intrinsic stationarity*.

Statistical analysis becomes much simpler if the process is assumed to be stationary. However,

> "... Analysing observations from a stochastic process as if the process were stationary -when it is not- can lead to erroneous inferences and conclusions."
>
> *Schabenberger and Gotway (2005, pp 42)*

In the case of regional (lattice) data, the spatial domain is fixed and discrete. This means that a regional spatial process in the plane has the following representation:

$$\{\mathbf{z}((s_{1i}, s_{2j}), \omega) | (s_{1i}, s_{2j}) \in D \text{ a countable set} \subset \mathbb{Z}^2, \omega \in \Omega : D \times \Omega \to \mathbb{R}\} \tag{2.3}$$

This representation is very abstract and reveals little about the structure of the random field under study. Thus, it is necessary to represent it in another way, specifically, the random field may be formulated in the spatial domain or in the frequency domain. The distinction between spatial and frequency representation depends on whether the process is expressed in terms of functions of the observed coordinates, or in terms of a random field contained in a space consisting of frequencies.

2.1 Bi-dimensional discrete random fields in the spatial domain

One of the most useful representations in the spatial domain is the *simultaneous spatial autoregressive* (SAR) model.[1] This model has been extensively studied by Whittle (1954); Ord (1975); Anselin (1988); Haining (1990); Fingleton (1999); Lauridsen and Kosfeld (2004, 2006, 2007) and Beenstock and Felsenstein (2008). One possibility of its popularity is that the simultaneous spatial autoregressive model is narrowly related to the Autoregressive time series model. For instance, the latter model is not stationary if the autocorrelation coefficient is equal to one; this property is present in the spatial analogue. Another important point that supports the idea of studying the SAR representation is that a simultaneous spatial autoregressive model or a *simultaneous spatial moving average* model might cause spatial autocorrelation. However, a manifestation of spatial non-stationarity is only attributed to a SAR representation, and this phenomenon is of principal interest in this thesis. Specifically, the spatial unit root autoregressive model.

The formulation of a simultaneous spatial autoregressive model is the following. $\mu(s_{1i}, s_{2j})$ denotes the mean of the discrete random spatial process at location (s_{1i}, s_{2j}). Thus $z(s_{1i}, s_{2j})$

[1] In the following a given realization $\omega \in \Omega$ is assumed.

is thought to consist of the mean contribution, contributions of the neighbouring sites, and a random noise $\epsilon(s_{1i}, s_{2j})$, which is uncorrelated. Then for a row-column lattice,

$$z(s_{1i}, s_{2j}) = \mu(s_{1i}, s_{2j}) + \sum_{h_1}\sum_{h_2} b_{ij,h_1h_2}(z(s_{1i\pm h_1}, s_{2j\pm h_2}) - \mu(s_{1i\pm h_1}, s_{2j\pm h_2})) + \epsilon(s_{1i}, s_{2j}) \quad (2.4)$$

where $h_1, h_2 = 0, 1, \dots,$ and coefficients b_{ij,h_1h_2} describe the spatial connectivity of the sites. These coefficients govern the spatial autocorrelation structure, but not directly. The response at sites (s_{1i}, s_{2j}) and $(s_{1i\pm h_1}, s_{2j\pm h_2})$ can be correlated, even if $b_{ij,h_1h_2} = 0$.

A weak stationary random field is characterized with a constant mean, but if the mean changes with location, this one is called the large-scale structure in data, and by definition the random field is not stationary. The idea is not to associate stationarity properties with the attribute $\mathbf{z}(\mathbf{s_1}, \mathbf{s_2})$, but with its de-trended version, i.e., the process without the large-scale structure. Then, it is supposed that $\mu(s_{1i}, s_{2j}) = \mathbf{x}(s_{1i}, s_{2j})'\boldsymbol{\beta}$ and $[b_{ij,h_1h_2}] = \rho\mathbf{W}$. Where $\mathbf{x}(s_{1i}, s_{2j})'$ is a $1 \times k$ vector of regressors, $\boldsymbol{\beta}$ is a $k \times 1$ vector of parameters, ρ is a scalar that has to be estimated and $\mathbf{W}$ is a user defined spatial connectivity matrix whose dimension is $n \times n$, where $n = rc$, r and c are number of rows and columns, respectively. This matrix induces the spatial covariance structure of the model. Thus,

$$z(s_{1i}, s_{2j}) = \mathbf{x}(s_{1i}, s_{2j})'\boldsymbol{\beta} + \rho \sum_{z(s_{1l}, s_{2k}) \in N(i,j)} w_{ij,lk}(z(s_{1l}, s_{2k}) - \mathbf{x}(s_{1l}, s_{2k})'\boldsymbol{\beta}) + \epsilon(s_{1i}, s_{2j}) \quad (2.5)$$

where $N(i, j)$ is the set of neighbours of $z(s_{1i}, s_{2j})$, $i = 1, 2, \dots, r$ and $j = 1, 2, \dots, c$.[2]

If the following notation is adopted:

$$\mathbf{z} = \begin{bmatrix} z(s_{11}, s_{21}) \\ z(s_{11}, s_{22}) \\ \vdots \\ z(s_{12}, s_{21}) \\ \vdots \\ z(s_{1n}, s_{2n}) \end{bmatrix}, \mathbf{X} = \begin{bmatrix} x_1(s_{11}, s_{21}) & x_2(s_{11}, s_{21}) & \dots & x_k(s_{11}, s_{21}) \\ x_1(s_{11}, s_{22}) & x_2(s_{11}, s_{22}) & \dots & x_k(s_{11}, s_{22}) \\ \vdots & \vdots & \dots & \vdots \\ x_1(s_{12}, s_{21}) & x_2(s_{12}, s_{21}) & \dots & x_k(s_{12}, s_{21}) \\ \vdots & \vdots & \dots & \vdots \\ x_1(s_{1n}, s_{2n}) & x_2(s_{1n}, s_{2n}) & \dots & x_k(s_{1n}, s_{2n}) \end{bmatrix}, \boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix},$$

[2] Observe that this last representation is not only for regular lattices, it can also represent regional data where the number of regions are n.

$$\mathbf{e} = \begin{bmatrix} e(s_{11}, s_{21}) \\ e(s_{11}, s_{22}) \\ \vdots \\ e(s_{12}, s_{21}) \\ \vdots \\ e(s_{1n}, s_{2n}) \end{bmatrix}, \boldsymbol{\epsilon} = \begin{bmatrix} \epsilon(s_{11}, s_{21}) \\ \epsilon(s_{11}, s_{22}) \\ \vdots \\ \epsilon(s_{12}, s_{21}) \\ \vdots \\ \epsilon(s_{1n}, s_{2n}) \end{bmatrix}$$

The process can be expressed in matrix form as

$$\mathbf{z} = \mathbf{X}\boldsymbol{\beta} + \rho\mathbf{W}(\mathbf{z} - \mathbf{X}\boldsymbol{\beta}) + \boldsymbol{\epsilon} \tag{2.6}$$

then

$$(\mathbf{I} - \rho\mathbf{W})(\mathbf{z} - \mathbf{X}\boldsymbol{\beta}) = \boldsymbol{\epsilon} \tag{2.7}$$

which implies that

$$\mathbf{z} = \mathbf{X}\boldsymbol{\beta} + (\mathbf{I} - \rho\mathbf{W})^{-1}\boldsymbol{\epsilon} \tag{2.8}$$

or equivalently

$$\mathbf{z} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e} \tag{2.9}$$

where

$$\mathbf{e} = \rho\mathbf{W}\mathbf{e} + \boldsymbol{\epsilon} \tag{2.10}$$

It follows that $E[\mathbf{z}] = \mathbf{X}\boldsymbol{\beta}$ and $Var[\mathbf{z}] = \sigma^2(\mathbf{I} - \rho\mathbf{W})^{-1}(\mathbf{I} - \rho\mathbf{W}')^{-1}$ given that $E[\epsilon] = \mathbf{0}$ and $Var[\epsilon] = \sigma^2\mathbf{I}$. Note that $(\mathbf{I} - \rho\mathbf{W})$ should be non-singular, this imposes restrictions on the value of ρ. If $\mathbf{W}$ is row standardized, so that the influence of neighbours can be represented in terms of averages, then $1/\omega_{min} < \rho < 1$, where ω_{min} is the smallest negative eigenvalue of $\mathbf{W}$ (Ord, 1975; Anselin, 1982). This condition is strongly related to weak stationarity of a random field. Specifically, ρ in this range is necessary to ensure weak stationarity but is not sufficient due to edge and corner effects (Haining, 1990). Additionally, we can observe that the impulse response function of the spatial autoregressive process is the following,

$$\frac{\partial \mathbf{z}}{\partial \boldsymbol{\epsilon}} = \begin{bmatrix} \frac{\partial z(s_{11},s_{21})}{\partial \epsilon(s_{11},s_{21})} & \frac{\partial z(s_{11},s_{21})}{\partial \epsilon(s_{11},s_{22})} & \cdots & \frac{\partial z(s_{11},s_{21})}{\partial \epsilon(s_{1n},s_{2n})} \\ \frac{\partial z(s_{11},s_{22})}{\partial \epsilon(s_{11},s_{21})} & \frac{\partial z(s_{11},s_{22})}{\partial \epsilon(s_{11},s_{22})} & \cdots & \frac{\partial z(s_{11},s_{22})}{\partial \epsilon(s_{1n},s_{2n})} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial z(s_{12},s_{21})}{\partial \epsilon(s_{11},s_{21})} & \frac{\partial z(s_{12},s_{21})}{\partial \epsilon(s_{11},s_{22})} & \cdots & \frac{\partial z(s_{12},s_{21})}{\partial \epsilon(s_{1n},s_{2n})} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial z(s_{1n},s_{2n})}{\partial \epsilon(s_{11},s_{21})} & \frac{\partial z(s_{1n},s_{2n})}{\partial \epsilon(s_{11},s_{22})} & \cdots & \frac{\partial z(s_{1n},s_{2n})}{\partial \epsilon(s_{1n},s_{2n})} \end{bmatrix} = (\mathbf{I} - \rho \mathbf{W})^{-1} \tag{2.11}$$

If there is a spatial unit root process, i.e., $\rho = 1$, the impulse response function will not tend to zero as the distance between a pair of locations tends to infinity (Beenstock and Felsenstein, 2008). Thus, to analyse the case $\rho = 1$ is important because this condition implies that the random field is not stationary. Observe that although $w_{ij,lk} = 0$, it could happen that $Cov[z(s_{1i}, s_{2j}), z(s_{1l}, s_{2mj})] \neq 0$.

With respect to the mechanisms to estimate spatial models, we have that the Ordinary Least Square estimate of ρ in a pure first order spatial autoregressive model, i.e., $\mathbf{z} = \rho\mathbf{W}\mathbf{z} + \boldsymbol{\epsilon}$ is the following (Anselin, 1988):

$$\hat{\rho} = ((\mathbf{Wz})'(\mathbf{Wz}))^{-1}(\mathbf{Wz})'\mathbf{z} \tag{2.12}$$

thus

$$\hat{\rho} = \rho + ((\mathbf{Wz})'(\mathbf{Wz}))^{-1}(\mathbf{Wz})'\boldsymbol{\epsilon} \tag{2.13}$$

Asymptotically, the consistency of this estimator depends on the following two conditions:

$$n^{-1}((\mathbf{Wz})'(\mathbf{Wz})) \xrightarrow{p} \mathbf{Q} \tag{2.14}$$

$$n^{-1}((\mathbf{Wz})'(\boldsymbol{\epsilon})) \xrightarrow{p} 0 \tag{2.15}$$

where $\mathbf{Q}$ is a finite and non-singular matrix.

Whereas the first condition can be satisfied with the proper structure of the spatial weight matrix, the second condition does not hold in the spatial case. Indeed,

$$n^{-1}((\mathbf{Wz})'(\boldsymbol{\epsilon})) = n^{-1}\boldsymbol{\epsilon}'\mathbf{B}\boldsymbol{\epsilon} \tag{2.16}$$

where $\mathbf{B} = (\mathbf{I} - \rho \mathbf{W}')^{-1}\mathbf{W}'$.

The presence of the spatial weight matrix in this expression results in a quadratic form in the error term. Therefore, except in the trivial case where $\rho = 0$, this expression does not converge in probability to zero (Anselin, 1988). Consequently, the Ordinary Least Square estimator is inconsistent, irrespective of the properties of the error term.

On the other hand, if it is assumed $\epsilon \sim \mathcal{N}(0, \sigma^2\mathbf{I})$, then using the change of variable technique (see equation 2.8)

$$f(\mathbf{z} : \mathbf{X}, \mathbf{W}, \boldsymbol{\beta}, \sigma^2, \rho) = (2\pi\sigma^2)^{-n/2}|\mathbf{I} - \rho\mathbf{W}| Exp\left\{\frac{-1}{2\sigma^2}(\mathbf{z} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{I} - \rho\mathbf{W})'(\mathbf{I} - \rho\mathbf{W})(\mathbf{z} - \mathbf{X}\boldsymbol{\beta})\right\} \tag{2.17}$$

Note that the use of the change of variable technique requires the non-singularity of $(\mathbf{I} - \rho\mathbf{W})$. In fact, if $(\mathbf{I} - \rho\mathbf{W})$ is singular, $|(\mathbf{I} - \rho\mathbf{W})| = \prod_i (1 - \rho\omega_i) = 0$ and the probability distribution function degenerates to zero.[3]

The log-likelihood function is obtained as:

$$\mathcal{L}(\boldsymbol{\beta}, \sigma^2, \rho : \mathbf{z}, \mathbf{X}, \mathbf{W}) = -\frac{n}{2}Ln(2\pi) - \frac{n}{2}Ln(\sigma^2) + Ln|\mathbf{I} - \rho\mathbf{W}| - \frac{1}{2\sigma^2}(\mathbf{z} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{I} - \rho\mathbf{W})'(\mathbf{I} - \rho\mathbf{W})(\mathbf{z} - \mathbf{X}\boldsymbol{\beta}) \tag{2.18}$$

It is therefore necessary to ensure that $|(\mathbf{I} - \rho\mathbf{W})| > 0$.

The first order conditions for Maximum Likelihood estimators are obtained by taking the partial derivatives of the log-likelihood with respect to $\boldsymbol{\beta}, \sigma^2$ and ρ (equation 2.18). Consequently, the estimation process can proceed according to the following stages:

- Carry out Ordinary Least Squares of $\mathbf{X}$ on $\mathbf{z}$
- Compute an initial set of residuals
- Calculate the concentrated log-likelihood (Anselin, 1988)
- Given these residuals, find ρ that maximizes the concentrated log-likelihood
- Given $\hat{\rho}$, carry out Estimated Generalized Least Squares
- Compute a new set of residuals
- If convergence criterion is met, continue the following stage, else return to stage four
- Compute $\hat{\sigma}^2$

[3] ω_i are the eigenvalues of $\mathbf{W}$.

The stage four necessitates a nonlinear optimization routine. This can be carried out through a direct search approach. Other possibilities can be a Gauss Newton approach or a Davidon Fletcher Powell procedure.

Magnus (1978) shows that the Maximum Likelihood estimator of ρ is consistent, asymptotically efficient and asymptotically normal.

2.2 Bi-dimensional discrete random fields in the frequency domain

In order to represent the structure of a random field in the frequency domain (spectral representation) it is necessary to use the Fourier transform. The Fourier transform is an operation that transforms one complex-valued function of a real variable into another. In this context, the domain of the original function is a countable subset $D \subset \mathbb{Z}^2$, and that of the new function is a frequency, a subset $S \subset \mathbb{C}^2$.

The covariance function $C(h_1, h_2)$ and the spectral density function $s(\xi_1, \xi_2)$ form a Fourier transform pair, where $(\xi_1, \xi_2) \in S$ represents the frequency domain. Assuming that the covariances are absolutely summable, the autocovariance generating function of a discrete random field is given by (Besag and Kooperberg, 1995)

$$\sum_{h_1=-\infty}^{\infty} \sum_{h_2=-\infty}^{\infty} C(h_1, h_2) z_1^{h_1} z_2^{h_2} \tag{2.19}$$

where z_1 and z_2 are complex scalars. If this expression is divided by $(2\pi)^2$ and evaluated at some z_1 and z_2 represented by $z_1 = e^{-i\xi_1}$ and $z_2 = e^{-i\xi_2}$ where $i = \sqrt{-1}$ and ξ_1, ξ_2 are real values, the result is called the *population spectrum* of z (Hamilton, 1994). Specifically,

$$\begin{aligned}
s(\xi_1, \xi_2) &= \frac{1}{(2\pi)^2} \sum_{h_1=-\infty}^{\infty} \sum_{h_2=-\infty}^{\infty} C(h_1, h_2) z_1^{h_1} z_2^{h_2} \\
&= \frac{1}{(2\pi)^2} \sum_{h_1=-\infty}^{\infty} \sum_{h_2=-\infty}^{\infty} C(h_1, h_2) e^{-i\xi_1 h_1} e^{-i\xi_2 h_2} \\
&= \frac{1}{(2\pi)^2} \sum_{h_1=-\infty}^{\infty} \sum_{h_2=-\infty}^{\infty} C(h_1, h_2) e^{-i(\xi_1 h_1 + \xi_2 h_2)} \\
&= \frac{1}{(2\pi)^2} \sum_{h_1=-\infty}^{\infty} \sum_{h_2=-\infty}^{\infty} C(h_1, h_2)(\cos(\xi_1 h_1 + \xi_2 h_2) - i \sin(\xi_1 h_1 + \xi_2 h_2))
\end{aligned}$$

$$= \frac{1}{(2\pi)^2}\left\{C(0,0) + 2\sum_{h_2=1}^{\infty} C(0,h_2)\cos(h_2\xi_2) + 2\sum_{h_1=1}^{\infty}\sum_{h_2=-\infty}^{\infty} C(h_1,h_2)\cos(h_1\xi_1 + h_2\xi_2)\right\}$$

We use De Moivre's theorem and the last equality uses the facts that $C(h_1, h_2)$ is an even function, that is $C(h_1, h_2) = C(-h_1, -h_2)$, and some trigonometric properties like $\cos(0) = 1, \sin(0) = 0, \sin(\theta) = -\sin(-\theta)$ and $\cos(\theta) = \cos(-\theta)$.

Then for a rectangular rc row-column lattice,

$$s(\xi_1, \xi_2) = \frac{1}{(2\pi)^2}\sum_{h_1=-\infty}^{\infty}\sum_{h_2=-\infty}^{\infty} C(h_1,h_2)cos(\xi_1 h_1 + \xi_2 h_2) \tag{2.20}$$

where

$$C(h_1, h_2) = \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} cos(\xi_1 h_1 + \xi_2 h_2)s(\xi_1, \xi_2)d\xi_1 d\xi_2 \tag{2.21}$$

Intuitively, the spectral density captures the frequency content of a random field and helps to identify periodicities. Like the covariance function, the spectral density function is an even function.

An empirical estimator of $s(\xi_1, \xi_2)$ is the *periodogram*, $I(\xi_1, \xi_2)$. It can be established that for $\xi_1 \neq 0$ and $\xi_2 \neq 0$, the periodogram turns out to be the Fourier transform of the *sample covariance function*, $\hat{C}(h_1, h_2)$. The former has a considerable advantage due to the periodogram values being -at least asymptotically- independent, while the sample covariance does not satisfy this condition because the sampling variance depends on $C(h_1, h_2)$ and neighbouring values of the covariance are substantially correlated (Schabenberger and Gotway, 2005).

$$\hat{C}(h_1, h_2) = \frac{1}{n}\sum_{l}^{L}\sum_{m}^{M}(z(s_{1l}, s_{2m}) - \bar{z})(z(s_{1l+h_1}, s_{2m+h_2}) - \bar{z}) \tag{2.22}$$

where $h_1 = -r+1, -r+2, ..., 0, 1, 2, .., r-1$, $h_2 = -c+1, -c+2, ..., 0, 1, 2, ..., c-1$, $n = rc$, $l = max(1, 1-h_1)$, $L = min(r, r-h_1)$, $m = max(1, 1-h_2)$, $M = min(c, c-h_2)$ and $\bar{z} = n^{-1}\sum_{i=1}^{r}\sum_{j=1}^{c} z(s_{1i}, s_{2j})$. $\hat{C}(h_1, h_2)$ can be thought as the average covariance over all pairs of observations whose coordinates differ by (h_1, h_2). For large n, the sample covariance function is an approximately unbiased estimate of $C(h_1, h_2)$.

On the other hand, the periodogram is given by

$$I(\xi_1,\xi_2)=\frac{1}{(2\pi)^2}\sum_{h_1=-r+1}^{r-1}\sum_{h_2=-c+1}^{c-1}\hat{C}(h_1,h_2)cos(\xi_1 h_1+\xi_2 h_2) \tag{2.23}$$

and $I(0,0)=0$ (Ripley, 2004). This expression suggests a simple method to obtain the periodogram. Compute the sample covariance function for the combinations of lags $h_1 = -r+1,...,r-1$, $h_2=-c+1,...,c-1$. Once the sample covariance has been obtained for all relevant lags, cycle through the set of frequencies $S=\{(\xi_1,\xi_2):\xi_1=\{-\frac{2\pi}{r}\lfloor\frac{r-1}{2}\rfloor,\ldots,\frac{2\pi}{r}\lfloor\frac{r}{2}\rfloor\}$ and $\xi_2=\{-\frac{2\pi}{c}\lfloor\frac{c-1}{2}\rfloor,\ldots,\frac{2\pi}{c}\lfloor\frac{c}{2}\rfloor\}\}$ where $\lfloor.\rfloor$ is the greatest integer (floor) function. These frequencies, which are multiples of $2\pi/r$ and $2\pi/c$, are known as the *Fourier frequencies*.
Asymptotically ($r\to\infty$, $c\to\infty$ and r/c converges to a non-zero constant), at the non-zero Fourier frequencies, $2I(\xi_{1i},\xi_{2j})/s(\xi_{1i},\xi_{2j})\xrightarrow{d}\chi_2^2$, where χ_2^2 is a Chi-square distribution with two degrees of freedom. Hence, asymptotically, $E[I(\xi_{1i},\xi_{2j})]=s(\xi_{1i},\xi_{2j})$ and $Cov[I(\xi_{1i},\xi_{2j}),I(\xi_{1l},\xi_{2m})]=[s(\xi_{1i},\xi_{2j})]^2$ if $\xi_{1i}=\xi_{1l},\xi_{2j}=\xi_{2m}$ and 0 in another case, i.e., the periodogram is asymptotically independent at different Fourier frequencies.

2.3 SAR process and its representation in the frequency domain

In order to highlight the relation of a pure SAR process and its representation in the frequency domain, we express an SAR model on a regular lattice by

$$z(s_{1i},s_{2j})=\rho\sum_{h_1}\sum_{h_2}w_{ij,h_1h_2}z(s_{1i\pm h_1},s_{2j\pm h_2})+\epsilon(s_{1i},s_{2j}) \tag{2.24}$$

where $h_1,h_2=0,1,2,\ldots,$ $E[z(s_{1i},s_{2j})]=0, Var[z(s_{1i},s_{2j})]=C(0,0), w_{ij,00}=0, w_{ij,h_1h_2}=w_{ij,-h_1-h_2}, \rho\sum\sum w_{ij,h_1h_2}\cos(\xi_1h_1+\xi_2h_2)<1$ and the number of nonzero w_{ij,h_1h_2}'s is finite. The autocovariance generating function of this process is (Besag, 1972)

$$\sum_{l=-\infty}^{\infty}\sum_{m=-\infty}^{\infty}C(l,m)z_1^lz_2^m=\sigma_{\mathbf{z}}^2(1-\rho\sum_{h_1}\sum_{h_2}w_{ij,h_1h_2}z_1^{h_1}z_2^{h_2})^{-1}(1-\rho\sum_{l}\sum_{m}w_{ij,h_1h_2}z_1^{-h_1}z_2^{-h_2})^{-1} \tag{2.25}$$

Thus, the population spectrum is given by

$$s(\xi_1,\xi_2)\;=\;\frac{1}{(2\pi)^2}\sum_{l=-\infty}^{\infty}\sum_{m=-\infty}^{\infty}C(l,m)z_1^lz_2^m$$

$$
\begin{aligned}
&= \frac{1}{(2\pi)^2} C(0,0)(1-\rho\sum_{h_1}\sum_{h_2} w_{ij,h_1h_2} z_1^{h_1} z_2^{h_2})^{-1}(1-\rho\sum_{h_1}\sum_{h_2} w_{ij,h_1h_2} z_1^{-h_1} z_2^{-h_2})^{-1} \\
&= \frac{1}{(2\pi)^2} C(0,0)(1-\rho\sum_{h_1}\sum_{h_2} w_{ij,h_1h_2} e^{i(\xi_1 h_1+\xi_2 h_2)})^{-1}(1-\rho\sum_{h_1}\sum_{h_2} w_{ij,h_1h_2} e^{-i(\xi_1 h_1+\xi_2 h_2)})^{-1} \\
&= \frac{1}{(2\pi)^2} C(0,0)(1-\rho\sum_{h_1}\sum_{h_2} w_{ij,h_1h_2}(\cos(\xi_1 h_1+\xi_2 h_2)+i\sin(\xi_1 h_1+\xi_2 h_2)))^{-1} \\
&\quad (1-\rho\sum_{h_1}\sum_{h_2} w_{ij,h_1h_2}(\cos(\xi_1 h_1+\xi_2 h_2)-i\sin(\xi_1 h_1+\xi_2 h_2)))^{-1} \\
&= \frac{1}{(2\pi)^2} C(0,0)(1-\rho\sum_{h_1}\sum_{h_2} w_{ij,h_1h_2}\cos(\xi_1 h_1+\xi_2 h_2))^{-2}
\end{aligned}
$$

There are two important facts in the spectral representation of an SAR model that should be mentioned. First, the population spectrum for a pure SAR random field is nonnegative for all (ξ_1, ξ_2), and depends on the spectral frequencies. And second, if there is no kind of spatial autocorrelation, i.e. $\rho = 0$, the spectral density is constant in the frequency domain. These facts will be used to create a statistical test to check the null hypothesis of a spatial unit root process (see Chapter 4).

3 Unconditional simulation of a Gaussian random field with a simultaneous spatial autoregressive structure: consequences of a spatial unit root

This section is strongly based on Cressie (1993), Schabenberger and Gotway (2005) and Bivand et al. (2008). On the other hand, the R package (R Development Core Team, 2011) was used to build all the algorithms in this thesis, specifically the libraries developed by Bivand (2011); Venables and Ripley (2011); Adler and Murdoch (2011); Furrer et al. (2010); Finley and Banerjee (2010).

