2026 | Faculty of Sciences

UHASSELT
KNOWLEDGE IN ACTION

Doctoral dissertation submitted to obtain the degree of
Doctor of Sciences: Statistics, to be defended by

Yessika Adelwin Natalia

DOCTORAL DISSERTATION

Exploring sociodemographic drivers of the COVID-19 pandemic in Belgium: Insights from hierarchical and fractal-based spatiotemporal models

Promoter: Prof. Dr Geert Molenberghs | UHasselt – KU Leuven

Co-promoters: Prof. Dr Christel Faes | UHasselt
Prof. Dr Thomas Neyens | UHasselt – KU Leuven

www.uhasselt.be
Hasselt University
Martelarenlaan 42 | BE-3500 Hasselt

2026| Faculty of Sciences.

UHASSELT
KNOWLEDGE IN ACTION

Doctoral dissertation submitted to obtain the degree of
Doctor of Sciences: Statistics, to be defended by

Yessika Adelwin Natalia

DOCTORAL DISSERTATION

Exploring sociodemographic drivers of the COVID-19 pandemic in Belgium: Insights from hierarchical and fractal-based spatiotemporal models

Promoters: Prof. Dr Geert Molenberghs | UHasselt – KU Leuven

Co-promoters: Prof. Dr Christel Faes | UHasselt
Prof. Dr Thomas Neyens | UHasselt – KU Leuven

DECLARATION

The PhD researcher and the UHasselt supervisor hereby formally declare that the research conducted for the purpose of this PhD thesis was executed in accordance with the principles of good scientific conduct, as stipulated in the UHasselt Integrity charter, the UHasselt charter supervisor - PhD Researcher, the UHasselt Integrity Policy and the UHasselt guidelines for the use of (generative) AI in research.

The author asserts that this PhD thesis is made Open Access immediately upon submission. The full text is publicly available without restrictions.

During the preparation of this PhD thesis the author used ChatGPT (OpenAI, GPT-5.2) to improve language, readability, and grammar. After using this tool, the author reviewed and edited the content as needed and take full responsibility for the content of this thesis.

Fear of the Lord is the foundation of true knowledge,
but fools despise wisdom and discipline.
— Proverbs 1:8

ACKNOWLEDGMENT

First and foremost, I want to praise God, Jesus Christ, for His grace and mercy that have enabled me to finish this race. To God be the glory.

I would like to sincerely thank my wonderful supervisors: Prof. Dr. Geert Molenberghs, Prof. Dr. Christel Faes, and Prof. Dr. Thomas Neyens, who guided me through thick and thin and shared their valuable knowledge and expertise. Back in 2020, I began this PhD journey like a baby, with little to no knowledge of Bayesian spatiotemporal analysis, longitudinal models, or fractal dimension. Their valuable insights really helped me charting this old and new territories. While I am still far from an expert, I am truly grateful to have learned this side of statistics under their guidance. In addition, my gratitude goes to all other jury members: Prof. Dr. Niel Hens, Prof. Dr. Johan Verbeeck, Prof. Dr. Marta Blangiardo, dr. Naïma Hammami, Prof. Dr. Philippe Beutels, Prof. Dr. Mathias Dewatripont, Prof. em. Dr. Noel Veraverbeke for their constructive comments and feedback that greatly improved an earlier version of this thesis.

Aan mijn (ex-)collega assistenten: Maren, Annelies, Leyla, Helene, Kato, Sara, Wouter, Hanne. Bedankt voor jullie steun en samenwerking tijdens mijn assistenttraject. Het doet deugd om te weten dat ik bij jullie altijd terechtkan voor een goede babbel na een zware werkzitting.

To (ex-)office mates turned friends who ever shared a cozy working space with me: Elisa, Stijn, Neilshan, Astrid, Liz, Bryan. Thank you for the good times and small chats while waiting for my models to run or when I faced writer's block.

Ik ben heel erg dankbaar voor mijn lieve familie in België: mijn man

Martijn, dochter Meiko, papa Erik, mama Ghislaine en zus Elien. Het was een zwaar traject van zes jaar met bloed, zweet en tranen (letterlijk en figuurlijk). Zonder jullie steun had ik dit niet kunnen volbrengen.

Untuk Papa, Mama, William, dan Bryan. Meskipun kalian jauh di mata, tapi selalu dekat di hati. Aku yakin dan percaya bahwa doa kalian selalu menyertai perjalananku di sini. Sahabat setiaku Ala, terima kasih telah menjadi bagian dari support system aku sejak zaman SMA hingga kini kita sudah berkeluarga. Mari kita rencanakan sesi curcol berikutnya.

Last but not least, I would like to thank all members of DSI, whom I cannot mention individually. From the very first day, I felt genuinely welcomed and throughout the years, DSI has been a warm and supportive environment to grow both professionally and personally. The openness, collaboration, and sense of community have made it a truly comfortable place to work and learn. I am deeply grateful to have been part of this environment and I will always look back at the past six years with great appreciation. DSI has truly felt like a home away from home.

Yessika Adelwin Natalia
Bree, 29 June 2026

PUBLICATIONS

Materials presented throughout this thesis are based on the following publications:

Natalia, Y. A., Faes, C., Neyens, T., Molenberghs, G. (2022). The COVID-19 wave in Belgium during the Fall of 2020 and its association with higher education. *PLoS ONE* 17(2): e0264516. DOI: 10.1371/journal.pone.0264516.

Natalia, Y. A., Faes, C., Neyens, T., Chys, P., Hammami, N., Molenberghs, G. (2023). Fractal dimension based geographical clustering of COVID-19 time series data. *Sci Rep* 13, 4322. DOI: 10.1038/s41598-023-30948-7.

Natalia, Y. A., Faes, C., Neyens, T., Hammami, N., Molenberghs, G. (2023). Key risk factors associated with fractal dimension based geographical clustering of COVID-19 data in the Flemish and Brussels region, Belgium. *Front. Public Health* 11:1249141. DOI: 10.3389/fpubh.2023.1249141.

Natalia, Y. A., Molenberghs, G., Faes, C., Neyens, T. (2024). Geospatial patterns of excess mortality in Belgium: Insights from the first year of the COVID-19 pandemic. *Spatial and Spatio-temporal Epidemiology* 49:100660. DOI: 10.1016/j.sste.2024.100660.

Natalia, Y. A., Verbeeck, J., Faes, C., Neyens, T., Molenberghs, G. (2024). Unraveling the impact of the COVID-19 pandemic on the mor-

tality trends in Belgium between 2020–2022. *BMC Public Health* 24, 2916. DOI: 10.1186/s12889-024-20415-x.

Natalia, Y. A., Molenberghs, G., Neyens T., Hens, N., Faes, C. (2025) Empirical analysis of COVID-19 confirmed cases, hospitalizations, vaccination, and international travel across Belgian provinces in 2021. *PLoS One* 20(5): e0322017. DOI: 10.1371/journal.pone.0322017

Related publications briefly discussed in this dissertation:

Brackx, F., Vanongeval, F., **Natalia, Y. A.**, Molenberghs, G., Steenberghen, T. (2022). The effect of transborder mobility on COVID-19 incidences in Belgium. *International Journal of Environmental Research and Public Health* 19(16), 9968. DOI: 10.3390/ijerph19169968.

Natalia, Y. A., Delporte, M., De Witte, D., Beutels, P., Dewatripont, M, Molenberghs, G. (2023). Assessing the impact of COVID-19 passes and mandates on disease transmission, vaccination intention, and uptake: a scoping review. *BMC Public Health* 23, 2279. DOI: 10.1186/s12889-023-17203-4.

Publications not covered in this dissertation:

Natalia, Y. A., Herzog, S.A., Neyens, T., Zurl, C.J., Strenger, V., Molenberghs, G., Faes, C. (2025). COVID-19 RT-qPCR-based screening in Austrian schools and incidences in the general population: a Bayesian spatiotemporal analysis. *Arch Public Health* 83, 152. DOI: 10.1186/s13690-025-01655-8.

Natalia, Y. A., De Cauwer, H., Neyens, T., Goniewicz, K., Somville, F., Molenberghs, G. (2025). A Bayesian network analysis of aviation terrorism attack risks. *J Transp Secur* 18, 19. DOI: 10.1007/s12198-025-00315-w.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

This is a list of abbreviations used in the thesis.

AI	:	Artificial intelligence
AIC	:	Akaike information criterion
ARIMA	:	Autoregressive integrated moving average
CAR	:	Conditional autoregressive
CCA	:	Canonical correlation analysis
COVID-19	:	Coronavirus disease 2019
DIC	:	Deviance Information Criterion
ECDC	:	European Centre for Disease Prevention and Control
GIS	:	Geographic information system
ICD	:	International statistical classification of diseases and related health problems
INLA	:	Integrated nested Laplace approximation
ISO	:	International Organization for Standardization
LOESS	:	Locally estimated scatterplot smoothing
MCMC	:	Monte Carlo Markov chain
ML	:	Machine learning
PC	:	Penalized complexity
RSV	:	Respiratory syncytial virus
SARS-CoV-2	:	Severe acute respiratory syndrome coronavirus 2
SE	:	Standard error
SI	:	Stringency index
SMR	:	Standardized mortality ratio
SS	:	Spatially structured
SU	:	Spatially unstructured
WAIC	:	Watanabe-Akaike Information Criterion

WHO	:	World Health Organization
YLL	:	Years of life lost

VARIABLES AND RESOLUTIONS

This is a summary of the variables and resolutions used throughout the thesis.

Chapter	Study period	Outcome	Explanatory variables	Spatial unit	Temporal unit
2	1 June – 31 December 2020	COVID-19 new cases	Age, sex, income, population density, proportion of higher education student	Municipality	15 days
	1 January – 31 December 2020	Excess mortality	Age, sex, COVID-19 incidence	Municipality	6 months
3	1 January – 31 December 2021	COVID-19 new cases & hospitaliza-tions	Vaccination coverage, incoming travel rate, positivity rate, stringency index	Province	7 days
	1 January 2020 – 31 December 2022	Excess mortality	Age, sex, heatwave	Country	7 days
	1 January 2020 – 31 December 2021	Cause-specific mortality	Age, sex, stringency index	Country	7 days

Chapter	Study period	Outcome	Explanatory variables	Spatial unit	Temporal unit
4	1 January – 31 December 2021	COVID-19 new cases	Shannon index, income, population size, population density, proportion of older age population, vaccination coverage, trust in the federal government, trust in the regional government, satisfaction with the healthcare system	Statistical sector	Daily

CHAPTER 1

GENERAL INTRODUCTION

The year 2020 holds special significance for me, as it marked the intersection of two transformative events in my life: the beginning of my journey as a doctoral candidate and the onset of the global coronavirus disease 2019 (COVID-19) pandemic. The unprecedented nature of this health crisis reshaped social norms, research priorities, and public health perspectives. In this context, my work naturally evolved to address pressing questions related to disease transmission and monitoring methodologies.

To provide a foundation for understanding this research, a brief historical overview of the COVID-19 pandemic is presented in Section 1.1, highlighting its timeline, key developments, and implications for public health systems worldwide. Section 1.2 offers an introduction to the current statistical methodologies used to evaluate disease transmission dynamics, particularly in the context of spatiotemporal analysis. Recognizing the importance of monitoring disease trends not only for COVID-19 but also for other emerging and endemic diseases, Section 1.3 introduces the use of fractals and fractal dimensions as an alternative approach to capture the complexity inherent in disease transmission patterns and improve our capacity for timely public health interventions.

Finally, Section 1.4 provides an overview of the core topics covered in this thesis, setting the stage for a deeper exploration of the methodologies, findings, and implications of this research. Through this work, I aim to contribute novel insights to the field of infectious disease surveillance (with

possible application in other health related fields) and offer alternative solutions for navigating future health crises.

1.1 SARS-CoV-2 outbreak and COVID-19 pandemic

The COVID-19 pandemic, caused by the novel severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), has profoundly affected global health, economies, and general daily life. The first known cases emerged in Wuhan, China, in December 2019, presenting as atypical pneumonia (Wu et al., 2020). The virus spread rapidly, which led to a declaration of public health emergency of international concern by the World Health Organization (WHO) on January 30, 2020, followed by the declaration of a pandemic on March 11, 2020 (World Health Organization, 2020b). In response to the escalating crisis, countries around the world implemented measures such as lockdowns, travel restrictions, and social distancing to curb transmission. Healthcare systems faced unprecedented challenges, with many overwhelmed by the surge in new cases and hospitalizations.

