A spatial temporal analysis of the relation between omicron variant's incidence and foreigners' mobility in Portugal

Maria Emília Alves Matos Folque

Thesis to obtain the Master of Science Degree in

Biomedical Engineering

Supervisor(s): Professor Leonardo Azevedo Guerra Raposo Pereira
Doutor André Peralta Santos

Examination Committee

Chairperson: Professor Rita Homem de Gouveia Costanzo Nunes
Supervisor: Doutor André Peralta Santos
Member of the Committee: Professor Manuel Luís Castro Ribeiro

June 2023

Dedicated to my mom,
who always believed in me.

Declaration

I declare that this document is an original work of my own authorship and that it fulfills all the requirements of the Code of Conduct and Good Practices of the Universidade de Lisboa.

Acknowledgments

First, I would like to thank all the teachers I have had over the years, especially my co-supervisors, Prof. Leonardo Azevedo, and Doutor André Peralta Santos, for their support and guidance throughout this dissertation.

I would also like to express my gratitude to the Direção-Geral da Saúde and SIBS for providing the data needed for the development of this work.

Last, but not least, I would like to thank all my family and friends who have helped me every step of the way. It has been a long and arduous journey but I wouldn't have had it any other way.

Resumo

Em situações de emergência, as autoridades necessitam de métodos rapidamente adaptáveis que possibilitem melhorar a sua capacidade de deteção e de resposta a crises. Adicionalmente, as decisões tomadas pelas autoridades necessitam de ser monitorizadas, de forma a garantir que as medidas implementadas funcionam conforme planeado. Este trabalho pretende avaliar a eficácia da proibição de viagens imposta por Portugal à África do Sul entre 29 de novembro e 13 de dezembro de 2021 na redução da disseminação da variante omicron, através de análises de correlação cruzada espacial, utilizando rastreamento de transações com cartões eletrónicos. Apesar da queda no número de transações verificada durante o período em que as restrições de viagens foram impostas, os resultados deste trabalho sugerem que a medida foi ineficaz em diminuir ou retardar a disseminação da variante omicron. Neste trabalho, nenhuma das análises de correlação cruzada espacial forneceu qualquer correlação estatisticamente significativa entre os movimentos de transações com cartão dos Sul-Africanos em Portugal e a disseminação da variante no país. Contudo, são necessários estudos mais aprofundados para compreender se existe a possibilidade de as transações com cartões eletrónicos serem consideradas um bom proxy de previsão da proliferação em futuras pandemias.

Palavras-chave: Correlação cruzada espacial, Incidência omicron, Transações com cartões eletrónicos, Proibição de viagens

Abstract

In emergency situations, decision-makers need quickly adaptable methods that enhance their ability to detect and respond to crises. These actions also need to be monitored to make sure the measures implemented are working as planned. This work attempts to assess the effectiveness of the travel ban imposed by Portugal on southern African countries between November 29^{th} and December 13^{th}, 2021. A spatial cross-correlation analysis was utilized to understand if South Africans' mobility, tracked using card transactions, had any relationship with the spread of the omicron variant. Despite the fall in the number of transactions during the period when travel restrictions were imposed, the measure was ineffective in decreasing or delaying the spread of omicron. Neither spatial cross-correlation study provided any statistically significant correlation between South Africans' card transactions movements in Portugal and the spread of the variant in the country. However, more in-depth research needs to be conducted to fully understand if there is a way card transactions could ever be considered a good proxy to predict the proliferation of future pandemics.

Keywords: Spatial cross-correlation, Omicron incidence, Card transactions, Travel ban

Contents

List of Tables

List of Figures

Nomenclature

COVID-19 Coronavirus Disease-2019

CSV Comma-Separated Values

DGS Direção Geral da Saúde

EWDS Early Warning Detection Systems

INSA Instituto Nacional de Saúde Dr. Ricardo Jorge

ITS Interrupted time series

PSCC Partial Spatial Cross-correlation Coefficient

SCC Simple Correlation Coefficient

SCI Spatial Cross-correlation Index

SIBS Sociedade Interbancária de Serviços

TXT Text File

VOC Variants of Concern

WHO World Health Organization

Chapter 1

Introduction

1.1 Motivation

On December 31st, 2019, Wuhan Municipal Health Commission reported the first cluster of cases of what would eventually be identified as a novel coronavirus. In the following day, the World Health Organization (WHO) set up the Incident Management Support Team, readying itself to deal with a possible outbreak. The first reported case outside China was confirmed on January 13th, 2020 in Thailand. The day after the virus's genetic sequence of this disease was shared and two days after the first confirmed death [1]. It reached Europe on January 24st, 2020, when three infections were detected in France [2]. The first cases in Portugal appeared on March 2nd, 2020 [3]. By that day, over 94.000 cases had been registered worldwide [4].

WHO stated that respiratory droplets were the most common method of transmission of coronaviruses [5]. Therefore the new virus, now known as COVID-19 (Coronavirus Disease-2019), would be highly transmitted in crowded environments, especially ones with poor ventilation.

The presence of asymptomatic cases was a pivotal contributor to the spread of the virus. Since people might not show any symptoms, and still infect others [6]. Making it difficult to implement effective measures to contain the virus propagation.

All these factors lead to the rapid emergence and spread of COVID-19, and a pandemic was declared by WHO on March 11th, 2020. Governments worldwide were then forced to take various preventive measures.

In response to the situation, quarantines were implemented, mainly on the infected, alongside restrictions on the movement of people and a wide mandate on the use of masks [7][8]. In Portugal, various mobility constraints were also implemented. Some of these included the closure of schools and businesses, limitations of land, sea, and air border activities, and mandatory lockdowns. The main objective of all measures was to delay the spread of the virus by reducing the likelihood of infections. These measures were deemed necessary to try to prevent the collapse of the national health service whilst vaccines and possible treatments were being developed.

As is common in viruses, COVID-19 changed over time creating variants. Some of these changes altered the virus's properties, affecting its transmissibility, the severity of the resulting illness, the effectiveness of treatments, and other public health and social measures [9]. In late 2020, certain variants posed a heightened threat to public health which prompted WHO to classify some as Variants of Concern (VOC). This was done to prioritize monitoring and research efforts and to better adjust the response to COVID-19. As of the beginning of 2022, the most significant variants were: alpha, beta, gamma, delta, and omicron [10].

In the backend of 2021, the omicron variant was first discovered in South Africa and it rapidly spread worldwide, reaching Portugal on November 17th [11]. This variant quickly rose to prominence, becoming the dominant variant in Portugal by the second to last week of the year (from December 20th to 26th, 2021) [12].

On November 17th, the day before the first recorded omicron case in Portugal, the total number of new confirmed cases was 1423. By December 26th, 2021 that figure had risen to 7689 and it would continue to rise until the end of the following month [13]. Figures 1.1 and 1.2 further illustrate the evolution of the spread of the virus and its variants in Portugal.

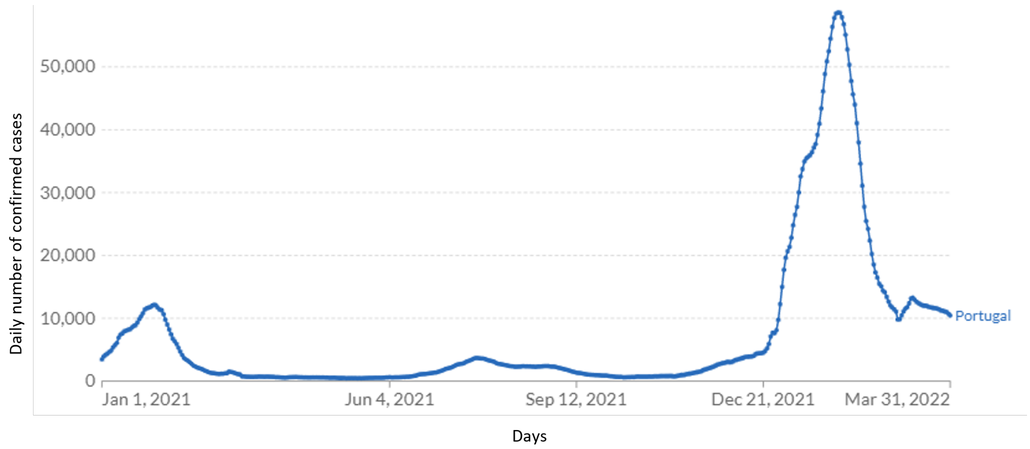


Figure 1.1: Daily new confirmed COVID-19 cases in Portugal. Note: Adapted May 10, 2023 from Our World in Data (n.d.), Daily new confirmed COVID-19 cases.

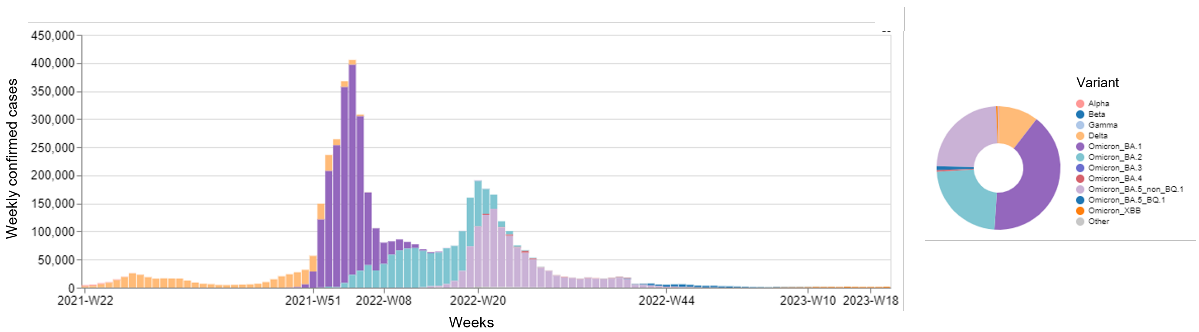


Figure 1.2: Weekly COVID-19 cases divided by the corresponding variant. Note: Adapted May 10, 2023 from INSaFLU (2023), Número de casos semanais de SARS-CoV-2 divididos pelas frequências projetadas das variantes.

To try to mitigate this rapid spread, governments all around the world based their decisions on previous studies, emphasizing the current reliance on data to create the best possible course of action [8]. Ergo, it's important to assess the effectiveness of measures implemented, as well as explore alternative approaches that can successfully protect the population while minimizing negative impacts on society. By utilizing COVID-19 data, researchers can employ statistical methods in order to better understand the spread of the virus. Thus, it's crucial to analyze the available data, examine how it might relate to other variables, and enhance our readiness for future pandemics.