> "Real data are important for the development of statistical methods, and, ideally, their analysis also stimulates research in statistical theory. Simulated data have a different role. They may be used to validate or establish properties of a statistical method under assumed model, which includes checking the validity of asymptotic properties in finite samples."
>
> *Cressie (1993, pp 568)*

Constructing a realization of a random field is not a trivial task, since it requires knowledge of the spatial distribution. From a particular dataset, it may be inferred, under stationarity assumptions, the first and second moment structure of the field. Inferring the joint distribution from the mean and covariance functions is not possible unless the random field is a *Gaussian Random Field* (GRF).[1] Even if the spatial distribution is known, it is usually not possible to construct a realization via simulation from it (Schabenberger and Gotway, 2005).

Several methods are available to simulate GRFs unconditionally.[2] The simplest method relies on the reproductive property of the multivariate Gaussian distribution and the fact that a positive definite matrix $\mathbf{\Sigma}$ can be represented as $\mathbf{\Sigma} = \mathbf{\Sigma}^{1/2}\mathbf{\Sigma}'^{1/2}$. If $\mathbf{z} \sim \mathcal{N}_n(\boldsymbol{\mu}, \mathbf{\Sigma})$,

[1] A random field is a Gaussian Random Field, if the cumulative distribution function is that of a k-variate Gaussian random variable for all k.

[2] A simulation method that honours the data in the sense that the simulated value at an observed location agrees with the observed value is termed a conditional simulation. Simulation methods that do not honour the data are called unconditional simulations.

$\mathbf{x} \sim \mathcal{N}_n(\mathbf{0}, \mathbf{I})$, then $\boldsymbol{\mu} + \boldsymbol{\Sigma}^{1/2}\mathbf{x}$ has a $\mathcal{N}_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.

In the Cholesky decomposition, if $\boldsymbol{\Sigma}_{n\times n}$ is a positive definite matrix, thus there exists an upper triangular matrix $\mathbf{U}_{n\times n}$ such that $\boldsymbol{\Sigma} = \mathbf{U}'\mathbf{U}$ where $\mathbf{U}$ is unique. Since $\mathbf{U}'$ is lower triangular and $\mathbf{U}$ is upper triangular, this decomposition is often referred to as the lower-upper or LU decomposition (Greene, 2003). This suggests a simple method of generating data from a $\mathcal{N}_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ distribution. Generate n independent standard Gaussian random variables, calculate the Cholesky root $\mathbf{U}'$ of the covariance matrix $\boldsymbol{\Sigma}$ and a $(n \times 1)$ vector of means $\boldsymbol{\mu}$. Return $\mathbf{z} = \boldsymbol{\mu} + \mathbf{U}'\mathbf{x}$ as a realization from a $\mathcal{N}_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. This method works well for small to moderate problems (Schabenberger and Gotway, 2005). However when n is larger than say, 1,000, and Σ is sparse, various numerical inaccuracies may result (Cressie, 1993).

In order to make the simulations of lattice data, the first step is to define the spatial connectivity matrix$\mathbf{W}_{n\times n}$, that induces the spatial covariance structure of the model. In the context of data arranged in a regular rectangular grid, the original construction of this matrix is based on the notion of binary contiguity between spatial units. According to this notion, the underlying structure of neighbours is expressed by $\{0, 1\}$ values. The most used criteria are: the rook style where two spatial units that have a common edge are considered to be contiguous, the queen style where a common vertex or edge are considered, and the torus style that is useful for simulations, because, since all spatial units have equal numbers of neighbours, and there are no edges, the structure of the graph is as neutral as can be achieved. By construction, the principal diagonal of the contiguity matrix is filled with zeros. Once the contiguity matrix based on $\{0, 1\}$ criteria is built, this matrix is row standardised.

Assuming that the structure of the random field is established by an SAR model whose stochastic errors are not correlated and homocedastic, $\boldsymbol{\Sigma} = \sigma^2(\mathbf{I} - \rho\mathbf{W})^{-1}(\mathbf{I} - \rho\mathbf{W}')^{-1}$ is obtained. Note that $(\mathbf{I} - \rho\mathbf{W})$ should be non-singular, this imposes restrictions on the value of ρ. If $\mathbf{W}$ is row standardised, then $1/\omega_{min} < \rho < 1$, where ω_{min} is the smallest negative eigenvalue of $\mathbf{W}$ (Anselin, 1982). Specifically, the dominant eigenvalue of a row standardised contiguity matrix is 1, and given that $|(I - \rho\omega_i)| = \prod_i(1 - \rho\omega_i)$, then ρ has to satisfy this condition. Additionally, the Wold representation of a SAR field is $\mathbf{z} = (I - \rho\mathbf{W})^{-1}\boldsymbol{\epsilon}$, then if there is a spatial unit root process, the impulse response function will not tend to zero as the distance between a pair of locations tends to infinity (Beenstock and Felsenstein, 2008). Thus, to analyse the case $\rho = 1$ is important because this condition implies that the random field is not stationary, so in this case, the *Moore-Penrose inverse* of $\boldsymbol{\Sigma}$ is used, if $\boldsymbol{\Sigma}$ is non-singular, the Moore-Penrose inverse is the familiar ordinary inverse (Asmar, 1995).[3] Lauridsen and Kosfeld (2006, 2007) follow this strategy.

[3]A Moore-Penrose inverse of a matrix $\boldsymbol{\Sigma}$ is a unique matrix $\boldsymbol{\Sigma}^+$ that satisfies the following requirements: $\boldsymbol{\Sigma}\boldsymbol{\Sigma}^+\boldsymbol{\Sigma} = \boldsymbol{\Sigma}$, $\boldsymbol{\Sigma}^+\boldsymbol{\Sigma}\boldsymbol{\Sigma}^+ = \boldsymbol{\Sigma}^+$, $\boldsymbol{\Sigma}^+\boldsymbol{\Sigma}$ is symmetric and $\boldsymbol{\Sigma}\boldsymbol{\Sigma}^+$ is symmetric.

Figure 3.1: Variances of simulated SAR GRF with $\rho = 0.9$ on a lattice 10×10.

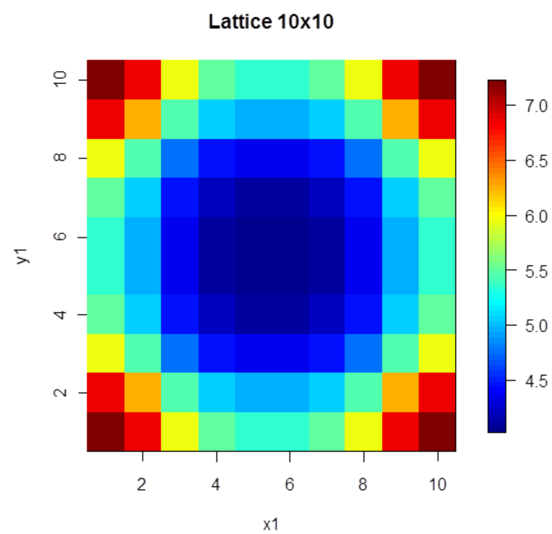
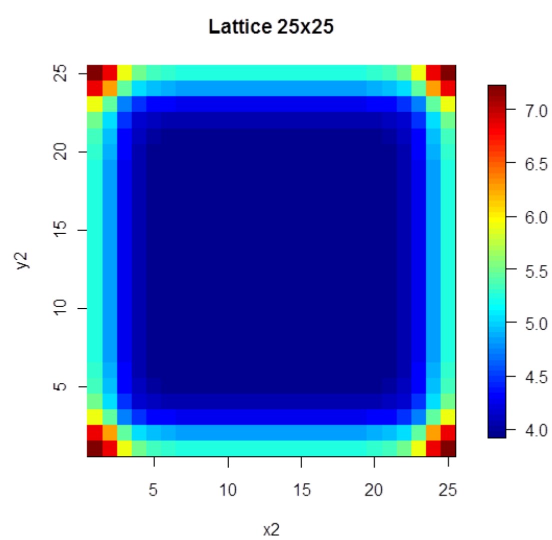


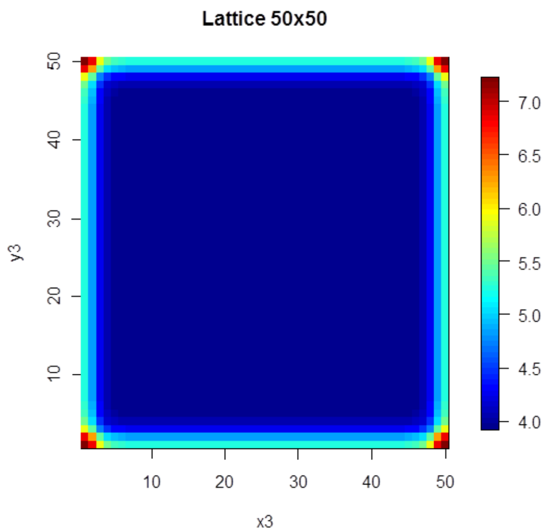
Figure 3.2: Variances of simulated SAR GRF with $\rho = 0.9$ on a lattice 25×25.



As can be seen in Figures 3.1, 3.2 and 3.3, $|\rho| < 1$ is a necessary but not a sufficient condition to generate a stationary Spatial Autorregresive Gaussian Random Field; because edge and corner effects cause variances to change with spatial location,[4] but it asymptotically

[4]These variances are obtained from the principal diagonal of Σ.

Figure 3.3: Variances of simulated SAR GRF with $\rho = 0.9$ on a lattice 50×50.



approaches to stationarity when $|\rho| < 1$ and $n \to \infty$ (Fingleton, 1999). Similar findings are shown by Haining (1990).

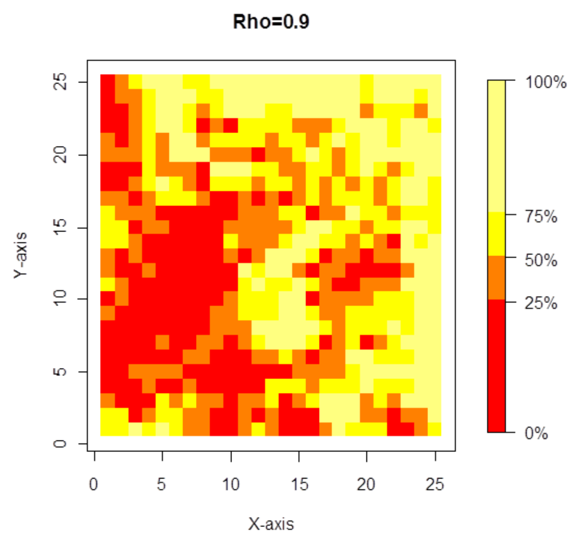
Figures 3.4, 3.5 and 3.6 show three different realizations of simulated SAR GRFs with $|\rho| < 1$ on a regular lattice of 25×25 units with a row standardised rook style spatial connectivity matrix; we can see in the figures that different spatial autocorrelation parameters imply different configuration of the spatial process. On the other hand, Figure 3.7 displays a realization of a non-stationary spatial autoregressive process, i.e., $\mathbf{z} = \rho\mathbf{W}\mathbf{z} + \boldsymbol{\epsilon}$ where $\rho = 1$ and $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I})$. The Moore-Penrose inverse is used in order to get $(\mathbf{I} - \rho\mathbf{W})^{-1}$. In 3.7, we observe that the field exhibits a clear trend.

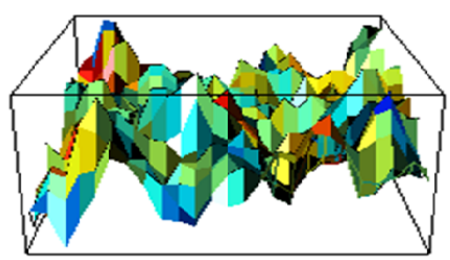
As can be seen in Figures 3.8, 3.9 and 3.10, the variances of a unit root spatial autoregressive Gaussian Random Field are not stable, although dimension of the lattice increases. A fact that is corroborated in Table 3.1 where the mean, maximum and minimum variances of the processes associated with different ρ's are displayed. As can be seen in Table 3.1, the variance tends to stabilize when n increases, except in the case of a spatial unit root process. Similar findings are found by Fingleton (1999) and Lee and Yu (2009).[5]

In order to highlight similar characteristics of a spatial unit root process and a random walk process in time, the spatial correlogram function is estimated, where spatial lag is built based on higher order contiguities, that is, spatial correlograms are constructed by taking an input

[5] Lee and Yu (2009) find similar outcomes but they use $\rho_{n0} = 1 - 1/\psi_n, \psi_n \to \infty, n \to \infty$.

Figure 3.4: Simulated SAR fields with $\rho = 0.9$.

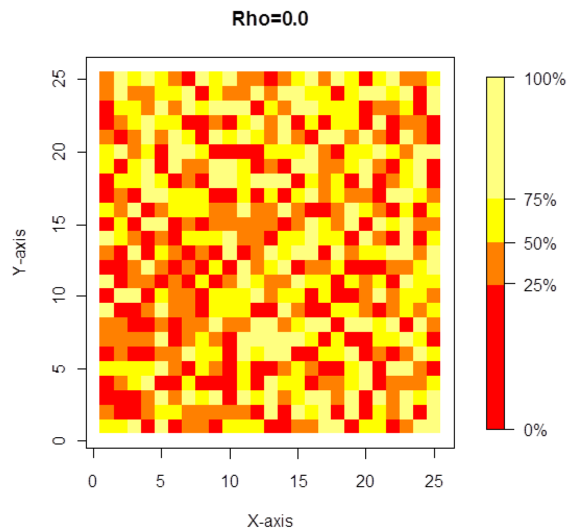


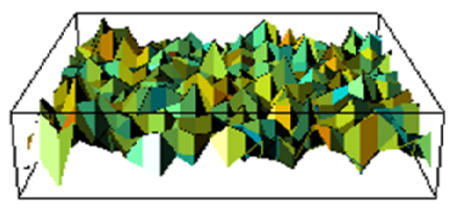


list of neighbours as the first order set, and stepping out across the graph to second, third, and higher order neighbours based on the number of links traversed, but not permitting cycles, which could risk making $z(s_{1i}, s_{2j})$ a neighbour of $z(s_{1i}, s_{2j})$ itself (Bivand et al., 2008). Figure 3.11 shows how the spatial correlogram decreases slower as the spatial correlation coefficient increases; this characteristic resembles what is observed when there is a unit root process in time (Enders, 1995).

Another common statistic for testing $H_0 : \rho = 0$ is the Moran's test (Moran, 1950). This statistical test pertains to the close interval $[-1, 1]$, where -1 indicates perfect dispersion, 0 random spatial pattern and 1 perfect correlation. Figure 3.12 shows that Moran's test detects spatial correlation at different spatial lags but is more persistent under spatial unit root processes.

Figure 3.5: Simulated SAR fields with $\rho = 0.0$.

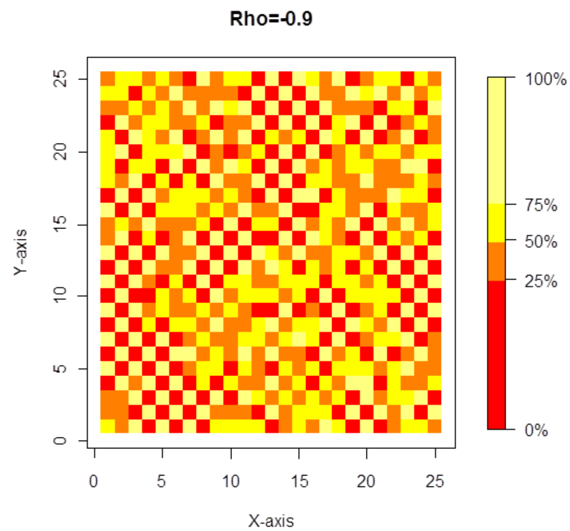


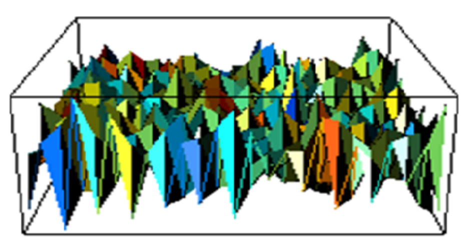


The assumptions of constant mean and constant variance of $\mathbf{z}$ must not be taken lightly when testing for spatial autocorrelation with Moran's test. Values in close spatial proximity may be similar, not because of spatial autocorrelation but because the values are independent realizations from distributions with similar mean (Schabenberger and Gotway, 2005). Figures 3.13 and 3.14 are built to show differences between stochastic and deterministic non-stationary random fields. Independent extractions draw from a Gaussian Random Field with $E[z(s_{1i}, s_{2j})] = 1.4 + 0.1s_{1i} + 0.2s_{2j} + 0.002(s_{1i})^2 + 0.003(s_{2j})^2$ were assigned to the sites of a 25×25 regular lattice. These data do not exhibit any spatial autocorrelation but are not mean-stationary; although the Moran's I statistics detected spurious autocorrelation.

Moran's test rejects $H_0 : \rho = 0$ for both processes, i.e., the spatial unit root process and the deterministic non-stationary process. Specifically, the I statistic is equal to 0.95 and 0.84 for the unit root process and deterministic process, respectively. These values indicate positive spatial autocorrelation. In both cases, the p-value is near to zero. Figure 3.15

Figure 3.6: Simulated SAR fields with $\rho = -0.9$.



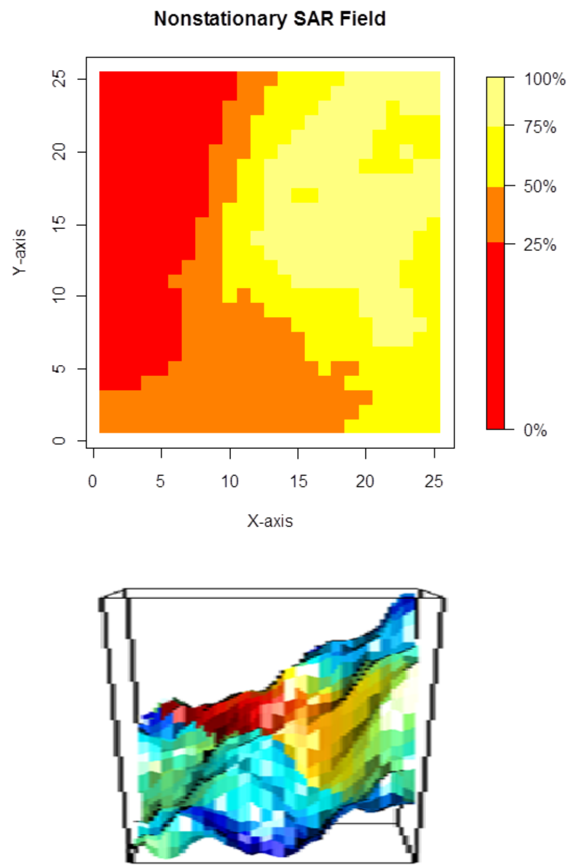


shows histograms of Moran's test under random permutations associated with each process. Random permutations are recommended because standard Moran's test is based on strong distributional assumptions. Equivalently, the p-values in both cases are equal to zero. As cited by Schabenberger and Gotway (2005), the impact of heterogeneous means on the interpretation of Moran's I is both widely ignored and completely confused throughout the literature.[6] In order to perform the Moran's test, it is advisable to fit a mean model to the data and examine whether the residuals from the fit exhibit spatial autocorrelation.

We estimate the periodogram for SAR fields which have different values of spatial autocorrelation coefficients to characterize them on the frequency domain. Figure 3.16 shows the sample covariance and the periodogram functions associated with an SAR model with $\rho = 0$. As can be seen in Figure 3.16, the sample covariance function is close to zero everywhere, except for $(s_{1i}, s_{2j}) = (0, 0)$, where the sample covariance function estimates the variance

[6] Mur and Trívez (2003) analyse the consequences of deterministic trends in spatial econometrics.

Figure 3.7: Simulated non-stationary SAR GRF, $\rho = 1$.



of the process. On the other hand, the periodogram is more or less evenly distributed, a characteristic that can be observed from the population spectrum because $\rho = 0$ implies $s(\xi_1, \xi_2) = C(0,0)/(2\pi)^2$. This means that high and low ordinates occur for large and small frequencies (ξ_1, ξ_2).

As can be seen in Figure 3.17, where the sample covariance and the periodogram functions are shown for a SAR process with $\rho = 0.9$ on a 10×10 lattice, sample covariances are substantial for small s_{1i} and s_{2j} indicating high spatial correlation. The strong spatial autocorrelation coefficient is associated with large periodogram ordinates for small frequencies. As can be seen from the population spectrum of an SAR field, when $\rho > 0$, $(1 - \rho \sum_{h_1} \sum_{h_2} w_{ij,h_1h_2} \cos(\xi_1 h_1 + \xi_2 h_2))$ is a monotonically increasing function in (ξ_1, ξ_2) over $[0, \pi]^2$, meaning that $s(\xi_1, \xi_2)$ is monotonically decreasing.

Figure 3.8: Variances of non-stationary SAR GRF, $\rho = 1$ on a lattice 10×10.

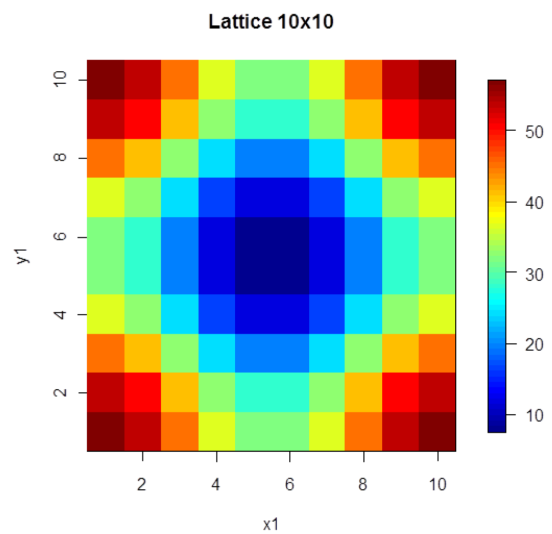


Figure 3.9: Variances of non-stationary SAR GRF, $\rho = 1$ on a lattice 25×25.

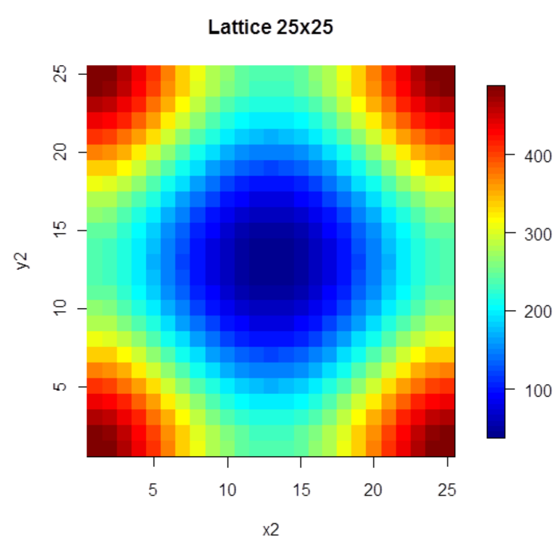


Figure 3.18 shows the sample covariance and the periodogram functions associated with a SAR model with $\rho = 1$. As can be seen in Figure 3.18, sample covariances are large even for large distances. And the periodogram shows that the field is explained by low frequencies, but large frequencies are also important. However, it must be pointed out that the population spectrum is built for a stationary process, which is not a good assumption in this case.

Figure 3.10: Variances of non-stationary SAR GRF, $\rho = 1$ on a lattice 50 $\times$ 50.

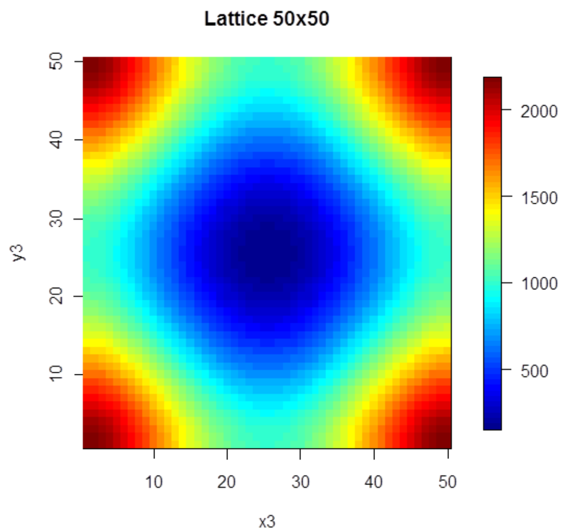


Table 3.1: Simulated SAR models: Mean, Maximum and Minimum variances.

		Lattice		
ρ		10×10	25×25	50×50
0.2	Mean	1.035	1.032	1.031
	Max	1.048	1.048	1.048
	Min	1.031	1.031	1.031
0.5	Mean	1.286	1.261	1.253
	Max	1.402	1.402	1.402
	Min	1.245	1.245	1.245
0.9	Mean	5.216	4.395	4.153
	Max	7.224	7.222	7.222
	Min	4.029	3.925	3.925
0.95	Mean	11.602	8.741	7.936
	Max	15.961	15.875	15.875
	Min	8.386	7.201	7.200
1	Mean	32.652	238.562	1,006.720
	Max	57.036	487.504	2,186.233
	Min	7.567	38.410	154.674

Source: Author's estimations.

Figure 3.11: Spatial correlogram.