The heavy burden of the COVID-19 pandemic in Europe started in January 2020 when the first cases were identified in France (World Health Organization, 2020a). The virus spread rapidly, with Italy reporting a significant outbreak in February 2020, followed by other countries including Spain, France, and Germany. The first confirmed case in Belgium was reported on February 4, 2020, involving a repatriated individual from Wuhan, China (FPS Health, Food Chain Safety and Environment, 2020). Shortly in early March, community transmission was confirmed which linked to travelers returning from Northern Italy. The country experienced a rapid increase in cases during March and April 2020, leading to the implementation of strict containment measures, including the closure of schools, non-essential businesses, and restrictions on gatherings. A national lockdown was enforced to mitigate the spread.

By the end of 2020, vaccines against COVID-19 had reached the market and could be added as a prophylactic measure. Europe started the vaccination campaign in December 2020 and, as reported by the European Centre for Disease Prevention and Control (ECDC), more than 981 million vaccine doses had been administered in this region per October 5, 2023 (European Centre for Disease Prevention and Control, 2023a). Despite these efforts, the virus continued to evolve, with new variants such as Delta and Omicron emerging, necessitating ongoing public health

responses.

As of November 23, 2025, Belgium had reported approximately 4.9 million confirmed COVID-19 cases (both symptomatic and asymptomatic) and 34,339 deaths since the onset of the pandemic (Our World in Data, 2025). Like other countries around the world, Belgium continues to navigate the aftermath of the COVID-19 pandemic, focusing on vaccination efforts, public health measures, and economic recovery initiatives.

1.2 Concept of disease mapping and spatiotemporal analysis

Infectious diseases often exhibit substantial geographic variation due to differences in population or environmental characteristics. Understanding this geographical heterogeneity is a central objective of spatial epidemiology, a field that integrates epidemiological methods with spatial statistical techniques to study the geographical distribution of health outcomes and their determinants. Within this framework, disease mapping aims to estimate and visualize spatial patterns of disease risk across geographical units while accounting for differences in various population characteristics such as population size and demographic composition (Elliott and Wartenberg, 2004; Lawson, 2018). Such methods are particularly important for public health surveillance because they help identify areas of elevated risk, guide targeted interventions, and support the allocation of healthcare resources.

The origins of disease mapping can be traced back to the cholera outbreak of 1854, when British physician John Snow used spatial mapping techniques to identify the outbreak source in London. By plotting cholera deaths on a map, Snow demonstrated that the cases clustered around the Broad Street pump, providing early evidence of waterborne transmission and illustrating the value of spatial analysis in epidemiology (Tulchinsky, 2018). His work laid the foundation for modern spatial epidemiology, which was further developed by many scholars such as Moran (1950) who introduced the concept of spatial autocorrelation, and Tobler (1959) who formulated the well-known first law of geography ("everything is related to everything else, but near things are more related than distant things"). Their principles provide the conceptual basis for many spatial statistical models that explicitly incorporate spatial dependence.

Spatial epidemiology is commonly used to analyze areal data where disease counts are aggregated within predefined administrative units. Let

Y_i denote the observed number of cases in the area i and E_i the expected number of cases in that area. A common measure used in disease mapping is the relative risk θ_i, defined as observed values divided by expected values $\left(\theta_i = \frac{Y_i}{E_i}\right)$. However, disease counts may be highly unstable due to random variation, especially in small areas. Therefore, statistical models are often used to stabilize risk estimates through spatial smoothing which borrows strength from neighboring areas through the use of spatial random effects. Generally, a spatial model can be written as

$$\begin{aligned} Y_i &\sim \text{Poisson}(E_i\theta_i) \\ \log(\theta_i) &= \alpha + \nu_i + \upsilon_i \end{aligned} \tag{1.1}$$

where α represents the overall log-risk, ν_i is a spatially structured random effect capturing spatial autocorrelation between neighboring areas, and υ_i is an unstructured random effect representing independent heterogeneity (Besag et al., 1991; Lawson, 2018).

With the rise of computing technology in the 1960s and 1970s, geographic information system (GIS) emerged, allowing researchers to analyze spatial data with greater precision. During the same period, statistical methods for analyzing temporal data also advanced substantially. In particular, the autoregressive integrated moving average (ARIMA) framework introduced by Box et al. (2015) provided a systematic approach for modeling and forecasting time series data. Although these methods were initially developed for purely temporal data, they laid the groundwork for integrating temporal dynamics into spatial analytical frameworks.

While traditional disease mapping focuses primarily on spatial variation at a single point in time, many public health questions require understanding how disease patterns evolve simultaneously across space and time. This has led to the development of spatiotemporal disease mapping, which extends spatial models by incorporating temporal components and their interaction with spatial effects. These models allow researchers to analyze dynamic processes such as infectious disease spread, identify areas experiencing change in risk over time, and detect emerging clusters or outbreaks (Lawson, 2018; Blangiardo and Cameletti, 2015). Advances in statistical computing during the 1980s and 1990s considerably expanded the capabilities of spatiotemporal modeling. Developments in spatial econometrics and hierarchical Bayesian modeling enabled the construction of

more flexible models capable of handling complex dependence structures (Paelinck and Klaassen, 1979; Good, 1980). The emergence of big data, machine learning (ML), and artificial intelligence (AI) further revolutionized the field, making spatiotemporal analysis more powerful and widely accessible. Modern techniques, including spatiotemporal deep learning models, agent-based simulations, and remote sensing data analysis, have significantly improved predictive capabilities in various domains, such as climate change modeling, urban planning, and infectious disease surveillance.

1.3 Concept of fractals and fractal dimension

In recent years, the application of fractal concepts, particularly the fractal dimension, has gained increasing attention in spatiotemporal analysis. Fractals are intricate geometric structures characterized by self-similarity in which their patterns repeat on different scales (Avnir et al., 1998). Unlike traditional Euclidean shapes, which are defined by integer dimensions, fractals often exhibit complex, irregular structures that require a more nuanced mathematical description (see an example in Figure 1.1). This complexity extends the notion of dimensionality beyond whole numbers, capturing how fractals scale differently from standard geometric figures and providing a powerful framework to describe a wide range of complex real-world phenomena (Mandelbrot, 1975).

Figure 1.1. A fern is a classic example of a fractal structure, where individual leaflets exhibit self-similarity with the overall frond. The fractal dimension of a fern typically ranges between 1.7 and 1.9, reflecting its high structural complexity.

Within this context, fractal dimension plays a crucial role as a statistical measure of complexity, quantifying how details within a fractal structure change with scale (Mandelbrot, 1982), reflecting the extent to which a

fractal fills space as one zooms in on it. For example, while a line is one-dimensional and a plane is two-dimensional, a fractal curve may exhibit a fractal dimension between 1 and 2, indicating that it is more intricate than a simple curve but does not fully occupy a plane (Peitgen et al., 2004). This non-integer dimension effectively captures the space-filling capacity of fractals, making it a valuable tool for analyzing complex spatial and temporal patterns.

Historically, fractal dimension was applied mainly in fields such as geography, mathematics, and engineering (Lovejoy, 1982; Paumgartner et al., 1981; Hatlee and Kozak, 1981; Weitz et al., 1985). However, recent advances have broadened its application in diverse scientific domains, including economics, medicine, and epidemiology (Bianchi and Frezza, 2017; Kesić and Spasić, 2016; Lemmens et al., 2020). In epidemiology, the fractal dimension has been explored as a means of extracting meaningful patterns from seemingly noisy epidemic data. The underlying principle is that even highly irregular or chaotic information may have a discernible structure when analyzed through a fractal perspective. Păcurar and Necula (2020) demonstrated that the application of fractal-based approaches to epidemic curves revealed novel insights into disease outbreaks and improved predictive modeling (see Figure 1.2). This growing body of research highlights the potential of fractal analysis as a complementary tool to understand complex epidemiological dynamics and improve disease surveillance strategies.

(a) Map of Covid-19 cases in Italy (b) Map of Covid-19 cases in Spain (c) Map of Covid-19 cases in Germany

Figure 1.2. Map of potential viral dissemination over each country as reported by Păcurar and Necula (2020). The corresponding fractal dimension values are (a) 1.2193, (b) 1.2865, and (c) 1.3222. Higher values indicate more complex dissemination patterns.

1.4 Chapter overview

This thesis aims to investigate the relationship between key demographic factors, such as socioeconomic status, ethnicity, or mobility patterns, and the dynamics of the COVID-19 pandemic, with a particular focus on Belgium. The relationship is explored using both traditional spatiotemporal methods and a modified fractal-based approach. By assessing the strengths and limitations of conventional spatiotemporal models, we can identify potential gaps in existing methodologies and evaluate the added value of fractal-based techniques to improve public health interventions.

In Chapter 2, the application of spatial data analysis to COVID-19 data from 2020 is explored, with an emphasis on Bayesian methods. Bayesian approaches, which integrate prior distributions with likelihood functions to derive posterior distributions (Lesaffre and Lawson, 2012; Lawson, 2018), offer a flexible and robust framework to model spatial dependencies and uncertainties in epidemiological data. This chapter provides a practical implementation of the Bayesian spatial model in analyzing the incidence patterns of COVID-19 related to the higher education system in Belgium, as well as the impact of COVID-19 on excess mortality trends. Using these methods, spatial heterogeneity in disease spread was identified, in addition to the contribution of demographic and structural factors to variations in health outcomes in different administrative units.

Chapter 3 shifts the focus to the temporal and longitudinal aspects of the COVID-19 pandemic in 2021, particularly in relation to mobility patterns and the first year of the COVID-19 vaccination campaign. Understanding how mobility behaviors influenced viral transmission is crucial for evaluating the effectiveness of public health interventions and policy decisions. In addition, this chapter includes the three-year mortality trends in Belgium during the COVID-19 pandemic, offering a comprehensive assessment of its long-term impact on public health and population dynamics.

One of the critical challenges that emerged during an outbreak was the heavy workload faced by healthcare and data collection systems, which in turn introduced additional noise and heterogeneity into the data (Dureau et al., 2013). In Chapter 4, a fractal-based approach is proposed as a complementary methodological framework to address these challenges, especially after the vaccination campaign in 2021. Fractal analysis provides a powerful tool for examining the complexity of disease transmission dynamics, particularly at the lowest administrative levels, where spatial and

temporal variations are most pronounced. By applying fractal-based techniques, this chapter seeks to improve the characterization of noisy and fragmented data, offering deeper insights into the underlying patterns of epidemic spread and their implications for public health monitoring.

Finally, Chapter 5 synthesizes the key findings of this thesis, highlighting the methodological contributions and epidemiological insights gained through the integration of Bayesian spatial models, longitudinal analyses, and fractal-based approaches. It also outlines possible directions for future research, emphasizing the potential for further refinement and application of these methods in the study of infectious disease dynamics and beyond.

CHAPTER 2

UNCOVERING THE FIRST YEAR OF THE COVID-19 PANDEMIC: EVIDENCE FROM BAYESIAN HIERARCHICAL MODELS

This chapter is based on the following publications:

- **Natalia, Y. A.**, Faes, C., Neyens, T., Molenberghs, G. (2022). The COVID-19 wave in Belgium during the Fall of 2020 and its association with higher education. *PLoS ONE* 17(2): e0264516. DOI: 10.1371/journal.pone.0264516.
- **Natalia, Y. A.**, Molenberghs, G., Faes, C., Neyens, T. (2024). Geospatial patterns of excess mortality in Belgium: Insights from the first year of the COVID-19 pandemic. *Spatial and Spatio-temporal Epidemiology* 49:100660. DOI: 10.1016/j.sste.2024.100660.

2.1 Introduction

March 2020 remains an unforgettable moment for many people in Belgium, as it marked the first wave of the COVID-19 pandemic, which disproportionately affected the elderly population (Sciensano, 2025a). Fortunately, through collective public health efforts that include strict lockdown measures, social distancing policies, and widespread awareness campaigns, the number of reported cases and deaths declined considerably as the

country approached the summer months, particularly between late May and early July 2020.

Despite this initial improvement, a resurgence of cases was observed in August 2020. The 14-day incidence rate increased from 11 per 100,000 on July 5 to 71 per 100,000 on August 9, signaling the beginning of a second wave. Unlike the first wave in March that primarily affected older individuals, this resurgence was driven by a rise in infections among younger adults, particularly those aged 20–29 years. This shift in demographic trends aligned with the global patterns observed during the same period (Our World in Data, 2025).

Figure 2.1 illustrates the rolling seven-day sum of confirmed cases increases or decreases relative to the immediately preceding seven-day period in the second half of 2020. The figure highlights a sharp increase in early September, likely driven in part by returning travelers at the end of the summer holiday. This surge also coincided with the start of compulsory education on September 1; however, the subsequent slowdown in case growth during the second half of September provides indirect evidence that school reopening had no immediate or strong impact on transmission.

