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

Maria Emília Alves Matos Folque
maria.folque@tecnico.ulisboa.pt

Instituto Superior Técnico, Lisboa, Portugal

June 2023

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

1. Introduction

1.1. Motivation

On December 31^{st}, 2019, Wuhan Municipal Health Commission reported the first cluster of cases of what would eventually be identified as a novel coronavirus [1]. COVID-19 (Coronavirus Disease-2019), transmitted by respiratory droplets, quickly spread worldwide, reaching Portugal on March 2^{nd}, 2020 [2][3]. On March 11^{th}, the World Health Organization (WHO) declared a pandemic, and governments worldwide were forced to take various preventive measures.

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, and the effectiveness of treatments [4]. In late 2020, certain variants posed a heightened threat to public health which prompted WHO to classify some as Variants of Concern (VOC). As of the beginning of 2022, the most significant variants were: alpha, beta, gamma, delta, and omicron.

In the backend of 2021, the omicron variant was first discovered in South Africa and it rapidly spread worldwide, reaching Portugal on November 17^{th} [5]. This variant quickly rose to prominence, becoming the dominant variant in Portugal by the second to last week of the year (from December 20^{th} to 26^{th}, 2021).

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 [6]. 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.

Previous works show 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. 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 are 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 transactions. By analyzing spatial 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 29^{th}, Portugal decided to ban air traffic coming from South African countries.

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. 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

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% [7][8]. 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 [9]. Even a thoroughly focused travel ban might also not be enough given the highly interconnected world we live in nowadays [10]. On account of this, border restrictions were found to be more effective when coupled with mandatory quarantine and screening [11].

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 [12].

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 [12]. 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 [13] and even the social media platform, Twitter [14], 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 [15][16].

Understanding the spatiotemporal intricacies of COVID-19 became critical in the effort to continue with a flexible strategy to mitigate its spread.

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 [17].

Geographic analyses that mapped information to all administrative levels were carried out worldwide. The data was used to create dashboards [18], monitor epidemiology trends [19], help find hotspots and outbreaks [20][21], 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.

Moran´s I was used to try to map the spatial distribution of COVID-19 cases and the pandemic spread rate [22][23]. In Brazil, Moran's I was 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 [24].

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 [25]. With these types of examinations, authorities can focus on more at-risk communities and politicians can implement specific measures that help alleviate their circumstances.

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 [26].

In 2022, a study accurately predicted daily COVID-19 cases by employing aggregate mobility statistics collected from Google's Community Mobility Report [27]. The year prior, researchers had found evidence that movement patterns are correlated with the level of transmission [28]. 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 there 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 [29].

2. Dataset Description

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. It spans between November 15^{th}, 2021, and December 26^{th}, 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.

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 per 100,000 inhabitants was calculated for each municipality.

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. 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. Once again, it is important to reiterate, that just like the COVID-19 Data, there are no individual identifiers, and all transfers are anonymous.

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. Once again, 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, interrupted time series and spatial cross-correlation.

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

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.

Lastly, that number of transfers was transformed into a rate per 100,000 inhabitants.

3. Methodology

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 29^{th} to December 13^{th}, 2021 . 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 [30].

In order to conduct this methodology three variables are required. 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). X_t, a dummy variable indicating the pre or post-intervention periods, it has the value of 1 for the two weeks the travel ban was in effect and 0 for the weeks before and after that. And Y_t, the outcome at time t.

The following regression model was used:

$$Y_t = \beta_0 + \ \beta_1 T \ + \beta_2 X_t + \beta_3 T X_t \,. \tag{1}$$

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.

3.2. Spatial Cross-Correlation

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

A paper by Chen, proposing 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 [31].

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

$$X_C = X - \mu_x \,, \tag{2}$$

$$Y_C = Y \ - \ \mu_y \ \,, \tag{3}$$

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

$$\mu_x = \frac{1}{n} \sum_{i=1}^{n} x_i \,, \tag{4}$$

$$\mu_y = \frac{1}{n} \sum_{i=1}^{n} y_i \,, \tag{5}$$

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

$${\sigma_x}^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i \ - \ \mu_x)^2 = \frac{1}{n} \ X_c^T \ \ X_c \,, \tag{6}$$

$${\sigma_y}^2 = \frac{1}{n} \sum_{i=1}^{n} (y_i \ - \ \mu_y)^2 = \frac{1}{n} \ Y_c^T \ \ Y_c \,, \tag{7}$$

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

$$x = \frac{X - \mu_x}{\sigma_x} = \frac{X_c}{\sigma_x} \,, \tag{8}$$

$$y = \frac{Y - \mu_y}{\sigma_y} = \frac{Y_c}{\sigma_y} \,. \tag{9}$$

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

$$W = [w_{ij}]_{nn} \,. \tag{10}$$

This matrix was created using a spatial contiguity matrix and it has three main properties: symmetry; zero diagonal elements, meaning the entries in the diagonal are all 0; and unitization condition, meaning the sum of all matrix entrances is 1.