After a literature review was conducted, it became evident that travel restrictions, alongside quarantine protocols, and social distancing guidelines, were one of the most common measures implemented worldwide during the pandemic. In addition, the research also showed that spatial correlation models were widely used and were deemed quite useful for understanding the spread of the virus.

Furthermore, early detection of COVID-19 VOCs was critical for adapting non-pharmacological interventions aiming to mitigate the virus's impacts on public health. Timely identification of VOCs allows for prompt adjustment of preventative measures to better target the specific threats posed by each variant. Additionally, Early Warning Detection Systems (EWDS) enable healthcare organizations to prepare for potential surges in cases and allow researchers to study the efficacy of current vaccines and treatments against these new strains. As such, effective surveillance and monitoring of VOCs were essential to maintaining an agile and effective response to the ongoing pandemic.

Incorporating new and unconventional data in EWDS can significantly enhance their ability to detect and respond to the emergence of VOCs. One such data source is electronic payment systems, which can provide valuable information about the country of origin of international transactions. By analyzing patterns in electronic payments, it might be possible to identify potential correlations between the introduction of VOCs and increased economic activity from individuals originating from affected regions. This information can serve as an early warning signal for public health officials, allowing for the implementation of targeted interventions and mitigating the risks associated with the introduction of new VOCs.

1.2 Objectives

For the reasons portrayed above, a deeper spatial-temporal analysis of the evolution of COVID-19 in Portugal needed to be conducted. For it, the time period concerning the rapid emergence of the omicron variant was chosen. During that span, on November 29th, Portugal decided to ban air traffic coming from South Africa, Botswana, Eswatini, Lesotho, Mozambique, Namibia, and Zimbabwe. Since only Mozambique has direct flights to Portugal, testing and quarantine were made mandatory for any passenger arriving from the southern African nations [14].

As such, this dissertation aims to answer two questions:

- Was the air traffic ban imposed by Portugal on South Africa effective in preventing the spread of omicron?

- Can South Africans' tracked movements through card transactions be used to predict the spatiotemporal spread of omicron?

This study focuses on South African citizens since, as previously mentioned, that is where the first case of omicron was discovered. One of the deciding factors in choosing to utilize foreigners' movements as one of the bases of our correlation was a study performed in February 2021, which successfully described the association between transmission and mobility [15]. An interrupted time series (ITS) and spatial cross-correlation analysis were the methods chosen in pursuance of the answers to the two questions posed above.

1.3 Literature Review

The COVID-19 pandemic prompted unprecedented governmental action around the world as it became imperative to contain the spread of the virus. To do so, several policies such as the usage of masks, travel and overall mobility restrictions, and full quarantines were enacted to try to keep people safe [16].

Numerous studies tried to determine to which extent international travel bans actually impacted the initial spread of the virus. It was estimated these measures led to a reduction of cases exported globally between 70% and 77% [17][18]. As it pertains to the European Union, both its internal and external border restrictions contributed notably to a decrease in the speed at which the virus spread across member states [19].

However, some research shows that in order for travel bans to be justifiable for a nation they need to be strictly targeted. Meaning that some unfocused restrictions have little to no impact on the overall evolution of the pandemic. Unless the country in question has a low COVID-19 incidence and a large number of infected travelers arriving from other countries, these measures are ineffective in stopping the spread [20].

A thoroughly focused travel ban might also not be enough given the highly interconnected world we live in nowadays. With the lofty number of indirect connections in the global flight network, attempts to solely postpone the introduction of COVID-19 would be of limited benefit, unless the objective is only to slow down the transmission by a few days, or if the delay is coupled with additional measures taken domestically [21].

On account of this, border restrictions were found to be more effective when coupled with mandatory quarantine and screening [22]. With regard to the advantages of screening, it was found that only particularly effective processes actually help reduce the risk of COVID-19 importation and exportation. In certain countries, a 90% effective screening of asymptomatic travelers could delay the average time of the epidemic by up to 20 days [23].

After the initial set of proposed measures, authorities required mechanisms to monitor the ever-changing landscape of the pandemic in order to adapt and retool their strategy. To do so, several EWDS were utilized to detect early-stage changes in the territorial spread of infections [24]. EWDS are integrated systems of threat monitoring, forecasting, and risk assessment processes that enable decision-makers to take timely and measured actions to reduce disaster likelihood in advance of hazardous events [25].

A study suggests that the ability to aggregate large datasets from different sources with the tracing of confirmed cases and the prediction of the local dynamics of contagion through early indicators can be an effective way to combat future pandemics [24]. This rapid interpretation of evolving data can be imperative in the delineation of the best course of action to prevent and contain infection outbreaks. Furthermore, more timely measures may diminish the need for prolonged restrictions that carry more drastic socio and economic implications.

The EWDS may be coupled with the monitoring of other relevant datasets. Successful research was done using data from Google searches [26] and even the social media platform, Twitter [27], to uncover early-warning signals of COVID-19 outbreaks. Other non-conventional datasets such as wastewater measurements and animal surveillance, were also deemed successful in the early detection of VOCs [28][29].

Understanding the spatiotemporal intricacies of COVID-19 became critical in the effort to continue with a flexible strategy to mitigate its spread. This grasp on geographical data served to illuminate the magnitude and consequences of the pandemic, facilitating informed decision-making and strategic planning [30].

A study in China ascertained the distribution of COVID-19 cases and their correlation with the migration of Wuhan's population in the early phases of the pandemic by using a spatial-temporal model. Such discovery can be critical for early warning and prevention of future outbreaks [31]. Characteristics of temporal and spatial distribution from Wenzhou were also analyzed. The findings revealed the main reason for the outbreak of COVID-19 in the city, the high number of people who flooded in from Hubei Province, where Wuhan is located [32].

Since COVID-19 quickly became an international affair, researchers outside China also began conducting their own investigations. Geographic analyses that mapped information to all administrative levels, including counties, were carried out worldwide. The data was used to create dashboards [33], monitor epidemiology trends [34], help find hotspots and outbreaks [35][36], and overall better track the evolution of the pandemic. It was also crucial for resource management as it offered decision-makers the necessary aid to balance the supply and demand of limited materials.

A study used a geographical information system-based methodology to examine the relationship between the reported incidence of COVID-19 and spatial patterns analyzed in Iraqi provinces. In it, the local Moran's I was applied to measure the spatial distribution of coronavirus and examine if there were clusters. This spatial analysis of the distribution of the virus incidence can assist in detecting exposure and help mitigate the impact [37].

Plenty of research using Moran´s I was done to try to map the spatial distribution of COVID-19 cases and the pandemic spread rate [38][39]. In Brazil, Moran's I was also utilized in an attempt to analyze the spatial correlation between confirmed cases and the intensive care unit beds exclusive to the disease in the municipalities of Paraná. This study allowed the finding of priority areas of care in the state regarding the dissemination and control of the disease [40].

Additionally, space-time statistics are helpful in the development of more targeted measures. For instance, in Melbourne, a spatial analysis was carried out to determine areas of priority that have a substantial number and proportion of elderly individuals aged 65 or above with disabilities, as well as significant obstacles in accessing primary healthcare services [41]. With these types of examinations, authorities can focus on more at-risk communities and politicians can implement specific measures that help alleviate their circumstances.

Furthermore, research was also conducted to determine the effectiveness of implemented measures such as lockdowns. A study in Spain determined that the lockdown effectively prevented the spread of the virus both within and between provinces, reducing possible cases by nearly 83%. Nevertheless, if the lockdown had been brought forward a week, it would have resulted in a further reduction of 11.6% in the number of COVID-19 infections-[42]. This means that for the measure to be even more fruitful, it should be implemented at the earliest possible stage.

Statistical models not only grant decision-makers with accurate existing data but also allow for predictions and forecasts to be made. It is proven that the use of spatiotemporal models significantly enhances the predictions of the number of people infected by the virus thus helping create better health policy proposals [43].

In 2022, a study accurately predicted daily COVID-19 cases by employing aggregate mobility statistics collected from Google's Community Mobility Report [44]. The year prior, researchers had found evidence that movement patterns are correlated with the level of transmission [15]. This makes mobility a good investigation proxy for infection propagation and opens the door to numerous research forecasting possibilities.

In the last decade, card transactions have been used as a proxy for human mobility. Research conducted suggests the is substantial value in using datasets of bank card transactions for studies of various aspects of domestic and foreign people's behavior inside the country. It also concludes such data has the potential to be used to support policy decisions [45].

A Spanish study highlighted the need to find alternative mobility proxies, especially in countries with scarcely available users' mobile phone data. The research, which used card transaction data, gathered information from payments for essential goods and transportation means, used to commute to and from work, to track the impact the crisis had in the country [46].

1.4 Thesis Outline

The remainder of this document is organized as follows:

- Chapter 2 describes the datasets used to develop the work.
- Chapter 3 details the methodology implemented.
- Chapter 4 presents the obtained results, discusses their validity, and outlines some of the limitations faced during its development.

- Finally, chapter 5 highlights the main conclusions from this research.

Chapter 2

Dataset Description

The materials, which served as the fundamental groundwork of the entire research, can be divided into two categories:

i. COVID-19 Daily Confirmed Cases Data

ii. Electronic Payments Data

In order to proceed with the work, some pre-processing of the data provided was required. The pre-processing methods as well as detailed information about these datasets, including their source and structure, can be found below. This information is crucial to understand the entire process of developing this thesis.

2.1 COVID-19 Daily Confirmed Cases Data

This work focuses on COVID-19 infection data, specifically the daily confirmed cases of COVID-19 in each of the 278 municipalities in Continental Portugal. The data spans between November 15th, 2021 and December 26th, 2021, the period when the omicron variant entered the country and gained prominence. In that time frame two variants, delta, and omicron, were active in Portugal. Within the dataset, there are no individual identifiers, meaning all cases remained anonymous.

The data was provided by the Direção Geral da Saúde (DGS) in a CSV (Comma-Separated Values) file. In the dataset, every row represented the daily number of reported cases of each variant in each municipality, as portrayed in the table below. Table 2.1, q(d,m,v) illustrates the total number of infections that occurred for every single specific day (d), municipality (m), and variant (v).

Date	Municipality	Variant	Number of Cases
d	m	v	q(d,m,v)

Table 2.1: Data structure of infections data provided by DGS.

During pre-processing, all data concerning the delta variant was deleted as the focus of the study was solely on omicron. The daily cases were then aggregated into a 7-day cumulative. This was done to enable a comparison with the data from card transactions. Lastly, the weekly incidence rate (i(w,m)) per 100,000 inhabitants was calculated for each municipality.