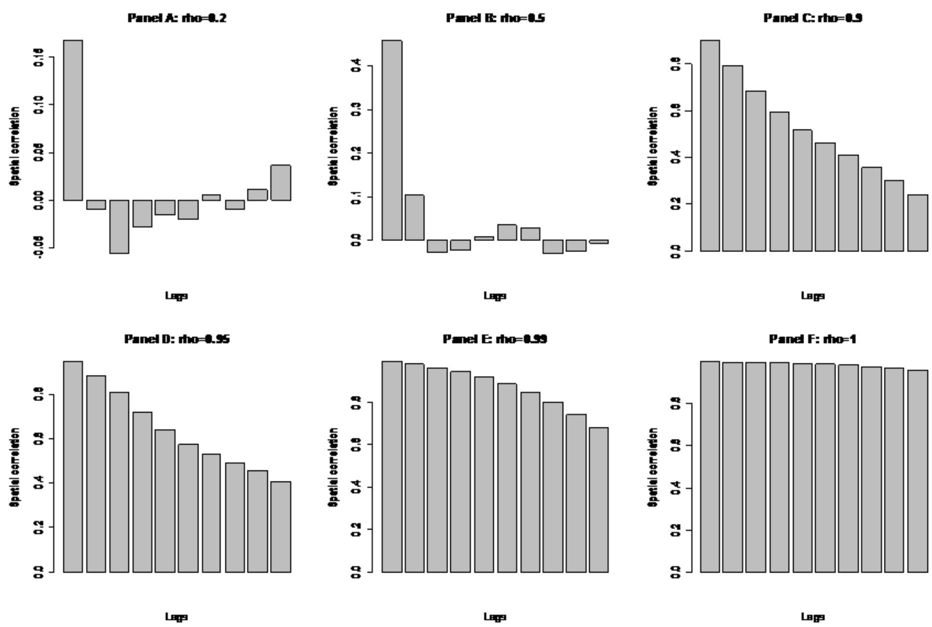


Figure 3.12: Moran's test at different spatial lags.

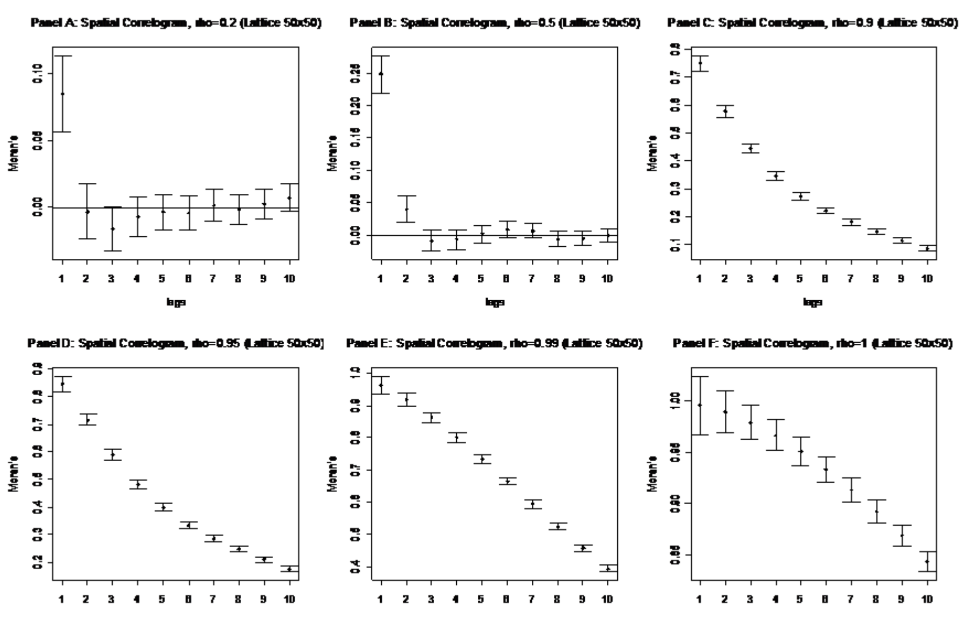


Figure 3.13: Simulated spatial unit root GRF.

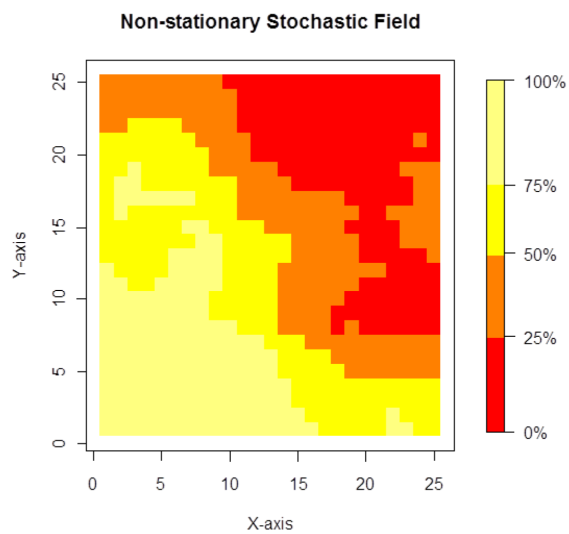
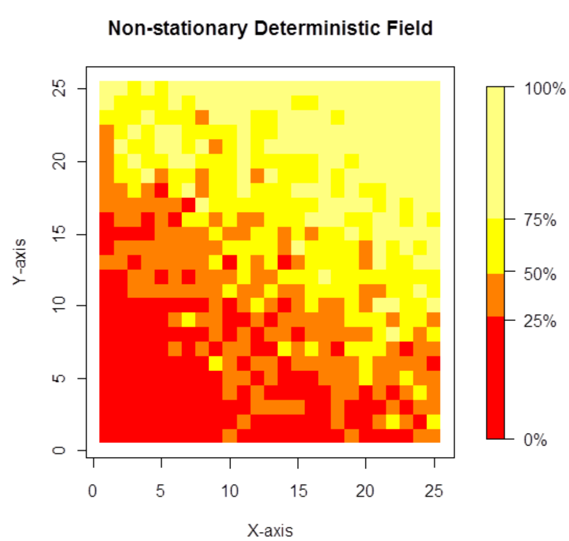


Figure 3.14: Simulated deterministic non-stationary GRF .



As can be seen from Figures 3.16, 3.17 and 3.18, the sample covariance and periodogram functions are symmetric functions, so we can just plot the right half-plane; the left half-plane can be found by a half-turn rotation.

Figure 3.15: Histogram of simulated non-stationary GRF.

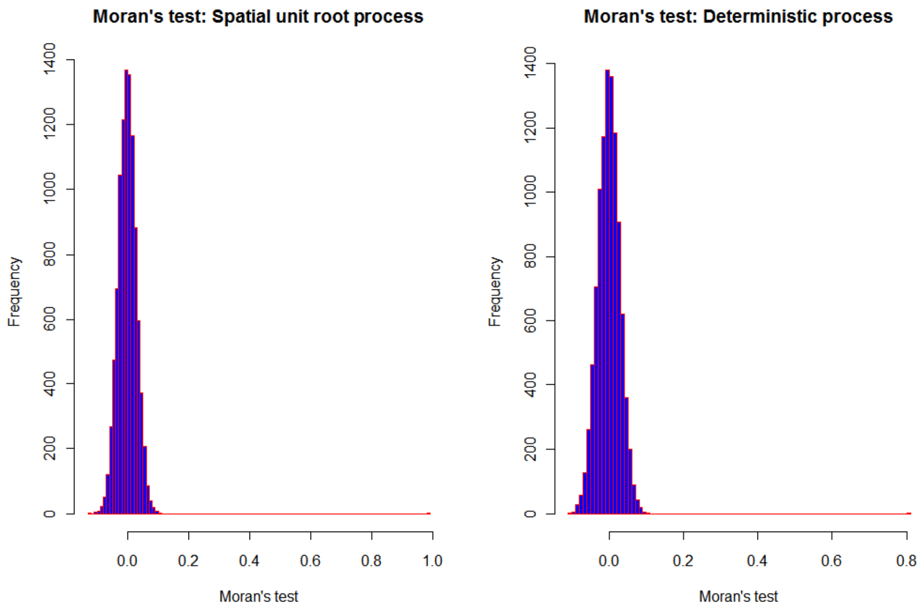
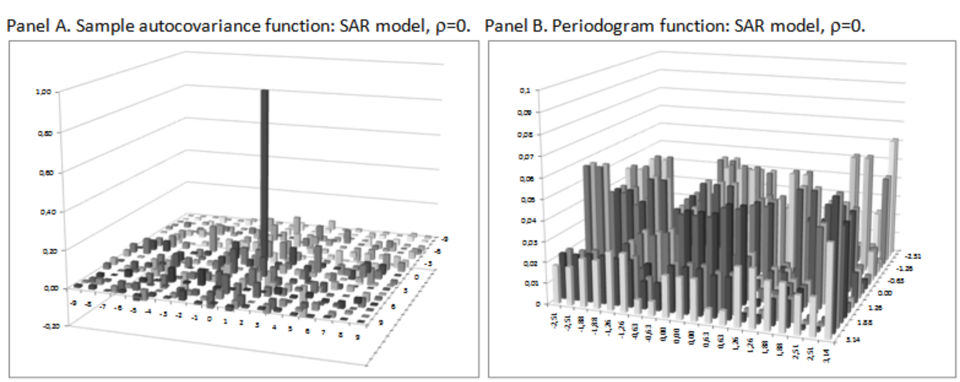


Figure 3.16: Spatial sample covariance and periodogram functions: SAR field with $\rho = 0$ on a Lattice 10×10.



Two simulation exercises are performed to analyse implications on statistical inference associated with a spatial unit root process. Specifically, 10,000 SAR models are simulated under different spatial autocorrelation coefficients on a 25×25 regular lattice where $\rho = \{0.2, 0.5, 0.9, 0.95, 0.99, 1\}$. Thus, the spatial autocorrelation coefficient is estimated for each simulation, and the empirical distribution function of $\hat{\rho}$ is estimated. As can be seen in

Figure 3.17: Spatial sample covariance and periodogram functions: SAR field with $\rho = 0.9$ on a Lattice 10×10.

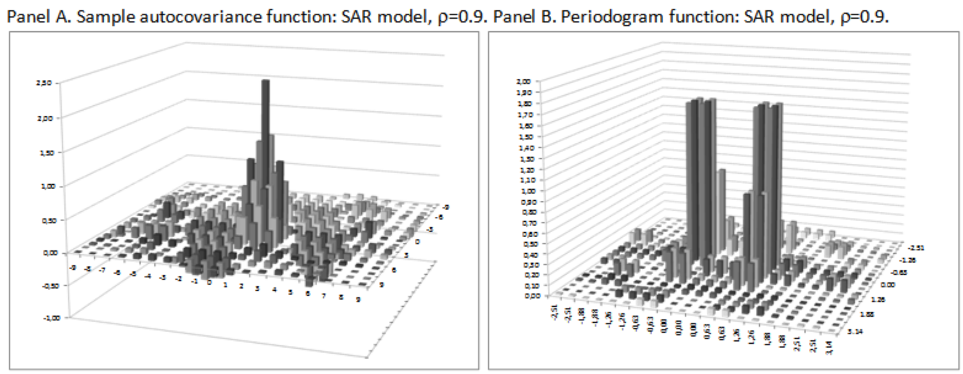


Figure 3.18: Spatial sample covariance and periodogram functions: SAR field with $\rho = 1$ on a Lattice 10×10.

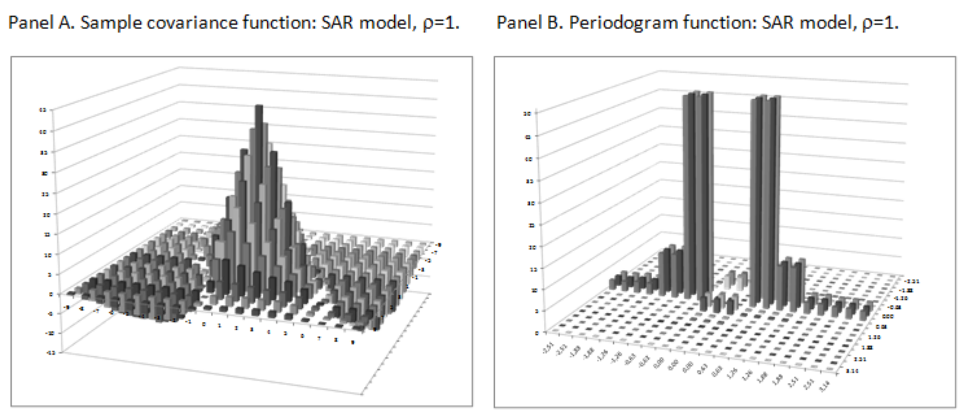


Figure 3.19, $E[\hat{\rho}] = \rho$, except when $\rho = 1$.[7]

Additionally, two independent realizations of spatial autoregressive models with row standardised rook style spatial connectivity matrix are simulated 10,000 times. Specifically, $\mathbf{z} = \rho\mathbf{W}\mathbf{z} + \boldsymbol{\epsilon_1}$ and $\mathbf{x} = \rho\mathbf{W}\mathbf{x} + \boldsymbol{\epsilon_2}$ where $\boldsymbol{\epsilon_1}$ and $\boldsymbol{\epsilon_2}$ are independent. After this, Ordinary Least Square regressions are conducted with each pair, and given a nominal size of 5% for the null hypothesis $H_0 : \beta = 0$, the empirical size is calculated. This process is replicated for

[7] Lee and Yu (2007) show that in the model $\mathbf{z} = \rho_{n0}\mathbf{W}\mathbf{z} + \boldsymbol{\epsilon}$ where $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I})$ and $\rho_{n0} = 1 - 1/\psi_n, \psi_n \to \infty, n \to \infty$, the QMLE $\hat{\rho}_n$ is n-consistent. This means that $\psi_n(\hat{\rho}_n - \rho_{n0}) = o_p(1)$ which implies that the estimate is asymptotically unbiased. However, there is a practical paradox because $\rho_{n0} = 1 - 1/\psi_n, \psi_n \to \infty, n \to \infty$ is very similar to $\rho = 1$. In fact, there are numerical problems when matrix $(\mathbf{I} - \rho_{n0}\mathbf{W})$ requires to be inverted; although in theory, there should not be any problem.

Figure 3.19: Densities of ρ estimates.

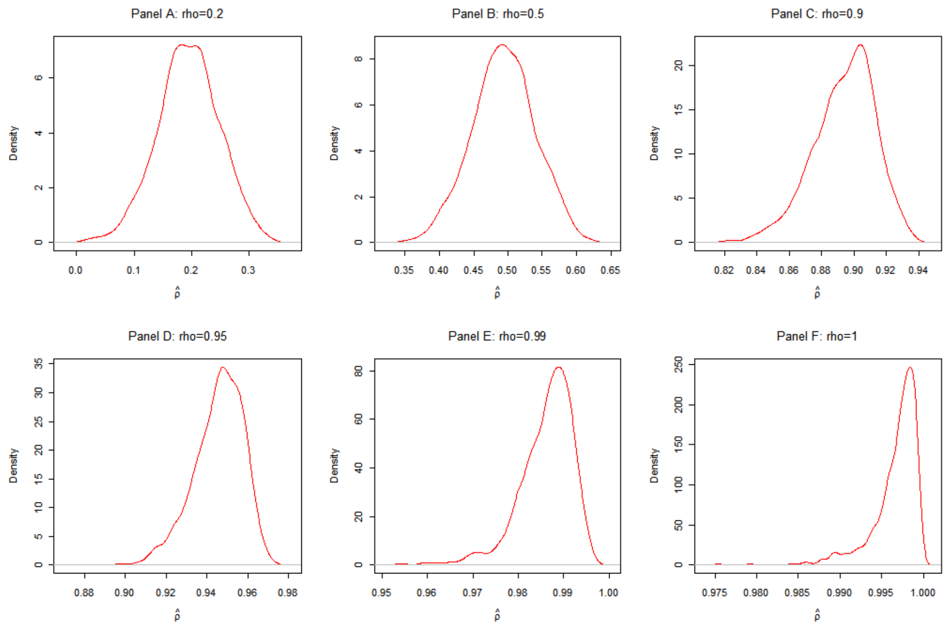


Table 3.2: Empirical size of $H_0 : \beta = 0$ for $\mathbf{z} = \beta\mathbf{x}+\epsilon$ given independent spatial autoregressive Gaussian random fields.*

	Lattice		
ρ	10×10	25×25	50×50
0.2	0.049	0.049	0.054
0.5	0.060	0.0595	0.077
0.9	0.150	0.149	0.156
0.95	0.218	0.219	0.218
1	0.367	0.766	0.926

**Nominal size is 0.05.*

Source: Author's estimations.

$\rho = \{0.2, 0.5, 0.9, 0.95, 1.0\}$ on regular lattices of different dimensions, specifically 10×10, 25×25 and 50×50. Table 3.2 displays the empirical size results, and as can be seen, if the spatial correlation coefficient is low, the empirical size is near to the nominal size, but this fact worsens when the autocorrelation coefficient increases, the t-statistic rejects H_0 when in fact this one is correct, i.e., the type I error increases. Given $\rho = 1$, the empirical size is considerably bigger than the nominal size, with the additional problem that when the dimension of the lattice increases, the empirical size increases. The fact that the t-statistic detects an apparent relationship of $\mathbf{z}$ and $\mathbf{x}$ when the variables exhibit a stochastic trend is known as nonsense spatial regression (Beenstock and Felsenstein, 2008). Fingleton (1999) and Lee and Yu (2009) show similar findings, although the latter analyse the case $\rho_{n0} = 1 - 1/\psi_n, \psi_n \to \infty, n \to \infty$.

4 A statistical test for a unit root process in SAR models

In order to use theory that is developed to test stationarity in time domain, we analyse four of the most useful tests and their implications if they are applied in the spatial domain. Specifically, Augmented Dickey and Fuller (1979) and Phillips and Perron (1988) tests to check stationarity in time domain are strongly based on a consistent estimation of the autoregressive coefficient obtained by OLS procedure. This estimator is used to build the asymptotic distribution of the statistical test for a unit root process. However, OLS estimation of the spatial autoregressive coefficient is not consistent. A consistent estimation can be obtained by Maximum Likelihood or Feasible Generalized Least Square but these estimators are not lineal, this implies great difficulty in building a spatial unit root test. Additionally, under the null hypothesis of a spatial unit root process the probability distribution function of a spatial autoregressive field degenerates to zero (Anselin, 1982). Specifically, the asymptotic ML estimates will only hold if the regularity conditions for the log-likelihood function are satisfied. This statement is not fulfilled under the null hypothesis. However, Lauridsen and Kosfeld (2004) and Beenstock and Felsenstein (2008) propose a statistical test to check stationarity under this consideration. Kwiatkowski et al. (1992) test for a unit root process in time domain bases its asymptotic distribution on the results of MacNeill (1978a) and MacNeill (1978b), who establishes convergence of partial sums of residuals in the time domain. Unfortunately, at this time, we have not found a theorem that establishes this outcome in spatial domain. Guan (2008) develops a homogeneity test for point process following the approach of Kwiatkowski et al. (1992), the asymptotic distribution of his statistical test is based on a result given by Ivanoff (1982). Finally, Priestley and Rao (1969) use evolutionary spectrum to test stationarity in time domain. This idea is taken by Mateu and Juan (2004) and Fuentes (2005), who develop a test for non-stationarity of geostatistical data.

It is known that the OLS estimation of a Gaussian $AR(1)$ process, $y_t = \rho y_{t-1} + \epsilon_t$ where $\epsilon_t \overset{i.i.d}{\sim} \mathcal{N}(0, \sigma^2)$ and $y_0 = 0$, has the following asymptotic outcome: $\sqrt{T}(\hat{\rho} - \rho) \xrightarrow{d} \mathcal{N}(0, (1 - \rho^2))$. This statement is also valid if $\rho = 1$, but it is not very helpful for hypothesis testing. Given this outcome, the most useful unit root tests build their asymptotic distributions as functions of a standard Brownian motion, which is a continuous time stochastic process, associating each date $r \in [0, 1]$ with the scalar $W(r)$ such that:

- $W(0) = 0$.

- For any dates $0 \leq r_1 < r_2 < \cdots < r_k \leq 1$, the changes $[W(r_2) - W(r_1)], [W(r_3) - W(r_2)], \ldots [W(r_k) - W(r_{k-1})]$ are independent multivariate Gaussian with $[W(s) - W(r)] \sim \mathcal{N}(0, s-t)$, $s > t$.

- For any given realization, $W(r)$ is continuous in r with probability 1.

Paelinck et al. (2004) caution about using the concept of Brownian motion in spatial domain. Specifically, because of a multilateral dependence structure in the space domain, it is difficult to determine an initial starting point and have independent increments, conditions that are required by a Brownian motion process.

4.1 A statistical test for a unit root in SAR models based on differences in the frequency domain

To avoid the use of Brownian motion process our approach takes elements from the representation of a random field in spatial domain and its spectral decomposition (see Chapter 2). As it is noticed by Mateu and Juan (2004), spectral analysis of stationary processes is particularly advantageous in the analysis of large data sets because the use of traditional techniques implies inversion of a large covariance matrix to compute the likelihood function. The use of a Fast Fourier Transform algorithm for spectral densities can be a good solution for this problem. Additionally, the periodogram, a nonparametric estimate of the spectral density, has asymptotic properties that facilitate working in the frequency domain.

In order to get some intuition about the statistical test that we propose, we depict in Figure 4.1 the periodogram functions of $\Delta z(s_{1i}, s_{2j})$ for different spatial autocorrelation parameters associated with a pure SAR model $z(s_{1i}, s_{2j})$.

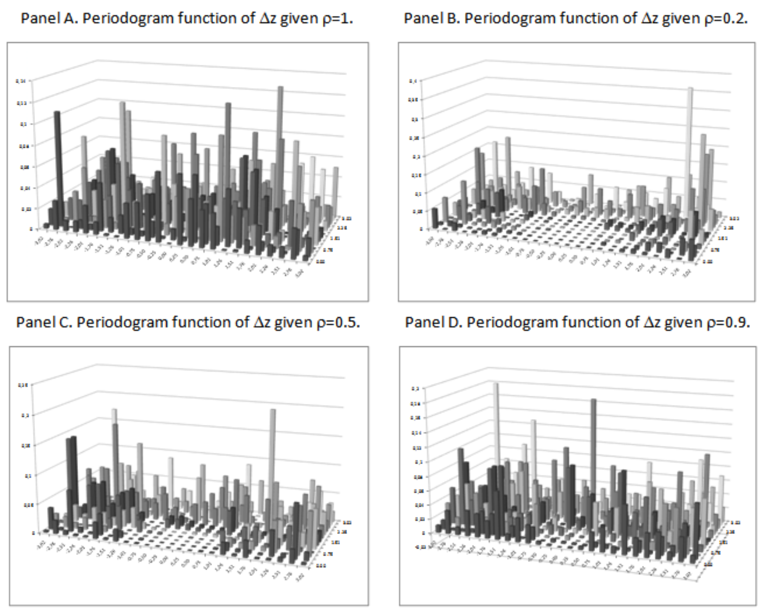
$$z(s_{1i}, s_{2j}) = \rho \sum_{h_1} \sum_{h_2} w_{ij,h_1 h_2} z(s_{1i \pm h_1}, s_{1j \pm h_2}) + \epsilon(s_{1i}, s_{2j}) \tag{4.1}$$

$h_1 = 0, 1, 2, ...$ and $h_2 = 0, 1, 2, ...$
And

$$\Delta z(s_{1i}, s_{2j}) = z(s_{1i}, s_{2j}) - \sum_{h_1} \sum_{h_2} w_{ij,h_1 h_2} z(s_{1i \pm h_1}, s_{1j \pm h_2}) \tag{4.2}$$

The important fact is that for $\rho = 1$, the periodogram is more or less evenly distributed in the space of frequencies (see Figure 4.1, Panel A), while for $\rho = \{0.2, 0.5, 0.9\}$ there is a monotonically increasing tendency over (ξ_1, ξ_2) whose origin is $(0, 0)$. However, this pattern is less clear as ρ increases to 1 (see Figure 4.1, Panels B, C and D).[1]

Figure 4.1: Spatial periodogram functions: $\Delta z(s_{1i}, s_{2j}), \rho = \{0.2, 0.5, 0.9, 1\}$.



Using this intuitive result, an *ad hoc* statistical test is proposed to check $H_0 : \rho = 1$, i.e, non-stationarity of a pure SAR representation of a random field. The idea behind is to take advantage of the behaviour of the periodograms associated with different spatial autocorrelation structures. Therefore, we compare low and high frequencies that are obtained from a spectral density estimate.

The following theorem, which is one of the most important theoretic developments of this dissertation, establishes the asymptotic distribution of the statistical test for a unit root in a spatial autoregressive model in regional data.

Theorem 1. Asymptotic distribution of a statistical test for a unit root in a spatial autoregressive model in regional data based on differences in the frequency domain.

Given $z(s_{1i}, s_{2j}) = \rho \sum_{h_1} \sum_{h_2} w_{ij,h_1h_2} z(s_{1i\pm h_1}, s_{1j\pm h_2}) + \epsilon(s_{1i}, s_{2j})$ *where* $\epsilon(s_{1i}, s_{2j}) \overset{i.i.d}{\sim} \mathcal{N}(0, \sigma^2)$. *Then, under* $H_0 : \rho = 1$, $\Delta z(s_{1i}, s_{2j}) \equiv z(s_{1i}, s_{2j}) - \sum_{h_1} \sum_{h_2} w_{ij,h_1h_2} z(s_{1i\pm h_1}, s_{1j\pm h_2}) =$

[1] Evenness is a property of the periodogram. It is thus sufficient to compute the periodogram only for the set of frequencies which removes from the space of frequencies the points with $\xi_{2j} < 0$. So we plot only the right half-plane; the left half-plane can be found by a half-turn rotation.

$\epsilon(s_{1i}, s_{2j})$, this implies that the spectral density function of $\Delta z(s_{1i}, s_{2j})$ is constant and equals $\sigma^2/(2\pi)^2$ $\forall \xi_{1i}, \xi_{2j}$. Given that $\frac{2I(\xi_{1i},\xi_{2j})}{s(\xi_{1i},\xi_{2j})} \xrightarrow{d} \chi^2_2$ where χ^2_2 is a Chi-square distribution with two degrees of freedom, $\forall \xi_{1i} \neq 0, \xi_{2j} \neq 0$, and asymptotically $Cov[I(\xi_{1i}, \xi_{2j}), I(\xi_{1l}, \xi_{2k})] = 0$, $\forall \xi_{1i} \neq \xi_{1l}, \xi_{2j} \neq \xi_{2k}$, then under H_0, we have that $\hat{\psi} = \frac{\frac{\sum_{\xi_{1i},\xi_{2j}\neq[0,0]} I(\xi_{1i},\xi_{2j})}{m_1}}{\frac{\sum_{\xi_{1l},\xi_{2k}\neq[0,0]} I(\xi_{1l},\xi_{2k})}{m_2}} \xrightarrow{d} F_{(2m_1,2m_2)}$ where $(\xi_{1i}, \xi_{2j}) \neq (\xi_{1l}, \xi_{2k})$, m_1 and m_2 are the number of periodogram ordinates in the sums, $m = m_1 + m_2$ is the total number of periodogram ordinates in the frequency domain and $F_{(2m_1,2m_2)}$ is a F-Snedecor distribution with $2m_1$ and $2m_2$ degrees of freedom.