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

$$R_c = x^T W y = y^T W x \,. \tag{11}$$

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

$$f^{(xy)} = x y^T W x \,, \tag{12}$$

$$f^{(yx)} = y x^T W y \,, \tag{13}$$

$$f^{(xx)} = x x^T W y \,, \tag{14}$$

$$f^{(yy)} = y y^T W x \,, \tag{15}$$

$$f^{(x)} = n W x \,, \tag{16}$$

$$f^{(y)} = n W y \,. \tag{17}$$

Scatterplot	Abscissa	Ordinate	Trend Line
1^{st} plot	x	$f^{(y)}$	$f^{(xy)}$
2^{nd} plot	x	$f^{(y)}$	$f^{(xx)}$
3^{rd} plot	y	$f^{(x)}$	$f^{(yx)}$
4^{th} plot	y	$f^{(x)}$	$f^{(yy)}$

Table 1: Variables relations for spatial cross-correlation scatterplots.

Table 1 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.

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

$$R_0 = x^T W_0 y = y^T W_0 x\,, \tag{18}$$

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

$$R_p = R_0 - R_c\,. \tag{19}$$

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.

4. Results and discussion

As this research was conducted using aggregated weekly data, Table 2 shows each week's corresponding dates.

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

Table 2: 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.

Figures 1 and 2 show the evolution curves of both the weekly number of confirmed omicron cases and the weekly number of South African card transactions in continental Portugal.

Figure 1: Omicron cases curve.

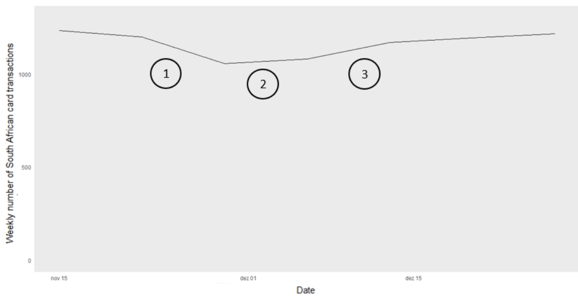


Figure 2: South African card transactions curve.

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 steep and quick rise (sections 2 and 3). The transactions curve displays a decrease in the number of transactions made in the middle of the 6-week period (section 2). This drop coincides with the travel ban imposed by the country on Southern African flights.

To further comprehend the spread of both variables, heat maps of continental Portugal were made. 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 3 and 4.

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.

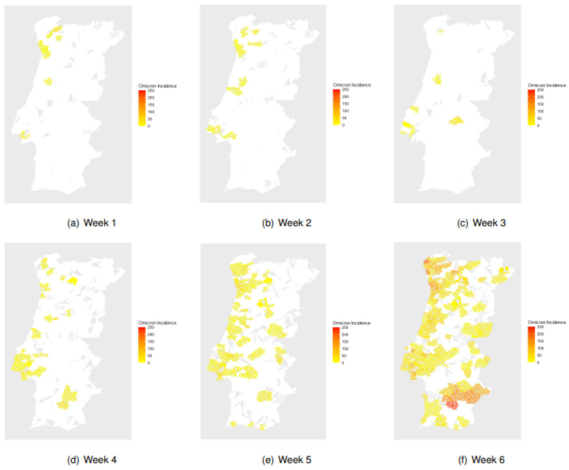


Figure 3: Evolution of omicron incidence.

Figure 4: Evolution of transactions incidence.

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 5.

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}.

There was an abrupt fall in the number of transactions during the period when travel restrictions were imposed. However, 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 5 than in the previous figures because for this travel

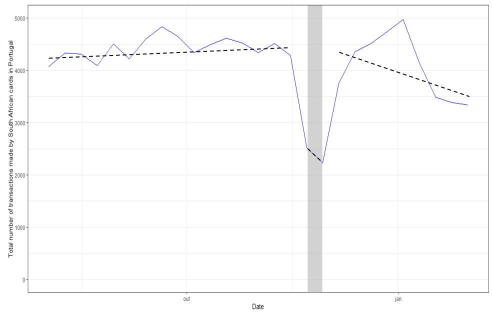


Figure 5: Interrupted time series of the total number of transactions from South African cards in Portugal, with the travel ban highlighted from November 29^{th} to December 13^{th}, 2021.

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 even after the ban was lifted.

Although Figure 5 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 try 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.

Figures 6 to 9 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 3 through 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.

Since COVID-19 can take some days to develop and detect this analysis was carried out over a period of time larger than a week. 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, the ones that recorded transactions in at least three of the weeks.

4.3.1 Robustness Test 1

Figures 6 and 7 and Tables 3 and 4 show the results of the study conducted on all 278 municipalities of continental Portugal.

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

Since the first omicron hotspot worldwide was in South Africa it would be plausible to assume there could be a connection between South Africans' card 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 found in the results.

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. 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.

4.3.2 Robustness Test 2

Figures 8 and 9 and Tables 5 and 6 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.