The number of inhabitants for each municipality was taken from a public dataset provided by the Portuguese national statistics institute, the Instituto Nacional de Estatística, with data from the 2021 Census [47].

Week	Municipality	Rate
w	m	i(w,m)

Table 2.2: Data structure of incidence data used for analysis.

For the purpose of this study, it is important to note these data might not be thoroughly accurate in depicting the total amount of South Africans' COVID-19-positive cases. Since there is a higher testing hesitancy in foreigners due to a disconnection from healthcare and social services when compared to local residents. This imbalance in the target population can lead to ascertainment bias and misguiding results.

The Instituto Nacional de Saúde Dr. Ricardo Jorge (INSA) was crucial in obtaining this dataset. INSA is a state laboratory that aims to better public health through research and technological development [48]. The institute had a vital role in epidemiological monitoring and was tasked to track the continuous genetic diversity of COVID-19. For the duration of the study, INSA was conducting an average of 538 genomic sequences a week. Samples were collected randomly for laboratories in all 18 Portuguese districts [49].

At the time, the institute had implemented several different methods to identify the omicron variant:

- Random weekly sample of national scope for whole-genome sequencing;
- Verification on samples associated with suspected cases through a direct search for mutations (using PCR tests);
- Real-time monitoring of S gene target failure with help from laboratories using TaqPath - ThermoFisher tests.

The last measurement was the main one used to infer the frequency and geo-temporal dispersion of the variant [49].

2.2 Electronic Payments Data

The data for card transactions was provided in a TXT (Text File) format by SIBS (Sociedade Interbancária de Serviços), a financial services company that manages the integrated banking network in Portugal, including ATM and POS terminals. SIBS Analytics possesses aggregate data on all payments with bank cards in Portugal, including with national and foreign bank-issued cards. Therefore, SIBS was able to provide all weekly payments done by SIBS card owners from August 2^{nd}, 2021, to February 6^{th}, 2022. It is important to reiterate, that just like the COVID-19 Data, there are no individual identifiers, and all transfers are anonymous.

The dataset contained numerous information regarding card payments. For the purpose of this research, only details regarding the week of the transaction, the municipality where it occurred, and the country the card was registered were taken into account. The research only focuses on data from omicrons' origin country, South Africa, so all other information was deleted.

It is important to note that every data point represents a transaction by a card. Meaning the same card could be responsible for several transactions on the same day, week, or month. So what this dataset portrays is the number of transactions made by South African cards and not the number of South African people that made card transactions.

This dataset was used in two different processes during this study, interrupted time series and spatial cross-correlation.

For the ITS methodology, the entire span of the dataset was considered and the municipalities were disregarded. Meaning every single South African card transfer that occurred in the same week in Portugal, t(w), was grouped together.

Week	Number of Transfers
w	t(w)

Table 2.3: Data structure of card transaction data used for interrupted time series analysis.

Regarding the spatial cross-correlation analysis, it is important to note that since it can only be executed using also COVID-19 data, only transfers made between November 15^{th}, 2021 and December 26^{th}, 2021 were taken into account.

To conduct this methodology, the weekly number of transfers of interest for each municipality was required. To do so, South African transactions made in the same municipality were aggregated. This process culminated in a table where each line represented all transfers that occurred in a specific week (w) at a certain municipality (m).

Week	Municipality	Number of Transfers
w	m	t(w,m)

Table 2.4: Intermediate data structure of card transaction data used for spatial cross-correlation analysis.

Lastly, that number of transfers was transformed into a rate(i(w,m)) per 100,000 inhabitants. The number of inhabitants for each municipality was taken from a public dataset provided by the Portuguese national statistics institute, the Instituto Nacional de Estatística, with data from the 2021 Census [47].

Week	Municipality	Rate
w	m	i(w,m)

Table 2.5: Data structure of card transaction data used for spatial cross-correlation analysis.

Chapter 3

Methodology

A detailed description of the methodology utilized in this dissertation can be found in the sub-chapters below.

3.1 Interrupted time series

One of the goals of this dissertation is to determine the success of the air traffic ban imposed by Portugal on South Africa from November 29th to December 13th, 2021 [50]. To do so, an interrupted time series analysis was conducted. This process was chosen since it has been proven valuable in evaluating the efficacy of health interventions implemented at a clearly defined point in time [51].

In order to conduct this methodology three variables are required:

i. T: the time elapsed since the start of the study, in the unit representing the frequency with which observations are taken (in this case, weeks);

ii. X_t; a dummy variable indicating the pre or post-intervention period;

iii. Y_t: the outcome at time t.

For this research, T is expressed in weeks, as it must represent the same time unit as the observations were taken. X_t value is 1 for the two weeks the travel ban was in effect and is 0 for the weeks before and after that.

The following regression model was used:

$$Y_t = \beta_0 + \ \beta_1 T \ + \beta_2 X_t + \beta_3 T X_t \,. \tag{3.1}$$

Where β_0 constitutes the baseline level at the beginning of the study (T=0), β_1 is the change in outcome allied to the passage of time, β_2 represents the level change succeeding the travel ban, and β_3, depicts the slope change following the restrictions.

Week	Time Elapsed (T)	Travel Ban (X_t)	Number of Transfer (Y_t)
		...	
08/11/2021	15	0	4343
15/11/2021	16	0	4519
22/11/2021	17	0	4291
29/11/2021	18	1	2525
06/12/2021	19	1	2228
13/12/2021	20	0	3766
20/12/2021	21	0	4357
27/12/2021	22	0	4519
		...	

Table 3.1: Excerpt from the dataset used for interrupted time series analysis.

3.2 Spatial Cross-Correlation

This dissertation attempts to ascertain if South Africans' tracked movements through card transactions could be used to predict the spread of omicron in Portugal. To do so, a correlation methodology was needed.

A paper by Yanguang Chen, that proposes a novel set of models and analytical procedures for spatial cross-correlation analysis, was chosen. This new theoretical framework is derived for geographical cross-correlation modeling and shows a rethinking of Moran's index. Firstly, some spatial cross-correlation coefficients are defined, then, a pair of scatterplots of spatial cross-correlation is used to visually reveal the causality behind spatial systems [52].

Figure 3.1 illustrates the paths to the 2-Dimensional spatial cross-correlation, starting from simple cross-correlation and by way of autocorrelation [52].

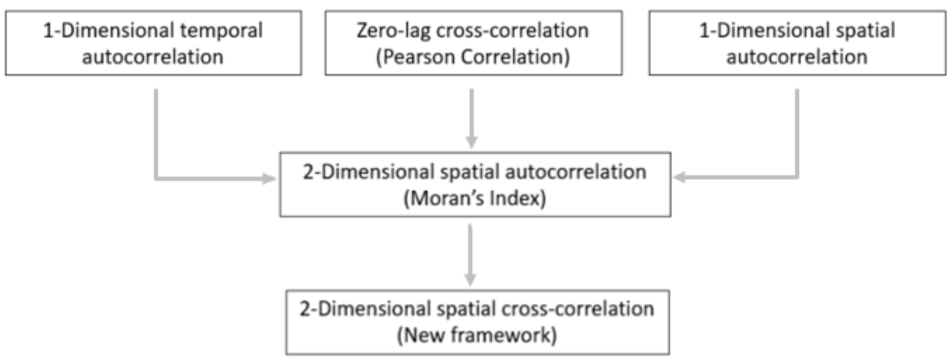


Figure 3.1: The path to the 2-Dimensional spatial cross-correlation.

For this work, the two variables measured were the weekly incidence of omicron and the weekly number of South African card transactions. Throughout this section, they will be referred to as a pair of vectors, **X** and **Y** respectively.

The centralized variables can be described by

$$X_C = X - \mu_x \,, \tag{3.2}$$

$$Y_C = Y \ - \ \mu_y \ \ , \tag{3.3}$$

where μ_x and μ_y represent the average values of the variables x_i and y_i, that can also be expressed as

$$\mu_x = \frac{1}{n}\sum_{i=1}^{n} x_i \,, \tag{3.4}$$

$$\mu_y = \frac{1}{n}\sum_{i=1}^{n} y_i \,, \tag{3.5}$$

in which n is the total number of elements in a system. Meaning that in this study, n is the number of municipalities.

The variables' variances are calculated by

$${\sigma_x}^2 = \frac{1}{n}\sum_{i=1}^{n} (x_i \ - \ \mu_x)^2 = \frac{1}{n} \ X_c^T \ \ X_c \,, \tag{3.6}$$

$${\sigma_y}^2 = \frac{1}{n}\sum_{i=1}^{n} (y_i \ - \ \mu_y)^2 = \frac{1}{n} \ Y_c^T \ \ Y_c \,, \tag{3.7}$$

where σ_x and σ_y reflect the standard deviations of x_i and y_i respectively, and the ”T” signifies transpose.

With this, a pair of standardized vectors x and y can be obtained

$$x = \frac{X - \mu_x}{\sigma_x} = \frac{X_c}{\sigma_x} \,, \tag{3.8}$$

$$y = \frac{Y - \mu_y}{\sigma_y} = \frac{Y_c}{\sigma_y} \,. \tag{3.9}$$

Both x and y length is equal to n.

As the model is based on spatial distance, an n-by-n unitary spatial weights matrix needed to be defined

$$W = [w_{ij}]_{nn} \,. \tag{3.10}$$

This matrix, which is actually a unitized spatial weights matrix, was created from a spatial contiguity matrix using the inverse distance between municipalities. It has three main properties: symmetry

$$w_{ij} = w_{ji}\,, \tag{3.11}$$

zero diagonal elements, which implies that the entries in the diagonal are all 0

$$|w_{ii}| = 0\,, \tag{3.12}$$

unitization condition, meaning the sum of all matrix entrances is 1

$$\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij} = 1\,. \tag{3.13}$$

According to Chen, a coefficient of spatial cross-correlation can be calculated by taking into account an improved formula of Moran's index for spatial auto-correlation. The new coefficient, the Spatial Cross-correlation Index (SCI), is as follows

$$R_c = x^T W y = y^T W x\,. \tag{3.14}$$

SCI values fall between -1 and 1.