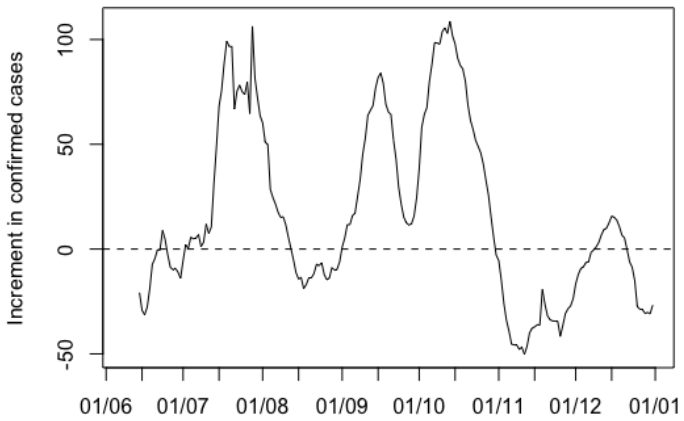


Figure 2.1. Increment function of COVID-19 confirmed cases. The % increase or decrease in the rolling seven-day sum of the number of cases is relative to the immediately preceding seven-day period. The curve starts on September 1, 2020 and runs through the end of 2020.

In contrast to compulsory education, higher education institutions reopened a little later (on September 14 in the southern Walloon Region and on September 21 in the northern Flemish Region). Some studies in the

United States identified universities as hotspots for increased transmission of COVID-19 among young adults (Wilson et al., 2020; Vang et al., 2021). Transmission in these settings was likely facilitated by the congregate living arrangements on campus and outside the campus, as well as social activities (Vang et al., 2021; Leidner et al., 2021). It is widely known that a large fraction of higher education students in Belgium reside in student accommodation during the weekdays and return home on weekends, creating frequent inter-regional mobility. Given this commuting pattern of Belgian students, it may have contributed to the acceleration of virus spread. By October 1, 2020, a substantial increase in the incidence of COVID-19 was observed, suggesting a potential link between higher education reopening and the subsequent surge in new cases.

On the other hand, the dynamic complexity of COVID-19 transmission also altered the mortality pattern in 2020. During this period, the cumulative COVID-19 mortality rate in Belgium reached 1,680.24 per million inhabitants, making it one of the highest globally (Our World in Data, 2025). However, this should be seen against the background of very extensive reporting in the country, with different reporting practices in some countries (Aron et al., 2020; Verbeeck et al., 2021). Taking into account different preventive measures and testing strategies during the pandemic, this number was not suitable to compare different countries in this period (Molenberghs et al., 2022). Therefore, it is important to consider excess mortality, i.e., the mortality above and beyond deaths that would have occurred in normal, pre-pandemic conditions (Checchi and Roberts, 2005), as an alternative metric to evaluate the impact of COVID-19 in different settings.

Demographic and socioeconomic factors are widely recognized as key determinants in the spread and impact of many diseases, including COVID-19. Evidence from various countries highlights the disproportionate burden of COVID-19 among populations with lower socioeconomic status. For instance, severely distressed communities in Hong Kong and the United States experienced significantly higher rates of COVID-19 cases and deaths compared to areas with higher socioeconomic standing (Siu, 2020; Hawkins et al., 2020). Contrary to these findings, New Zealand presented a different pattern, where nearly half of COVID-19 cases were imported and disproportionately associated with individuals of higher socioeconomic status (Jefferies et al., 2020). This can be attributed to the exceptionally low

overall incidence during the pandemic in this country, such that the inclusion of a relatively small number of cases substantially affects the overall incidence. On the other hand, population density also plays a critical role in the transmission dynamics of infectious diseases. In highly populated countries such as India and Algeria, a higher population density was directly related to an increase in COVID-19 transmission (Bhadra et al., 2020; Kadi and Khelfaoui, 2020). A comparative study among countries in Europe, East Asia, Australia, and the United States reported similar results (Garland et al., 2020).

This chapter aims to contextualize the patterns observed during the first year of the COVID-19 pandemic, focusing on both the incidence of COVID-19 and the resulting excess mortality. The analysis of COVID-19 incidence covers the period from June to December 2020 and explores the influence of various demographic and structural factors, including the presence of higher education students, sex differences, socioeconomic status, and population density. In parallel, excess mortality patterns are examined with respect to sex, age group, and COVID-19 incidence.

For both analyses, spatial conditional autoregressive (CAR) Poisson models were employed, which account for the spatial dependence between neighboring administrative units. This approach enhances the accuracy of the estimates by incorporating the spatial structure of the data, allowing for a more nuanced understanding of how demographic and spatial factors influenced the pandemic's impact across different areas.

2.2 Materials and methods

2.2.1 Study area

Belgium is a federal state characterized by a complex political and administrative structure composed of three Regions and three Communities. The three territorial Regions are the Flemish Region in the north, the Walloon Region in the south, and the Brussels-Capital Region in the center. In parallel, the country consists of three linguistic Communities: the Flemish Community, the French Community, and the German-speaking Community. While the Flemish and French Communities largely correspond to the Flemish and Walloon Regions, respectively, the institutional structure is not strictly territorial. The Brussels-Capital Region is unique in that it is officially bilingual and hosts both the Flemish and French Communities. In

contrast, the German-speaking Community constitutes a small but constitutionally recognized entity located within the Walloon Region, primarily in the eastern part of the country near the German border.

The Flemish and Walloon Regions are each subdivided into five provinces. The Flemish Region comprises Antwerp, East Flanders, Flemish Brabant, Limburg, and West Flanders, while the Walloon Region consists of Hainaut, Liège, Luxembourg, Namur, and Walloon Brabant. At a finer administrative level, Belgium is further divided into 581 municipalities (300 municipalities are located in the Flemish Region, while 262 and 19 municipalities are located in the Walloon Region and the Brussels-Capital Region, respectively) with responsibilities ranging from civil registration to public health management. These municipalities vary widely in terms of population size, geographic area, socioeconomic status, and population density, which are crucial factors in understanding heterogeneous patterns of disease transmission and public health outcomes throughout the country. A detailed map of these divisions is presented in Figure 2.2.

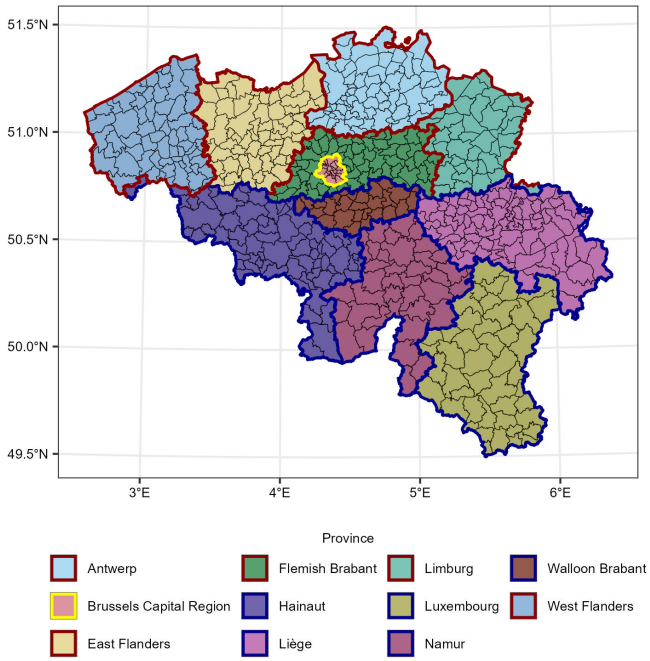


Figure 2.2. Administrative units in Belgium. Thicker border around each province indicates the regional division: Flemish Region = red, Walloon Region = blue, Brussels-Capital Region = yellow.

2.2.2 Data

Individual-level data on daily confirmed cases of COVID-19, including information on sex, age, and residential municipality, were obtained from Sciensano, the Belgian institute of public health (Sciensano, 2025b). Additional population data at the municipality-level, such as the number of deaths, the number of students in higher education, total inhabitants, and mean income (used as a proxy for socioeconomic status), were provided by Statbel, the Belgian statistical office (Statbel, 2025). The geographical maps used to visualize the results were adapted from the statistical sector map publicly available through the website of Statbel.

2.2.3 Statistical methods

The following methods were performed in R through package R-INLA (Lindgren and Rue, 2015).

2.2.3.1 Modeling COVID-19 incidence

The trends in COVID-19 incidence between June 1 and December 31, 2020, were descriptively analyzed across eight age groups: $\leq$17 years, 18–29 years, 30–39 years, 40–49 years, 50–59 years, 60–69 years, 70–79 years, and $\geq$80 years. In addition, the spatial distributions of key explanatory variables incorporated in the spatial model were also examined to better understand their potential associations with regional variation in COVID-19 incidence.

Given the shared language spoken in both Walloon and Brussels-Capital Regions, as well as sociocultural ties, common institutional structures, media environments, and public health communication strategies, the latter is grouped with Walloon and referred to as Wallo-Brux in this study. In Belgium, testing strategies changed substantially during 2020, with periods of restricted testing (e.g., prioritization of symptomatic individuals and healthcare workers) in March–April 2020 followed by phases of expanded capacity and broader eligibility in May–August 2020 before again limited to symptomatic cases due to capacity crisis (Sciensano, 2025a). Therefore, it is possible that the observed incidence rates only partially reflect changes in the true transmission dynamics. To account for this, a 15-day interval was selected for each region (Flemish and Wallo-Brux),

allowing for a flexible examination of the COVID-19 dynamics at certain time points within these geographical areas and partially capture temporal shifts in testing intensity.

Within each region, consisting of municipalities i ($i = 1, \ldots, 300$ for Flemish and $i = 1, \ldots, 281$ for Wallo-Brux), and per time period of 15 days, O_{ijk} denotes the number of observed COVID-19 cases in municipality i, for age group $j = 1, \ldots, 8$ and sex $k =$ male, female. N_{ijk} represents the population size per municipality, age group, and sex. A Poisson regression model was assumed for O_{ijk} given by:

$$O_{ijk}|\theta_{ijk} \sim \text{Poisson}(N_{ijk}\theta_{ijk}) \tag{2.1}$$

with θ_{ijk} the incidence of COVID-19 cases in municipality i, for age group k, and sex l, modeled on the logarithmic scale by the linear predictor:

$$\begin{aligned}\log(\theta_{ijk}) = \alpha_{1j} + \alpha_{2k} + \beta_1 \text{income}_i + \beta_2 \log(\text{popdens})_i + \gamma_j * \text{sturatio}_i \\ + \nu_i + \upsilon_i\end{aligned} \tag{2.2}$$

with parameters α_{1k} and α_{2l} representing the means of age and sex groups. Municipality-specific covariates used in the model are: standardized mean income denoted as income_i, logarithm of population density per km^2 denoted as $\log(\text{popdens})_i$, and standardized student ratio, i.e., the number of higher education students per 100 inhabitants in municipality i denoted as sturatio_i. Note that an age-specific parameter γ_k is assumed for the student ratio, as interest is in the association of higher education students with COVID-19 incidence in each age group. This variable was included as a pragmatic proxy for contact intensity and social mixing patterns due to the increase in COVID-19 cases in the younger adult population. While more detailed variables such as occupation-specific exposure, workplace mobility, or individual-level contact patterns would provide a more direct measure of transmission risk, such data were not available in routinely collected surveillance data. Given the short time interval (15 days) in each model, it was assumed that the age and sex-specific population N_{ijk} would remain constant because population changes during this short period are typically negligible compared to the variability observed in disease incidence.