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

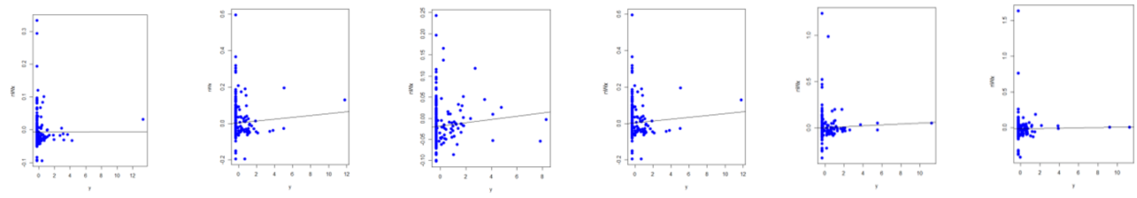
	Week 2	**Week 3**	**Week 4**	**Week 5**	**Week 6**
SCC	-1,39E^{-1}	1,83E^{-2}	5,44E^{-2}	-9,79E^{-2}	1,38E^{-1}
SCI	8,64E^{-4}	-7,62E^{-4}	7,77E^{-4}	3,91E^{-4}	-2,71E^{-4}
PSCC	-1,39E^{-1}	1,90E^{-2}	5,36E^{-2}	-9,83E^{-2}	1,38E^{-1}
R^2	2,33E^{-2}	2,30E^{-2}	1,52E^{-2}	8,15E^{-3}	2,72E^{-3}
F	6,61	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 6: Correlation coefficients values for 64 municipalities and lag=1.

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.

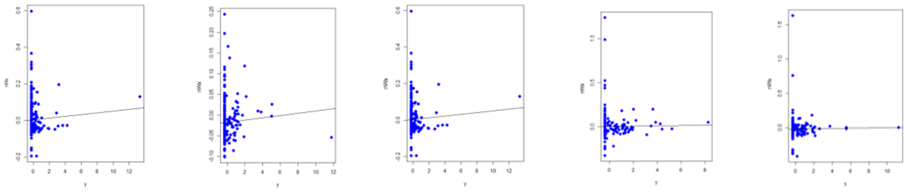
The same fluctuation between positive and negative values that occurred in the first study can be seen in the second one. However, the relation between a negative SCC and PSCC and the overall decrease in South African transactions is not as clear as it was for all 278 municipalities. Unlike in the previous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

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



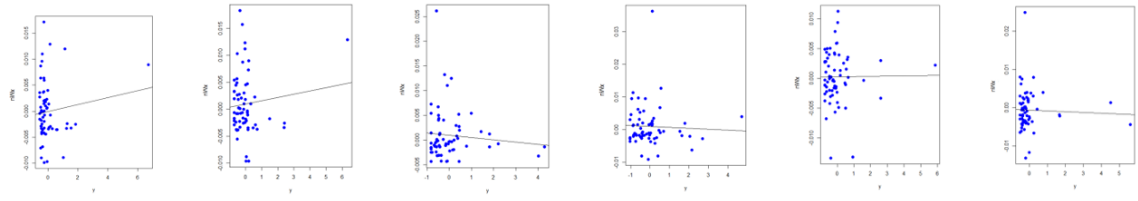
(a) Week 1 (b) Week 2 (c) Week 3 (d) Week 4 (e) Week 5 (f) Week 6

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



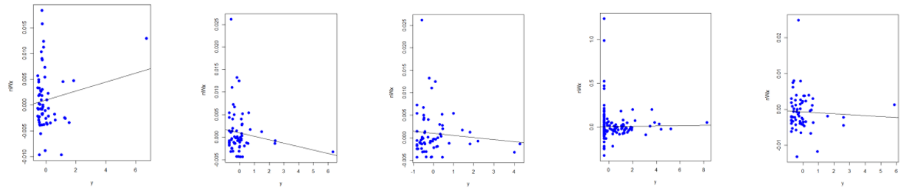
(a) Week 2 (b) Week 3 (c) Week 4 (d) Week 5 (e) Week 6

Figure 7: Spatial cross-correlation scatterplots for omicron incidence, for 278 municipalities, from weeks 1 to 5 with lag=1.



(a) Week 1 (b) Week 2 (c) Week 3 (d) Week 4 (e) Week 5 (f) Week 6

Figure 8: Spatial cross-correlation scatterplots for omicron incidence, for 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 9: Spatial cross-correlation scatterplots for omicron incidence, for 64 municipalities, from weeks 1 to 5 with lag=1.

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).

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.

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.

After the literature review, 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 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. 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. 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.

Due to all the limitations expressed, 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

5. Conclusions

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? An ITS and spatial cross-correlation analyses were the methodologies chosen to do so.

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 card transactions could ever be considered a good method to predict the proliferation of future pandemics.

References

[1] World Health 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] World Health 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.

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

[5] World Health 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.

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

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

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

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

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

[11] 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.

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

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

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

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

[16] 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.

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

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

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

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

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

[22] 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.

[23] 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.

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

[25] 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.

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

[27] 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.

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

[29] 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.

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

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