Two pairs of scatterplots are needed to visually portray spatial cross-correlations. To generate these plots, the six following variables were required

$$f^{(xy)} = xy^T W x\,, \tag{3.15}$$

$$f^{(yx)} = yx^T W y\,, \tag{3.16}$$

$$f^{(xx)} = xx^T W y\,, \tag{3.17}$$

$$f^{(yy)} = yy^T W x\,, \tag{3.18}$$

$$f^{(x)} = nWx\,, \tag{3.19}$$

$$f^{(y)} = nWy\,. \tag{3.20}$$

The subsequent Table 3.2 shows how the variables are matched in order to make the cross-correlation scatterplots.

When an overlapping of the trendline and the scattered data points is made, a scatter diagram for spatial cross-correlation analysis can be obtained. And, since the first and second plots are the same, and so are the third and the fourth, only two scatterplots are needed to fully illustrate spatial cross-correlation.

Scatterplot	Abscissa	Ordinate		Effect
		Scattered Points	Trend Line	
First plot	x	f$^{(y)} = nWy$	f$^{(xy)} = xy^TWx$	x acts on y
Second plot	x	f$^{(y)} = nWy$	f$^{(xx)} = xx^TWy$	x acts on y
Third plot	y	f$^{(x)} = nWx$	f$^{(yx)} = yx^TWy$	y acts on x
Fourth plot	y	f$^{(x)} = nWx$	f$^{(yy)} = yy^TWx$	y acts on x

Table 3.2: Variables relations for spatial cross-correlation scatterplots.

An example of two scatterplots representing a positive correlation between variables can be found in Figure 3.2 [52].

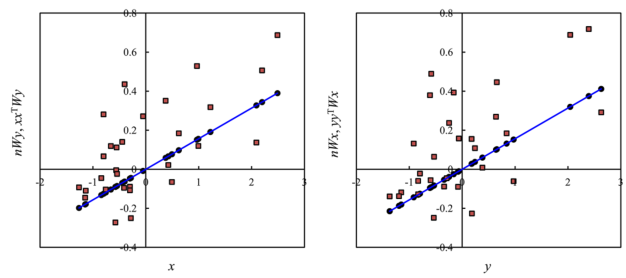


Figure 3.2: Example of the dual scatterplots of spatial cross-correlation.

The model also defines two other variables: the Simple Correlation Coefficient (SCC) and the Partial Spatial Cross-correlation Coefficient (PSCC).

The SCC can be treated as a special case of SCI when spatial distance is left out of account

$$R_0 = x^TW_0y = y^TW_0x\,, \tag{3.21}$$

where W_0 represents a unitary identity matrix, which takes the place of the unitized spatial weights matrix. R_0 is just a Pearson's correlation coefficient, which indicates a simple cross-correlation between x and y.

Chen also defines the PSCC as

$$R_p = R_0 - R_c = x^TW_0y \;-\; x^TWy = y^TW_0x \;-\; y^TWx\,. \tag{3.22}$$

Bearing all that in mind, the spatial correlation coefficients can be clarified in the following way. The spatial cross-correlation index, R_c, portrays the correlation between x and y taking into account spatial distances. The partial spatial cross-correlation coefficient, R_p, depicts the cross-correlation between x and y, free from the spatial distances. And, the simple correlation coefficient, R_0, which conveys both cross-correlations.

SCI	(R_c)	cross-correlation between x and y, taking into account spatial distances
PSCC	(R_p)	cross-correlation between x and y, free from the spatial distances
SCC	(R_0)	both the cross-correlations

Table 3.3: Spatial correlation coefficients overview.

Other coefficients were used to gauge the statistical significance of the results obtained. The coefficient of determination (R^2) denotes the proportion of the variation in the dependent variable that is predictable from the independent variable. R^2 indicates the goodness of fit for the regression of nWx depending on y. The higher this value is (from 0 to 1), the better the fit. The F statistics can be used to judge the cause and effect since it indicates the extent to which an independent variable can explain the corresponding dependent variable. For a result to be considered statistically significant, it needs an F value higher than 3.95 and a P value lower than 0,05.

Chapter 4

Results and Discussion

The results of performing both the ITS and the spatial cross-correlation analysis on the datasets provided by DGS and SIBS can be found in this chapter. As this research was conducted using aggregated weekly data, Table 4.1 shows each week's corresponding dates.

Weeks	Dates
Week 1	November 15th to 21st, 2021
Week 2	November 22nd to 28th, 2021
Week 3	November 29th to December 5th, 2021
Week 4	December 6th to 12th, 2021
Week 5	December 13th to 19th, 2021
Week 6	December 20th to 26th, 2021

Table 4.1: Dates of each week of the study.

4.1 Data Exploration

A visual overview was performed at the beginning of the study to help with the familiarization and better understanding of the data available.

Table 4.2 shows the total number of omicron cases and South African card transactions per week that occurred in continental Portugal.

Week	Number of omicron cases	Number of South African card transactions
Week 1	25	1235
Week 2	44	1200
Week 3	20	1057
Week 4	235	1081
Week 5	1298	1171
Week 6	4714	1194

Table 4.2: Number of omicron cases and South African card transactions per week.

The number of omicron cases grew exponentially during the 6-week span of the study, whilst the number of South African card transactions stayed relatively steady. The average for the omicron dataset was 1056 cases per week and the variance was 3454794. For the card dataset, the average was 1156 weekly South African transactions and the variance was 5054.

Figures 4.1 and 4.2 show the temporal evolution curves of both the weekly number of confirmed omicron cases and the weekly number of South African card transactions in continental Portugal.

The omicron curve shows a steady but relatively mild increase in cases in the first three weeks (section 1 of the graph) followed by a quick rise (section 2) and an even steeper growth (section 3) by the backend of the study.

The transactions curve displays a decrease in the number of transactions made right at the beginning of the study (section 1) followed by an increase in such transactions (section 3) making it that the number of transactions in week 1 is very similar to the number of transactions in week 6. The lowest point in the graph coincides with the travel ban imposed by the country on Southern African flights (section 2).

Figure 4.1: Omicron cases curve.

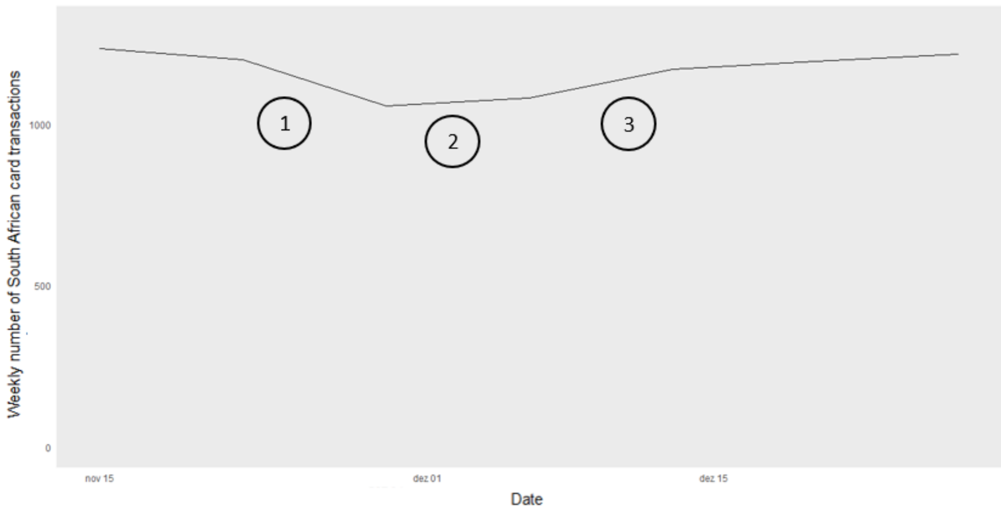


Figure 4.2: South African card transactions curve.

To further comprehend the spread of both variables, heat maps of continental Portugal were made using the programming language R[2]. The spatial evolution of both omicron and South African card transactions' rates in the six-week span of the study can be found below in Figures 4.3 and 4.4. In those maps, the highest rates are represented in red and the lowest in lighter yellow, municipalities without available data are painted in white.

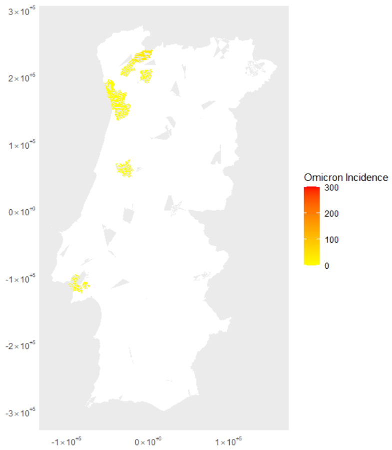
In them, the rapid spread of the omicron variant is apparent. It is also clear that the first hotspots appeared around Lisbon and in the northwest region of continental Portugal. These are some of the country's most populous areas and are near the two busiest airports. From there, omicron proliferated to the majority of all Portuguese land territory in a matter of weeks. By the final week of the study, there was also a very high concentration of omicron incidence in Alentejo.

Despite changes in mobility restrictions, the majority of the transaction rate numbers remained quite consistent throughout the entire time.

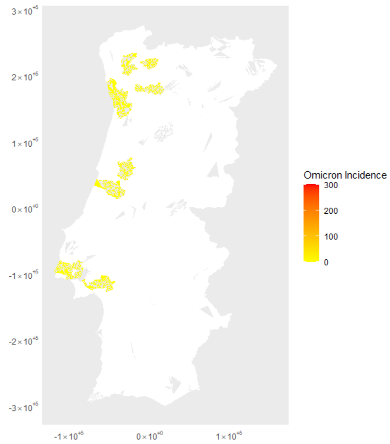
A closer look was taken at the initial omicron hotspots. Figure 4.5 depicts the evolution of the two variables in continental Portugal, and the municipalities of Lisbon and Braga. These areas were chosen as they represent the two biggest cities surrounded by the municipalities that showed the highest incidence in the first three weeks of the study.

The aim of this analysis was to try to see if there was any significant difference in the trendline between the COVID-19 hotspots when compared to the rest of the country. Perhaps, right before the virus outburst, there would be a surge in foreign transactions. However, this is not what the graphs show, with the transaction line keeping quite steady throughout the six weeks, except for a slight decrease in week 3 in Lisbon, possibly due to the travel ban.

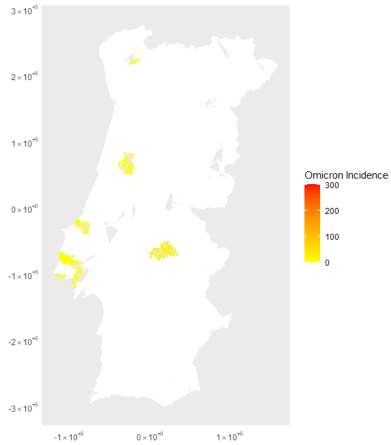
Despite also being a hotspot, Alentejo was not chosen to be a part of this analysis since omicron only strongly appear by week 6, and by that point, most of the transmission occurring in the country was from Portuguese to Portuguese.