The regression coefficients were given a normal prior distribution with zero mean and a small precision parameter of 0.001. The model allows

for both spatially structured heterogeneity and unstructured heterogeneity (Besag et al., 1991). The spatially-unstructured random effect υ_i is assumed to follow an independent normal distribution $\upsilon_i \sim N(0, \sigma_\upsilon^2)$. The spatially-structured random effect ν_i is defined as a Gaussian random field, accounting for the spatial autocorrelation. Each municipality i is characterized by a set of neighbors $\mathcal{N}_i$, defined as the areas that share boundaries with municipality i. The intrinsic CAR model specifies the distribution of ν_i, conditional on all the other values ν_j for $j \neq i$, given by :

$$\begin{aligned} \nu_i|\nu_{-i} &\sim N\left(\bar{\nu}_i, \sigma_i^2\right) \\ \bar{\nu}_i &= \frac{1}{|\mathcal{N}_i|}\sum_{j=1}^{n} a_{ij}\nu_j, \\ \sigma_i^2 &= \frac{\sigma_\nu^2}{|\mathcal{N}_i|}, \end{aligned} \tag{2.3}$$

where the weights a_{ij} are defined as 1 if areas i and j are adjacent and 0 otherwise; and $|\mathcal{N}_i|$ denotes the number of neighbors of area i. These CAR convolution models were fitted using integrated nested Laplace approximation (INLA), which provides a wide range of flexible models, including generalized linear mixed models, as well as spatial and spatio-temporal models. INLA simplifies the computation by performing approximate Bayesian inference in latent Gaussian models, offering a more efficient alternative to traditional Monte Carlo Markov chain (MCMC) methods (Rue et al., 2009). Each spatial random effect is associated with its own precision parameter, $\tau_\nu = \frac{1}{\sigma_\nu^2}$ and $\tau_\upsilon = \frac{1}{\sigma_\upsilon^2}$. By default, INLA assigns weakly informative log-Gamma priors to these precision parameters, typically specified as $\tau_\nu \sim$ LogGamma$(1, 0.00005)$ and $\tau_\upsilon \sim$ LogGamma$(1, 0.00005)$.

2.2.3.2 Modeling excess mortality

Mortality data from 2017 to 2020 were aggregated by five age groups: $<$25 years, 25–44 years, 45–64 years, 65–84 years, and $\geq$85 years, as well as by sex (male and female). To capture the temporal dynamics of the pandemic, particularly the two distinct COVID-19 waves that affected Belgium in 2020, the data were stratified into two time periods: January 1 to June 30 and July 1 to December 31. The mortality trends were then examined across age and sex within each interval to assess changes in excess mortality patterns associated with the pandemic.

In this study, baseline mortality was calculated as the average of observed mortality over the period 2017–2019. Using a recent pre-pandemic baseline allows for a direct and interpretable comparison with observed mortality during 2020, while minimizing the influence of longer-term demographic or epidemiological changes (Kontis et al., 2020). However, a simple average over three years might be sensitive to short-term anomalies, such as severe influenza seasons or extreme weather events (e.g., heat waves). It has been verified that no extreme mortality shocks occurred in Belgium during the selected reference period that would substantially distort the baseline. Although alternative approaches have been proposed in the literature to estimate expected mortality more robustly using a modification of Serfling model (Nielsen, Mazick, et al., 2013), overdispersed Poisson model (Msemburi et al., 2023), or quasi-Poisson model (Scortichini et al., 2021), most of them are performed for time series with daily or weekly time points. Since the data were aggregated per six months period, these methods would not be suitable to calculate the baseline mortality.

The standardized mortality ratio (SMR), defined as the number of deaths in 2020 divided by the average number of deaths in 2017–2019, was calculated separately for each age group and sex. To quantify the uncertainty around the estimates, 95% confidence intervals were calculated using the Delta method, which approximates the probability distribution for a function of an asymptotically normal estimator (Liu, 2012). The SMR logarithm ($\log \theta_i$) was assumed to follow a normal distribution with variance $\text{Var}(\log \theta_i) = \frac{1}{\theta_i^2}\text{Var}(\theta_i) = \frac{1}{Y_i}$, where Y_i represents the observed mortality in a given period for each age group and sex. With a back-transformation, the 95% confidence interval for θ_i is further given by $\left[\theta_i \exp(-Z_{1-\frac{\alpha}{2}}\frac{1}{\sqrt{Y_i}}); \theta_i \exp(Z_{1-\frac{\alpha}{2}}\frac{1}{\sqrt{Y_i}})\right]$, where $Z_{1-\frac{\alpha}{2}}$ is the critical value from the standard normal distribution.

A spatially discrete geostatistical model was fitted to the mortality data originating from the two time intervals. It should be noted that, based on the exploratory data analysis, a different mortality pattern was identified in age groups younger than 45 years. To avoid any misinterpretation of the results, the younger age groups were excluded from the following spatial analysis.

For each municipality i within an interval of six months, O_{ijk} denotes the number of deaths in municipality i, for age group $j = 1, 2, 3$ correspond-

ing to age groups 45–64 years, 65–84 years, and ≥85 years, respectively, with sex $k =$ male, female. E_{ijk} represents the expected average mortality based on the mortality rate in the years 2017–2019 per municipality, age group, and sex within each interval. A Poisson regression model is assumed for O_{ijk} given by:

$$O_{ijk}|\theta_{ijk} \sim \text{Poisson}(E_{ijk}\theta_{ijk}), \tag{2.4}$$

with θ_{ijk} the relative risks of mortality in 2020 in municipality i, for age group j, and sex k, compared to the pre-pandemic reference period 2017–2019, which is modeled on the logarithmic scale as the linear predictor:

$$\eta_{ijk} = \text{log}(\theta_{ijk}) = \alpha + \beta_{1j} + \beta_{2k} + \beta_3 \text{log(covir)}_{ijk} + \nu_i + \upsilon_i, \tag{2.5}$$

with parameters β_{1j} and β_{2k} representing the effect of age and sex groups. Considering the large heterogeneity in COVID-19 incidence per age group and sex, the logarithmic scale of COVID-19 incidence rate, denoted as log(covir)_{ijk}, was used as a covariate in the model.

The regression coefficients were given a normal prior distribution with zero mean and a small precision parameter of 0.001. To account for spatial heterogeneity, a spatially-structured random effect ν_i and a spatially-unstructured random effect υ_i are included in the model as described in Section 2.2.3.1. To compare the variability explained by the spatially structured random effect, the precision matrix is scaled as suggested by Sørbye and Rue (2014). Initially, the spatial random effects were fitted using default priors, i.e., $\tau_\nu \sim \text{LogGamma}(1, 0.00005)$ and $\tau_\upsilon \sim \text{LogGamma}(1, 0.00005)$. As a sensitivity analysis for these priors, the results were compared to models with a penalized complexity (PC) prior as proposed by Simpson et al. (2017), which is defined as:

$$\pi(\tau) = \tfrac{\lambda}{2}\tau^{-3/2}\text{exp}(-\lambda\tau^{-1/2}), \tau > 0, \lambda > 0 \tag{2.6}$$

with $\lambda = -\text{log}(a)/U$. A value of one was considered a reasonable upper bound for the marginal standard deviation U with weight $a = 1$. The prior for τ_ν and τ_υ is then expressed as $P\left(\frac{1}{\sqrt{\tau}} > 1\right) = 0.1$.

To assess unusual elevations in mortality risks, the exceedance probability, defined as the proportion of the relative risk's posterior probability that exceeds a given threshold value, was utilized. The probability can be

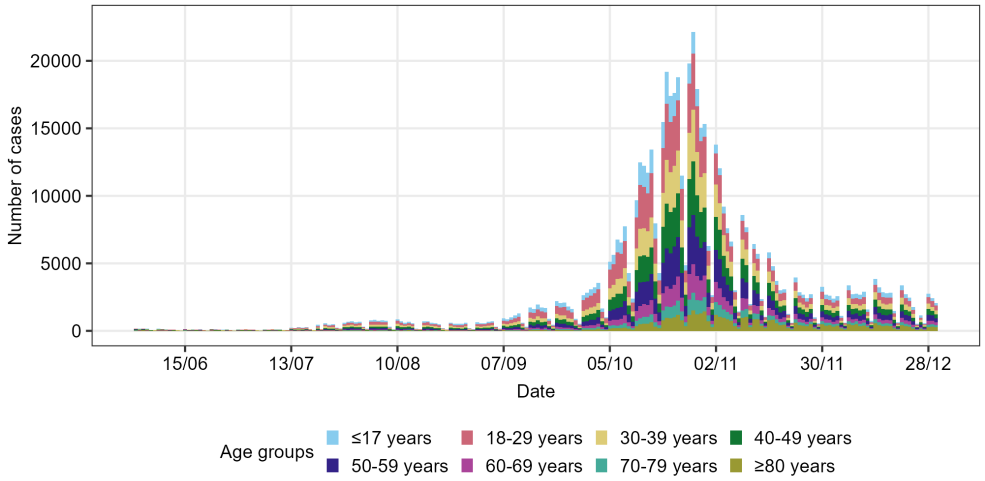


Figure 2.3. Distribution of daily COVID-19 cases per age group.

calculated using the marginal posterior distribution of θ_{ijk}. In the context of this study, a threshold relative risk of 1.2 was chosen to define unusually high mortality risk levels.

2.3 Results

2.3.1 Incidence of new COVID-19 cases

2.3.1.1 Exploratory data analysis

Approximately 577,900 cases with known residential, age, and sex information were reported from June 1 to December 31, 2020. The number of daily reported cases increased strongly, especially in age groups younger than 50 years old, starting from mid September 2020 with the highest incidence at the end of October 2020 (Figure 2.3).

To provide a different perspective, the number of reported COVID-19 cases was presented alongside the preventive measures implemented in both regions. Despite the similarity in public health interventions across the two regions as depicted in Figure 2.4, a considerable difference in the number of reported cases could be observed before and after September 2020. The number of COVID-19 cases in the Wallo-Brux Region increased earlier and reached a higher level compared to the Flemish Region, after the reopening of higher education.

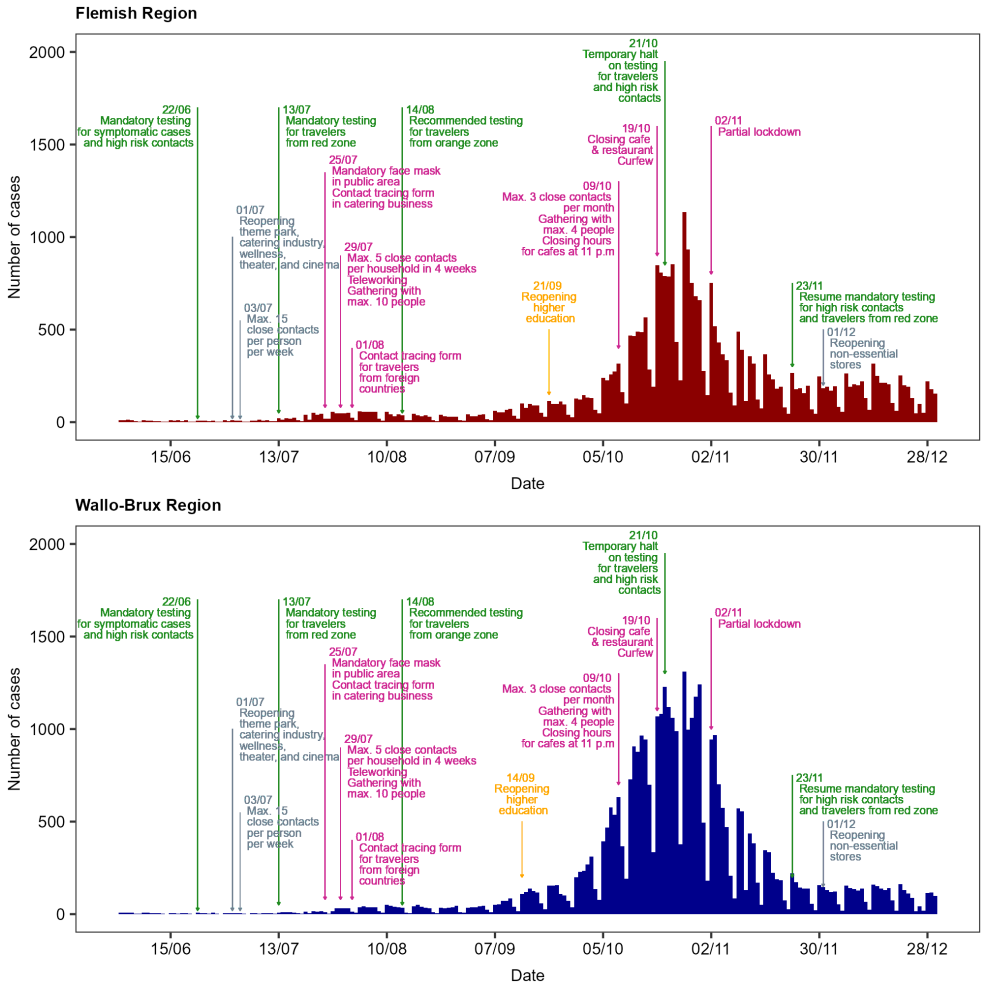


Figure 2.4. COVID-19 epidemic curve. Different colors indicate different type of strategies taken in this period. Green = testing strategy, grey = relaxation strategy, purple = preventive measures.