Proof. Given that $\epsilon(s_{1i}, s_{2j}) \overset{i.i.d}{\sim} \mathcal{N}(0, \sigma^2)$ then $C(h_1, h_2) = \sigma^2$ if $h_1 = 0, h_2 = 0$ and 0 in another case, this implies that $s(\xi_{1i}, \xi_{2j}) = \frac{1}{(2\pi)^2} \sum_{h_1=-\infty}^{\infty} \sum_{h_2=-\infty}^{\infty} C(h_1, h_2) cos(\xi_1 h_1 + \xi_2 h_2) = \sigma^2/(2\pi)^2$ $\forall \xi_{1i}, \xi_{2j}$ (see Chapter 2). Additionally, under $H_0 : \rho = 1$, $\Delta z(s_{1i}, s_{2j}) = \epsilon(s_{1i}, s_{2j})$ thus the population spectrum of $\Delta z(s_{1i}, s_{2j})$ is $\sigma^2/(2\pi)^2$ $\forall \xi_{1i}, \xi_{2j}$. And due to asymptotically $Cov[I(\xi_{1i}, \xi_{2j}), I(\xi_{1l}, \xi_{2k})] = 0$, $\xi_{1i} \neq \xi_{1l}, \xi_{2j} \neq \xi_{2k}$ and $\frac{2I(\xi_{1i},\xi_{2j})}{s(\xi_{1i},\xi_{2j})} \xrightarrow{d} \chi^2_2$ where χ^2_2 is a Chi-square distribution with two degrees of freedom, then $\frac{\sum_{\xi_{1i},\xi_{2j}\neq[0,0]} 2I(\xi_{1i},\xi_{2j})}{s(\xi_{1i},\xi_{2j})} \xrightarrow{d} \chi^2_{2m_1}$ and $\frac{\sum_{\xi_{1l},\xi_{2k}\neq[0,0]} 2I(\xi_{1l},\xi_{2k})}{s(\xi_{1l},\xi_{2k})} \xrightarrow{d} \chi^2_{2m_2}$ where m_1 and m_2 are the number of periodogram ordinates in the sums, $m = m_1 + m_2$ is the total number of periodogram ordinates in the frequency domain and $\chi^2_{2m_1}$ and $\chi^2_{2m_2}$ are Chi-square distributions with $2m_1$ and $2m_2$ degrees of freedom, respectively. Given that $(\xi_{1i}, \xi_{2j}) \neq (\xi_{1l}, \xi_{2m})$ then $\frac{\sum_{\xi_{1i},\xi_{2j}\neq[0,0]} 2I(\xi_{1i},\xi_{2j})}{s(\xi_{1i},\xi_{2j})}$ and $\frac{\sum_{\xi_{1l},\xi_{2k}\neq[0,0]} 2I(\xi_{1l},\xi_{2k})}{s(\xi_{1l},\xi_{2k})}$ are asymptotically independent thus $\hat{\psi} = \frac{\frac{\sum_{\xi_{1i},\xi_{2j}\neq[0,0]} 2I(\xi_{1i},\xi_{2j})}{s(\xi_{1i},\xi_{2j})2m_1}}{\frac{\sum_{\xi_{1l},\xi_{2k}\neq 0} 2I(\xi_{1l},\xi_{2k})}{s(\xi_{1l},\xi_{2k})2m_2}} = \frac{\frac{\sum_{\xi_{1i},\xi_{2j}\neq 0} 2I(\xi_{1i},\xi_{2j})}{(\sigma^2/(2\pi)^2)2m_1}}{\frac{\sum_{\xi_{1l},\xi_{2k}\neq 0} 2I(\xi_{1l},\xi_{2k})}{(\sigma^2/(2\pi)^2)2m_2}} = \frac{\frac{\sum_{\xi_{1i},\xi_{2j}\neq[0,0]} I(\xi_{1i},\xi_{2j})}{m_1}}{\frac{\sum_{\xi_{1l},\xi_{2k}\neq[0,0]} I(\xi_{1l},\xi_{2k})}{m_2}} \xrightarrow{d} F_{(2m_1,2m_2)}$ by the continuous mapping theorem, where $F_{(2m_1,2m_2)}$ is a F-Snedecor distribution with $2m_1$ and $2m_2$ degrees of freedom. $\square$

Obviously, we chose (ξ_{1i}, ξ_{2j}) equal to the m_1 frequencies around the origin $(0, 0)$ (low frequencies) and (ξ_{1l}, ξ_{2k}) are the other frequencies surrounding the central area (high frequencies). This is due to using the pattern of the periodogram associated with $\Delta z(s_{i1}, s_{2i})$ in order to build our statistic.

This test can be carried out by comparing $\hat{\psi}$ against the confidence bounds $(F_{\alpha/2,2m_1,2m_2}, F_{1-\alpha/2,2m_1,2m_2})$. Values of this statistic outside of these bounds suggest $\rho \neq 1$, which implies that the field does not have a spatial unit root.

To highlight the finite sample properties of the proposed statistical test and compare its properties with Likelihood Ratio, Wald and Lagrange Multiplier (Score) statistical tests, i.e., tests based on principled inference, we conduct some simulation exercises to calculate the size and power of different tests to check an spatial unit root process. Specifically, we simulate spatial unit root pure SAR models on lattices of dimensions 10×10 and 25×25, each one 1,000 times, using rook and queen contiguity criteria. Then we fix a nominal size of 0.05 to test the null hypothesis of a spatial unit root field using different tests.[2]

In particular, we calculate the size of the statistical tests proposed by Lauridsen and Kosfeld (2006) and Lauridsen and Kosfeld (2004). Lauridsen and Kosfeld (2006) propose a two stage strategy to test $H_0 : \rho = 1$ using the Lagrange Multiplier error statistic developed by Anselin (1988) (see Chapter 1). On the other hand, it can be intuitive to estimate the SAR model, $z(s_{1i}, s_{2j}) = \rho \sum_{h_1} \sum_{h_2} w_{ij,h_1h_2} z(s_{1i\pm h_1}, s_{1j\pm h_2}) + \epsilon(s_{1i}, s_{2j})$, and test the null hypothesis of a spatial unit root process using a Wald or Likelihood Ratio test. However, it is well known that $\frac{1-\hat{\rho}}{\widehat{s.e(\hat{\rho})}}$ does not adhere to a normal or t distribution. Then, Lauridsen and Kosfeld (2004) use a Wald test to check the null hypothesis of a spatial unit root process, but they build the distribution of the statistic by Monte Carlo simulation (see Chapter 1). Moreover, we calculate the size of the Likelihood Ratio and Wald tests to test $H_0 : \gamma = 1 - \rho = 0$ in the model $\Delta z(s_{1i}, s_{2j}) = \alpha + \gamma \sum_{h_1} \sum_{h_2} w_{ij,h_1h_2} \Delta z(s_{1i\pm h_1}, s_{1j\pm h_2}) + \epsilon(s_{1i}, s_{2j})$ where this model is estimated using Maximum Likelihood. As can be seen in Table 4.1, the empirical size is larger than the nominal size in a 10×10 regular lattice, except in the case of queen criterion with the Lauridsen and Kosfeld (2006) test. The size of the statistical test based on the differences between low and high frequencies is the highest, 0.095 and 0.092 in the cases of rook and queen criteria, respectively. While the test proposed by Lauridsen and Kosfeld (2006) and the Likelihood Ratio test are near the nominal size. As the dimension of the lattice increases, there is a reduction in the gap between the test proposed by Lauridsen and Kosfeld (2006) and the test based on the differences between frequencies, i.e., there is a low probability of type I error, and we get sensible outcomes with the test in the frequency domain. Finally, the worst outcomes as the dimension of the lattice increases are gotten by the test proposed by Lauridsen and Kosfeld (2004).

To analyse the power of the statistical test, we conduct another kind of simulation exercise. Specifically, we simulate stationary SAR fields on lattices of 10×10 and 25×25 with different contiguity criteria and toroidal boundary restrictions. Additionally, we use different levels of spatial autocorrelation to analyse the performance of the tests. As can be seen in Table 4.2, the power improves if the spatial autocorrelation coefficient decreases, but there is lower discrimination, i.e., there is higher probability of type II error when the spatial root is near one, but this problem is solved as the dimension of the lattice increases. We can observe that the power of the test based on differences between low and high frequencies is the worst when the lattice is of dimension 25×25. On the other hand, the Lauridsen and Kosfeld (2006) test obtains the worst outcomes with 10×10 lattices. Finally, the best power is obtained with the test proposed by Lauridsen and Kosfeld (2004).

[2] It should be kept in mind that the simulation standard error of each estimate is approximately equal to 0.007 (the usual standard error in computing a binomial proportion is equal to $\sqrt{(0.05)(0.95)/1,000}$).

Table 4.1: Empirical size of statistical tests for a pure unit root in a spatial autoregressive process in lattice data.*

Rook		Queen	
10×10	25×25	10×10	25×25
Frequency domain			
0.095	0.054	0.092	0.057
Lauridsen and Kosfeld (2006)			
0.053	0.052	0.046	0.045
Lauridsen and Kosfeld (2004)			
0.068	0.065	0.09	0.087
LR			
0.053	0.048	0.057	0.054
Wald			
0.060	0.049	0.072	0.054

**Nominal size is 0.05.*

Source: Author's estimations.

As a conclusion, the test proposed by Lauridsen and Kosfeld (2004) has the worst size but the best power under the parametric specification that is assumed. On the other hand, the size and power of the frequency domain test that we propose gets better as the sample increases. However, principle inference tests overcome our test, an outcome that is expected under the parametric specification. The test proposed by Lauridsen and Kosfeld (2006) obtains an excellent size, although its power is exceeded by the Likelihood Ratio and Wald tests. With regard to these tests, we obtain sensible outcomes. Finally, all these results are robust in the face of different contiguity criteria or boundary restrictions.

In order to ensure accuracy at the null hypothesis, we propose a Monte Carlo test. Specifically, the test is implemented by simulating unit root SAR processes under rook contiguity criterion without toroidal restrictions, and calculating values of the statistical test $\hat{\psi}$ and comparing them to each statistic calculated from the simulated data under alternative hypothesis. Following Cressie (1993), let $\hat{\psi}$ denote a test statistic and let $\left\{\hat{\psi}_i : i = 1, 2, \ldots, k-1\right\}$ denote $(k-1)$ values of the statistic generated by independently simulating data of size n under $H_0 : \rho = 1$. To see the power of this Monte Carlo test, call $\hat{\psi}_k = \hat{\psi}_{(\rho)(s)}$ where $\rho = \{0.2, 0.5, 0.9, 0.95\}, s = \{1, 2, ..., S\}$ and S the number of simulations, the calculated value of $\hat{\psi}$ with simulated realizations of random fields under the different alternative hypothesis. Ask whether $\hat{\psi}_k$ is equal to $\hat{\psi}_i$'s; this is accomplished by ordering $\hat{\psi}_{(1)} \leq \hat{\psi}_{(2)} \leq \hat{\psi}_{(3)} \cdots \leq \hat{\psi}_{(k)}$ and not rejecting the null hypothesis if $\hat{\psi}_k$ is in a given interval. Specifically, given that we

Table 4.2: Empirical power of statistical tests for a pure unit root in a spatial autoregressive process in lattice data.

	Rook		Queen		Rook (Torus)		Queen (Torus)	
ρ	10×10	25×25	10×10	25×25	10×10	25×25	10×10	25×25
				Frequency domain				
0.2	0.956	1.000	0.648	0.966	0.978	1.000	0.635	0.952
0.5	0.819	1.000	0.506	0.908	0.867	1.000	0.522	0.886
0.9	0.286	0.606	0.180	0.319	0.251	0.580	0.154	0.293
0.95	0.164	0.287	0.161	0.287	0.144	0.266	0.111	0.135
				Lauridsen and Kosfeld (2006)				
0.2	1.000	1.000	0.995	1.000	1.000	1.000	0.982	1.000
0.5	0.984	1.000	0.932	1.000	0.973	1.000	0.871	1.000
0.9	0.293	0.972	0.232	0.947	0.239	0.954	0.135	0.901
0.95	0.156	0.682	0.124	0.682	0.100	0.597	0.075	0.513
				Lauridsen and Kosfeld (2004)				
0.2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.5	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.9	0.960	1.000	0.793	1.000	0.694	1.000	0.440	1.000
0.95	0.632	1.000	0.406	1.000	0.342	1.000	0.251	1.000
				LR				
0.2	1.000	1.000	1.000	1.000	1.000	1.000	0.997	1.000
0.5	0.990	1.000	0.982	1.000	0.984	1.000	0.953	1.000
0.9	0.330	0.978	0.348	0.947	0.280	0.965	0.209	0.918
0.95	0.176	0.706	0.172	0.706	0.109	0.626	0.131	0.562
				Wald				
0.2	1.000	1.000	1.000	1.000	1.000	1.000	0.999	1.000
0.5	0.996	1.000	0.991	1.000	0.992	1.000	0.969	1.000
0.9	0.397	0.988	0.391	0.964	0.330	0.973	0.230	0.925
0.95	0.208	0.726	0.197	0.726	0.132	0.656	0.151	0.581

Source: Author's estimations.

set $\alpha = 0.05$ and $S = 1,000$, we do not reject $H_0 : \rho = 1$ if $\hat{\psi}_k \in \left[\hat{\psi}_{(26)}, \hat{\psi}_{(975)}\right]$. To see the size of the Monte Carlo test, we simulate independent spatial unit root processes S times, and calculate $\hat{\psi}_k$ for each realization. Ask whether $\hat{\psi}_k$ is not equal to $\hat{\psi}_i$'s; this is accomplished by ordering $\hat{\psi}_{(1)} \leq \hat{\psi}_{(2)} \leq \hat{\psi}_{(3)} \cdots \leq \hat{\psi}_{(k)}$ and rejecting the null hypothesis if $\hat{\psi}_k$ is one of the smaller or one of the larger order statistics. Specifically, we reject $H_0 : \rho = 1$ if $\hat{\psi}_k \in \left\{\left[\hat{\psi}_{(1)}, \hat{\psi}_{(2)}, \ldots, \hat{\psi}_{(25)}\right], \left[\hat{\psi}_{(976)}, \hat{\psi}_{(977)}, \ldots, \hat{\psi}_{(1,000)}\right]\right\}$.

The outcomes for the power and size of this statistic are shown in Tables 4.3 and 4.4, respectively.

Table 4.3: Empirical power of the Monte Carlo test for a pure unit root in a spatial autoregressive process in lattice data.

	Lattice	
ρ	10×10	25×25
0.2	0.941	1.000
0.5	0.817	1.000
0.9	0.243	0.580
0.95	0.135	0.279

Source: Author's estimations.

Table 4.4: Empirical size of the Monte Carlo test for a pure unit root in a spatial autoregressive process in lattice data.*

	Lattice	
ρ	10×10	25×25
1	0.044	0.040

**Nominal size is 0.05.*

Source: Author's estimations.

It is well known that a weak stationary random field is characterized with a constant mean, but if the mean changes with location, the idea is not to associate stationarity properties with the attribute $\mathbf{z}(\mathbf{s_1}, \mathbf{s_2})$, but with the process without the large-scale structure. The following theorem, which is another important theoretic development of this dissertation, establishes the asymptotic distribution of the statistical test for a unit root in a general spatial autoregressive model in regional data.

Theorem 2. Asymptotic distribution of a statistical test for a unit root in a general spatial autoregressive model in regional data based on differences in the frequency domain.

Given $z(s_{1i}, s_{2j}) = \mathbf{x}(s_{1i}, s_{2j})'\boldsymbol{\beta} + \rho \sum_{z(s_{1l}, s_{2k}) \in N(i,j)} w_{ij,lk}(z(s_{1l}, s_{2k}) - \mathbf{x}(s_{1l}, s_{2k})'\boldsymbol{\beta}) + \epsilon(s_{1i}, s_{2j})$ *where* $N(i,j)$ *is the set of neighbours of* $z(s_{1i}, s_{2j})$ *and* $\epsilon(s_{1i}, s_{2j}) \overset{i.i.d}{\sim} \mathcal{N}(0, \sigma^2)$*, then under* $H_0 : \rho = 1$, $\Delta z(s_{1i}, s_{2j}) = \Delta \mathbf{x}(s_{1i}, s_{2j})'\boldsymbol{\beta} + \epsilon(s_{1i}, s_{2j})$ *where* $\Delta z(s_{1i}, s_{2j}) \equiv z(s_{1i}, s_{2j}) -$

$\sum_{z(s_{1l},s_{2k})\in N(i,j)} w_{ij,lk}(z(s_{1l},s_{2k})$ *and* $\Delta\mathbf{x}(s_{1i},s_{2j}) \equiv \mathbf{x}(s_{1i},s_{2j}) - \sum_{z(s_{1l},s_{2k})\in N(i,j)} w_{ij,lk}\mathbf{x}(s_{1l},s_{2k})$. *Therefore, given a consistent estimate* $\hat{\boldsymbol{\beta}}$ *of* $\boldsymbol{\beta}$, $\hat{\epsilon}(s_{1i},s_{2j}) \xrightarrow{p} \epsilon(s_{1i},s_{2j})$ *where* $\hat{\epsilon}(s_{1i},s_{2j}) = \Delta z(s_{1i},s_{2j}) - \Delta\mathbf{x}(s_{1i},s_{2j})'\hat{\boldsymbol{\beta}}$. *This implies that the spectral density function of* $\hat{\epsilon}(s_{1i},s_{2j})$ *converge in probability to* $\sigma^2/(2\pi)^2\ \forall\xi_{1i},\xi_{2j}$. *Given that* $\frac{2I(\xi_{1i},\xi_{2j})}{s(\xi_{1i},\xi_{2j})} \xrightarrow{d} \chi^2_2$ *where* χ^2_2 *is a Chi-square distribution with two degrees of freedom,* $\forall\xi_{1i} \neq 0, \xi_{2j} \neq 0$, *and asymptotically* $Cov[I(\xi_{1i},\xi_{2j}), I(\xi_{1l},\xi_{2k})] = 0$, $\forall\xi_{1i} \neq \xi_{1l}, \xi_{2j} \neq \xi_{2k}$, *then under* H_0, *we have that* $\hat{\psi} = \frac{\frac{\sum_{\xi_{1i},\xi_{2j}\neq[0,0]} I(\xi_{1i},\xi_{2j})}{m_1}}{\frac{\sum_{\xi_{1l},\xi_{2k}\neq[0,0]} I(\xi_{1l},\xi_{2k})}{m_2}} \xrightarrow{d} F_{(2m_1,2m_2)}$ *where* $(\xi_{1i},\xi_{2j}) \neq (\xi_{1l},\xi_{2k})$, m_1 *and* m_2 *are the number of periodogram ordinates in the sums,* $m = m_1 + m_2$ *is the total number of periodogram ordinates in the frequency domain and* $F_{(2m_1,2m_2)}$ *is a F-Snedecor distribution with* $2m_1$ *and* $2m_2$ *degrees of freedom.*

In this context, $I(\xi_{1i},\xi_{2j})$ is the periodogram of the residuals from regression of $\Delta z(s_{1i},s_{2j})$ on $\Delta\mathbf{x}(s_{1i},s_{2j})$.

Proof. Given $z(s_{1i},s_{2j}) = \mathbf{x}(s_{1i},s_{2j})'\boldsymbol{\beta} + \rho\sum_{z(s_{1l},s_{2k})\in N(i,j)} w_{ij,lk}(z(s_{1l},s_{2k}) - \mathbf{x}(s_{1l},s_{2k})'\boldsymbol{\beta}) + \epsilon(s_{1i},s_{2j})$ where $N(i,j)$ is the set of neighbours of $z(s_{1i},s_{2j})$ and $\epsilon(s_{1i},s_{2j}) \overset{i.i.d}{\sim} \mathcal{N}(0,\sigma^2)$.
Then, under $H_0 : \rho = 1$,
$z(s_{1i},s_{2j}) = \mathbf{x}(s_{1i},s_{2j})'\boldsymbol{\beta} + \sum_{z(s_{1l},s_{2k})\in N(i,j)} w_{ij,lk}(z(s_{1l},s_{2k}) - \mathbf{x}(s_{1l},s_{2k})'\boldsymbol{\beta}) + \epsilon(s_{1i},s_{2j})$,
thus
$z(s_{1i},s_{2j}) - \sum_{z(s_{1l},s_{2k})\in N(i,j)} w_{ij,lk} z(s_{1l},s_{2k}) = \mathbf{x}(s_{1i},s_{2j})'\boldsymbol{\beta} - \sum_{z(s_{1l},s_{2k})\in N(i,j)} w_{ij,lk}\mathbf{x}(s_{1l},s_{2k})'\boldsymbol{\beta} + \epsilon(s_{1i},s_{2j})$, i.e., $\Delta z(s_{1i},s_{2j}) = \Delta\mathbf{x}(s_{1i},s_{2j})'\boldsymbol{\beta} + \epsilon(s_{1i},s_{2j})$.
Therefore, given a consistent estimate $\hat{\boldsymbol{\beta}}$ of $\boldsymbol{\beta}$, i.e., $Lim_{n\to\infty}P\left\{||\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}|| \geq \delta\right\} = 0, \forall\delta > 0$ then
$Lim_{n\to\infty}P\left\{|\hat{\epsilon}(s_{1i},s_{2j}) - \epsilon(s_{1i},s_{2j})| \geq \delta\right\} =$
$Lim_{n\to\infty}P\left\{|(\Delta z(s_{1i},s_{2j}) - \Delta\mathbf{x}(s_{1i},s_{2j})'\hat{\boldsymbol{\beta}}) - (\Delta z(s_{1i},s_{2j}) - \Delta\mathbf{x}(s_{1i},s_{2j})'\boldsymbol{\beta})| \geq \delta\right\} =$
$Lim_{n\to\infty}P\left\{|\Delta\mathbf{x}(s_{1i},s_{2j})'(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})| \geq \delta\right\} = 0$ by the Slustky theorem, i.e., $\hat{\epsilon}(s_{1i},s_{2j}) \xrightarrow{p} \epsilon(s_{1i},s_{2j})$.
Again by Slustky theorem, the spectral density function of $\hat{\epsilon}(s_{1i},s_{2j})$ converge in probability to $\sigma^2/(2\pi)^2\ \forall\xi_{1i},\xi_{2j}$.
The rest of the proof is given in proof of Theorem 1. □

Observe that if we follow the notation given in Section 2.1, specifically equation 2.8, we get the following outcomes.

Given $\mathbf{z} = \mathbf{X}\boldsymbol{\beta} + (\mathbf{I} - \rho\mathbf{W})^{-1}\boldsymbol{\epsilon}$, under $H_0 : \rho = 1$, $\mathbf{z} - \mathbf{W}\mathbf{z} = \mathbf{X}\boldsymbol{\beta} + (\mathbf{I} - \mathbf{W})^{-1}\boldsymbol{\epsilon} - \mathbf{W}\mathbf{X}\boldsymbol{\beta} - \mathbf{W}(\mathbf{I} - \mathbf{W})^{-1}\boldsymbol{\epsilon}$ which implies that $(\mathbf{I} - \mathbf{W})\mathbf{z} = (\mathbf{I} - \mathbf{W})\mathbf{X}\boldsymbol{\beta} + (\mathbf{I} - \mathbf{W})(\mathbf{I} - \mathbf{W})^{-1}\boldsymbol{\epsilon}$, i.e, $\Delta\mathbf{z} = \Delta\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$ where $\Delta = \mathbf{I} - \mathbf{W}$ is the spatial difference operator. On the other hand, if $\rho \neq 1$, then $\Delta\mathbf{z} = \Delta\mathbf{X}\boldsymbol{\beta} + (\mathbf{I} - \mathbf{W})(\mathbf{I} - \rho\mathbf{W})^{-1}\boldsymbol{\epsilon}$ where $(\mathbf{I} - \mathbf{W})(\mathbf{I} - \rho\mathbf{W})^{-1}\boldsymbol{\epsilon}$ is autocorrelated. This implies that the periodogram of the residuals will not be evenly distributed.

In order to analyse the finite sample properties of the statistical test for a unit root in a general spatial autoregressive model, we simulate the process $z(s_{1i}, s_{2j}) = 2+0.5x_1(s_{1i}, s_{2j}) - x_2(s_{1i}, s_{2j}) + e(s_{1i}, s_{2j})$ where $\mathbf{e} = (\mathbf{I} - \rho\mathbf{W})^{-1}\boldsymbol{\epsilon}$, $\epsilon(s_{1i}, s_{2j}) \sim \mathcal{N}(0,1)$, $x_1(s_{1i}, s_{2j}) \sim \mathcal{P}(100)$ and $x_2(s_{1i}, s_{2j}) \sim \mathcal{U}(10, 50)$. This process is simulated 1,000 times on lattices of dimensions 10×10 and 25×25, using the rook and queen contiguity criteria with and without toroidal boundary conditions and $x_1(s_{1i}, s_{2j})$ and $x_2(s_{1i}, s_{2j})$ fixed. To analyse the power of different tests to check a unit root in a general spatial autoregressive model, we set $\rho = \{0.2, 0.5, 0.9, 0.95\}$ and fix a nominal size of 0.05. Then, we calculate one minus the type II error for five tests: the test in the frequency domain that we propose, the tests proposed by Lauridsen and Kosfeld (2006) and Lauridsen and Kosfeld (2004), and the Likelihood Ratio and Wald tests.

To implement our test, we calculate $\Delta\mathbf{z}$, $\Delta\mathbf{x_1}$ and $\Delta\mathbf{x_2}$ and obtain by Ordinary Least Square a consistent estimate of $\boldsymbol{\beta}' = [\beta_0, \beta_1, \beta_2]$ in the regression $\Delta\mathbf{z} = \beta_0 + \beta_1\Delta\mathbf{x_1} + \beta_2\Delta\mathbf{x_2} + \boldsymbol{\epsilon}$. Then, we take the residuals of this regression, which are a consistent estimate of $\boldsymbol{\epsilon}$, and apply the procedure that is used in the case of a pure spatial autoregressive process. Additionally, these residuals are used to implement the Likelihood Ratio and Wald tests for the null hypothesis $H_0 : \gamma = 1 - \rho = 0$ on the model $\hat{\epsilon}(s_{1i}, s_{2j}) = \alpha + \gamma \sum_{\hat{\epsilon}(s_{1l}, s_{2k}) \in N(i,j)} w_{ij,lk}\hat{\epsilon}(s_{1i}, s_{2j}) + v(s_{1i}, s_{2j})$. Also, we calculate the power of the tests developed by Lauridsen and Kosfeld (2006) and Lauridsen and Kosfeld (2004) following the stages that they propose.