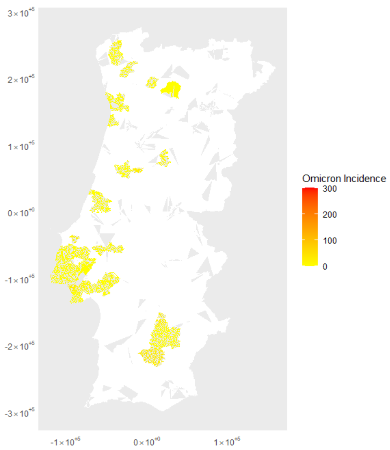
(a) Week 1



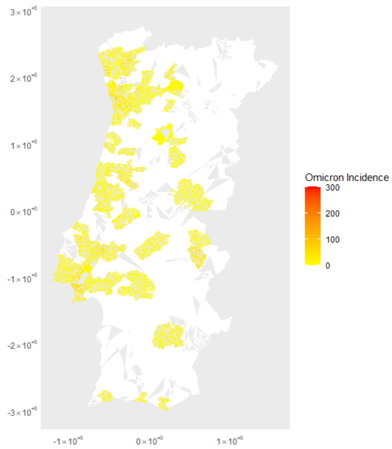
(b) Week 2



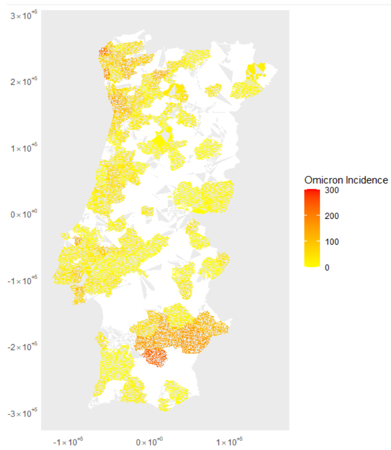
(c) Week 3



(d) Week 4

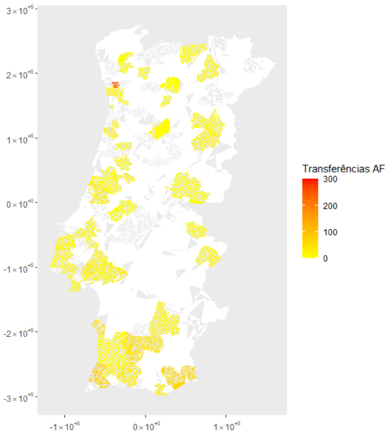


(e) Week 5

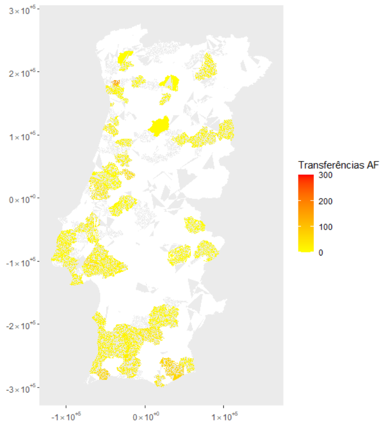


(f) Week 6

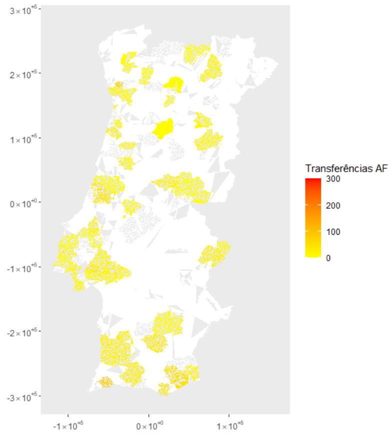
Figure 4.3: Evolution of omicron incidence in continental Portugal.



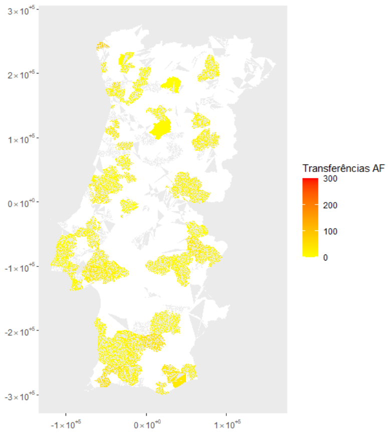
(a) Week 1



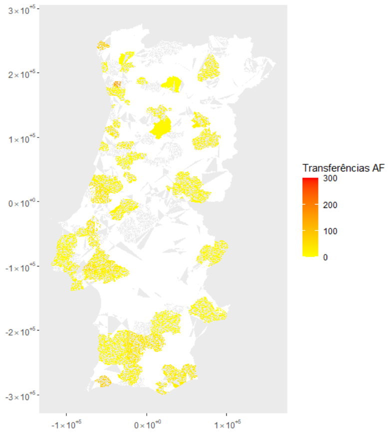
(b) Week 2



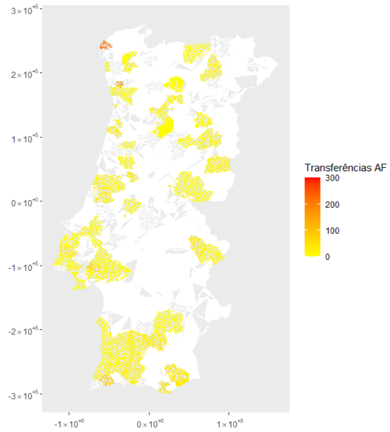
(c) Week 3



(d) Week 4

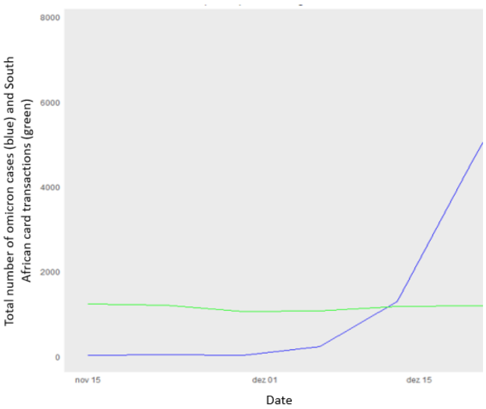


(e) Week 5

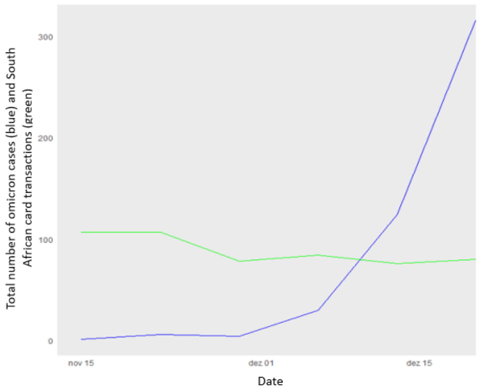


(f) Week 6

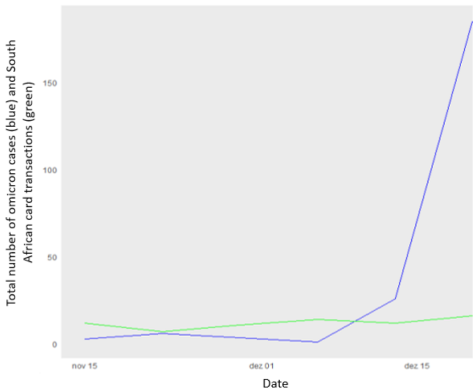
Figure 4.4: Evolution of South African transactions rate in continental Portugal.



(a) Portugal



(b) Lisbon



(c) Braga

Figure 4.5: Evolution of confirmed omicron cases (in blue) and South African card transactions (in green) in continental Portugal, Lisbon, and Braga.

4.2 Travel Ban Analysis: Interrupted Time Series

To assess the effectiveness of the travel ban imposed by Portugal on southern African countries, an ITS methodology was employed. In it, the total number of transactions made by South African cards anywhere in Portugal was taken into account. The results of the analysis performed are depicted in Figure 4.6.

The travel ban was put in place on November 29^{th}, 2021 but only lasted 15 days as it was lifted on December 13^{th}. This two-week span corresponds to weeks 3 and 4 of the spatial cross-correlation analysis.

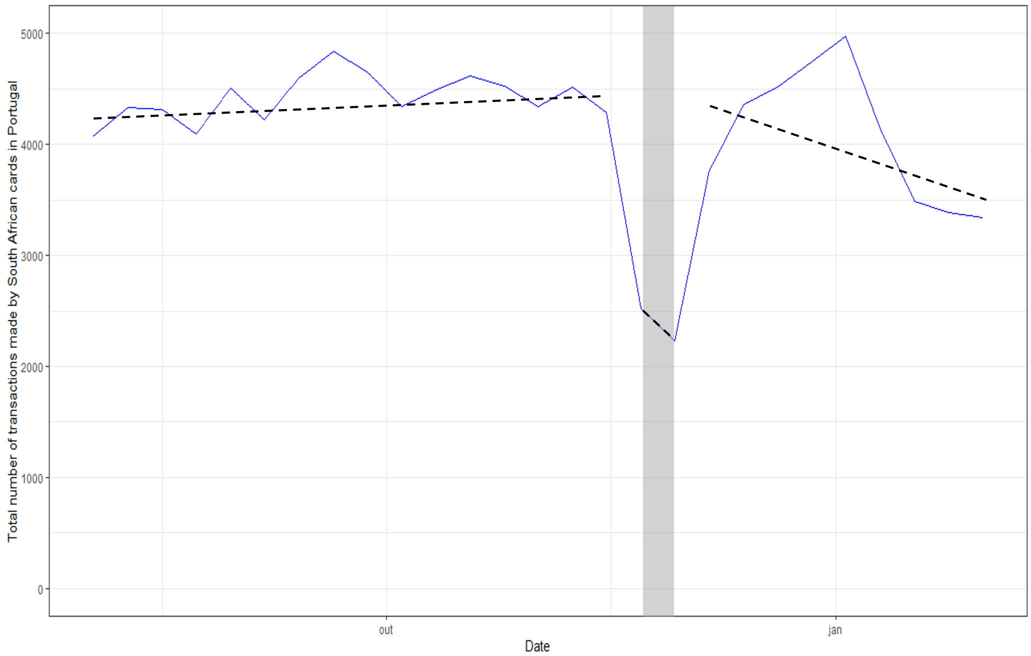


Figure 4.6: Interrupted time series of the total number of transactions from South African cards in Portugal, with the travel ban highlighted from November 29th to December 13th, 2021.

