Statistical approach and territorial analysis on the Covid-19 emergency: the case study of the municipality of Tito

2023 IEEE International Workshop on Metrology for Living Environment (MetroLivEnv) | 978-1-6654-5693-7/23/$31.00 ©2023 IEEE | DOI: 10.1109/MetroLivEnv56897.2023.10164005

Fausto Bisaccia
Department of Mathematics, Computer Science and Economics
University of Basilicata
Potenza, Italy
fausto.bisaccia@studenti.unibas.it

Valentina Santarsiero
School of Engineering, Laboratory of Urban and Regional System Engineering (LISUT)
University of Basilicata
Potenza, Italy
valentina.santarsiero@unibas.it

Simone Corrado
School of Engineering, Laboratory of Urban and Regional System Engineering (LISUT)
University of Basilicata
Potenza, Italy
simone.corrado@unibas.it

Luigi Santopietro
School of Engineering, Laboratory of Urban and Regional System Engineering (LISUT)
University of Basilicata
Potenza, Italy
luigi.santopietro@unibas.it

Antonella Iuliano
Department of Mathematics, Computer Science and Economics
University of Basilicata
Potenza, Italy line 4: City, Country
antonella.iuliano@unibas.it

Beniamino Murgante
School of Engineering, Laboratory of Urban and Regional System Engineering (LISUT)
University of Basilicata
Potenza, Italy
beniamino.murgante@unibas.it

***Abstract*—*This paper analyzes the impact of the Covid-19 pandemic on Tito, a small municipality in Basilicata, Italy. Descriptive and inferential statistical models were used to study the course of the pandemic and predict its future development. The paper examines the possible consequences of proximity to a Site of National Interest on the spread of infection. The linear regression model was used to analyze the above phenomena. The paper also illustrates the multiple linear regression model and how it was constructed, considering the qualitative explanatory variables and ANOVA. The models were applied to data collected in the Municipality of Tito to test the trend of infection within the municipal area, with particular attention to the variables of gender, age, area of residence and vaccination status.*

Keywords*—*Covid-19, Spatial planning, Simple Linear Regression, Interested National Site

I. INTRODUCTION

The propagation of Covid-19 brought dramatic consequences worldwide, causing the death of millions of people, overcrowding of hospitals, and long-term health problems (a phenomenon identified as 'long covid'). Furthermore, with the introduction of mass vaccination, despite an initial semblance of newfound normalcy, the effects of the pandemic are still present, and the implementation of restrictive measures to curb contagion, such as the implementation of lockdown, has caused social and economic hardship [1]. During this period, the media and daily reports have familiarized us with statistical terms and objects such as the total number of positives, cured, deceased, and swabs taken, or, to use terms such as positivity rate, mortality rate, and percentage vaccinated in the total population. The statistical models used in this work make it possible to analyze the trend of the phenomenon and try to make predictions about its possible future trends [2]. Statistical models make it possible to study how a certain factor may influence the trend of another and to understand how they are correlated. In particular, the regression model was used for this analysis to assess the course of the pandemic and predict its future developments to provide a general picture of the spread of the virus in the municipality of Tito.

The Covid-19 pandemic includes a series of other serious respiratory diseases that have appeared in recent years, all with a very high mortality rate and severe symptoms. To keep the evolution of the epidemic under control, reference was made to the R0 indicator, the average number of secondary infections produced by an already infected individual who came into contact with a population that had never been exposed to a similar pathogen before. For Covid-19, on the other hand, the R0 index fluctuates from a value of 1, in which the epidemic is considered controllable, up to 3 in periods of maximum virus concentration. This means that an infected individual can infect two or three other people.

This work aims to apply the models of descriptive and inferential statistics to a limited context, such as a small municipality in Basilicata. This paper aims to apply the models of descriptive and inferential statistics to a limited context, such as that of a small municipality in Basilicata. For the dispicline of urban and regional planning, it is important to analyze small and disadvantaged contexts. Understanding the dynamics in these contexts is useful for targeting funding and developing place-based policies, e.g., SNAI [3–5].