We can see in Table 4.5 that the best power is obtained using the test proposed by Lauridsen and Kosfeld (2004). Additionally, it can be seen in this table that we have similar outcomes with the Likelihood Ratio, the Wald and the frequency domain tests when these are used on lattices of dimension 10×10. Under this circumstance, the test proposed by Lauridsen and Kosfeld (2006) shows the worst results. On the other hand, if we analyse lattices of dimension 25×25, we obtain the worst outcomes with our test. These results are robust because we get them independently of the contiguity criterion or boundary restriction.

To analyse the size of the tests, we simulate the same process 1,000 times but fix $\rho = 1$, then with a nominal size of 0.05, we calculate the type I error. As can be seen in Table 4.6, we have the worst outcomes with the test proposed by Lauridsen and Kosfeld (2004). We can see in this table that we get sensible results with the frequency domain test; although when we apply this test on lattices of dimension 10×10, we obtain an empirical size bigger than the sizes obtained with the Likelihood Ratio, Wald and Lauridsen and Kosfeld (2006) tests. On the other hand, if we analyse lattices of dimension 25×25, it can be seen in Table 4.6 that we have the best outcome with the frequency domain test under the assumption of a rook contiguity criterion.

Table 4.5: Empirical power of statistical tests for a unit root in a general spatial autoregressive process in lattice data.

	Rook		Queen		Rook (Torus)		Queen (Torus)	
ρ	10×10	25×25	10×10	25×25	10×10	25×25	10×10	25×25
				Frequency domain				
0.2	0.949	1.000	0.622	0.971	0.965	1.000	0.655	0.957
0.5	0.818	1.000	0.553	0.907	0.853	1.000	0.503	0.887
0.9	0.263	0.612	0.190	0.316	0.228	0.548	0.162	0.270
0.95	0.174	0.282	0.116	0.144	0.150	0.240	0.098	0.163
				Lauridsen and Kosfeld (2006)				
0.2	0.999	1.000	0.994	1.000	1.000	1.000	0.983	1.000
0.5	0.969	1.000	0.899	1.000	0.960	1.000	0.828	1.000
0.9	0.268	0.971	0.212	0.943	0.246	0.937	0.123	0.901
0.95	0.152	0.670	0.099	0.599	0.101	0.568	0.062	0.519
				Lauridsen and Kosfeld (2004)				
0.2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.5	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.9	0.993	1.000	0.841	1.000	0.770	1.000	0.565	1.000
0.95	0.872	1.000	0.466	1.000	0.500	1.000	0.276	0.999
				LR				
0.2	0.999	1.000	1.000	1.000	1.000	1.000	0.998	1.000
0.5	0.984	1.000	0.962	1.000	0.978	1.000	0.930	1.000
0.9	0.303	0.977	0.341	0.959	0.246	0.953	0.191	0.932
0.95	0.167	0.695	0.160	0.646	0.116	0.590	0.099	0.564
				Wald				
0.2	0.999	1.000	1.000	1.000	1.000	1.000	0.999	1.000
0.5	0.993	1.000	0.976	1.000	0.987	1.000	0.951	1.000
0.9	0.372	0.981	0.374	0.966	0.289	0.963	0.218	0.942
0.95	0.192	0.715	0.176	0.657	0.137	0.616	0.121	0.578

Source: Author's estimations.

All the previous simulation exercises assume a spatial autoregressive process, and we get sensible outcomes with our frequency domain test in general. However, our test is exceeded by the parametric tests many times; therefore, we perform other simulation exercises where a spatial autoregressive process is not assumed. Specifically, we perform two simulation exercises on lattices of dimension 25×25 where the process is not stationary due to non-constant mean. These processes are $z(s_{1i}, s_{2j})^1 = 10sin(0.5s_{1i})+10cos(0.5s_{2j})+\epsilon(s_{1i}, s_{2j})$ and

Table 4.6: Empirical size of statistical tests for a general unit root in a spatial autoregressive process in lattice data.*

Rook		Queen	
10×10	25×25	10×10	25×25
Frequency domain			
0.088	0.055	0.089	0.042
Lauridsen and Kosfeld (2006)			
0.051	0.058	0.038	0.046
Lauridsen and Kosfeld (2004)			
0.212	0.759	0.108	0.414
LR			
0.056	0.058	0.050	0.046
Wald			
0.066	0.060	0.061	0.047

**Nominal size is 0.05.*

Source: Author's estimations.

$z(s_{1i}, s_{2j})^2 = 1.4 + 0.1s_{1i} + 0.2s_{2j} + 0.002s_{1i}^2 + 0.003s_{2j} + \epsilon(s_{1i}, s_{2j})$ where in both $\epsilon(s_{1i}, s_{2j}) \sim \mathcal{N}(0,1)$. We calculate the size of the tests, and as we can see in Table 4.7, we get sensible outcomes with our test, the best sizes are gotten using the frequency domain and Wald tests. On the other hand, the worst sizes are obtained by the test of Lauridsen and Kosfeld (2004). As a conclusion, there is some evidence that our test performs better than the parametric tests when the data generating process is different from a spatial autoregressive process.

Table 4.7: Empirical size of statistical tests for the non-stationary process due to non-constant mean

	$z(s_{1i}, s_{2j})^1$	$z(s_{1i}, s_{2j})^2$
Frequency domain	0.084	0.096
Lauridsen and Kosfeld (2006)	0.202	0.216
Lauridsen and Kosfeld (2004)	0.000	0.000
LR	0.138	0.143
Wald	0.083	0.092

**Nominal size is 0.05.*

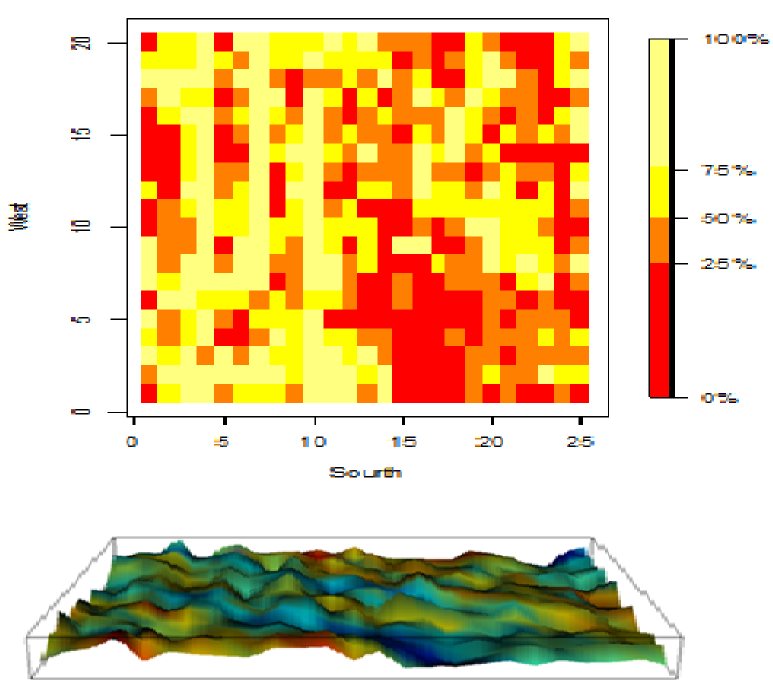
Source: Author's estimations.

5 Applications

5.1 Mercer-Hall Dataset

In order to apply the statistical test that is developed, we use the Mercer-Hall dataset (Mercer and Hall, 1911), which is widely used in lattice literature (Whittle, 1954; Cressie, 1993; Ripley, 2004).

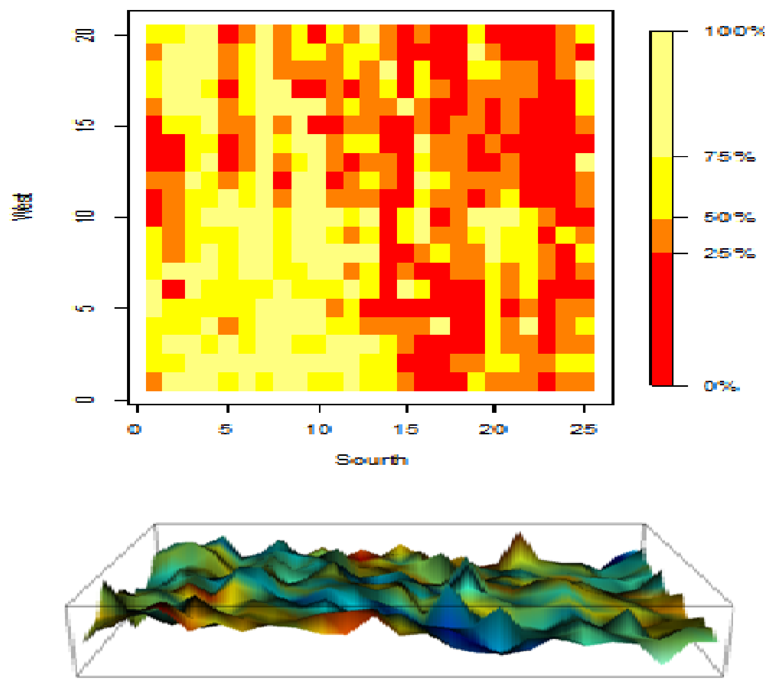
Figure 5.1: Mercer-Hall's dataset: grain.



"...Mercer and Hall (1911) were trying to determine the optimum plot size for agricultural yield trials:

- Plots that are too small will be too variable;

Figure 5.2: Mercer-Hall's dataset: straw.



- Plots that are too large waste resources (land, labour, seed); if the land area is limited, the number of treatments (varieties, fertilizers, herbicides ...) will be restricted.

So they performed a very simple experiment at the famous Rothamsted Experiment Station (Harpenden, Herts, England): an apparently homogeneous field was selected, prepared as uniformly as possible and planted to same variety of wheat. They attempted to treat all parts of the field exactly the same in all respects during subsequent farm operations. When the wheat had matured, the field was divided into 500 equally-sized plots. Each plot was harvested separately. Both grain and straw were air-dried, then hand-threshed and weighed to a precision of 0.01 lb (4.54 g = 0.00454 kg). The reported values are thus air-dry weight in pounds per plot.

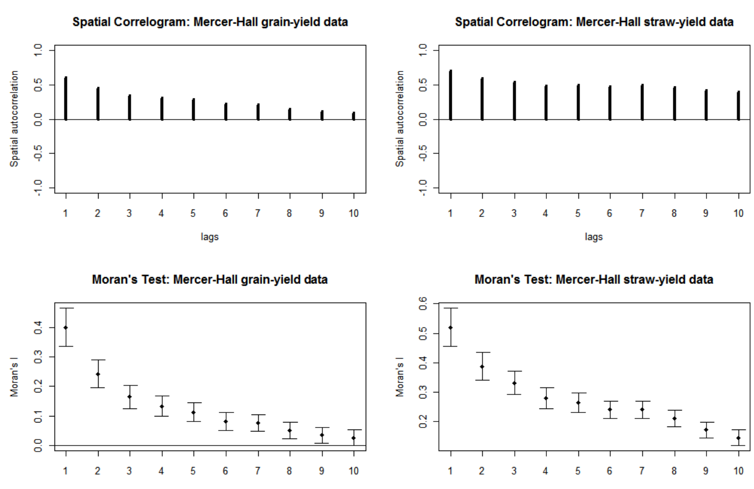
The field was a square of 1 acre, a historical English measure of land area which is equivalent to 0.40469 ha (4,046.9 m^2), or 63.615 m on a side. The field was divided into a 20 rows by 25 columns, giving 500 plots, each of 1/500 acre. Dividing the square by the number of rows and columns, we obtain plots 3.1807 m long $\times$ 2.5446 m wide, with an area of 8.0937 m^2. We do not have records of the original orientation of the field, so we assume that the rows ran W (west) to E (east), with 25 plots in each row, beginning at 1 on the W and running to 25 at the E. Then the columns run S (south) to N (north) with 20 plots in each, beginning at 1 on the S and running to 20 at the N. . . .

The first statistician to deal with the dataset was none other than "Student", the pen name of William Seely Gosset, who developed methods for statistical treatment of small samples, including the t-distribution, among many other accomplishments. He wrote an appendix (pp. 128-132) to Mercer and Hall's paper."

Rossiter (2010, pp 118)

The Mercer-Hall's dataset is displayed in Figures 5.1 and 5.2, as we can see, there are some signs of spatial autocorrelation because of ripples running West-East. Seemingly, spatial autocorrelation in straw-yield data is stronger than spatial autocorrelation in grain-yield data.

Figure 5.3: Mercer-Hall's dataset: spatial correlogram and Moran's I test.



As can be seen in Figure 5.3, evidence against not spatial correlation is stronger in straw-yield data than in grain-yield data because spatial correlogram and Moran's I statistics decrease slower for the former. In fact, Moran's I tests under randomisation are 0.40 and 0.52, respectively. At 5% significant level, we reject the null hypothesis of no spatial correlation for each variable. Given the strong distributional assumptions of the Moran's statistic, we test spatial correlation for each variable with Moran's I test under permutations and 10,000 simulations. Figure 5.4 shows the outcomes, and again we reject the null hypothesis at conventional significance levels due to their Moran statistic values being 0.41 and 0.52, respectively.

Figure 5.4: Mercer-Hall's dataset: Histograms of Moran's I test under permutations.

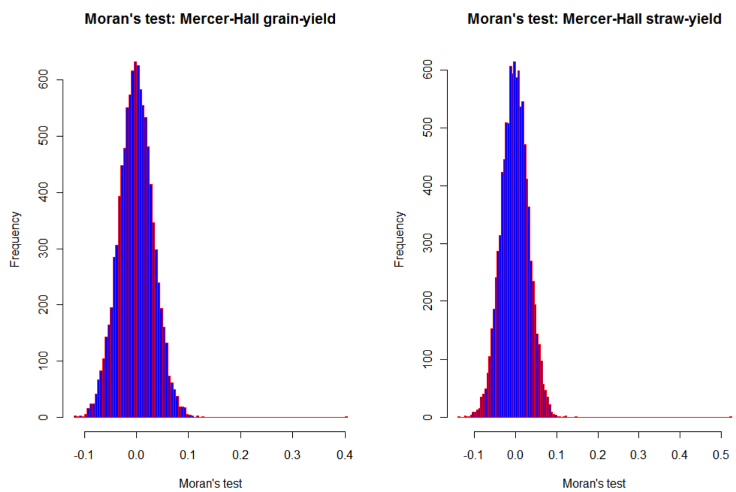
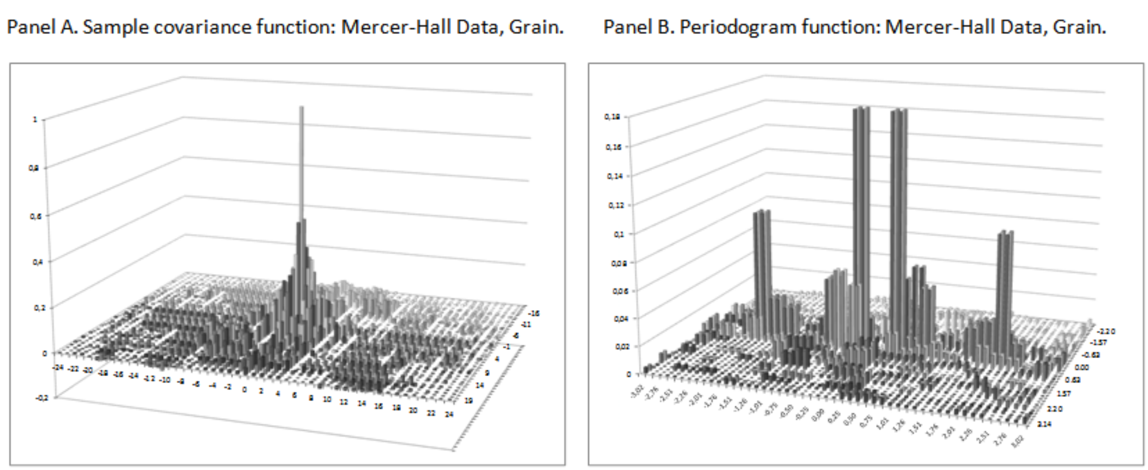


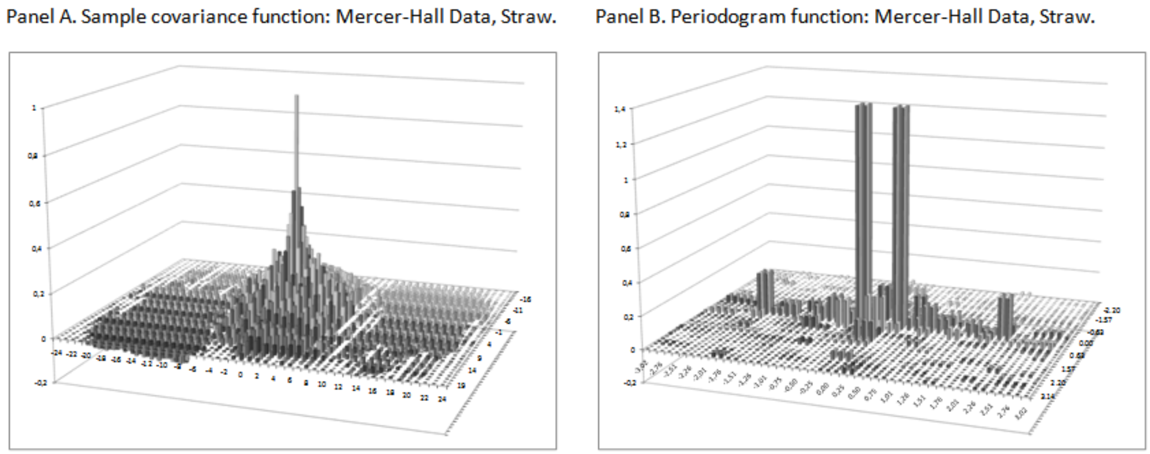
Figure 5.5: Mercer-Hall's dataset: spatial covariance and periodogram functions of grain-yield.



Figures 5.5 and 5.6 display the sample spatial covariance (Panel A) and the periodogram (Panel B) functions for grain-yield and straw-yield, respectively.[1] Due to ripples running

[1]The mean is removed from the original array.

Figure 5.6: Mercer-Hall's dataset: spatial covariance and periodogram functions of straw-yield.



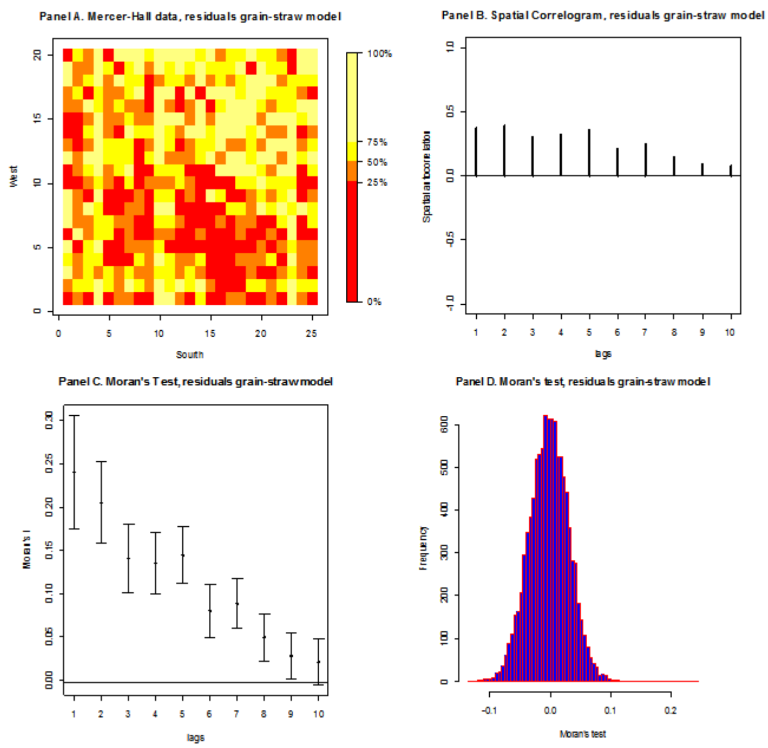
West-East, there are peaks for small coordinates that decrease slowly, especially for straw-yield data. In interpreting plots such as Panel B of Figures 5.5 and 5.6, it is often more convenient to think in terms of the period of a cycle function rather than its frequency. As can be seen in these figures, there are noticeable peaks on the periodograms at frequencies $(0, -2.01)$ and $(0, 2.01)$, corresponding to a wavelength of approximately 8 meters in the East-West direction.[2]

With regard to this last point, there are some studies that re-analyse the Mercer-Hall dataset using spectral analysis. Specifically, McBratney and Webster (1981) show that there is a periodicity in the East-West direction for both grain and straw data with a wavelength of three rows of plots. As they say, this wavelength is attributed to an earlier ridge and furrow system. Ripley (2004) computes a smoothed spectral density estimate using the Fast Fourier Transform. The necessity of using a smoothed estimate arises because the periodogram estimate is biased in a small sample due to leakage effect. In particular, Ripley (2004) uses a bivariate Normal density function as window in his estimations, and discovers a peak at frequencies $(10\pi/32, 0)$ and $(-10\pi/32, 0)$, corresponding to a wavelength of 35 feet in the East-West direction.

We estimate pure SAR models for grain and straw data, and obtain $\hat{\rho}_{grain} = 0.60$ and $\hat{\rho}_{straw} = 0.70$, both parameters are statistically significant at 5% level. These outcomes are consistent with the evidence that was shown earlier where the sample spatial autocorrelation function for the straw data decreases slower than the grain data.

[2]If the frequency of a cycle is ξ, the period of the cycle is $2\pi/\xi$, then for this case the period is $2.5446(2\pi/2.01)$ meters.

Figure 5.7: Grain-Straw model: Residuals.



After this analysis, we test the null hypothesis of a spatial unit root in the Mercer-Hall dataset with the statistic that we propose. First, the test based on the asymptotic distributions is applied to the grain and straw dataset. Specifically, we use the following algorithm.

Algorithm for the test

- Calculate $\Delta z(s_{1i}, s_{2j})^{grain} = z(s_{1i}, s_{2j})^{grain} - \sum_{h_1}\sum_{h_2} w_{ij,h_1h_2} z(s_{1i\pm h_1}, s_{1j\pm h_2})^{grain}$ and $\Delta z(s_{1i}, s_{2j})^{straw} = z(s_{1i}, s_{2j})^{straw} - \sum_{h_1}\sum_{h_2} w_{ij,h_1h_2} z(s_{1i\pm h_1}, s_{1j\pm h_2})^{straw}$
- Estimate the periodograms of $\Delta \mathbf{z}^{grain} = \left[\Delta z(s_{11}, s_{21})^{grain}, \Delta z(s_{11}, s_{22})^{grain}, \ldots, \Delta z(s_{1n}, s_{2n})^{grain}\right]'$ and $\Delta \mathbf{z}^{straw} = \left[\Delta z(s_{11}, s_{21})^{straw}, \Delta z(s_{11}, s_{22})^{straw}, \ldots, \Delta z(s_{1n}, s_{2n})^{straw}\right]'$
- Calculate the means of the periodograms associated with the $m_1 = m/2$ central frequencies around the origin in the space of frequencies, and the mean of the rest of the periodograms for each random field
- Estimate $\hat{\psi}$ for each process
- Compare $\hat{\psi}$ of each process against the confidence bounds $(F_{\alpha/2,2m_1,2m_2}, F_{1-\alpha/2,2m_1,2m_2})$

- Values of these statistics outside of bounds suggest $\rho \neq 1$, which implies that the fields do not have a spatial unit root

As can be seen in Table 5.1, the test rejects the null hypothesis of a spatial unit root process in both grain and straw data.

Table 5.1: Results of the spatial unit root test based on asymptotic distributions: Mercer-Hall's dataset.

	The asymptotic test		
	$\hat{\psi}$	$F_{0.025,500,496}$	$F_{0.975,500,496}$
Grain	0.348	0.838	1.192
Straw	0.286	0.838	1.192

Source: Author's estimations.

Additionally, we use the Monte Carlo test to check the null hypothesis of a spatial unit root. The algorithm that is used is the following.

Algorithm for the Monte Carlo test

- Calculate $\hat{\psi}_k$ for each process
- Estimate the variances of each random field, $\hat{\sigma}^2_{\Delta \mathbf{z}^{grain}}$ and $\hat{\sigma}^2_{\Delta \mathbf{z}^{straw}}$
- Simulate $k-1$ independent spatially uncorrelated autoregressive random fields with variances $\hat{\sigma}^2_{\Delta \mathbf{z}^{grain}}$ and $\hat{\sigma}^2_{\Delta \mathbf{z}^{straw}}$ and dimension 20×25 for each field. Remember, under $H_0 : \rho = 1$, $\Delta \mathbf{z} = \boldsymbol{\epsilon}$
- Calculate $\left\{ \hat{\psi}_i : i = 1, 2, \ldots, k-1 \right\}$ for each spatially uncorrelated autoregressive random field
- Order $\hat{\psi}_{(1)} \leq \hat{\psi}_{(2)} \leq \hat{\psi}_{(3)} \cdots \leq \hat{\psi}_{(k)}$
- Ask whether $\hat{\psi}_k$ for each field is different from the other $\hat{\psi}_i$'s. This is accomplished by rejecting the null hypothesis if $\hat{\psi}_k$ is one of the smaller or one of the larger order statistics

We can see in Table 5.2 that the Monte Carlo test rejects the null hypothesis of a spatial unit root in grain and straw data.

Table 5.2: Results of the spatial unit root tests based on Monte Carlo distributions: Mercer-Hall's dataset.

The Monte Carlo test			
	$\hat{\psi}$	$\hat{\psi}_{(25)}$	$\hat{\psi}_{(975)}$
Grain	0.348	0.792	1.267
Straw	0.286	0.762	1.299

Source: Author's estimations.

Summarizing the outcomes using our spatial unit root test for the Mercer-Hall dataset, we reject the null hypothesis of a spatial unit root process in these dataset. This implies that if we use these dataset in a regression analysis, we will not apparently have the problem of nonsense outcomes.