The distribution of other sociodemographic covariates per municipality is shown in Figure 2.5. The average student ratio was 3.51 students per 100 inhabitants and ranged from 0.82 to 10.49 students per 100 inhabitants, with higher ratios in areas closer to university towns (Figure 2.5A). On the other hand, there is a north-south trend, with lower student ratios in the southern part of the country. The mean income ranges from 8,835€ to 28,348€, with an overall average of 18,465€ (Figure 2.5B). In this map, differences between the Flemish and Wallo-Brux Regions are visible. The population density (Figure 2.5C) ranged from 0.05 tot 171.79

inhabitants per km^2, with an overall average of 7.44 inhabitants per km^2. This map clearly shows the more densely populated areas.

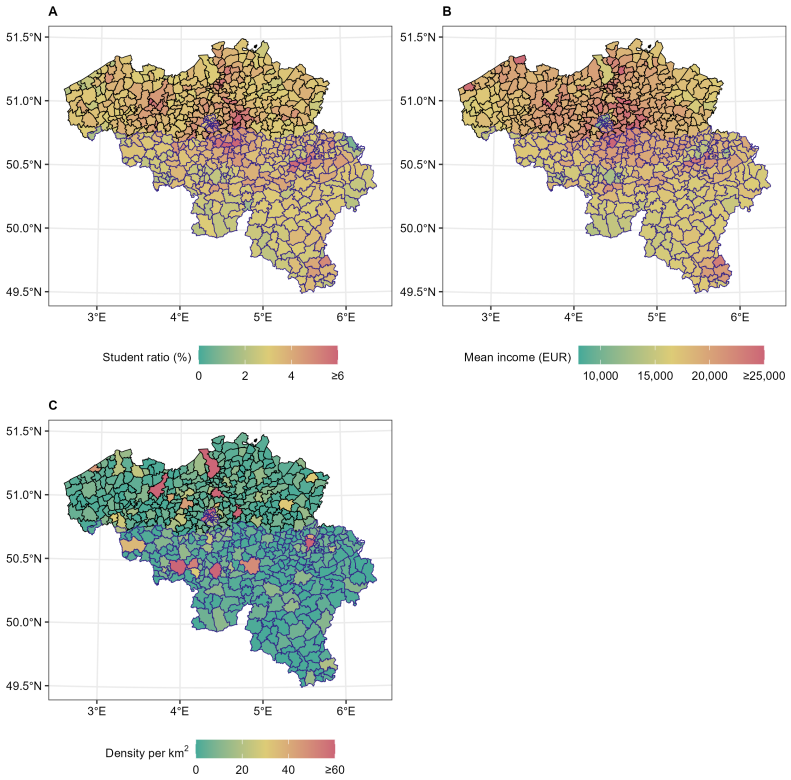


Figure 2.5. Spatial distribution of other sociodemographic covariates. (A) student ratio, (B) yearly mean income, and (C) population density. The black border indicates the Flemish Region and the blue border indicates the Wallo-Brux Region.

2.3.1.2 Age differences based on fitted models

The period- and region-specific overall incidence per age group, estimated by $\exp(\alpha_{1k})$, is presented in Figure 2.6. The age group of 18–29 years, followed by the age group 30–39 years had the highest predicted incidence of COVID-19 cases, particularly from mid-September 2020 until the end of October 2020, in both regions. From November 2020 onwards, the predicted incidence was highest within the group of those older than 80 years in both regions. This suggests that this epidemic wave started in younger individuals, and that infection spread from the younger individuals to the elderly population. While the trend looks similar in the Flemish and

Wallo-Brux Regions, note that the overall incidence was relatively higher in the latter region with considerably higher peak in the second half of October 2020.

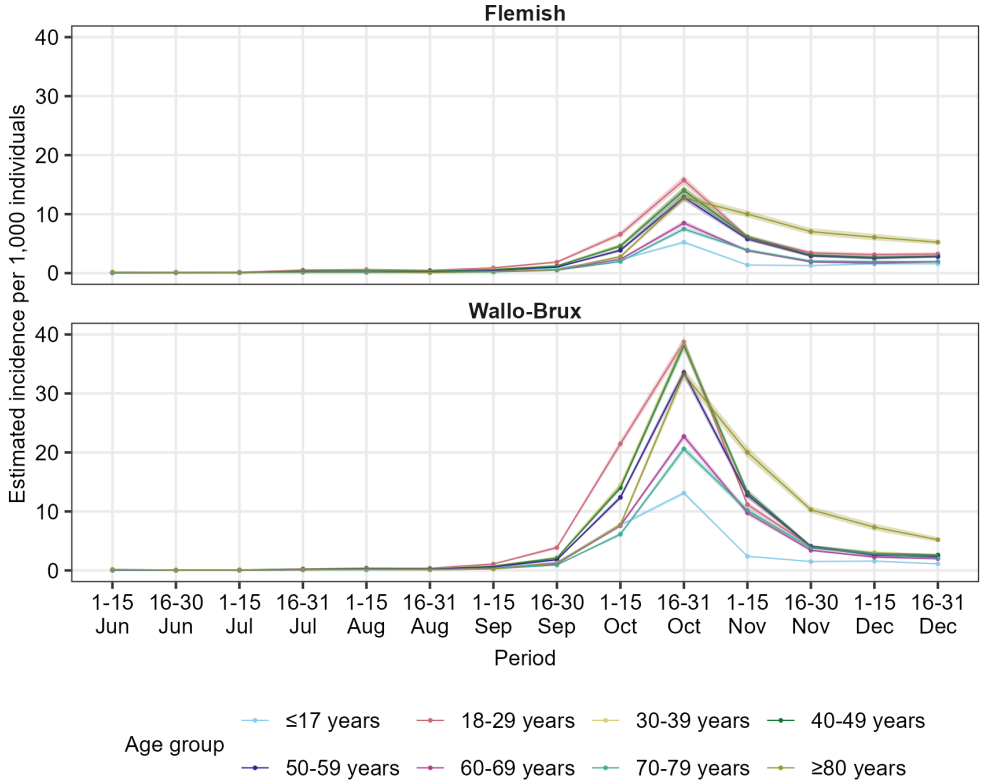


Figure 2.6. Estimated COVID-19 incidence in each age group.

2.3.1.3 Association with other sociodemographic factors

The association of other sociodemographic covariates included in the model is presented in Figure 2.7, top panels. In both regions, multiplicative effects greater than 1 (indicating a higher incidence rate of COVID-19) could be associated with females, especially from October 2020 onward. Similar trends of higher multiplicative effects could also be observed for population density until September 2020, particularly in the Wallo-Brux Region. The multiplicative effects then decreased to less than 1, indicating higher incidence rates in areas with a lower population density. Initially, municipalities with higher mean income were associated with a lower incidence of COVID-19 in both regions; however, this association reversed in October

2020 for the Wallo-Brux Region where the multiplicative effects increased slightly above 1, indicating a higher incidence in municipalities characterized by lower mean income levels.

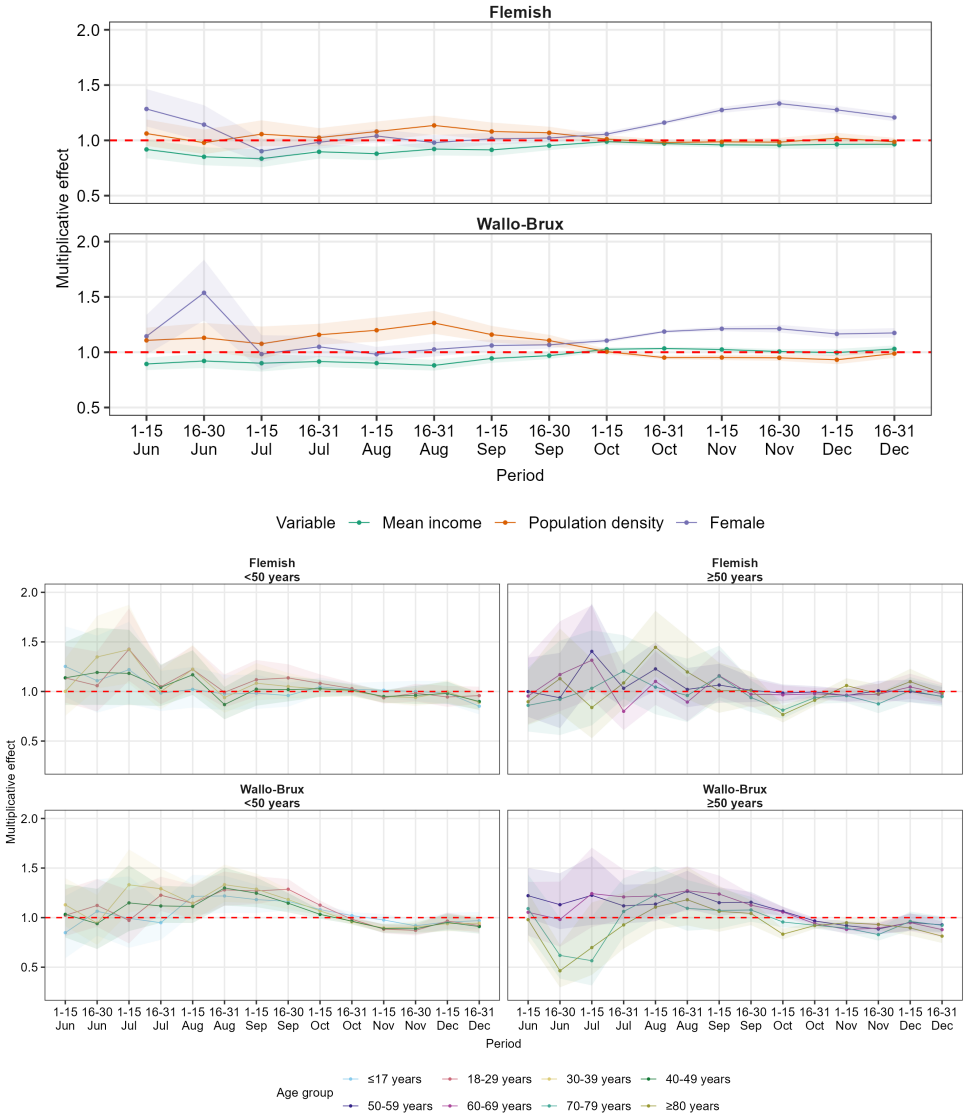


Figure 2.7. Association of other sociodemographic covariates with COVID-19 incidence. Top: Mean income, population density, and sex; Bottom: Interaction between proportion of higher education students and age groups.

The association of the student ratio with COVID-19 incidence in each

age group is illustrated in Figure 2.7 (bottom panels). In general, municipalities with a higher proportion of higher education students were associated with an increased incidence of new cases, particularly among individuals aged 30 to 69 years. The estimated multiplicative effect of the student ratio varied more substantially in the Flemish Region compared to the Wallo-Brux Region. In the Flemish Region, a pronounced effect was also observed among individuals aged 18–29 years during the summer months. In contrast, the association of student population in the Wallo-Brux Region followed a stable increasing trend in the summer months with a marked decline starting from mid-October.

2.3.1.4 Geospatial effects

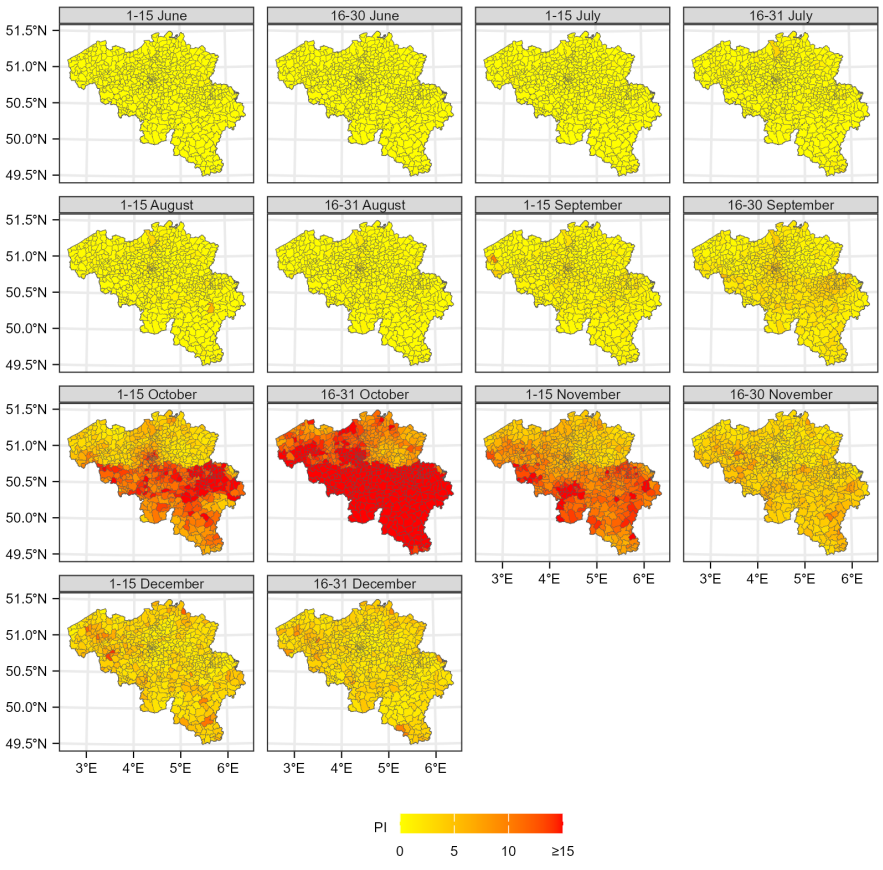


Figure 2.8. Predicted COVID-19 incidence per 1,000 individuals at municipality level. PI = predicted incidence.