There was an abrupt fall in the number of transactions during the period when travel restrictions were imposed. Although, that number quickly returned to around its average as soon as the regulations changed back.

The decrease in the total number of South African transactions is much sharper in Figure 4.6 than in the previous figures because for this travel ban analysis data from all of Portugal was utilized, not just from the continental part of the country. This was done because the ban was also imposed in the archipelagos airports.

The linear regression pre and post-ban go in different directions. However, since the post-ban period available is quite short, more data needed to be gathered to fully understand if this was a trend or if the numbers just needed some time to stabilize. Since other restrictions, like quarantines, were being enacted at the time, it is plausible that there might have been a slight decrease in the overall number of transactions at the backend of 2021 and the beginning of 2022.

Although Figure 4.6 shows that technically the travel ban was effective in diminishing the transactions when taking into account the omicron spread demonstrated above it seems the measure did not have the intended impact on COVID-19 propagation.

4.3 Omicron Spread Analysis: Spatial Cross-correlation

A spatial cross-correlation analysis was utilized to understand if South Africans' mobility, tracked using card transactions, had any influence on the spread of the omicron variant, therefore being useful to possibly predict the proliferation of future pandemics. The results of this study can be found in the figures and tables below.

Figures 4.7 to 4.10 display scatterplots that reflect the action y (card transaction incidence) has on x (omicron incidence). Only one scatterplot was used since it would not make sense to plot the effect COVID-19 cases might have on transactions. In Tables 4.3 through 4.6 correlation coefficients that provide a deeper understanding of the relation between omicrons' incidence and South African card transactions in continental Portugal can be found.

As previously mentioned, SCI depicts the correlation between both variables (x and y) considering spatial distances. SPCC portrays the cross-correlation between x and y, free from the spatial distances. Finally, the SCC can be obtained by adding SCI and SPCC, meaning it conveys both cross-correlations.

Since COVID-19 can take some days to transmit, develop and detect this analysis was carried out over a period of time larger than a week. So, for all the omicron data available, a test was conducted for the transactions made the same week (lag=0) and the week prior (lag=1). Due to a lack of data, it was not possible to perform the lag=0 analysis for week 1.

This study was first conducted for all 278 municipalities of continental Portugal. However, there were several municipalities without any omicron cases and/or South African card transactions that only added zeros to the model and therefore might have been skewing the results. For this reason, the same study was conducted but only for 64 specific municipalities. These were the municipalities that recorded transactions in at least three of the weeks, half the time span of the study.

The results obtained by both studies, as well as a discussion about them, can be found in the subchapters below.

4.3.1 Robustness Test 1

The figures and tables in this subsection show the results of the study conducted on all 278 municipalities of continental Portugal.

Since the first omicron hotspot worldwide was in South Africa it would be plausible to assume there could be a connection between South Africans' movements in Portugal and the spread of the variant in the country. Meaning, a strong direct correlation was excepted especially at the begging of the study before Portuguese people start being sick and spreading it themselves. However, that is not what can be found in the results exhibited below.

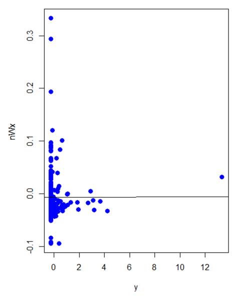
All correlation values are low and do not have any statistical significance. The cross-correlation values fluctuate between being positive and negative meaning no conclusion between the type of correlation (direct or indirect) can be drawn.

However, most of the negative SCC and PSCC values occurred when the overall number of transactions was decreasing, between weeks 1 and 3, and the number of omicron cases was slowly rising, as seen in Figures 4.1 and 4.2. For the last three weeks of the study, when both the number of transactions and the number of cases were increasing, the correlations were positive, which is what was expected. It is also important to note the SCI, which is the correlation factor that takes into account spatial distances, is almost always positive, except for week 4 with lag=0.

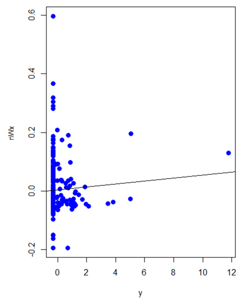
It does not appear to be any significant difference between lag=0 and lag=1 analysis.

The numbers show the PSCC values are constantly higher than SCI ones, except for week 2 lag=0. This is curious since SPCC portrays the cross-correlation between x and y, free from the spatial distances.

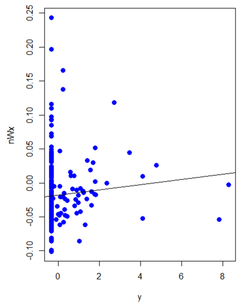
The outliers seen in the scatterplots represent the municipalities with the highest rate of transactions. For weeks 1, 2, and 5 that municipality is Trofa, for weeks 3 and 6 is Caminha, and for week 4 is Lagos.



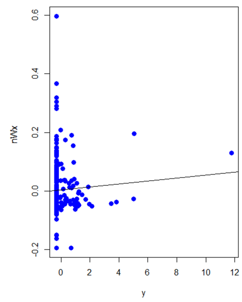
(a) Week 1



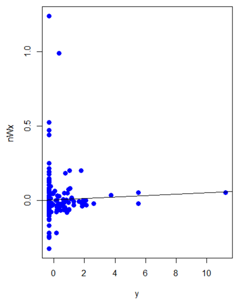
(b) Week 2



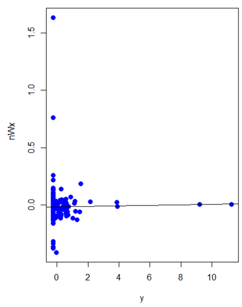
(c) Week 3



(d) Week 4

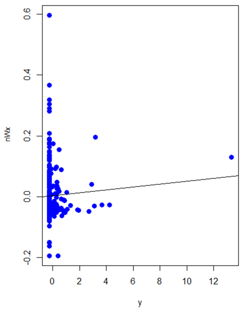


(e) Week 5

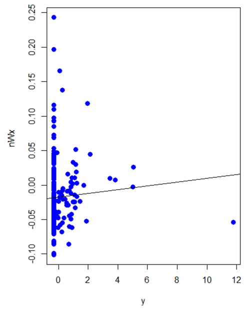


(f) Week 6

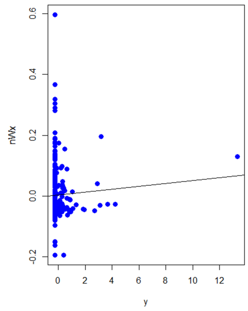
Figure 4.7: Spatial cross-correlation scatterplots for omicron incidence, for all 278 municipalities, from weeks 1 to 6 with lag=0.



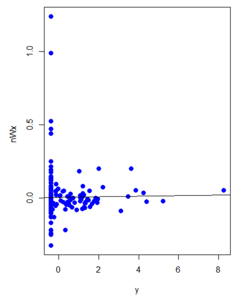
(a) Week 2



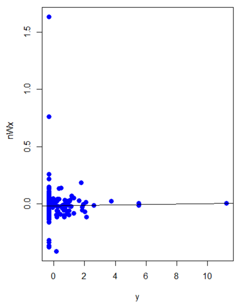
(b) Week 3



(c) Week 4



(d) Week 5



(e) Week 6

Figure 4.8: Spatial cross-correlation scatterplots for omicron incidence, for all 278 municipalities, from weeks 1 to 6 with lag=1.

	Week 1	**Week 2**	**Week 3**	**Week 4**	**Week 5**	**Week 6**
SCC	-1,72E^{-2}	2,04E^{-3}	-1,73E^{-2}	2,29E^{-1}	2,26E^{-1}	2,96E^{-1}
SCI	7,49E^{-5}	5,13E^{-3}	3,87E^{-3}	-1,35E^{-3}	4,68E^{-3}	2,18E^{-3}
PSCC	-1,72E^{-2}	-3,09E^{-3}	-2,12E^{-2}	2,31E^{-1}	2,21E^{-1}	2,94E^{-1}
R^2	2,71E^{-6}	3,45E^{-3}	6,58E-$^{-3}$	2,74E^{-4}	1,25E^{-3}	2,45E^{-4}
F	7,51E^{-4}	9,59E^{-1}	1,83	7,59E^{-2}	3,47E^{-1}	6,79E^{-2}
P	9,78E^{-1}	3,28E^{-1}	1,77E^{-1}	7,83E^{-1}	5,56E^{-1}	7,94E^{-1}

Table 4.3: Correlation coefficients values for 278 municipalities and lag=0.

	Week 2	**Week 3**	**Week 4**	**Week 5**	**Week 6**
SCC	-1,54E^{-2}	-1,58E^{-2}	9,27E^{-2}	2,29E^{-1}	2,34E^{-1}
SCI	4,75E^{-3}	2,74E^{-3}	1,94E^{-4}	2,14E^{-3}	1,69E^{-3}
PSCC	-2,01E^{-2}	-1,85E^{-2}	9,25E^{-2}	2,27E^{-1}	2,33E^{-1}
R^2	2,95E^{-3}	3,30E^{-3}	5,62E^{-6}	2,61E^{-4}	1,50E^{-4}
F	8,20E^{-1}	9,17E^{-1}	1,56E^{-3}	7,23E^{-2}	4,16E^{-2}
P	3,66E^{-1}	3,39E^{-1}	9,68E^{-1}	7,88E^{-1}	8,38E^{-1}

Table 4.4: Correlation coefficients values for 278 municipalities and lag=1.

4.3.2 Robustness Test 2

The figures and tables in this subsection convey the results of the study conducted on only the 64 municipalities of continental Portugal that had South African card transactions for at least three of the six weeks.

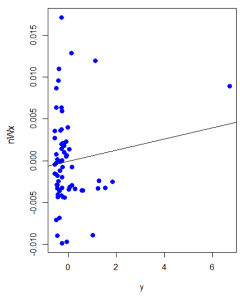
Correlation values, although still quite low, have slightly risen when compared to the first study. This shows municipalities without transactions might have been skewing the results. Now, some statistically significant f and p-values can be found for: week 1 lag=0, week 2 lag=1, week 3 lag=1, and week 4 lag=1. However, in those weeks half the correlation coefficients are positive half are negative making it hard to conclude anything.