After this preliminary analysis, we fit a model to the data and examine whether the residuals from the fit exhibit spatial autocorrelation. Thus, we postulate the linear model $z(s_1, s_2)^{grain} = \beta_0 + \beta_1 z(s_1, s_2)^{straw} + e(s_1, s_2)$ and obtain the Ordinary Least Square estimate of β_0 and β_1. Specifically, we get[3]

$$\hat{z}(s_1, s_2)^{grain} = \underset{(0.102)}{1.523} + \underset{(0.015)}{0.372} z(s_1, s_2)^{straw} \tag{5.1}$$

We can observe that there is a positive effect of straw-yield on grain-yield. Moreover, we find that Moran's I statistic is 0.24, this implies that there is spatial autocorrelation.

Figure 5.7 shows residuals plot (Panel A), spatial correlogram (Panel B), Moran's I test at different lags (Panel C) and the histogram of Moran's I test under Monte Carlo permutations (Panel D). As can be seen in this figure, there is strong evidence of spatial autocorrelation.

Additionally, we can see that there is evidence of spatial correlation of the residuals (Panel A) and there are noticeable peaks on the periodograms at frequencies $(0, -2.01)$ and $(0, 2.01)$, corresponding to a wavelength of about 8 meters in the East-West direction (see Figure 5.8).

Moreover, we perform Lagrange Multiplier diagnostics for spatial dependence in linear models. In Table 5.3 the outcomes can be seen.

[3]Standard deviation are in parenthesis.

Figure 5.8: Grain-Straw model: spatial covariance and periodogram functions of residuals.

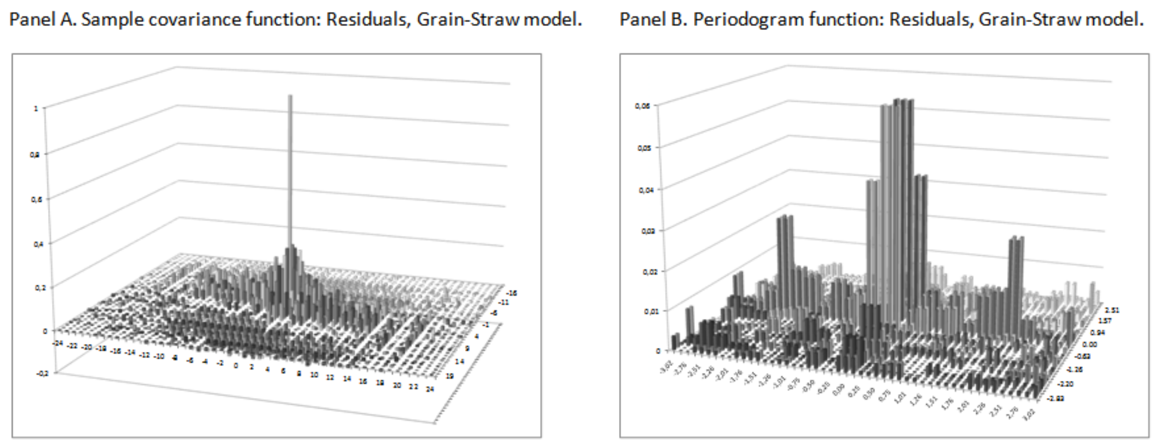


Table 5.3: Lagrange Multiplier diagnostics for spatial dependence in linear models: Mercer-Hall's dataset model.

LMerr	LMlag	RLMerr	RLMlag	SARMA
54.44	37.66	16.99	0.20	54.65

Source: Author's estimations.

These outcomes imply that there is spatial dependence in the model except in the RLMlag test. Therefore, we fit the spatial autoregressive model $z(s_1, s_2)^{grain} = \beta_0 + \beta_1 z(s_1, s_2)^{straw} + e(s_1, s_2)$ where $e(s_1, s_2) = \rho \mathbf{W} e(s_1, s_2) + \epsilon(s_1, s_2)$. We get

$$\hat{z}(s_1, s_2)^{grain} = \underset{(0.116)}{1.571} + \underset{(0.017)}{0.364} z(s_1, s_2)^{straw} + \hat{e}(s_1, s_2) \tag{5.2}$$

where

$$\hat{e}(s_1, s_2) = \underset{(0.055)}{0.364} \mathbf{W} \hat{e}(s_1, s_2) \tag{5.3}$$

We find that all variables are statistically significant at conventional levels. The Moran's statistical test in the residuals of this model is -0.036 which implies that there is no spatial autocorrelation. Additionally, we perform Lagrange Multiplier diagnostics for spatial dependency in this model, and find that the value of the LMerr test is 1.26 which implies that there is not spatial dependency. These outcomes apparently are not spurious due to grain and straw data not having a spatial unit root.

5.2 Demand of electricity in the Department of Antioquia, Colombia

In this section, we analyse the variables that would affect the electricity demand made by the representative agent of strata one in the Department of Antioquia, Colombia. In this context a representative agent is an average individual from strata one in each municipality. The residential consumers of utilities in Colombia are classified by strata. This is done in order to give subsidies to poor people who are in strata one and two. On the other hand, rich people, who are in strata five and six, have to pay contributions to the system in order to subsidize poor people. Finally, people who are in the medium strata (three and four) do not pay contribution or obtain any subsidy.

We collect information from 125 municipalities which are located in the Department of Antioquia. Specifically, we use the data of average per capita individuals classified in strata one. The analysis is focused on this segment of the population because despite obtaining subsidies, the expenditure of electricity of strata one is a big percentage of their income, a mean of 4.62%, while the average percentage of strata six is less that 1.00%. Thus, it is necessary to improve the mechanism to assign subsidies and achieve better equity. In order to take regulatory decisions based on a strong framework, it is mandatory to offer statistical tools which incorporate explicitly the spatial interaction of the municipalities because denying this aspect can cause bias and inconsistency of parameters estimates. The dataset contains information by municipality, specifically, the average annual per capita consumption of electricity (kWh), the price of electricity (COP$/kWh), the price of an electricity substitute (COP$/kWh) and the average annual per capita income (COP$/1.000).

5.2.1 Some empirical facts

As can be seen in Figure 5.9, the highest average electricity consumptions in strata one are located in municipalities with less than one thousand meters above sea level (see Figure 5.13) which face medium electricity prices (see Figure 5.10) and substitutes of electricity with the highest prices (see Figure 5.11). Occasionally, municipalities with the average highest electricity consumption are characterized by the highest average income (see Figure 5.12).

Table 5.4 shows some descriptive statistics. It can be seen that the annual average electricity consumption is 313.17 kWh with a standard deviation of 157.08 kWh. Additionally, the average annual per capita income is COP$825,280 with a standard deviation of COP$197,950. The municipality of Envigado exhibits the highest electricity consumption, and also, the highest income. On the other hand, Vigía del Fuerte has the lowest average annual per capita consumption of electricity. A reason that can explain this fact is that Vigía del

Table 5.4: Descriptive statistics: Electricity demand in the Department of Antioquia, Colombia.

Descriptive statistics				
Statistic	Electricity Consumption	Electricity Price	Substitute Price	Income
Mean	313.17kWh	126.83COP$/kWh	62.31COP$/kWh	COP$825,280
St.Dev.	157.08kWh	51.07COP$/kWh	12.47COP$/kWh	COP$197,950
Max.	785.25kWh	498.25COP$/kWh	117.31COP$/kWh	COP$1286,970
	Envigado	Entrerríos	Chigorodó	Envigado
Min.	35.46kWh	56.72COP$/kWh	33.89COP$/kWh	COP$479,090
	Vigía del Fuerte	San José de la Montaña	Venecia	El Bagre

Source: Author's estimations.

Fuerte is located in the Darién Gap, a large swath of undeveloped swampland and forest separating Panama's Darién Province in Central America from Colombia in South America. With regard to the price of electricity, the mean is 126.83 COP$/kWh and its standard deviation is 51.07 COP$/kWh. Entrerríos has the highest price and San José de la Montaña exhibits the lowest electricity price. Finally, the average price of the substitute is 62.31 COP$/kWh with its minimum located in Venecia and its maximum located in Chigorodó.

5.2.2 Spatial econometric analysis

First of all, a contiguity matrix standardized by rows based on the rook criterion is defined in order to perform the Moran test to check the presence of spatial autocorrelation in the variables that are studied. In this context, two regions are neighbours if at least two boundary points are within the snap distance of each other. The statistical evidence indicates that the null hypothesis of no spatial autocorrelation should be rejected at 5% significance level in all the variables (see Table 5.5). In order to check the robustness of these outcomes, a permutation test for Moran's I statistic is calculated by using 10,000 random permutations. These exercises confirm again that the null hypothesis should be rejected (see Figure 5.14).

Table 5.5: Moran's test: electricity demand, Antioquia (Colombia).

Electricity consumption	Electricity price	Substitute price	Income
0.421	0.148	0.384	0.650

Source: Author's estimations.

Figure 5.9: Average annual per capita electricity consumption of strata one (kWh): quintile spatial distribution, Antioquia (Colombia).

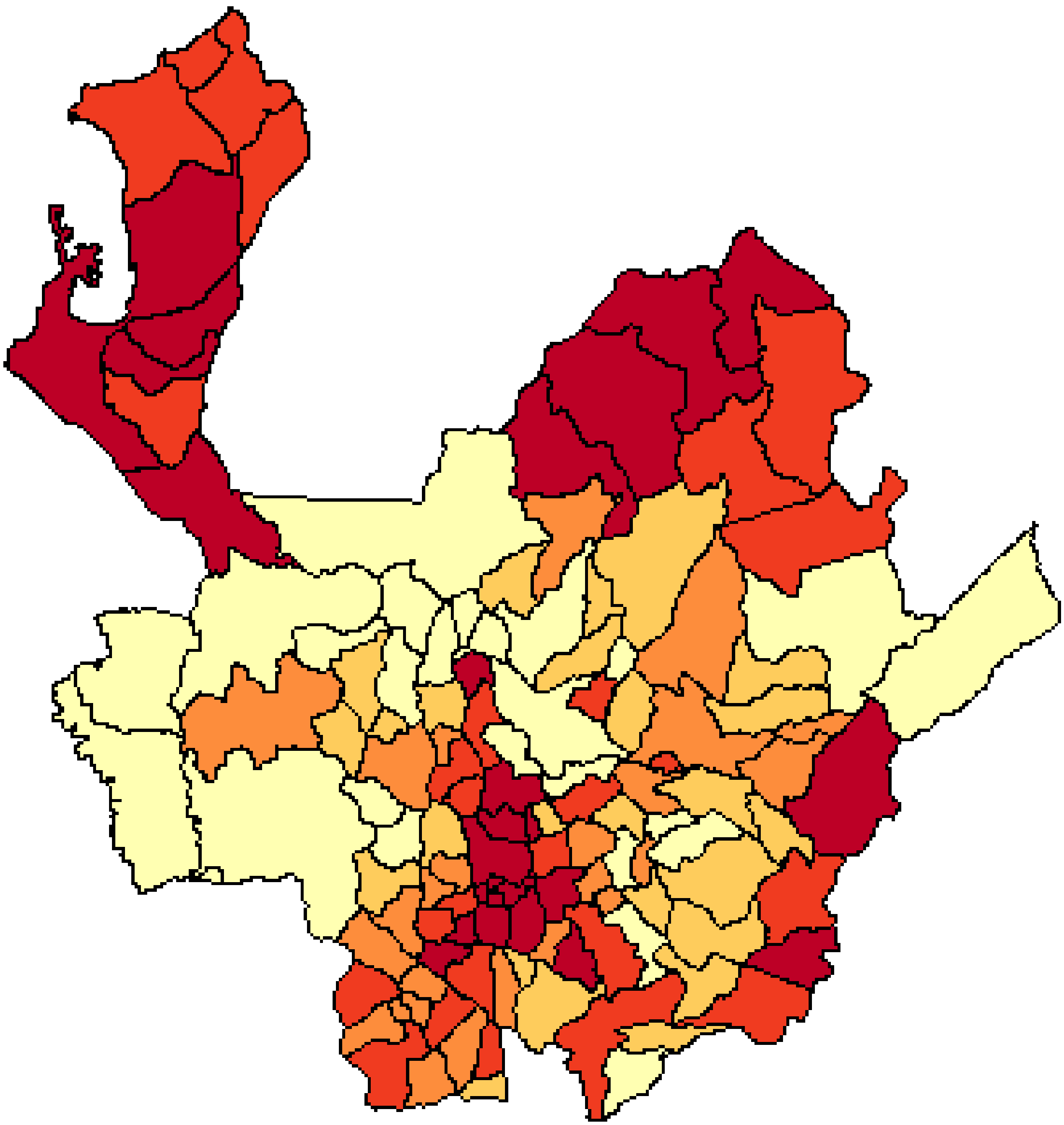


Figure 5.10: Average electricity price of strata one (COP$/kWh): quintile spatial distribution, Antioquia (Colombia).

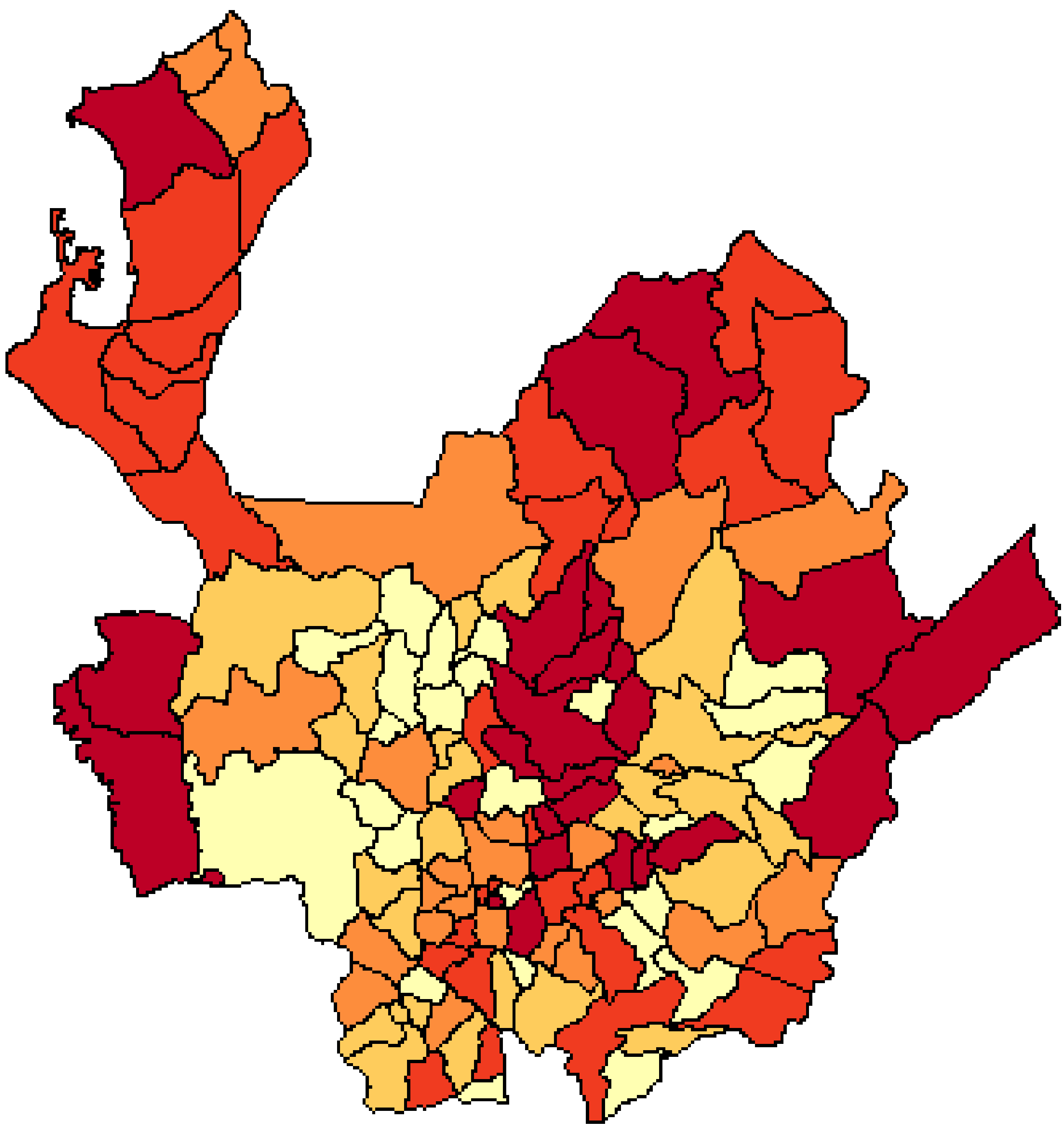


Figure 5.11: Average electricity substitute price of strata one (COP$/kWh): quintile spatial distribution, Antioquia (Colombia).

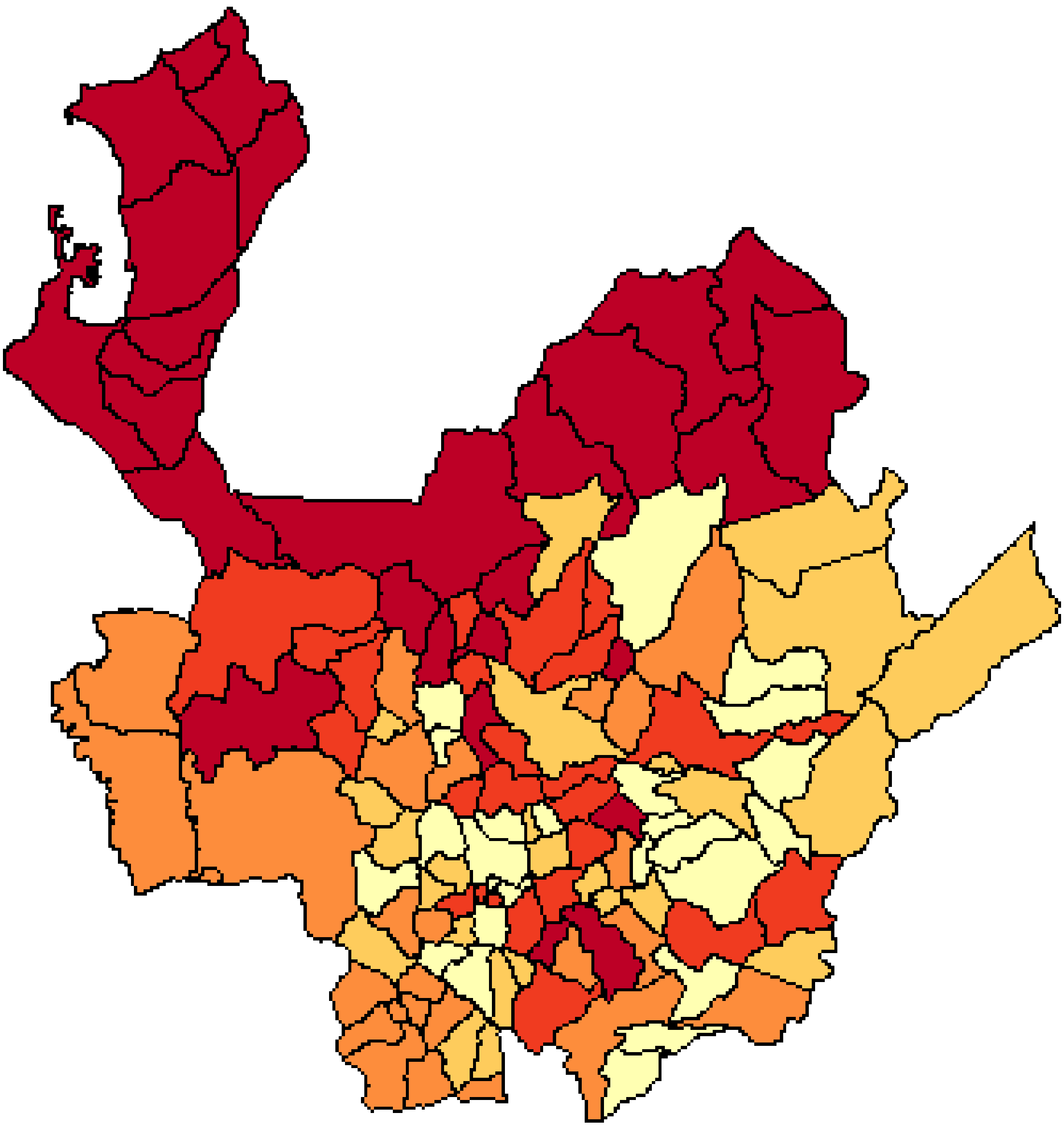


Figure 5.12: Average annual per capita income of strata one (COP$/1.000): quintile spatial distribution, Antioquia (Colombia).

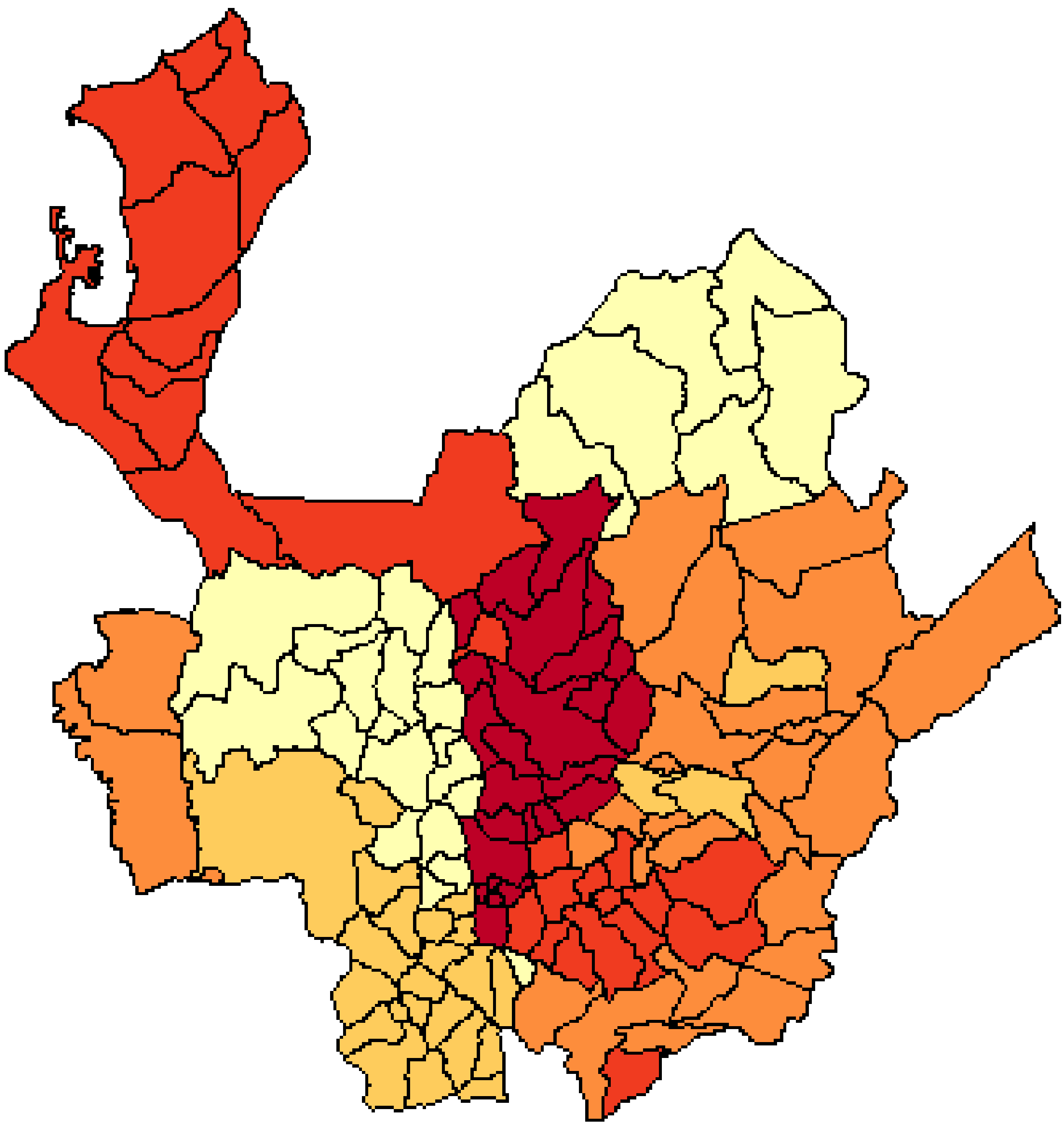


Figure 5.13: Less than one thousand meters above sea level: spatial distribution, Antioquia (Colombia).