During the summer months, the predicted COVID-19 incidence remained relatively low in both regions. The incidence increased sharply in October 2020 (Figure 2.8). The highest incidence was observed in the Wallo-Brux Region during the second half of October 2020, with predicted values ranging from 4.19 to 59.89 cases per 1,000 individuals and a regional average of 21.04 cases per 1,000 individuals. Although the predicted incidence decreased substantially in the latter half of November, several municipalities in both the Flemish and Wallo-Brux Regions retained relatively higher incidence rates through the end of the year.

2.3.2 Mortality pattern in Belgium

Between 2017 and 2020, a total of 455,858 deaths were recorded in Belgium, with annual counts of 109,273 in 2017, 110,339 in 2018, 108,431 in 2019, and a marked increase to 127,815 in 2020. The weekly number of deaths showed a relatively stable trend between 2017 and 2019, with a higher number of deaths reported between weeks 1 –8 in 2017 and weeks 8–14 in 2018 (Figure 2.9). In contrast, 2020 exhibited a considerable deviation from this pattern, with a substantial increase in mortality during weeks 11–20, corresponding to the first wave of the COVID-19 pandemic. Additional, smaller increases were observed in weeks 32–34, followed by a broader and less intense peak spanning weeks 41–52, consistent with the second pandemic wave.

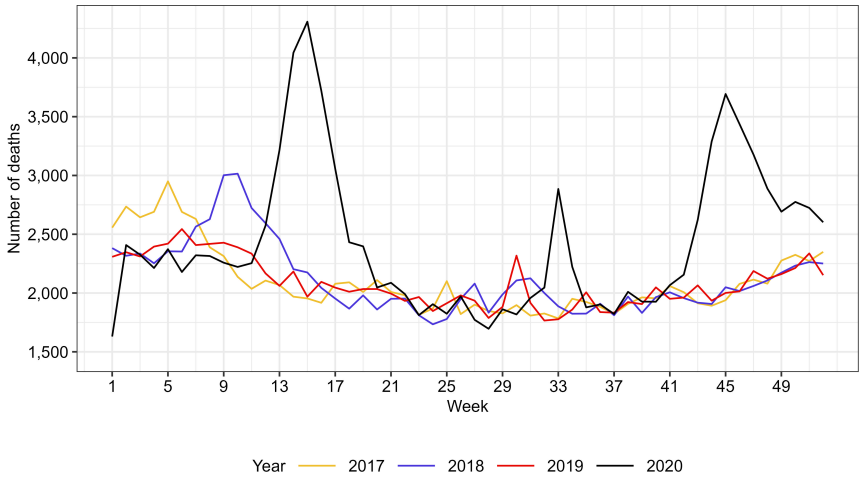


Figure 2.9. All-cause mortality in 2017–2020.

The stratified SMR within a period of six months is presented in Table 2.1. The standardized mortality ratios in age groups younger than 45 years were fairly close to or even lower than one, which means that the mortality in these age groups was lower compared to the pre-pandemic period. For this reason, the younger age groups were excluded from our subsequent spatio-temporal analysis.

Table 2.1. Observed standardized mortality ratio per age group and sex.

Period	Age group	Sex	Average mortality pre-pandemic	Mortality in 2020	SMR (95% confidence interval)
January 1 – June 30	$<$25 years	Male	302.33	239	0.79 (0.70,0.90)
		Female	188	158	0.84 (0.72,0.98)
	25-44 years	Male	720.67	737	1.02 (0.95,1.10)
		Female	379.67	373	0.98 (0.89,1.09)
	45-64 years	Male	4,258.33	4,425	1.04 (1.01,1.07)
		Female	2,602.33	2,605	1.00 (0.96,1.04)
	65-84 years	Male	13,628.33	15,352	1.13 (1.11,1.14)
		Female	10,552	11,746	1.11 (1.09,1.13)
	$\geq$ 85 years	Male	68,968.33	10,526	1.17 (1.15,1.20)
		Female	15,631	18,229	1.17 (1.15,1.18)
July 1 – December 31	$<$25 years	Male	317.67	219	0.69 (0.60,0.79)
		Female	194.33	142	0.73 (0.62,0.86)
	25-44 years	Male	721.33	738	1.02 (0.95,1.10)
		Female	367.33	372	1.01 (0.91,1.12)
	45-64 years	Male	4,036.33	4,486	1.11 (1.08,1.14)
		Female	2,548.67	2,601	1.02 (0.98,1.06)
	65-84 years	Male	12,613.33	15,314	1.21 (1.20,1.23)
		Female	9,633	11,473	1.19 (1.17,1.21)
	$\geq$ 85 years	Male	7,649.67	10,343	1.30 (1.28,1.33)
		Female	13,735	17,124	1.25 (1.23,1.27)

2.3.2.1 Association of age group, sex, and COVID-19 incidences

The prior sensitivity analysis indicated minimal changes in the estimated linear parameters, as well as the posterior density of each random effect (see Appendix A.3). However, the use of PC priors improved model performance, specifically by enhancing the proportion of variance explained by the spatially structured random effect. Consequently, the results presented below are based on the model incorporating PC prior.

The multiplicative effects of sex, age group, and COVID-19 incidences on the relative risks of mortality in 2020 compared to the pre-pandemic period were estimated by $\exp(\beta_{1j})$, $\exp(\beta_{2k})$, and $\exp(\beta_3)$, respectively (Figure 2.10). Elevated mortality risks were observed among older individuals, particularly those aged 65–84 and $\geq$85 years, as well as in areas with higher COVID-19 incidence. The female population was associated with significantly lower relative mortality risk. The association of these factors was more pronounced in the second half of 2020.

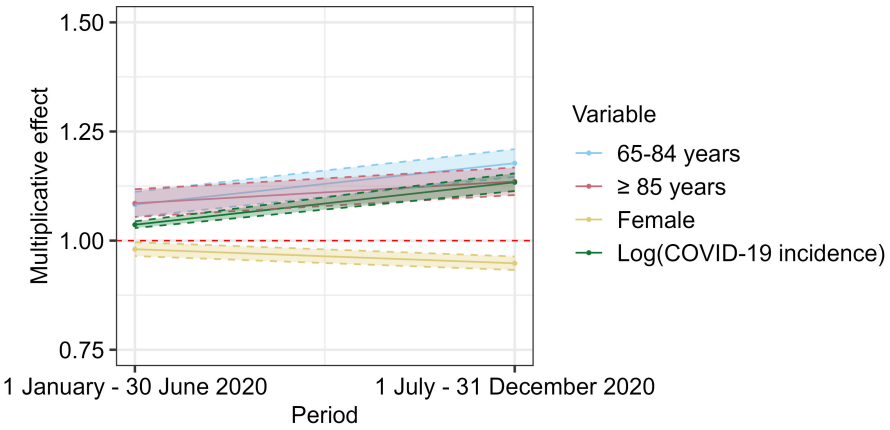


Figure 2.10. Association of age groups, sex, and COVID-19 incidence with excess mortality.

2.3.2.2 Geospatial effects

The map of posterior relative risks by age group, sex, and municipality is given in Figure 2.11. Across the country, the relative risks of mortality in 2020 were consistently higher than in the baseline time intervals of 2017–2019, with the most pronounced elevations observed during the second half of 2020. Elevated risks were evident across all age groups, although the magnitude of excess mortality was highest among the two oldest groups, underscoring their higher susceptibility to severe outcomes. Interestingly, some exceptions to this trend were identified in northeastern municipalities, where relative risks for individuals aged 65–84 years and $\geq$85 years were lower during the second half of 2020 compared to the first half.

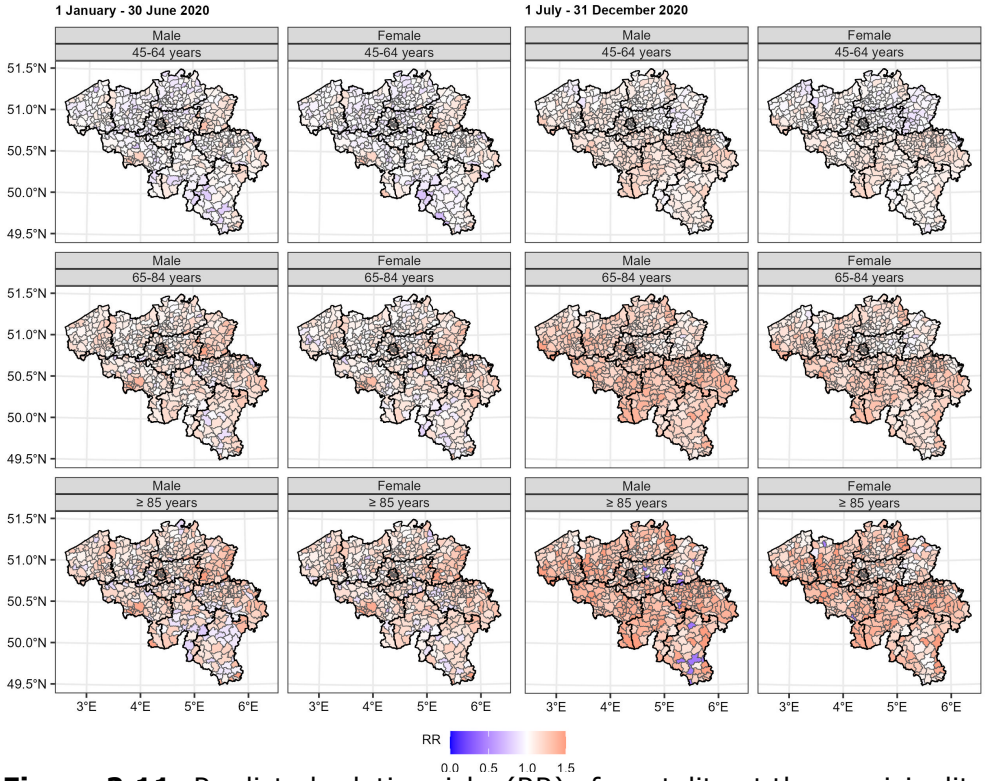


Figure 2.11. Predicted relative risks (RR) of mortality at the municipality level. The border of each province is indicated with black lines.

Elevated exceedance probabilities were predominantly observed in the oldest age groups (Figure 2.12). A clear temporal and spatial shift in municipalities with exceedance probabilities (RR > 1.2) above 80% was evident across the two time intervals. During the first half of 2020, higher exceedance probabilities for the age groups 65–84 years and ≥85 years were mainly concentrated in the northeastern and eastern municipalities, as well as in the Brussels-Capital Region and some larger cities in western Belgium. The pattern changed in the second half of 2020, where higher exceedance probabilities could be observed more in the northern, western, and southern parts of the country.