Not all countries have been affected in the same way by Covid-19. Therefore we illustrate how the health crisis has translated into economic terms with particular attention to the measures taken in Italy [6]. Finally, the trend of the Basilicata region is briefly presented, analyzing some of the main indicators of contagion. After the first phase of the emergency, the debate on the economic and social consequences that the pandemic would bring was ignited.

The economic repercussions of the pandemic depend on the direct and indirect effects of its supply and demand. One of the immediate effects of the Covid-19 pandemic has been

the sudden negative shock to supply, i.e., the shutdown of all production chains deemed non-essential. The lockdown, with the closure of plants and factories in most production sectors, led to the halt of production in certain geographical areas, which also led to negative consequences for other geographical areas and other production sectors. This negative shock also corresponds to the demand-side shock. With the halt in production activities, workers' incomes also suffered a downsizing, which led to a sharp drop in household consumption, mainly due to the uncertainty about the duration of the pandemic and the restrictive measures adopted. The epidemic has led investors to no longer take investment risks.

Consequently, banks are also feeling the pandemic's adverse effects because they can no longer give credit to their customers. Following the strict containment measures in Italy [7,8], the economic effects were immediately severe and with a very marked impact compared to other countries in the Euro area. This depended on various factors. The epidemic spread in Italy a few weeks earlier than in other European countries; this led to having to impose containment measures for a period longer with direct impacts on those sectors most affected by the shutdown of activities.

The phenomena illustrated above were analyzed through the application of a statistical model, in particular, the linear regression model [9].

II. Materials and Methodology

The Simple Linear Regression Model was first analyzed, where a quantitative phenomenon Y is explained by a single explanatory variable X. This model can be used for descriptive and interpretative purposes of a phenomenon, for simulation purposes (evaluate Y as a function of alternative scenarios of X) and for purposes of analysis of a real problem, such as the COVID-19 epidemic, starting only from the study of a variable. However, it is difficult to admit that a single quantitative variable X can explain a quantitative phenomenon Y; in fact, most real phenomena depend on and are influenced by several variables. The introduction of $p > 1$ explanatory variable leads to the Multiple Linear Regression Model. To describe how this latter model is built, all the phases already examined in the simple case are taken up, paying attention to all the differences and complications that this entails in all parts of the building process and verifying the validity of the Model. This first methodological phase concludes by considering the case in which the explanatory variables are qualitative and how they can influence the model construction and evaluation phase, specifying the Analysis of Variance (ANOVA). The models described were applied to data relating to the health emergency caused by Covid-19 in the Municipality of Tito to verify the infection trend within the municipal area [10]. The data collected derives from the database present at the municipal offices of the Municipality processed using descriptive and inferential statistical techniques.

The first part of the work focused on the descriptive and territorial analysis of the data collected to specify which part of the population was most affected by the virus. The variables taken into consideration are Gender, Age, the average duration of the disease, the area of residence of the individuals positive for the virus, and other available variables, with particular attention to the periods in which there was the highest number of infections. Furthermore, from the database available with Vaccinations data, it was possible to verify whether and to what extent a vaccinated individual (with one or more doses) has a shorter duration of the disease than an unvaccinated individual and, therefore, how the vaccines have influenced the progress of the infection. This led to an overall description of the epidemic trend in the municipality, considering all the variables described. Finally, after specifying the independent and dependent variables, the Multiple Linear Regression Model was built and applied to verify which variables most influenced the contagion's spread within the Municipality, confirming what was reported in the descriptive analysis.

III. Result and Discussion

The aim was to carry out a statistical investigation, based on the Multiple Linear Regression Model, on how the virus was introduced and spread in the population and how it spread, which part of the population was most affected, and if the vaccine is effective in decreasing the likelihood of the spread and severity of the disease.