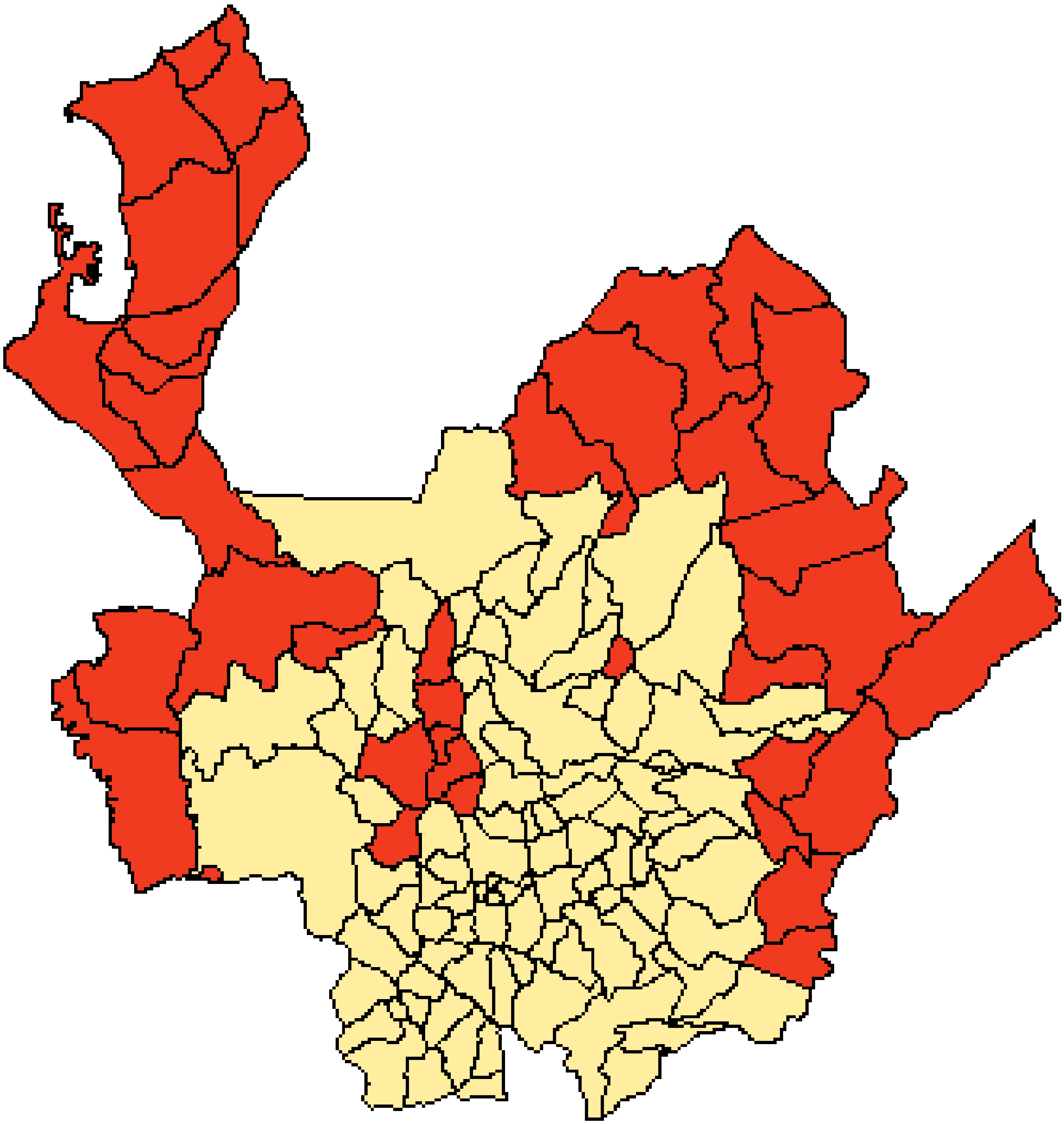
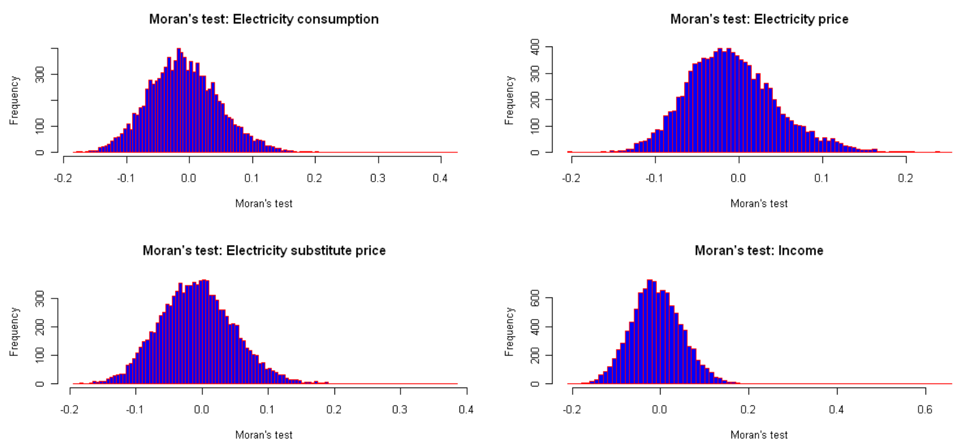
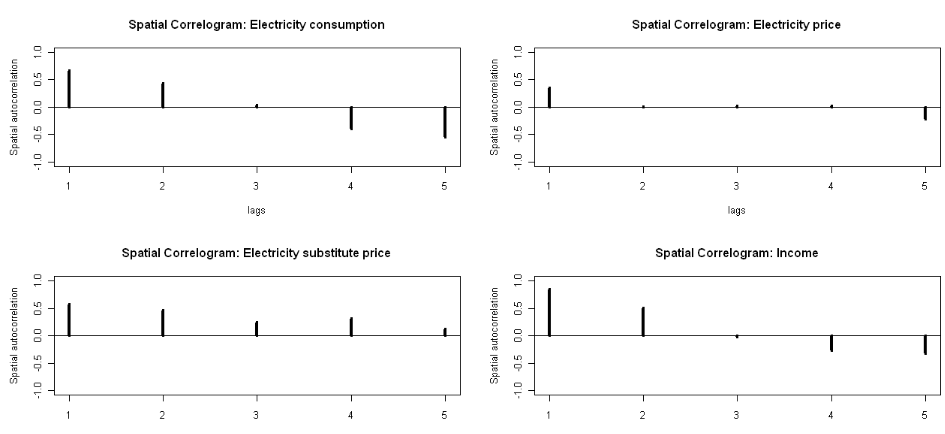


Figure 5.14: Moran's test: electricity demand, Antioquia (Colombia).



As can be seen in Figure 5.15, there is statistical evidence of spatial autocorrelation in the variables. The variable that exhibits the lowest spatial autocorrelation is the price of electricity.

Figure 5.15: Spatial autocorrelogram: electricity demand, Antioquia (Colombia).



However, there is a problem with the implementation of Moran's statistic because it assumes constant mean, but some economic foundations argue that the demand for electricity is a function of electricity price, substitute price, income and weather conditions. Therefore, we estimate the model $log(c(s_{1i}, s_{2j})) = \beta_0 + \beta_1 log(p(s_{1i}, s_{2j})) + \beta_2 log(sp(s_{1i}, s_{2j})) + \beta_3 log(y(s_{1i}, s_{2j})) + \beta_4 sl(s_{1i}, s_{2j}) + \mu(s_{1i}, s_{2j})$ where $c(s_{1i}, s_{2j})$ is the electricity consumption, $p(s_{1i}, s_{2j})$ is the electricity price, $sp(s_{1i}, s_{2j})$ is the substitute price, $y(s_{1i}, s_{2j})$ is the income and $sl(s_{1i}, s_{2j})$ is a categorical variable which takes the value of 1 if the municipality is located at less than one thousand meters above sea level and 0 in another case. We estimate the model in $log - log$ form because the parameter estimates can be interpreted as elasticities. An elasticity is a measure of sensitivity commonly used by economists, and it shows how an increment of 1% in an independent variable implies some percentage change in a dependent variable. We have the following results in this exercise.

$$log(\widehat{c(s_1, s_2)}) = \underset{(1.701)}{5.615} - \underset{(0.169)}{1.165} log(p(s_1, s_2)) + \underset{(0.224)}{0.148} log(sp(s_1, s_2)) + \underset{(0.197)}{0.735} log(y(s_1, s_2)) + \underset{(0.105)}{0.182} sl(s_1, s_2) \tag{5.4}$$

All variables are significant at 1% except $sl(s_1, s_2)$ which is significant at 10% and $log(sp(s_1, s_2))$ that is not significant.

These outcomes imply that an increment of 1% in electricity price implies a reduction in electricity demand of 1.165%. While an increment of 1% in income causes an increment of 0.735% in electricity consumption. Additionally, it is found that municipalities located at less than one thousand meters above sea level have an electricity demand 18.2% higher than municipalities without this characteristic. In general, all variables have the expected sign.

We perform spatial autocorrelation tests on residuals of this model. The value of the Moran's test is 0.49, this value implies the existence of spatial autocorrelation. Moreover, we can see in Table 5.6 that there is evidence of spatial autocorrelation except in the case of RLMlag statistic.

Table 5.6: Lagrange Multiplier diagnostics for spatial dependence in linear models: electricity demand, Antioquia (Colombia).

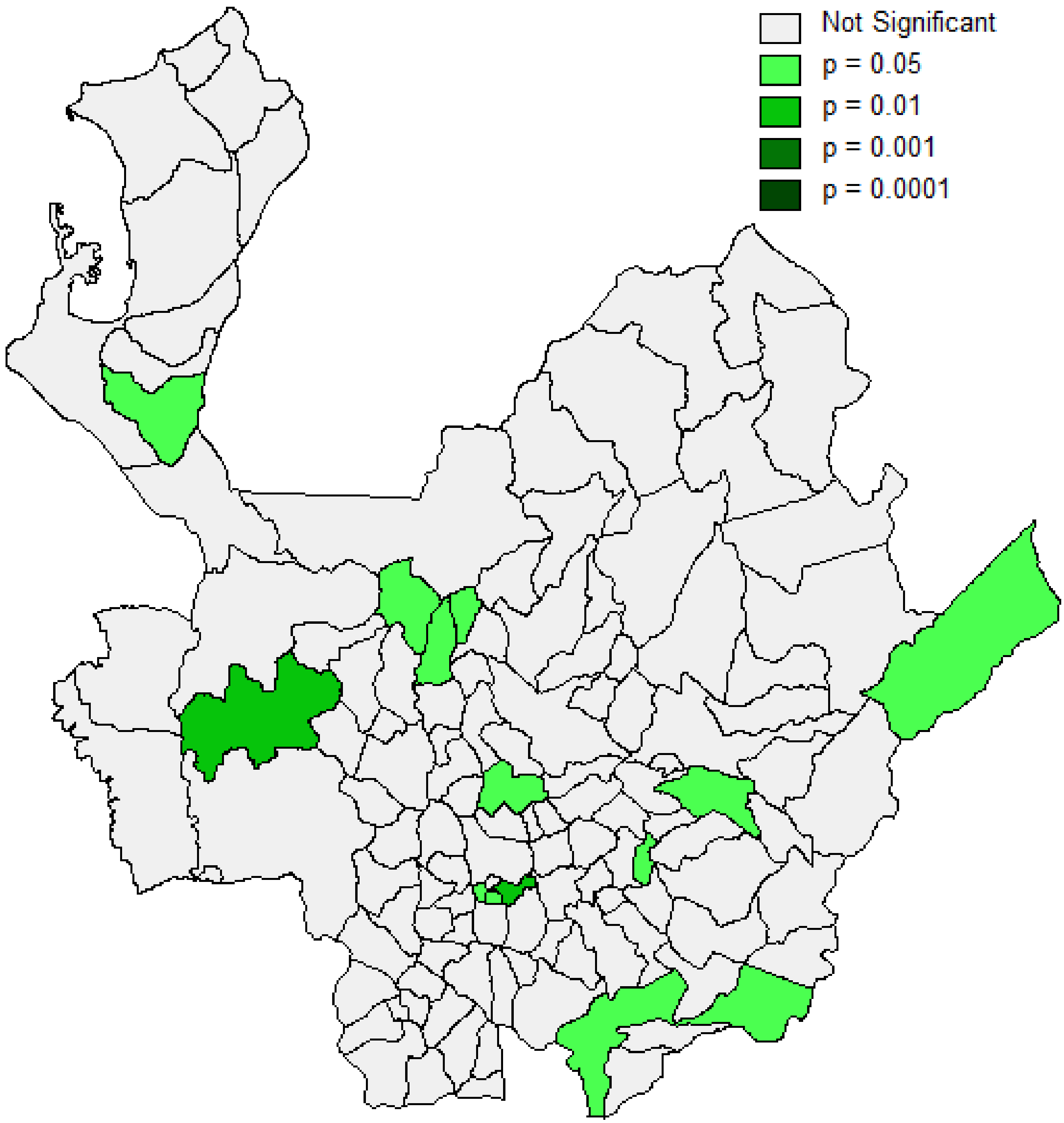
LMerr	LMlag	RLMerr	RLMlag	SARMA
73.30	68.79	5.59	1.08	74.39

Source: Author's estimations.

Additionally, we build the Local Index Spatial Autocorrelation map of the residuals to detect if there are zones with greater correlations than others.

We can see in Figure 5.16 that there are some regions that exhibit local spatial autocorrelation.

Figure 5.16: Local Index Spatial Autocorrelation Map: electricity demand, Antioquia (Colombia).



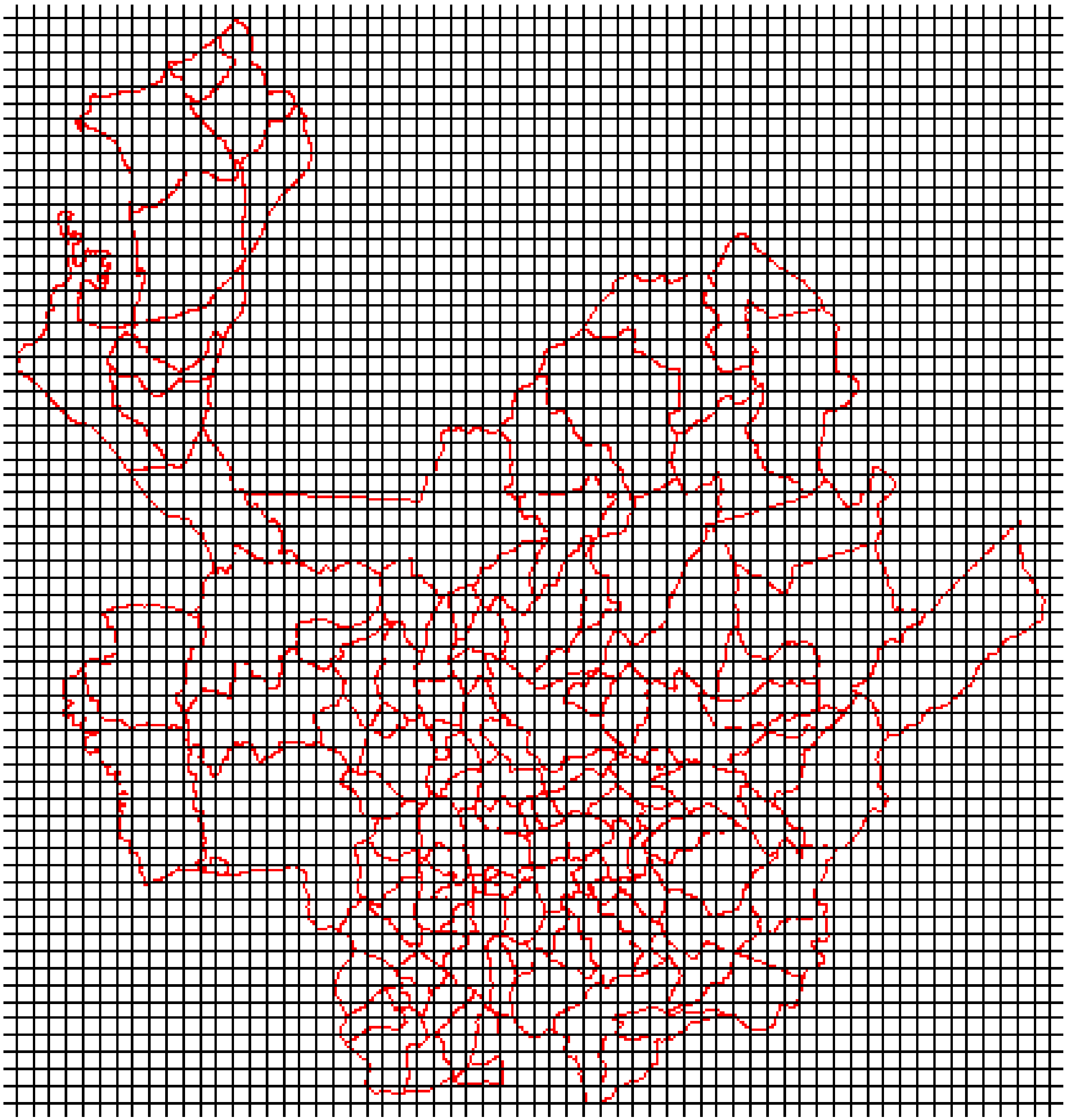
Summarizing the statistical evidence indicates that there are spatial effects that we do not take into account. Therefore, we might have a model with nonsense outcomes due to the presence of a spatial unit root in the data generating process of the electricity consumption. Thus, we implement the generalised spatial unit root test in the frequency domain to check this hypothesis. We implement the following algorithm.

Algorithm for the generalised test

- Calculate $\Delta z(s_{1i}, s_{2j}) = z(s_{1i}, s_{2j}) - \sum_{h_1} \sum_{h_2} w_{ij,h_1 h_2} z(s_{1i \pm h_1}, s_{1j \pm h_2})$ for each variable. The variables are in this case: $log(c(s_1, s_2))$, $log(p(s_1, s_2))$, $log(sp(s_1, s_2))$, $log(y(s_1, s_2))$ and $sl(s_1, s_2)$
- Estimate de model $\Delta log(c(s_{1i}, s_{2j})) = \beta_0 + \beta_1 \Delta log(p(s_{1i}, s_{2j})) + \beta_2 \Delta log(sp(s_{1i}, s_{2j})) + \beta_3 \Delta log(y(s_{1i}, s_{2j})) + \beta_4 \Delta sl(s_{1i}, s_{2j}) + \mu(s_{1i}, s_{2j})$
- Obtain the residuals of this model $\hat{\mu}(s_{1i}, s_{2j})$
- Estimate the periodogram of $\hat{\mu}(s_{1i}, s_{2j}) = \left[\hat{\mu}(s_{11}, s_{21}), \hat{\mu}(s_{11}, s_{22}), \ldots, \hat{\mu}(s_{1n}, s_{2n}\right]'$
- Calculate the mean of the periodogram associated with the $m_1 = m/2$ central frequencies around the origin in the space of frequencies, and the mean of the rest of the periodogram for each random field
- Estimate $\hat{\psi}$
- Compare $\hat{\psi}$ against the confidence bounds $(F_{\alpha/2,2m_1 2,m_2}, F_{1-\alpha/2,2m_1,2m_2})$
- Values of this statistic outside of bounds suggest $\rho \neq 1$, which implies that the field does not have a spatial unit root

However in this application, we have variables associated with a geometry that is very different compared to a regular lattice. Therefore, it is mandatory to handle some issues that arise in this case because the formulations of the spatial covariance and spectral density are based on regular lattices. This disadvantage can be handled easily because there are theoretical considerations that connect these concepts in both the regular lattice and irregular spatial polygons (Mateu and Juan, 2004). Specifically, we approximate the geometry of the regional data with a regular lattice with missing values (Fuentes, 2007). First of all, we overlay a spatial grid on the spatial polygons given by the map of the Department of Antioquia. The area of each rectangle in the grid is given by the smallest rectangle that contains completely the smallest polygon, i.e., 5,819×5,910. In Figure 5.17, the overlay of the grid on the spatial polygons can be seen.

Figure 5.17: Grid associated with Antioquia's dataset.



Second, we transfer the data in the spatial polygons to the spatial grid so that each cell has the average value of the overlaid polygons weighted by area. Renshaw (2002) establishes that this procedure is valid, but also warns that it is necessary to build a sufficiently fine covering mesh. The outcomes of these operations can be seen in Figure 5.18.

As can be seen in Figure 5.17, there are missing values in the spatial grid because of the irregular shape of the spatial polygons. Fuentes (2007) develops the theoretical framework associated with periodogram estimates for spatial regular lattices with missing values and irregularly spaced datasets. Specifically, it is demonstrated that asymptotically the bias is negligible, but in a finite sample the effect of missing data can create some impact. Additionally, the asymptotic variance of the spectrum estimates in presence of missing data is bigger than when there are no missing values.

Once the data in the spatial polygons to the spatial grid is transferred, we estimate the autocorrelation function and the periodogram of the residuals associated with the auxiliary regression that is used in the algorithm to test the spatial unit root in the electricity consumption. We omit the effect of missing values in this procedure (Fuentes, 2007). The outcomes are shown in Figures 5.19 and 5.20 where we can observe in this last figure that the periodogram function is concentrated at small frequencies, this probably implies that the process does not have a spatial unit root.

After this preliminary analysis, we follow the algorithm for the generalised test omitting the effect of missing values. We can see in Table 5.7 the outcomes of various spatial unit root tests. All the tests have as null hypothesis a spatial unit root process in the electricity consumption of the Department of Antioquia (Colombia). We reject this hypothesis with all tests including the frequency domain tests with different areas of the grids that are used to overlay the variable. Thus, the process is stationary once we take into account the large scale effects. Therefore, the estimation result are sensible.

However, there are spatial effects as we show at the beginning of this section. Then, we estimate a spatial autoregressive model, and obtain the following outcomes.

$$log(\widehat{c(s_1, s_2)}) = \underset{(1.876)}{7.578} - \underset{(0.122)}{1.150} log(p(s_1, s_2)) + \underset{(0.182)}{0.181} log(sp(s_1, s_2)) + \underset{(0.229)}{0.408} log(y(s_1, s_2)) + \underset{(0.100)}{0.033} sl(s_1, s_2) + \hat{\mu}(s_{1i}, s_{2j}) \tag{5.5}$$

$$\hat{\mu}(s_{1i}, s_{2j}) = \underset{(0.057)}{0.796} \sum_{h_1} \sum_{h_2} w_{ij,h_1h_2} \hat{\mu}(s_{1i\pm h_1}, s_{1j\pm h_2}) \tag{5.6}$$

Figure 5.18: Grid associated with residuals of auxiliary regression: Antioquia's variables of strata one.

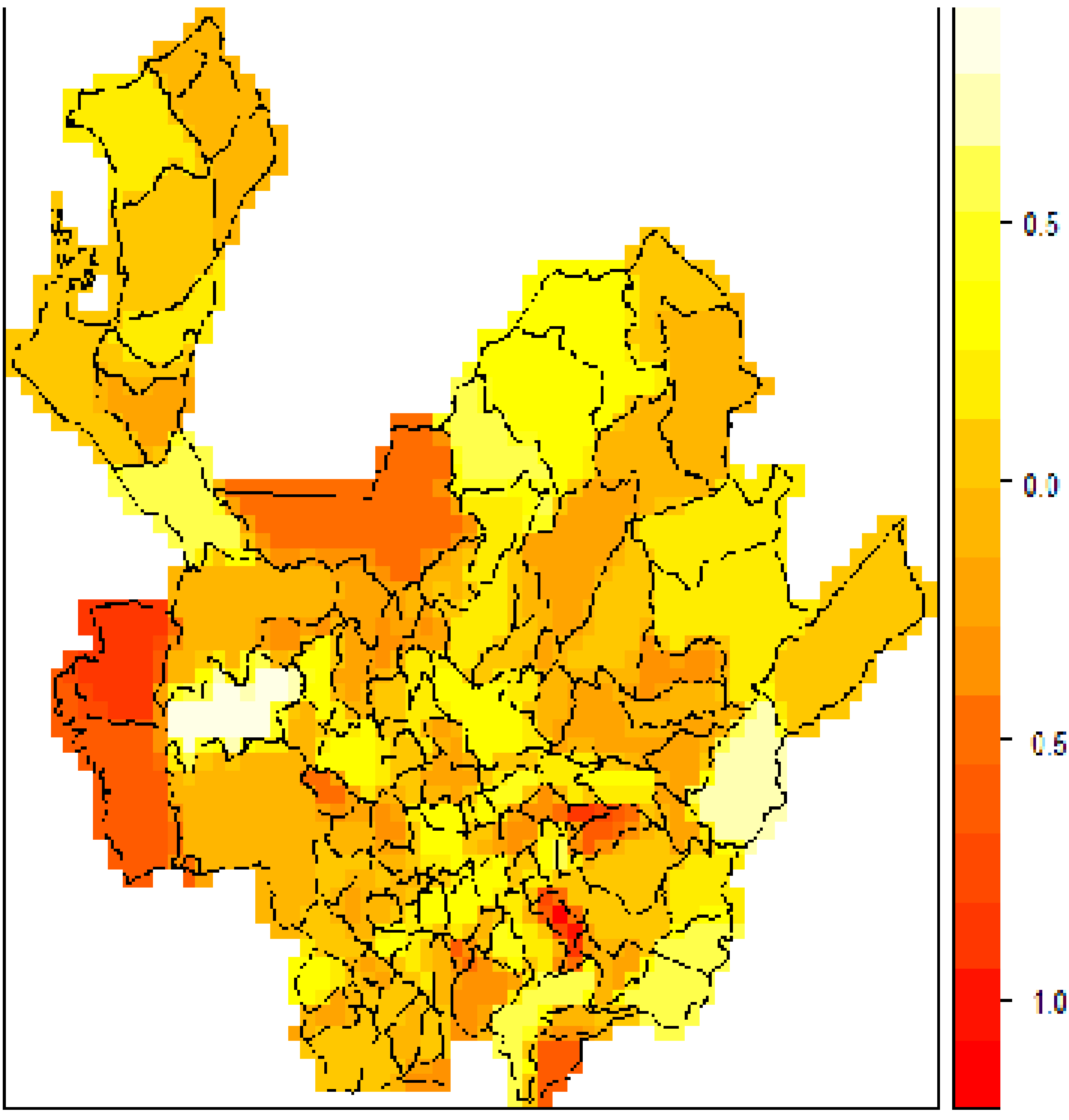
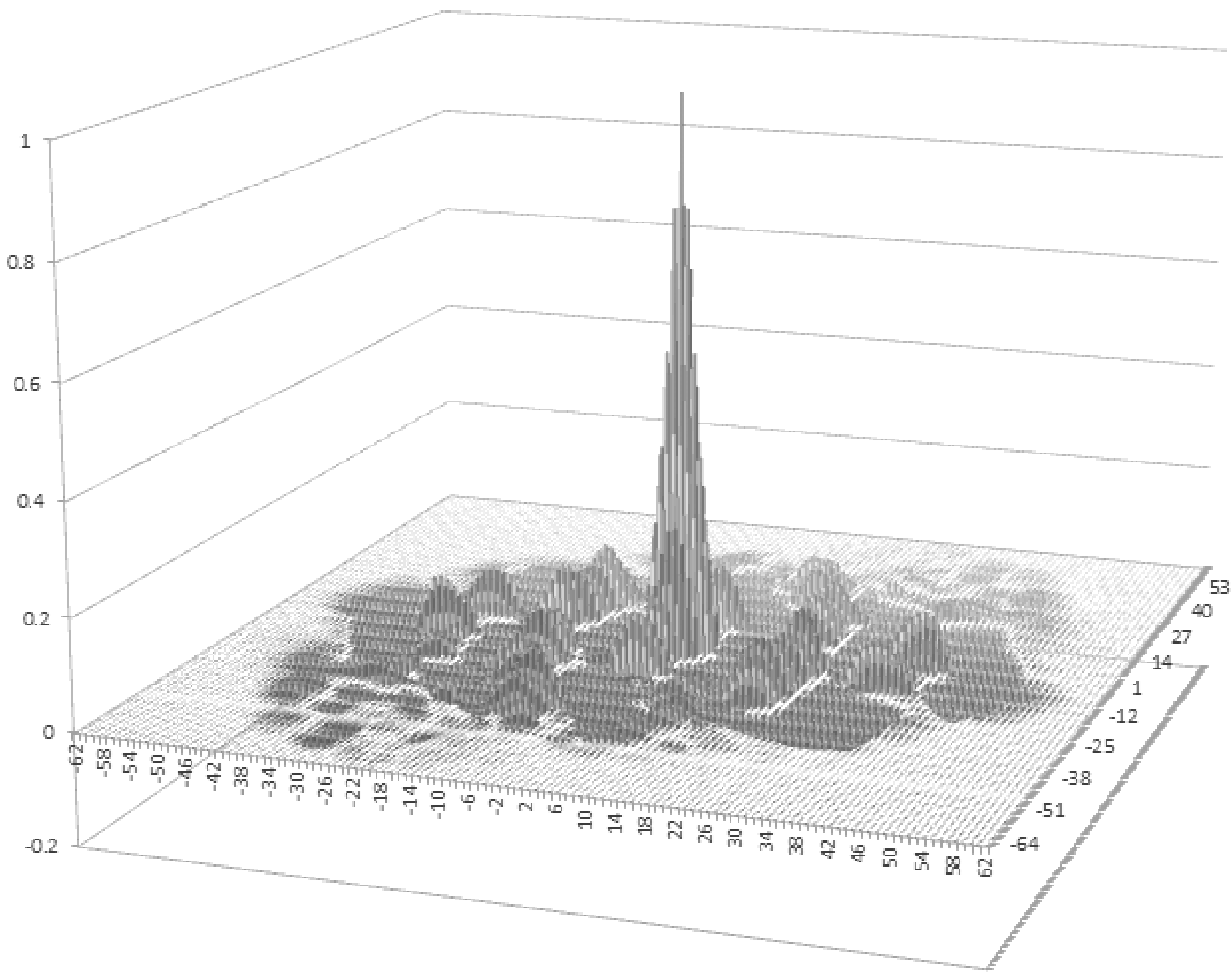