The same fluctuation between positive and negative values that occurred in the first study can be seen in the second one. However, the relation between a negative SCC and PSCC and the overall decrease in South African transactions is not as clear as it was for all 278 municipalities. Unlike in the previous analysis, the SCI is constantly switching from positive to negative.

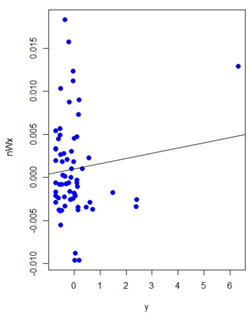
There are more statistically significant results for lag=1 than for lag=0.

All throughout the weeks, PSCC continues to be the dominant cross-correlation factor over SCI.

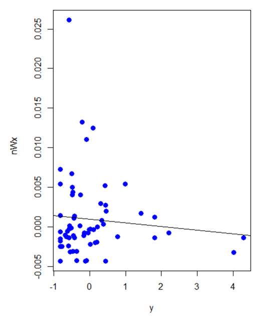
The same outliers are displayed in the scatterplots, the municipalities of Trofa (for all weeks except the 4^{th}), Caminha (weeks 3 and 6), and Lagos (in week 4).



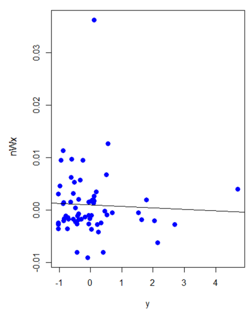
(a) Week 1



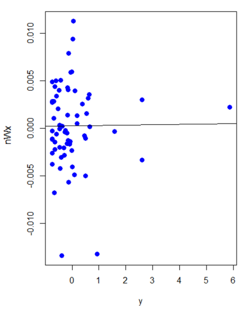
(b) Week 2



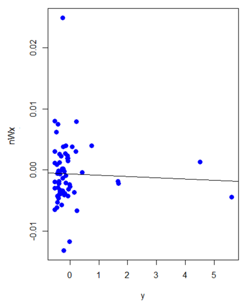
(c) Week 3



(d) Week 4

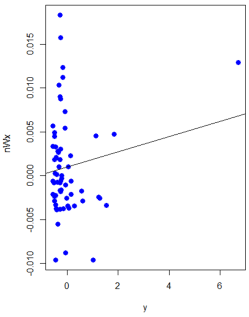


(e) Week 5

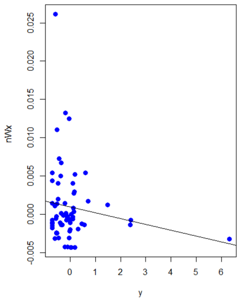


(f) Week 6

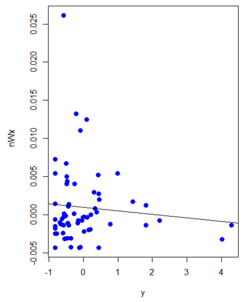
Figure 4.9: Spatial cross-correlation scatterplots for omicron incidence, for only 64 municipalities, from weeks 1 to 6 with lag=0.



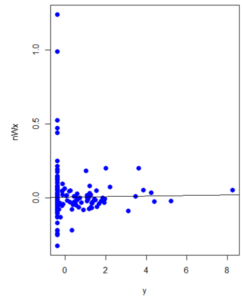
(a) Week 2



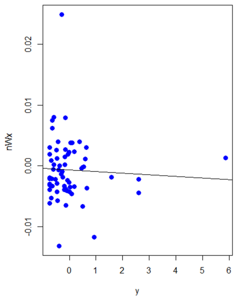
(b) Week 3



(c) Week 4



(d) Week 5



(e) Week 6

Figure 4.10: Spatial cross-correlation scatterplots for omicron incidence, for only 64 municipalities, from weeks 1 to 6 with lag=1.

	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6
SCC	-1,42E^{-1}	-1,20E^{-1}	3,79E^{-3}	-9,79E^{-2}	6,32E^{-2}	3,11E^{-1}
SCI	6,64E^{-4}	6,09E^{-4}	-4,63E^{-4}	-2,80E^{-4}	4,35E^{-5}	-2,03E^{-4}
PSCC	-1,42E^{-1}	-1,21E^{-1}	4,25E^{-3}	-9,76E^{-2}	6,31E^{-2}	3,11E^{-1}
R^2	1,50E^{-2}	1,16E^{-2}	8,49E^{-3}	1,98E^{-3}	1,00E^{-4}	1,54E^{-3}
F	4,22	3,25	2,37	5,50E^{-1}	2,77E^{-2}	4,27E^{-1}
P	4,41E^{-2}	7,62E^{-2}	1,29E^{-1}	4,61E^{-1}	8,68E^{-1}	5,16E^{-1}

Table 4.5: Correlation coefficients values for 64 municipalities and lag=0.

	Week 2	Week 3	Week 4	Week 5	Week 6
SCC	-1,39E^{-1}	1,83E^{-2}	5,44E^{-2}	-9,79E^{-2}	1,38E^{-1}
SCI	8,64E^{-4}	-7,62E^{-4}	7,77E^{-4}	3,91E^{-4}	-2,71E^{-4}
PSCC	-1,39E^{-1}	1,90E^{-2}	5,36E^{-2}	-9,83E^{-2}	1,38E^{-1}
R^2	2,33E^{-2}	2,30E^{-2}	1,52E^{-2}	8,15E^{-3}	2,72E^{-3}
F	6,61	6,52	4,28	2,28	7,55E^{-1}
P	1,25E^{-2}	1,31E^{-2}	4,28E^{-2}	1,36E^{-1}	3,88E^{-1}

Table 4.6: Correlation coefficients values for 64 municipalities and lag=1.

4.4 Discussion

An ITS analysis was performed to answer the first question posed in this dissertation: was the air traffic ban imposed by Portugal on South Africa effective in preventing the spread of omicron? The results show that although during the ban there was a decrease in the number the transactions it apparently did not have the intended impact on COVID-19 propagation.

The ineffectiveness of the measure might be due to the late introduction of the travel ban. By the time the country closed its border with South Africa, the variant was already in other countries, including Portugal, making this targeted ban redundant. Since other nationalities could freely enter the country and bring with them the virus [53]. This can be validated in literature since, as previously mentioned, even a thoroughly focused travel ban could not be effective given the highly interconnected world we live in nowadays.

The second question put forward in this work was: can South Africans' tracked movements through card transactions be used to predict the spread of omicron? The results of the spatial cross-correlation analysis performed show that this is not the case. As neither study revealed a significant correlation between both variables, unlike what was to be predicted.

After the literature review performed above these were not the results excepted. However, some explanations as to why that is the case can be as follows.

Firstly, it is feasible South Africans did not represent the biggest influence on the omicron spread in Portugal. Since, the two countries do not have a major flow of people between them and, as mentioned above, the variant quickly spread throughout the world meaning other nationalities could have been responsible for the proliferation of omicron in continental Portugal. Perhaps if the study focused on a broader group of people (not just South Africans) it would have been able to reach more statistically significant results.

Secondly, although the study contains the most crucial part, the introduction of omicron and the beginning of the spread, the time frame might have been too short and thus making it unable to fully capture the phenomenon.

It is also important to remember that every data point only represents a transaction by a card and not an individual. Meaning the same card could be responsible for several transactions on the same day, week, or month. This could also have impacted the results since a high rate of transactions in a specific location does not necessarily translate to a high rate of South Africans in that same region.

Another reason might be the possible ascertainment bias in the COVID-19 daily confirmed cases dataset. Although at first glance might not seem like much, if there is hesitancy by South Africans to get tested in Portugal, that can lead to a misrepresentation of reality and a reduced number of omicron cases especially in the beginning. This means the first emergence of the variant might not have occurred in those exact locations recorded in the dataset. As South African travelers could have infected other residents who themselves could have moved inside the country. This means there is a chance the reported case and the actual infection occurred in different municipalities.

Lastly, but contrary to the research found, perhaps card transactions are simply not a good proxy for mobility. Either that or mobility might not be as good at predicting the proliferation of pandemics as seen in other studies.

Due to all the limitations expressed above, future research needs to be conducted to extensively understand how successful card transactions can be as a proxy for mobility. Also, more in-depth analyses should be made to truly comprehend if mobility can ever be a useful tool in predicting the evolution of an epidemic.

As seen during the literature review, EWDS and spatial-temporal analyses have become crucially important to help decision-makers delineate the best course of action. These methodologies are extremely versatile and can be used in all facets of planning, prevention, mapping, control, and policy efficiency and impact assessments. With all the benefits highlighted throughout this work, it is clear that all further research on this topic can greatly improve public health and the overall well-being of the population.

It is also important to note that, with so much information available nowadays, it is becoming more and more imperative to think outside the box and find non-conventional data to explore and try to find connections. This can not only be done in healthcare but throughout the whole scientific spectrum, promoting interdisciplinarity and improving society.

Chapter 5

Conclusion

The aim of this dissertation was to try to answer two different questions. Firstly, was the air traffic ban imposed by Portugal on South Africa effective in preventing the spread of omicron? And secondly, can South Africans' tracked movements through card transactions be used to predict the spread of omicron?

These questions were chosen after a literature assessment. This review highlighted the common use of travel bans during the early stages of the pandemic as a preventive method to stop or delay the spread of COVID-19. The importance of EWDS was also prominent in literature as these systems allow for a rapid interpretation of constantly evolving data making them crucial tools for decision-makers facing everchanging scenarios. In addition, spatiotemporal analyses were deemed quite helpful in projecting the possible spread of the virus.

Two methodologies were used in order to answer the questions posed above. An ITS was utilized to assess the effectiveness of the travel ban on southern African countries. A spatial cross-correlation analysis was utilized to try to understand if South African card transactions had any influence on the spread of the omicron variant.

The ITS analysis indicated a fall in the number of transactions during the period when travel restrictions were imposed. However, that number quickly arouse as soon as the regulations changed back. Although technically the travel ban was effective in diminishing the transactions there did not appear to be any decrease or delay in the spread of omicron. This was perhaps due to the fact the variant, at that point in time, was already in several countries not included in the ban and had already reached Portugal.

The spatiotemporal analysis was performed in two datasets. Firstly, on one with all 278 municipalities from continental Portugal, and secondly, on one with just the 64 municipalities that recorded transactions in at least three of the weeks, half the time span of the study. Since COVID-19 can take some days to spread, this analysis was carried out over a period of time larger than a week. So, for all the omicron data available, a test was conducted for the transactions made the same week (lag=0) and the week prior (lag=1).