With the start of the vaccination campaign in the municipality by the Potenza Health Authority (ASP), the average duration of isolation has been slowly decreasing for those at least 80 years old. In fact, out of a sample of 273 individuals, we see that more people were vaccinated than unvaccinated. The number of people who got at least one dose of vaccine is 189, equal to 69 % of the sample, compared with several unvaccinated people, 84, equal to 31%. Of these, 105 are Women (56 % of the population), with an average age of around 41 years; 84 are Men (44 % of the population), with an average age of 39. The data concerning the frequencies of the variable Gender are shown in Table 1 below.

TABLE I. The number of people who got at least one dose of vaccine categorized by gender.

Frequencies of variable Gender			
Gender	***Absolute frequency***	***Relative frequency***	***Percentage***
Male	84	0.444	44.44%
Female	105	0.565	55.56%
Total	189	1	100%

Indeed, in January 2022, when there was a significant increase in the number of new positive cases, 415 positives were detected, including 192 men (46%) and 223 women (64%). Considering the isolation period factor, the data was also analyzed relative to the spatial localization. As depicted in the following map, the infected in "Zone D" show a more significant isolation period. This evidence may be influenced by the proximity to the industrial area in which the Tito Interested National Site (INS) resides, see Fig. 1.

Moreover, the results obtained from the study descriptive of the phenomenon confirm those obtained in the inferential analysis. The variables found to be most significant and with more extraordinary predictive ability on days of isolation are the "November 2020 period" and "Zone D" variable. In addition to the strong relationship between isolation days and the period of November 2020, in which the highest peak of Covid-19 cases was observed, "Zone D," in which some Rest Homes are located, was found to be one of the most significant variables, confirming how the virus was more

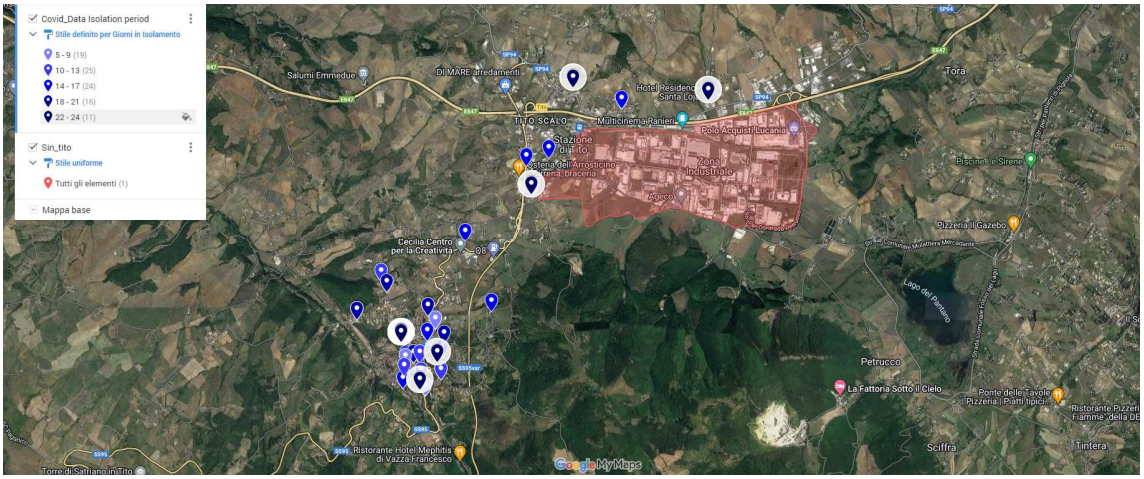

Fig. 1. Georeferenced Covid-19 isolation period data and in red the area of Tito Interested National Site.

concentrated in this zone leading to a more substantial number of isolation days.