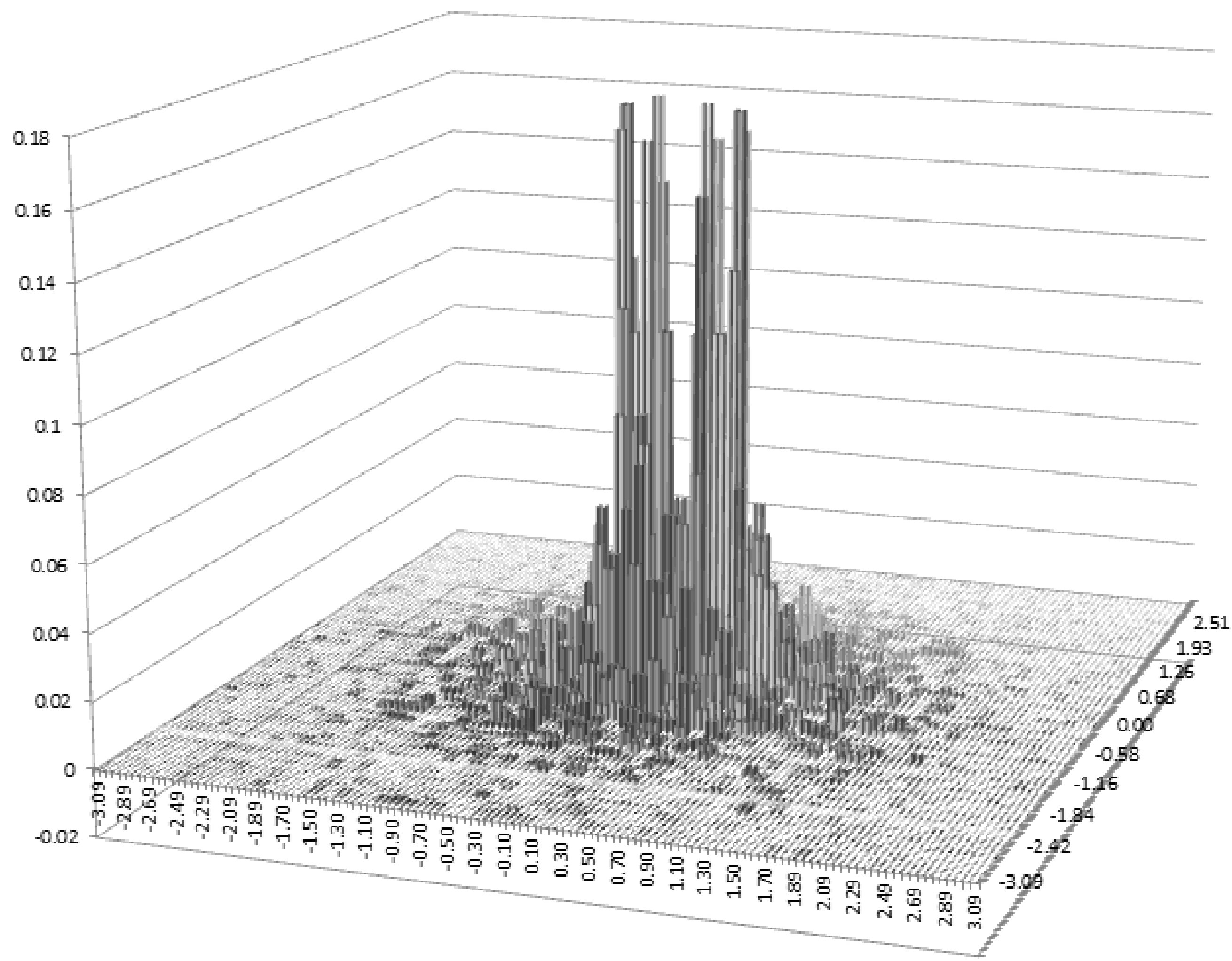
Figure 5.19: Residuals of the auxiliary regression to check a unit root in average electricity consumption of strata one: autocorrelation estimate function, Antioquia (Colombia).



We take into account the spatial effects in this estimation. Specifically, the likelihood ratio test value is 68.83 which implies that the spatial effect is relevant. Additionally, the value of Moran's tests is 0.009 which implies that there is no statistical evidence of spatial autocorrelation in the residuals of this model. Moreover, this fact is corroborated with the LMerr and LMlag tests.

In this estimation all variables have the expected sign. In particular, the price elasticity of electricity demand is significant and equal to -1.15 which implies that an increment of 1% in price means a reduction in electricity demand of 1.150%. Moreover, income elasticity is significant at 10% and

Figure 5.20: Residuals of the auxiliary regression to check a unit root in average electricity consumption of strata one: periodogram estimate function, Antioquia (Colombia).



implies that an increment of 1% in per capita income means an increment of 0.408% in electricity. On the other hand, once we take into account spatial effects, the dummy variable associated with sea level is not significant, this makes sense because this variable is related to geographical conditions, and these conditions are the core of a spatial autoregressive model. Finally, the substitute price is not significant. This result was also found in the first regression analysis that we did. A possible explanation of this fact, it is that electrical devices have few substitutes in rural areas which is the biggest area in Antioquia. Therefore a decrease in substitute prices does not imply a reduction in electricity consumption.

Table 5.7: Results of the spatial unit root tests: Electricity demand in the Department of Antioquia, Colombia.

Frequency domain test			
	$\hat{\psi}$	$F_{0.025,4094,4096}$	$F_{0.975,4094,4096}$
Lattice (5,819×5,910)	4.212	0.940	1.063
	$\hat{\psi}$	$F_{0.025,1482,1482}$	$F_{0.975,1482,1482}$
Lattice (10,000×10,000)	3.147	0.903	1.107
	$\hat{\psi}$	$F_{0.025,674,674}$	$F_{0.975,674,674}$
Lattice (15,000×15,000)	2.320	0.891	1.115
	$\widehat{LM}$	χ^2_1	
Lauridsen and Kosfeld (2006)	4.201	3.841	
	$\widehat{LM^{SIM}}$	$\widehat{LM^{SIM}_{Empirical}}$	
Lauridsen and Kosfeld (2004)	13.303	4.820	
	$\widehat{LR}$	χ^2_1	
LR	6.004	3.841	
	$\widehat{Wald}$	χ^2_1	
Wald	7.522	3.841	

Source: Author's estimations.

6 Conclusions, Recommendations and Future Research

6.1 Conclusions

Stationarity is a common assumption in applied work, especially in regional data analysis. However, this hypothesis should be checked. Therefore in this dissertation, the consequences on statistical inference when there is a non-stationary spatial autoregressive random field on a lattice are analysed. Specifically, we study the case of spatial unit root processes. This point is important because technically the collection of n indexed observations that make up a dataset in the spatial domain do not represent a sample of size n. Actually, they represent a single realization of a random experiment; a sample of size one from an n-dimensional distribution. We find that this phenomenon causes a tendency in the pattern that exhibits the process, and also that the variance of the process is not stable although dimension of the lattice increases. On the other hand, it is shown that many characteristics that present spatial unit root fields are the same as can be found in unit root processes in time domain. However, there is one difference that is very important, the Ordinary Least Square estimator of the autoregressive parameter is not consistent in the spatial domain. This fact gains relevance because many tests to check the presence of a unit root in the time domain are based on the consistency of the OLS estimator.

In order to build a spatial unit root test, we use the spectral representation of a spatial autoregressive random field. This strategy permits avoiding the use of Brownian motion in two dimensions, a concept which presents some complaints in the spatial domain due to the multilateral dependence structure in this space. Under the null hypothesis of a spatial unit root field, we propose a statistic based on the asymptotic properties of the periodogram function, and find its asymptotic distribution. Additionally, a Monte Carlo test strategy is proposed to improve the finite sample statistical properties of the test.

We perform some Monte Carlo experiments which indicate that the small sample properties of the frequency domain test are sensible. Its sample size converges to the nominal size and the power converges to one. However, if we compare these results with the ones obtained using other spatial unit root tests, we can see that these tests can over perform our test in some occasions when the data generating process is a spatial autoregressive model. This outcome is due to the parametric tests performing better under the assumptions of the parametric models. On the other hand, some simulation exercises show that our test gets the best size when the data generating process is not a spatial autoregressive model. Therefore, it is found that our test is less sensitive to model assumptions.

The Mercer-Hall data are used to apply the statistical methodology that is developed in this dissertation, and specifically, we test the null hypothesis of spatial unit root processes. We find that there is statistical evidence to reject the null hypothesis of a spatial unit root in the data. Additionally, the dataset of the demand of electricity in the Department of Antioquia is also used. Given that the geometry of these data is not a regular lattice, we transfer the data in the spatial polygons to a spatial grid in order to apply conventional statistical tools developed in regular lattice. It is found that there is statistical evidence to reject the null hypothesis of a spatial unit root in electricity consumption. This implies that the elasticities estimates are sensible, and thus, electricity demand reacts with price and income changes. Specifically, an increment of 1% in electricity price means a reduction of 1.150% in electricity consumption, and an increment of 1% in income implies an increment of 0.408% in the demand of this service. Additionally, we find that substitute electricity prices do not have an effect on electricity demand, and once we take into account spatial effects, there is no effect due to the sea level dummy variable.

Finally, we know that the concept of a spatial unit root is controversial because of many aspects (Florax and Vlist, 2003; Paelinck et al., 2004). In particular, the concept of integration associated with a unit root process and its interpretation as an accumulation of stochastic shocks, which is taken from time series analysis, sounds rare in spatial analysis using cross section data. This is due to cross section data being a picture in a moment of time. However, we can think beyond, and realise that a realisation of a spatial process in a moment of time is actually a picture of a realisation of a random field in three dimensions: latitude, longitude and time. Therefore, we can think the concept of integration in the context of spatial analysis with cross section data as an outcome of a non-stationary random field of three dimensions. Thus, the accumulation of shocks is not due to a current event, but a consequence of historical events. For instance, we can imagine a variable in some region in its initial moment, the region is fixed and time passes. Suppose that for any historical circumstance, the variable is non-stationary. Unfortunately, we do not normally have georeferenced historical data, although things are changing, so, we have a picture in a moment of time. Thus, evidence of a spatial unit root process is a consequence of a historical process, and we should not think in accumulation of shocks as an immediate propagation of these ones in a moment of time, but as evidence of non-stationarity due to historical events whose manifestation is present now.

Another controversial point about spatial unit root process is the fact that $\rho = 1$ is just a problem that is present when the contiguity matrix is standardised (Paelinck et al., 2004). Two things about this point; first of all, the contiguity matrix is standardised because this procedure gives some intuition about the spatial lag. Specifically, the spatial lag of a variable associated with a location can be seen as a weighted average of the same variable associated with its neighbours. Second, the spatial unit root process is present in other types of contiguity matrices. For example, in the case of a binary contiguity matrix, Beenstock and Felsenstein (2008) establish in a spatial autoregressive model that a spatial unit root process is generated by an autoregressive coefficient equal to the reciprocal of the number of neighbours. Therefore, the spatial unit root case is not just an algebraic problem.

Summarising, we know that there is a debate about the concept of unit roots and its application in the spatial context; however, we must keep in mind that spatial data are just a realisation of a random field. Therefore, checking stationarity is an important stage in any statistical analysis because the fulfilment of this characteristic facilitates enormously the inferential analysis. Unfortunately, there is a lot of empirical work that forgets this stage and assumes that stationarity is fulfilled, but this can lead to erroneous inferences (Schabenberger and Gotway, 2005). So, the central point in this dissertation from a philosophical perspective is to warn practitioners about the hypothesis of stationarity and its implications in applied work.

6.2 Recommendations and future research

There is a lot of future research that can be developed in order to improve the outcomes of this dissertation. For instance, it is well known that the formulation of the autocovariance that was used here is consistent but biased (Guyon, 1982). Therefore, it is a good idea to use the unbiased expression in order to improve the finite performance of the statistical tests. Additionally, Fuentes (2007) shows that $E[I(\xi_1,\xi_2)] = \frac{1}{rc(2\pi)^2}\int s(\xi_1,\xi_2)V(\theta_1-\xi_1,\theta_2-\xi_2)d\theta_1 d\theta_2$ where $V(\theta_1-\xi_1,\theta_2-\xi_2) = \left(\frac{sin^2(r\theta_1/2)}{sin^2(\theta_1/2)}\right)\left(\frac{sin^2(c\theta_2/2)}{sin^2(\theta_2/2)}\right)$ and (θ_1,θ_2) are Fourier frequencies. The periodogram bias arises due to leakage effect. Given this problem in the periodogram estimate, it is wise to use tapering or pre-whitening to reduce the bias in finite sample. Although, there is some information lost, it can be possible to use methodologies that minimize the loss. The implementation of these methodologies can improve the finite sample properties of the statistical test.

Although the mathematical basis of the statistical test that is proposed is the Fourier transform, it can be possible to explore the use of wavelets to formulate new statistical tests. This strategy has been implemented in the time series literature (Cardinali and Nason, 2007).

With regard to the simulation exercises, the statistical properties of the test were analysed in regular lattices, but in regional analysis, the datasets are normally irregular polygons. It is a good practice to perform Monte Carlo simulation exercises to analyse the finite sample properties of the statistical test in real maps where the data conserve some spatial autoregressive structures.

Unfortunately, the asymptotic distribution of the test under the alternative hypothesis that $\rho \neq 1$ is not demonstrated. This result is a necessary condition to show that the test is consistent. But this is non-trivial because, under the alternative hypothesis $\Delta\mathbf{z}$ does have a spectral representation that depends on ρ. Although that $\frac{2I(\xi_1,\xi_2)}{s(\xi_1,\xi_2)} \xrightarrow{d} \chi^2_2$ does not depend on ρ, $\sum_{\xi_1,\xi_2\neq 0} 2\frac{I(\xi_1,\xi_2)}{s(\xi_1,\xi_2)}$ does not converge to χ^2_{2m}, because now $s(\xi_1,\xi_2) \neq \sigma^2/(2\pi)^2$, and varies with (ξ_1,ξ_2). Actually, this is a sum of weighted Chi-Square variables, and we do not find the exact asymptotic distribution of this variable. However, there are approximations that might be used (Solomon and Stephens, 1977;

Castano and López, 2005).

Other ideas that can be used to build statistical tests to check stationarity on spatial random fields are based on the outcomes of Bickel and Wichura (1971) and Deo (1975). Bickel and Wichura (1971) show that the partial sum of independent random variables with zero means and finite variances in 2-dimensional time converges to a Brownian motion process on $[0, 1]^2$. The requirement for this outcome is that the random variables satisfy the Lindeberg's condition. On the other hand, Deo (1975) develops a functional central limit theorem for stationary φ-mixing random fields. These concepts play a fundamental role in the theory that is behind the statistical tests of a unit root process in time domain.

Additionally, there are some other ideas that can be taken from time series analysis to apply in the spatial context. For instance, Shitan (2008) introduces fractionally integrated separable spatial autoregressive models which are applied when there are processes with long memory in space. The effect of fractionally integrated models in space on the statistical test that is proposed can be analysed. Also, the effect of holes or structural changes in space can be studied, and their influence on the stationarity of the random field like it is done in time domain (Perron, 1989).

Actually, this dissertation is a grain in a theory that, although forgotten, can be fruitful due to the wide field in which it can be developed. There are huge opportunities to contribute to the theory.

Bibliography

Adler, D. and Murdoch, D. (2011). *rgl: 3D visualization device system (OpenGL). R package 0.92.798*.

Adler, R. and Taylor, J. (2007). *Random fields and geometry*. Springer, United States, first edition.

Anselin, L. (1982). A note of small sample properties of estimators in a first order spatial autoregressive model. *Environment and Planning*, 14:1023–1030.

Anselin, L. (1988). *Spatial econometrics: methods and models*. Kluwer academic publishers, United States, first edition.

Asmar, A. (1995). *Tópicos en teoria de matrices*. Universidad Nacional de Colombia, sede Medellín, Colombia, primera edition.

Baltagi, B. and Liu, L. (2010). Spurious spatial regression with equal weights. *Statistics and Probability Letters*, 80:1640–1642.

Baran, S. and Pap, G. (2011). Parameter estimation in a spatial unit root autoregressive model. Technical report, Faculty of Informatics, University of Debrecen.

Baran, S., Pap, G., and Zuijlen, M. (2004). Asymptotic inference for an unstable spatial AR model. *Statistics*, 38:465–482.

Beenstock, M. and Felsenstein, D. (2008). Testing spatial stationarity and spatial cointegration. Technical report, Hebrew University of Jerusalem. Working paper.

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

Besag, J. and Kooperberg, C. (1995). On conditional and intrinsic autoregressions. *Biometrika*, 82(4):733–746.

Bhattacharyya, B., Li, X., Pensky, M., and Richardson, G. (2000). Testing for unit roots in a nearly nonstationary spatial autoregressive process. *Ann. Inst. Statist. Math.*, 52(1):71–83.

Bhattacharyya, B., Richardson, G., and Franklin, L. (1997). Asymptotic inference for near unit roots in spatial autoregression. *The Annals of Statistics*, 25(4):1709–1724.

Bickel, J. and Wichura, M. (1971). Convergence criteria for multiparameter stochastic process and some applications. *The Annals of Mathematical Statistics*, 42(5):1656–1670.

Billingsley, P. (1995). *Probability and measure.* Jhon Wiley & Sons, United States, third edition.

Bivand, R. (2011). *spdep: Spatial dependence: weighting schemes, statistics and models. R package 0.5-31.*

Bivand, R., Pebesma, E., and Gomez-Rubio, V. (2008). *Applied spatial data analysis with R.* Springer, United States, first edition.

Bose, S. and Steinhardt, A. (1996). Invariant tests for spatial stationarity using covariance structure. *IEEE Transactions on Signal Processing*, 4(6):1523–1533.

Cardinali, A. and Nason, G. (2007). Costationarity and stationarity tests for stock index returns. Technical report, University of Bristol.

Castano, A. and López, F. (2005). Distribution of a sum of weighted non-central chi-square variables. *Test*, 14(2):397–415.

Cressie, N. (1993). *Statistics for spatial data.* Jhon Willey & Sons, INC, United States, revised edition.

Deo, C. (1975). A functional central limit theorem for stationary random fields. *The Annals of Probability*, 3(4):708–715.

Dickey, D. and Fuller, W. (1979). Distribution of the estimates for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74:427–431.

Enders, W. (1995). *Applied Econometric Time Series.* John Wiley & Sons inc.

Ephraty, A., Tabrikian, J., and Messer, H. (2001). Underwater source detection using a spatial stationarity test. *Journal of Acoustical Society of America*, 109(3):1053–1063.

Fingleton, B. (1999). Spurious spatial regression: some monte carlo results with a spatial unit root and spatial cointegration. *Journal of Regional Science*, 39:1–19.

Finley, A. and Banerjee, S. (2010). *MBA: Multilevel B-spline Approximation. R package 0.0-7.*

Florax, R. and Vlist, V. (2003). Spatial econometric data analysis: Moving beyond traditional models. *International Regional Science Review*, 26(3):223–243.

Fuentes, M. (2005). A formal test for nonstationarity of spatial stochastic processes. *Journal of Multivariate Analysis*, 96:30–54.

Fuentes, M. (2007). Approximate likelihood for large irregularly spaced spatial data. *Journal of the American Statistical Association*, 102(477):321–331.

Furrer, R., Nychka, D., and Sain, S. (2010). *Fields: Tools for spatial data. R package 6.3.*

Greene, W. (2003). *Econometric analysis.* PrenticeHall, United States, fifth edition.

Guan, Y. (2008). A kpss test for stationarity for spatial point process. *Biometrics*, 64(3):800–806.

Guyon, X. (1982). Parameter estimation for a stationary process on a d-dimensional lattice. *Biometrika*, 69(1):95–105.

Haining, R. (1990). *Spatial data analysis in the social and environmental sciences*. Cambridge University Press, United of Kingdom, first edition.

Hamilton, J. (1994). *Time Series Analysis*. Princeton University Press, first edition.

Ivanoff, G. (1982). Central limit theorems for point process. *Stochastic Process and their Applications*, 12:171–186.

Kwiatkowski, D., Phillips, P., Schmidt, P., and Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics*, 54:159–178.

Lauridsen, J. and Kosfeld, R. (2004). A wald test for spatial nonstationarity. *Estudios de Economía Aplicada*, 22(3):475–486.

Lauridsen, J. and Kosfeld, R. (2006). A test strategy for spurious spatial regression, spatial nonstationarity, and spatial cointegration. *Papers in Regional Science*, 85(3):363–377.

Lauridsen, J. and Kosfeld, R. (2007). Spatial cointegration and heteroscedasticity. *Journal of Geographical Systems*, 9(3):253–265.

Lee, L. and Yu, J. (2007). Near unit root in the spatial autoregressive model. Technical report, Department of Economics, The Ohio State University.

Lee, L. and Yu, J. (2009). Spatial non-stationarity and spurious regression: the case with a row-normalized spatial weights matrix. *Spatial Economic Analysis*, 4(3).

MacNeill, I. (1978a). Limit process for sequences of partial sums of regression residuals. *The Annals of Probability*, 6(4):695–698.

MacNeill, I. (1978b). Properties of sequences of partial sums of polynomial regression residuals with applications to test for changes of regression at unknown times. *The Annals of Probability*, 6(2):422–433.

Magnus, J. (1978). Maximum likelihood estimation of the gls model with unknown parameters in the disturbance covariance matrix. *Journal of Econometrics*, 7(3):281–312.

Martellosio, F. (2010). Eficiency of the ols estimator in the vicinity of a spatial unit root. Technical report, University of Reading.

Mateu, J. and Juan, P. (2004). A spectral test of nonstationarity for spatial processes. *Quantitative Geology and Geostatistics*, 13(5):343–354.

McBratney, A. and Webster, R. (1981). Detection of ridge and furrow pattern by spectral analysis of crop yield. *International Statistical Review*, 49:45–52.

Mercer, W. and Hall, A. (1911). The experimental error of field trials. *The Journal of Agricultural Science*, 4:107–132.

Moran, P. (1950). Notes on continuous stochastic phenomena. *Biometrika*, 37:17–33.

Mur, J. and Trívez, J. (2003). Unit root and deterministic trends in spatial econometric models. *International Regional Science Review*, 26:289–312.

Ord, K. (1975). Estimation methods for models of spatial interaction. *Journal of the American Statistical Association*, 70(349):120–126.

Paelinck, J., Mur, J., and Trívez, J. (2004). Econometría espacial: Más luces que sombras. *Estudios de Economía Aplicada*, 22(003):383–404.

Paulauskas, V. (2007). On unit root for spatial autoregressive models. *Journal of Multivariate Analysis*, 98:209–226.

Perron, P. (1989). The great crash, the oil price shock, and the unit root hypothesis. *Econometrica*, 57:1361–1401.

Phillips, P. and Perron, P. (1988). Testing for a unit root in time series regression. *Biometrika*, 75:335–346.

Priestley, M. and Rao, S. (1969). A test for non-stationarity of time-series. *Journal of the Royal Statistical Society*, 31(1):140–149.

R Development Core Team (2011). *R: A Language and Environment for Statistical Computing. R. Version 2.13.0.* R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.

Renshaw, E. (2002). Two-dimensional spectral analysis for marked point processes. *Biometrical Journal*, 44(6):718–745.

Ripley, B. (2004). *Spatial statistics.* Jhon Wiley & Sons, United States, second edition.

Roknossadati, S. and Zarepour, M. (2011). M-estimation for near unit roots in spatial autoregression with infinite variance. *Statistics*, 45(4):337–348.

Rossiter, D. (2010). Tutorial: Using the r environment for statistical computing. an example with the mercer-hall wheat yield dataset. Technical report, University of Twente, Faculty of Geo-Information Science and Earth.

Schabenberger, O. and Gotway, C. (2005). *Spatial methods for spatial data analysis.* Chapman & Hall CRC, United States, first edition.

Shitan, M. (2008). Fractionally integrated separable spatial autoregressive (fissar) model and some of its properties. *Communications in Statistics: Theory and Methods*, 37:1266–1273.

Solomon, H. and Stephens, M. (1977). Distribution of a sum of weighted chi-square variable. *Journal of the American Statistical Association*, 72(360):881–885.

Tobler, W. (1970). A computer movie simulating urban growth in the detroit region. *Economic Geography*, 46:234–240.

Venables, B. and Ripley, B. (2011). *MASS: Support Functions and Datasets for Venables and Ripley's MASS. R package 7.3-13*.

Whittle, P. (1954). On stationary processes in the plane. *Biometrika*, 41:434–449.