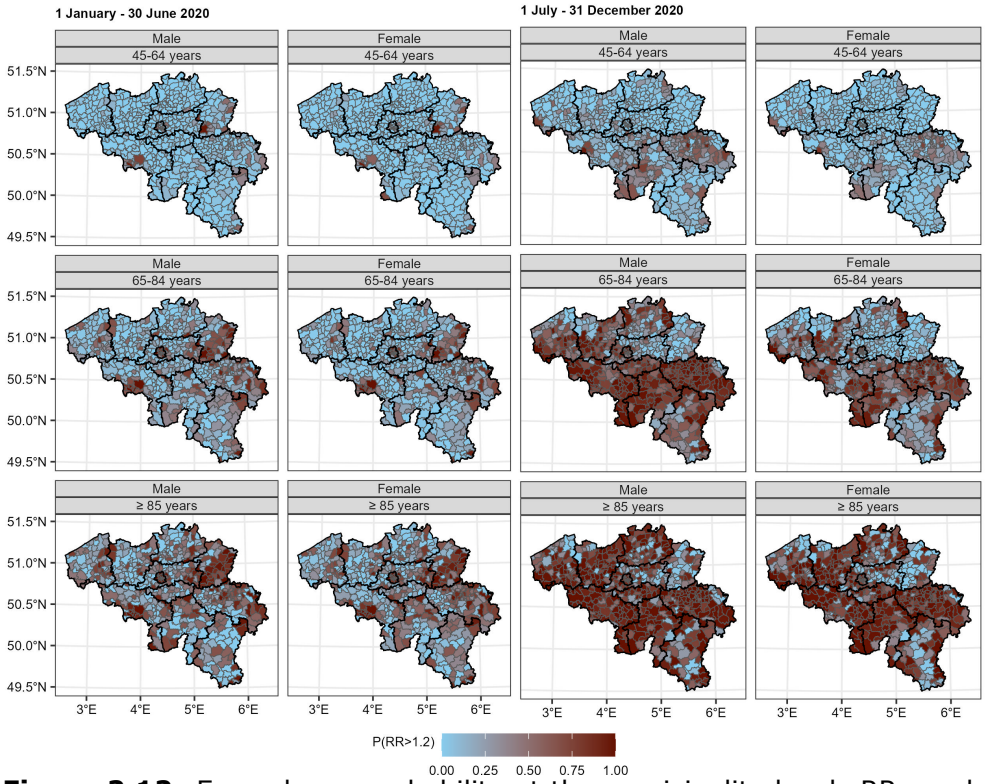


Figure 2.12. Exceedance probability at the municipality level. RR = relative risks. The border of each province is indicated with black lines.

2.4 Discussion

The findings presented in this chapter highlight the association of key sociodemographic factors, namely age, sex, income, population density, and the proportion of students in higher education, with the dynamics of COVID-19 incidence and mortality in Belgium during the first year of pandemic.

A pronounced demographic shift in the transmission of COVID-19 was observed in the second half of 2020, with an increase in incidence among younger age groups accompanied by fewer fatalities. Other studies in the United States, England, and Germany have documented a similar shift, reporting a higher proportion of infections among younger adults compared to the first wave in early 2020 (Boehmer et al., 2020; Pouwels et al., 2021; Lampl and Salzberger, 2020). Younger adults are more likely to have mild or no symptoms and they can unknowingly contribute to COVID-19 trans-

mission in the population, potentially jeopardizing people at increased risk of severe illness, as reported by Gandhi et al. (2020) in the United States or Hu et al. (2020) in China. A similar finding is partially reflected in the mortality analysis in this thesis, where significantly elevated relative risks of mortality were observed among older adults, particularly in the second half of 2020. It is further compounded by increased mortality in the nursing home population, which have been disproportionately affected during the pandemic in Belgium (Molenberghs et al., 2022; Torres et al., 2023; Basso et al., 2023). Moreover, in a nursing home setting, additional factors such as medical conditions or treatments administered to the residents (Vila-Corcoles et al., 2021), as well as the quality of care delivered by the facility (Cronin and Evans, 2022), can also influence mortality.

The higher multiplicative effect of COVID-19 among females may, in part, be attributed to the demographic composition of the Belgian population, where females slightly outnumber males. However, gendered occupational roles also play a role, with females being overrepresented in caregiving roles in both the formal and informal sectors, increasing their exposure risk (Gausman and Langer, 2020). Despite this, females were associated with lower relative mortality risks during the pandemic. In many European countries, higher excess mortality among males has been observed before and during the pandemic, which is arguably a general sex-specific phenomenon rather than specific to COVID-19 (Nielsen, Nørgaard, et al., 2021). The exact underlying mechanism is still under investigation, although results are starting to emerge (Haitao et al., 2020). However, factors such as differences in biological properties between males and females (Klein and Flanagan, 2016), the presence of underlying conditions that are more prevalent in men, e.g., cardiovascular or metabolic diseases (Di Angelantonio et al., 2015), or other health-related behaviors (D'Angelo et al., 2022; Thompson et al., 2016) could have an impact on health conditions and eventually, death.

Prior to October 2020, higher predicted incidences were associated with municipalities characterized by higher population density. This pattern is indicative of imported infections following the summer holiday period, echoing observations from Australia where returning international travelers contributed substantially to local transmission (Seemann et al., 2020). Following October 2020, however, local community transmission became more dominant in Belgium as areas with lower population density were as-

sociated with higher predicted incidences. On top of this, higher education students were found to be a key subpopulation in this context due to their increased mobility and complex social networks. Many university students in Belgium commute from home where they still live in the same environment with older people such as parents and grandparents, even during the period under investigation in our study. This increases the possibility of transmitting the infection to those who are at increased risk for severe disease (Venkatesan, 2020). Moreover, younger adults are more likely to engage in "high risk" behavior. Several studies in the United States and South Korea conducted during the initial phase of the COVID-19 pandemic (March–April 2020), reported that younger adults were more likely to leave their homes frequently (Alsan et al., 2020) and take part in social activities such as parties, fitness, or casual hang-outs (Nagata, 2020; Yasri and Wiwanitkit, 2020; Pedersen and Favero, 2020; Bae et al., 2020). In Poland, inappropriate practice of preventive measures such as wearing facial masks and social distancing, is more common in this age group as well (Matusiak et al., 2020). This is expected to render the preventive measures less effective and will likely boost further virus transmission.

Higher education students also belong to an age group that partly consists of working individuals, which a priori might complicate understanding dynamics in COVID-19 transmission in this age group, particularly after reopening of higher education in Belgium. However, relaxations prior to the start of the academic year did not address the workplace and teleworking had been recommended uninterruptedly since July 2020; further relaxation in economic activities did not occur until December 2020 (Figure 2.4). Based on this, the impact of behavior in the working sub-population within this age group on changes in COVID-19 incidence rates is assumed to be minimal.

Spatial analysis revealed clear regional disparities in COVID-19 incidence, with a distinct north-south gradient. In contrast, mortality patterns varied more dynamically. In the northeastern part of Belgium (i.e., Limburg province), the relative risks and exceedance probability were higher in the first half of 2020. This area was greatly affected by the first wave of COVID-19, yet it experienced comparatively milder incidences during the second half of 2020 (Sciensano, 2025a). Next to the high COVID-19 incidences in the first wave, Limburg province also had the highest COVID-19-related mortality among nursing home residents in this period

(Sciensano, 2023b). On the other hand, some provinces such as Liege, Luxembourg, and Hainaut had relatively lower incidences in the first half of 2020, followed by a surge during the second wave, leading to higher relative risks and exceedance probabilities for mortality in these areas. The Brussels Capital Region showed consistently high relative risk and exceedance probability throughout both waves, which was an indication of persistent high COVID-19 incidences over the year 2020.

Some limitations should be considered in the current studies. In general, analyzing data at the municipality level could minimize differences that might be observed at a smaller area level. On top of this, the completeness of data reporting warrants consideration, especially during the peak periods in 2020, when data collection capabilities may have reached their threshold. To flexibly model temporal changes in the COVID-19 incidences, separate spatial models were fitted within 15-day interval, which assumes the absence of temporal correlation at the municipality level. An alternative approach should be considered to model the local temporal trend. Finally, the study period was prior to the vaccination campaign which started in late December 2020. Future investigations should integrate vaccination data to better account for its impact on transmission dynamics and disease severity.

CHAPTER 3

UNCOVERING THE SUBSEQUENT YEARS OF THE COVID-19 PANDEMIC: EVIDENCE FROM LINEAR MIXED MODELS

This chapter is based on the following publications:

- **Natalia, Y. A.**, Molenberghs, G., Neyens, T., Hens, N., Faes, C. (2025). Empirical analysis of COVID-19 confirmed cases, hospitalizations, vaccination, and international travel across Belgian provinces in 2021. *PLoS One* 20(5): e0322017. DOI: 10.1371/journal.pone.0322017
- **Natalia, Y. A.**, Verbeeck, J., Faes, C., Neyens, T., Molenberghs, G. (2024). Unraveling the impact of the COVID-19 pandemic on the mortality trends in Belgium between 2020–2022. *BMC Public Health* 24, 2916. DOI: 10.1186/s12889-024-20415-x.

3.1 Introduction

In the early stages of the pandemic, many countries imposed nationwide lockdowns to limit the transmission and importation of SARS-CoV-2 (British Broadcasting Corporation, 2020). Subsequently, these measures were replaced by targeted travel restrictions, often based on the epidemiological situation in specific regions. In October 2020, the European Union introduced color-coded zones (green, orange, red, and gray)

based on the risk of COVID-19 infections in a specific country to facilitate free movement while maintaining safety (European Commission, 2020). Stricter measures were implemented for travelers originating from high-risk zones. Several studies reported that travel restrictions, particularly in international travel, had a consistent effect on slowing down the spread of COVID-19 (Bisanzio et al., 2022; Chen and Fu, 2022; Kwok et al., 2021).

By the end of 2020, vaccines against COVID-19 had reached the market and could be added as a prophylactic measure. Europe started the vaccination campaign in December 2020, and per February 19, 2023, more than 976 million vaccination doses had been administered in this region (European Centre for Disease Prevention and Control, 2023b). The first campaign focused on the older age population as the severity of COVID-19 increased with age and, consequently, hit the older age population harder (Molenberghs et al., 2022). By December 31, 2022, 75.65% of the Belgian population had completed the full primary COVID-19 vaccination, with 59.84% of these individuals also received a booster dose. Among adults 18 years and older, the full primary vaccination coverage was 82.06%, with 64.89% of this group had received a booster dose (Sciensano, 2025a). Many studies have reported the beneficial effect of COVID-19 vaccination in the population (Scobie et al., 2021; Greene et al., 2022).

Despite the wealth of literature on COVID-19 epidemiology, relatively few studies have examined the combined effects of human mobility, in particular international travel, together with COVID-19 vaccination within a single analytical framework. Zou et al. (2022) simulated the influence of vaccination coverage and daily mobility among provinces on COVID-19's effective reproduction number during the Chinese-Spring-Festival travel rush in 2021. Their findings demonstrated that vaccination decreases the reproductive number while elevated mobility exerted the opposite effect, potentially facilitating viral spread. However, these conclusions were derived primarily from simulated scenarios, highlighting the need for real-world data to validate and expand on these insights.

Moreover, as discussed in Chapter 2, mortality trends attributed to a novel infectious agent are highly influenced by numerous factors, including viral evolution, public health interventions, and the availability and quality of medical care (Islam et al., 2021; Kontis et al., 2020). The first year of the pandemic was particularly critical, marked by high uncertainty, evolving mitigation strategies, and the absence of vaccines. However, a

comprehensive understanding of the pandemic's trajectory requires continued assessment beyond this initial period, particularly after vaccination campaigns were actively deployed and public health policies were rapidly adapted in response to emerging variants and changing epidemiological conditions.

Therefore, this chapter is aimed to explore the evolving patterns of COVID-19 incidence (i.e., new cases and hospitalizations) and mortality until the third year of the pandemic, with a particular focus on the interplay between mitigation strategies and health outcomes. The incidence analysis covers the period from January to December 2021, a period characterized by intensive vaccination efforts in many countries, including Belgium. In parallel, mortality trends between 2020 and 2022 were examined across sex and age strata using both excess mortality and cause-specific mortality indicators. Although excess mortality provides a holistic measure of the pandemic's health burden, including indirect effects, cause-specific mortality allows for a more nuanced assessment of the long-term impact of COVID-19 on specific causes of death. These complementary approaches are essential to unravel the direct and indirect consequences of the pandemic and to evaluate the effectiveness of public health interventions over time.

3.2 Materials and methods

3.2.1 Data

Data on daily confirmed cases of COVID-19 and hospitalizations at the provincial level, along with weekly cumulative vaccinations at the municipality level, were made publicly available by Sciensano, the Belgian institute for public health (Sciensano, 2025b).

Until May 23, 2022, all incoming travelers to Belgium, regardless of their mode of transportation, were required to complete a Passenger Locator Form (PLF) prior to or upon arrival. During this period, travelers arriving from high-risk zones were required to undergo mandatory testing on both day 1 and day 7 post-arrival. Sciensano documented the volume of incoming travelers and their test results in weekly epidemiological reports (available in Dutch, French, and German) (Sciensano, 2022). Detailed weekly data on international arrivals and the day 1 post-arrival positivity rate at the provincial level are available up until early April 2022.