The statistical model shows that the variables "Age" and "Sex M" are also suitable for explaining the dependent variable variations. Furthermore, we have analyzed the phenomenon in terms of spatial diffusion and by observing the relationship between the epidemic and the critical environmental area. Thus, we can state that the study undertaken on the evolution of Covid-19 infection in the Municipality of Tito during the period from September 2020 to March 2022 has led to significant results that can summarize what happened in those months, focusing on the population groups most affected and the places most subjected, also for the nearby INS of industrial area of Tito. This study thus made it possible to give a more complete and accurate picture of the pandemic in the various areas of the municipality and the need of the evaluation of spatial variables for multiple applications [11]. In conclusion, spatial analysis of data associated with COVID-19 may be useful for:

- Outbreak identification: spatial analysis can help identify geographic areas where high numbers of COVID-19 cases occur, allowing public health authorities to focus on the most affected areas.
- Resource planning: spatial analysis can help understand where more health resources are needed, such as hospitals, health personnel, personal protective equipment, and medical supplies. This allows for better planning of resource distribution to meet patient needs.
- Policy evaluation: spatial analysis can help evaluate the effectiveness of health policies adopted to prevent the spread of COVID-19 in different geographic regions.
- Identification of trends: spatial analysis can help identify trends and patterns in the spread of COVID-19 in different geographic areas, enabling public health authorities to adopt effective strategies to prevent the spread of the virus in the future.

REFERENCES

1. Cartenì, A.; Di Francesco, L.; Martino, M. How mobility habits influenced the spread of the COVID-19 pandemic: Results from the Italian case study. Sci. Total Environ. 2020, 741, 140489, doi:10.1016/j.scitotenv.2020.140489.
2. Chu, J. A statistical analysis of the novel coronavirus (COVID-19) in Italy and Spain. PLoS One 2021, 16, e0249037, doi:10.1371/journal.pone.0249037.
3. Corrado, S.; Scorza, F. Machine Learning Based Approach to Assess Territorial Marginality. In Proceedings of the Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics); Springer Science and Business Media Deutschland GmbH, 2022; Vol. 13376 LNCS, pp. 292–302.
4. Las Casas, G.; Murgante, B.; Scorza, F. Regional local development strategies benefiting from open data and open tools and an outlook on the renewable energy sources contribution. Green Energy Technol. 2016, 0, 275–290, doi:10.1007/978-3-319-31157-9_14.
5. Barca, F.; Mccann, P.; Rodríguez-Pose, A. THE CASE FOR REGIONAL DEVELOPMENT INTERVENTION: PLACE-BASED VERSUS PLACE-NEUTRAL APPROACHES*. J. Reg. Sci. 2012, 52, 134–152, doi:10.1111/J.1467-9787.2011.00756.X.
6. Dettori, M.; Deiana, G.; Balletto, G.; Borruso, G.; Murgante, B.; Arghittu, A.; Azara, A.; Castiglia, P. Air pollutants and risk of death due to COVID-19 in Italy. Environ. Res. 2021, 192, doi:10.1016/j.envres.2020.110459.
7. Murgante, B.; Borruso, G.; Balletto, G.; Castiglia, P.; Dettori, M. Why Italy first? Health, geographical and planning aspects of the COVID-19 outbreak. Sustain. 2020, 12, 5064, doi:10.3390/su12125064.
8. Beniamino, M.; Balletto, G.; Giuseppe, B.; Lucia, S.; Angela, P.; Francesco, S.; Castiglia, P.; Antonella, A.; Dettori, M. A methodological proposal to evaluate the health hazard scenario from COVID-19 in Italy. Environ. Res. 2022, 209, 112873, doi:10.1016/j.envres.2022.112873.
9. Ricciardi, L.M.; Crescenzo, A. Elementi di statistica 2000.
10. Khan, M.A.; Khan, R.; Algarni, F.; Kumar, I.; Choudhary, A.; Srivastava, A. Performance evaluation of regression models for

COVID-19: A statistical and predictive perspective. Ain Shams Eng. J. 2022, 13, 101574, doi:10.1016/j.asej.2021.08.016.

11. Santarsiero, V.; Nolè, G.; Scorza, F.; Murgante, B. Evaluation of Spatial Variables Related to the Provision of Essential Services in the Basilicata Region. In Proceedings of the Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics); 2022; Vol. 13376 LNCS, pp. 344–353.