Neither study portrayed a statistically significant spatial cross-correlation between South Africans' card transaction movements in Portugal and the spread of the variant in the country. However, more in-depth research is recommended to fully understand how good a proxy card transactions are to citizens' mobility. And if there is a way card transactions could ever be considered a good method to predict the proliferation of future pandemics.

Bibliography

[1] W. H. Organization. Archived: WHO Timeline - COVID-19, year = 2020, url = https://www.who.int/news/item/27-04-2020-who-timeline—covid-19, urldate = 2020-04-27, .

[2] Stoecklin, S. B. et al. First cases of coronavirus disease 2019 (covid-19) in france: Surveillance, investigations and control measures, january 2020. *Euro Surveill*, 2020.

[3] Marques da Costa E. and Marques da Costa N. A pandemia covid-19 em portugal continental – uma anÁlise geogrÁfica da evoluÇÃo verificada nos meses de marÇo e abril. *Revista Brasileira de Geografia Médica e da Saúde*, 2020.

[4] Coronavirus Worldwide Graphs, year = , url = https://www.worldometers.info/coronavirus/worldwide-graphs/, urldate = 2023-04-16.

[5] W. H. Organization. Coronavirus disease (COVID-19): How is it transmitted?, year = 2021, url = https://www.who.int/news-room/questions-and-answers/item/coronavirus-disease-covid-19-how-is-it-transmitted, urldate = 2021-12-23, .

[6] Cavalcante, J. R. et al. Covid-19 no brasil: evolução da epidemia até a semana epidemiológica 20 de 2020. *Epidemiologia e Servicos de Saude: Revista Do Sistema Unico de Saude Do Brasil*, 2020.

[7] W. H. Organization. Mask use in the context of COVID-19, year = 2020, url = https://www.who.int/publications/i/item/advice-on-the-use-of-masks-in-the-community-during-home-care-and-in-healthcare-settings-in-the-context-of-the-novel-coronavirus-(2019-ncov)-outbreak, urldate = 2020-12-01, .

[8] Ayouni, I. et al. Effective public health measures to mitigate the spread of covid-19: a systematic review. *BMC Public Health*, 2021.

[9] W. H. Organization. Tracking SARS-CoV-2 variants, year = 2023, url = https://www.who.int/activities/tracking-SARS-CoV-2-variants, urldate = 2023-03-30, .

[10] V. Health. A Timeline of COVID-19 Variants, year = 2022, url = https://www.verywellhealth.com/covid-variants-timeline-6741198, urldate = 2022-10-25.

[11] W. H. Organization. One year since the emergence of COVID-19 virus variant Omicron, year = 2022, url = https://www.who.int/news-room/feature-stories/detail/one-year-since-the-emergence-of-omicron, urldate = 2022-10-25, .

[12] I. N. de Saúde Doutor Ricardo Jorge. Diversidade genética do novo coronavírus SARS-CoV-2 (COVID-19) em Portugal, year = , url = https://insaflu.insa.pt/covid19/, urldate = 2023, .

[13] O. W. in Data. Coronavirus (COVID-19) Cases, year = , url = https://ourworldindata.org/covid-cases, urldate = 2023.

[14] D. da República. Despacho n.º 11740-E/2021, de 26 de novembro, year = 2021, url = https://dre.pt/dre/detalhe/despacho/11740-e-2021-175043488, urldate = 2021-11-26.

[15] Nouvellet, P. et al. Reduction in mobility and covid-19 transmission. *Springer Nature*, 2021.

[16] Hale, T. et al. A global panel database of pandemic policies (oxford covid-19 government response tracker). *Nature Human Behaviour*, 2021.

[17] Anzai, A. et al. Assessing the impact of reduced travel on exportation dynamics of novel coronavirus infection (covid-19). *Journal of Clinical Medicine*, 2020.

[18] Chinazzi, M. et al. The effect of travel restrictions on the spread of the 2019 novel coronavirus (covid-19) outbreak. *Science*, 2020.

[19] Linka, K. et al. Outbreak dynamics of covid-19 in europe and the effect of travel restrictions. *Taylor and Francis*, 2020.

[20] Russell, T.W. et al. Effect of internationally imported cases on internal spread of covid-19: a mathematical modelling study. *Lancet Public Health*, 2020.

[21] Kucharski, A.J. et al. Travel measures in the sars-cov-2 variant era need clear objectives. *The Lancet*, 2022.

[22] Cowling, BJ. et al. Impact assessment of nonpharmaceutical interventions against coronavirus disease 2019 and influenza in hong kong: an observational study. *Lancet Public Health*, 2020.

[23] Mandal, S. et al. Prudent public health intervention strategies to control the coronavirus disease 2019 transmission in india: a mathematical model-based approach. *Indian Journal of Medical Research*, 2020.

[24] Vianello, C. et al. A perspective on early detection systems models for covid-19 spreading. *Biochemical and Biophysical Research Communications*, 2020.

[25] U. N. O. for Disaster Risk Reduction. Early warning system, year = , url = https://www.undrr.org/terminology/early-warning-system, urldate = .

[26] Bari, A. et al. Covid-19 early-alert signals using human behavior alternative data. *Social Network Analysis and Mining*, 2021.

[27] Lopreite, M. et al. Early warnings of covid-19 outbreaks across europe from social media. *Scientific Reports*, 2021.

[28] Subissi, L. et al. An early warning system for emerging sars-cov-2 variants. *Nature Medicine*, 2022.

[29] Pillay, L. et al. Potential and challenges encountered in the application of wastewater-based epidemiology as an early warning system for covid-19 infections in south africa. *American Chemical Society*, 2022.

[30] Franch-Pardo, I. et al. Spatial analysis and gis in the study of covid-19. a review. *Science of The Total Environment*, 2020.

[31] Chen, Z.L. et al. Distribution of the covid-19 epidemic and correlation with population emigration from wuhan, china. *Chinese medical journal*, 2020.

[32] Huang, H.. et al. Epidemic features and control of 2019 novel coronavirus pneumonia in wenzhou, china. *Chinese medical journal*, 2020.

[33] Dong, E. et al. An interactive web-based dashboard to track covid-19 in real time. *The Lancet Infectious Diseases*, 2020.

[34] Saha, A. et al. Monitoring and epidemiological trends of coronavirus disease (covid-19) around the world. *Matrix Science Medica*, 2020.

[35] Arab-Mazar, Z. et al. Mapping the incidence of the covid-19 hotspot in iran – implications for travellers. *Travel Medicine and Infectious Disease*, 2020.

[36] Rossman, H. et al. A framework for identifying regional outbreak and spread of covid-19 from one-minute population-wide surveys. *Nature Medicine*, 2020.

[37] Jaber, A. S. et al. . A moran's i autocorrelation and spatial cluster analysis for identifying coronavirus disease covid-19 in iraq using gis approach. *Caspian Journal of Environmental Sciences*, 2022.

[38] Rendana, M. et al. . Spatial distribution of covid-19 cases, epidemic spread rate, spatial pattern, and its correlation with meteorological factors during the first to the second waves. *Journal of Infection and Public Health*, 2021.

[39] Isazade, V. et al. . Integration of moran's i, geographically weighted regression (gwr), and ordinary least square (ols) models in spatiotemporal modeling of covid-19 outbreak in qom and mazandaran provinces, iran. *Modeling Earth Systems and Environment*, 2023.

[40] Covre, E. et al. . Spatial correlation of covid-19 with intensive care unit beds in paraná. *Revista de Saúde Pública*, 2022.

[41] Lakhani, A. et al. Which melbourne metropolitan areas are vulnerable to covid-19 based on age, disability, and access to health services? using spatial analysis to identify service gaps and inform delivery. *Journal of Pain and Symptom Management*, 2020.

[42] Orea, L. and Álvarez, I. C. . How effective has the spanish lockdown been to battle covid-19? a spatial analysis of the coronavirus propagation across provinces. *Health Econ*, 2022.

[43] Giuliani, D. et al. Modelling and predicting the spatio-temporal spread of covid-19 in italy. *BMC Infectious Diseases*, 2020.

[44] Boru, B. and Gursoy, M. E. Forecasting daily covid-19 case counts using aggregate mobility statistics. *Health Informatics in the Age of COVID-19*, 2022.

[45] Sobolevsky, S. et al. . Money on the move: Big data of bank card transactions as the new proxy for human mobility patterns and regional delineation. the case of residents and foreign visitors in spain. *2014 IEEE International Congress on Big Data*, 2014.

[46] Carvalho, V. M. et al. . Tracking the covid-19 crisis with high-resolution transaction data. *Royal Society Open Science*, 2021.

[47] I. N. de Estatística. Estimativas de População Residente, year = 2022, url = https://www.ine.pt/xportal/xmain?xpid=INExpgid=ine$_{destaquesDESTAQUESdest_boui} = 540837471DESTAQUESmodo = 2, urldate = 2022-07-15.$

[48] S. N. de Saúde. Instituto Nacional de Saúde Doutor Ricardo Jorge, IP, year = 2021, url = https://www.sns.gov.pt/entidades-de-saude/instituto-nacional-de-saude-dr-ricardo-jorge-ip/, urldate = 2021-03-11.

[49] I. N. de Saúde Doutor Ricardo Jorge. Diversidade genética do novo coronavírus SARS-CoV-2 (COVID-19) em Portugal. Relatório de situação: 21 de Dezembro de 2021, year = 2021, url = https://insaflu.insa.pt/covid19/relatorios/INSA$_{S}ARS_{C}oV_{2D}IVERSIDADE_{G}ENETICA_{r}elatorio_{s}ituacao_{2}021-12-21.pdf, urldate = 2021-12-21,.$

[50] TSF. Portugal retoma voos com Moçambique nos próximos dias. Quarentena e dois testes obrigatórios, year = 2021, url = https://www.tsf.pt/portugal/sociedade/portugal-retoma-voos-com-mocambique-nos-proximos-dias-quarentena-e-dois-testes-obrigatorios-14395836.html, urldate = 2021-12-10.

[51] Bernal, J. L. et al. Interrupted time series regression for the evaluation of public health interventions: a tutorial. *International Journal of Epidemiology*, 2017.

[52] Chen, Y. A new methodology of spatial cross-correlation analysis. *PLoS One*, 2015.

[53] V. of America. Omicron Variant Was in Western Europe Days Before First Identified in Southern Africa, year = 2021, url = ttps://www.voanews.com/a/omicron-variant-was-in-western-europe-days-before-first-identified-in-southern-africa/6333427.html, urldate = 2021-11-30.