StatBel, the Belgian national statistical institute, publicly released population data encompassing daily all-cause mortality, monthly cause-specific mortality, and annual population figures (Statbel, 2025). Data from 2009 to 2022 were used to model the mortality trend, aggregated by sex and age group (0–24 years, 25–44 years, 45–64 years, 65–74 years, 75–84 years, 85+ years).

Taking into account the potential influence of heatwave on the mortality pattern, data on heatwave occurrences in 2009–2022 were obtained from the Royal Meteorological Institute of Belgium (Het Koninklijk Meteorologisch Instituut van België, 2025). While the potential influence of cold waves should also be considered, Belgium experienced only two cold wave episodes during the study period (in 2009 and 2012, which lasted 18 and 12 days, respectively). Given the limited number and short duration of these events, cold waves were not included as a covariate in the analysis. The stringency index from Our World in Data (Our World in Data, 2025) was retrieved to assess the general influence of travel restriction and other NPIs. A higher stringency index indicates more stringent policies at the country level during a given period (Hale et al., 2021).

3.2.2 Statistical methods

The following methods were performed in R through package lme4 (Bates et al., 2015).

3.2.2.1 Modeling COVID-19 incidence

To ensure consistency across all data sources, the study period was confined to the calendar year 2021. Let Y_{ijc} be the COVID-19 incidence of confirmed cases per 100,000 individuals ($c = 1$) or the COVID-19 hospitalizations per 100,000 individuals ($c = 2$) at week $i = 1, \ldots, 52$ as defined by the International Organization for Standardization (ISO) 8601, i.e., each week begins on Monday and week 1 is the first week with at least four days in the new year that contains the first Thursday, in province $j = 1, \ldots, 11$. Given the heterogeneity of the outcome variables across provinces, a logarithmic transformation of Y_{ijc} was used with province j as a random effect.

The general proposed model is given by:

$$\log Y_{ijc} = (\beta_{0c} + b_{0jc}) + \sum_{m=1}^{n} \beta_{mc} \mathsf{X}_{mijc} + \varepsilon_{ijc}, \tag{3.1}$$

with $\varepsilon_{ijc} \sim N(0, \sigma_c^2)$, $b_{0jc} \sim N(0, \sigma_{Dc}^2)$, and ε_{ijc}, b_{0jc} mutually independent. Here, β_{0c} represents the outcome-specific fixed intercept and b_{0jc} denotes the province-specific random intercept. The term β_{mc} indicates the regression coefficient associated with each predictor, where n represents the total number of effects, including any interaction terms, incorporated into the model.

In general, four main explanatory variables were considered in our analysis:

1. The weekly vaccination coverage per province defined as the cumulative population percentage that received full primary vaccination with one or two-doses of vaccines. Taking into account that vaccination can have a delayed effect as immunity requires time to develop (H. Li et al., 2022), the effect of vaccination coverage at lag l week was explored as a potential predictor up to six months prior to the current week given the duration of our study period and the potential for waning immunity (Bobrovitz et al., 2023; COVID-19 Forecasting Team, 2023). This variable is denoted as $\mathsf{vfull}_{i-l,j}$ with $l = 0, 1, \ldots, 24$. On top of this, the non-linear effect of vaccination and a decrease in marginal benefit at high vaccination coverage (Fang et al., 2024; Yamana et al., 2023) was also considered by including the inverse of vaccination coverage at lag l week, denoted as $\frac{10}{\mathsf{vfull}_{i-l,j}}$, in our model selection.

2. The weekly incoming travel rate (travel_{ij}) calculated as the number of incoming travelers (regardless of zone and mandate to get tested) to province j at week i per 100 inhabitants of that province (also denoted as percentage).

3. The weekly positivity rate among travelers from high-risk zones (pos_{ij}) calculated as the number of positive tests per 100 conducted tests (also denoted as percentage).

4. The median stringency index (SI) at the corresponding week SI_i.

It is important to note that COVID-19 hospitalizations are a direct consequence of COVID-19 infections. For this reason, an alternative model for hospitalizations that incorporates the incidence of confirmed cases in the same week was evaluated as an explanatory variable. To avoid confusion, the logarithm of confirmed cases is denoted as $\log(\text{IC7})_{ij}$ when used as an explanatory variable in the model.

Considering different dynamics among these variables, several fixed effects variations which include different interaction terms were explored. To reduce the risk of overfitting, the variation was limited to include only two-way interaction terms. The most parsimonious model was selected based on the lowest Akaike information criterion (AIC) value.

As a sensitivity analysis, a subset of the data was extracted starting from week 20 (17–23 May), a time point at which the coverage of COVID-19 vaccination coverage had surpassed 10% in all provinces. The same model selection procedure applied to the full dataset was employed on this subset to evaluate the robustness of the primary findings.

3.2.2.2 Modeling excess and cause-specific mortality

To minimize daily variations and the weekend effect, the daily all-cause mortality data at the country level were arranged continuously and then aggregated based on sex and age groups in weekly periods according to the ISO 8601 definition. Week 53 was excluded because this week is not consistently present in every year. To accommodate heterogeneity in weekly mortality between age groups, a logarithmic transformation was applied to the mortality rate per 100,000 individuals, Y_{it}, with week $i = 1, \ldots, 52$ in the reference years $t = 2009, \ldots, 2019$. Taking into account the relative changes in each demographic group, a cubic spline interpolation was used to calculate the population in week i and year t (McNeil et al., 2011). Using a modification of the general linear mixed model proposed by Verbeeck et al. (2021), Y_{it} was modeled considering the effect of sex and age group $k = 1, \ldots, l$, with l denotes the number of age groups used in a specific model. The potential interaction between sex and age groups is also evaluated, as well as the effect of heatwave, which indicates whether the corresponding week is a part of a heatwave period. The proposed

model is given by:

$$\begin{aligned}\log Y_{itk} = (\beta_0 + b_{0t}) + \beta_1 \text{female}_{itk} + \sum_{k=1}^{l} \beta_{2k} \text{age}[k]_{itk} + \beta_3 \text{heatwave}_{it} \\ + \sum_{k=1}^{l} \beta_{4k} \text{female} \times \text{age}[k]_{itk} + \sum_{m=1}^{n} \alpha_m \sin\left(\frac{2m\pi i}{52}\right) \\ + \sum_{m=1}^{n} \gamma_m \cos\left(\frac{2m\pi i}{52}\right) + \varepsilon_{itk}, \end{aligned} \tag{3.2}$$

with $\varepsilon_{itk} \sim N(0, \sigma^2)$, $b_{0t} \sim N(0, \sigma_D^2)$, and ε_{itk}, b_{0t} mutually independent. This model captures the annual changes in the mortality rate for each age and sex group through the random intercept b_{0j}, as well as the cyclic mortality pattern within a year through n Fourier terms. Due to the possibility of a non-significant effect, the Fourier terms were allowed to be non-sequential during the model selection. For each model, up to six Fourier terms ($n = 3$) were explored, which correspond to a yearly sine wave (α_1), a yearly cosine wave (γ_1), a half-yearly sine wave (α_2), a half-yearly cosine wave (γ_2), a quarterly sine wave (α_3), and a quarterly cosine wave (γ_3). The most parsimonious model was selected based on sequential exclusion of non-significant terms and minimal AIC value. The weekly all-cause mortality rate in 2020–2022 was predicted from the final model and compared the 95% prediction interval (transformed back to the original scale) with the observed all-cause mortality rate.

The cause-specific mortality was evaluated using a descriptive time-trend analysis stratified by sex and age groups. The monthly data were available from 2009 until 2021. StatBel compiled the causes of death based on the international statistical classification of diseases and related health problems (ICD) version 11 (World Health Organization, 2024) and reported these causes in 21 groups. To maintain the comprehensiveness of our analysis, these groups were further divided into seven broader groups, namely COVID-19, external causes (including road accidents and suicide), heart and vascular diseases, infectious diseases, mental and behavioral disorder, neoplasms, and other causes (see Appendix B.1). Similar to the all-cause mortality, the cause-specific mortality rate was calculated as the number of cause-specific deaths divided by the interpolated monthly population. Cause-specific mortality trends for 2020 and 2021 were compared against the baseline period from 2009 to 2019.

Let Y_{itkc} be the logarithm of the monthly mortality rate for month $i = 1, \ldots, 12$, year $t = 2020, 2021$, and age group $k = 1, \ldots, 6$ modeled separately for each group of mortality cause $c = 1, \ldots, 7$. Given the monthly aggregation of the data, heatwave effects were not included, as their incorporation would implicitly assume that heatwave conditions persisted throughout the entire month. The median value of the stringency index for each corresponding month was included to account for the intensity of non-pharmaceutical interventions. To accommodate the granularity of cause-specific mortality data and ensure comparability across different causes, the model specified in (3.2) was adjusted to become:

$$\begin{aligned}\log Y_{itkc} = (\beta_{0c} + b_{0tc}) + \beta_1 \text{female}_{itkc} + \sum_{k=1}^{l} \beta_{2k} \text{age}[k]_{itkc} + \beta_3 \text{stringency}_{itc} \\ + \sum_{k=1}^{l} \beta_{4k} \text{female} \times \text{age}[k]_{itkc} + \beta_5 \text{female} \times \text{stringency}_{itc} \\ + \sum_{k=1}^{l} \beta_{6k} \text{age}[k] \times \text{stringency}_{itkc} \\ + \sum_{k=1}^{l} \beta_{7k} \text{female} \times \text{age}[k] \times \text{stringency}_{itkc} + \varepsilon_{itkc}, \end{aligned} \tag{3.3}$$

with $\varepsilon_{itkc} \sim N(0, \sigma^2)$, $b_{0tc} \sim N(0, \sigma^2_{Dc})$, and ε_{itkc}, b_{0tc} mutually independent. Similar to the all-cause mortality, the most parsimonious model was selected based on sequential exclusion of non-significant terms and minimal AIC value.

Finally, the cause-specific mortality burden was further assessed using years of life lost (YLL), which quantifies the impact of mortality by multiplying the number of deaths by the remaining life expectancy at the age of death, thereby capturing the potential years of life lost due to each cause of death (Bonneux, 2002). Since the cause-specific mortality data were already aggregated into specific age brackets, the median age of each group was used as a proxy for the age of death. Residual life expectancy estimates were obtained from the Belgian life table published by Statbel, and these values were multiplied by the corresponding number of deaths for each cause and year to compute the total YLL.

3.3 Results

3.3.1 COVID-19 incidences in 2021

The weekly time trend of variables described in Subsection 3.2.2.1 are presented in Figure 3.1. Multiple waves of confirmed cases were observed across all provinces (Figure 3.1A). A marked increase in the incidence of confirmed cases occurred between weeks 9 (March 1–7) and 17 (April 26–May 2), with the highest incidence recorded in Namur. A larger wave followed between weeks 41 (October 11–17) and 51 (December 20–26), peaking in West Flanders. An exception was seen in Brussels, where a modest increase was noted from week 25 (June 21–27), eventually developing into a more substantial wave by week 41. In the same periods, similar waves could be observed in hospitalizations (Figure 3.1B). Between weeks 9 and 17, a considerable increase in hospitalizations was observed in Brussels, Hainaut, and Namur, followed by a decline in line with trends in other provinces. From week 25 onward, hospitalizations began to rise again, particularly in Brussels. A sharp increase was subsequently observed in West Flanders starting around week 41, reaching a peak in week 47 (November 22–28), before declining again toward the end of 2021.

The percentage of population who received full primary vaccination increased considerably from week 17 (April 26–May 2), as shown in Figure 3.1C. Every province reached 60% or higher at the end of week 33 (August 16–22), except for Brussels. By the end of 2021, West Flanders had the highest coverage of full primary vaccination (77.16%). A notable increase could also be observed in the rate of incoming international travelers in all provinces starting from week 21 (May 24–30) as illustrated in Figure 3.1D. However, there was a marked difference in Brussels, where it maintained a consistently higher incoming travel rate throughout the year compared to other provinces.

The COVID-19 test positivity rate among travelers displayed variability throughout the year, with a similar general trend across the 11 provinces (Figure 3.1E). An exception could be observed in week 13 (March 29–April 4) when a pronounced peak in positivity rates among travelers who visited Namur was recorded. While the early months of 2021 were characterized by more stringent measures in Belgium (Figure 3.1F), particularly between week 13 (March 29–April 4) and 15 (April 12–18, median stringency index = 75.93), some of these restrictions were relaxed when conditions

improved into the summer months, allowing greater mobility and fewer constraints on travel. The median stringency index reached its lowest levels between week 41 (October 11–17) and 45 (November 8–14, median stringency index = 43.98).

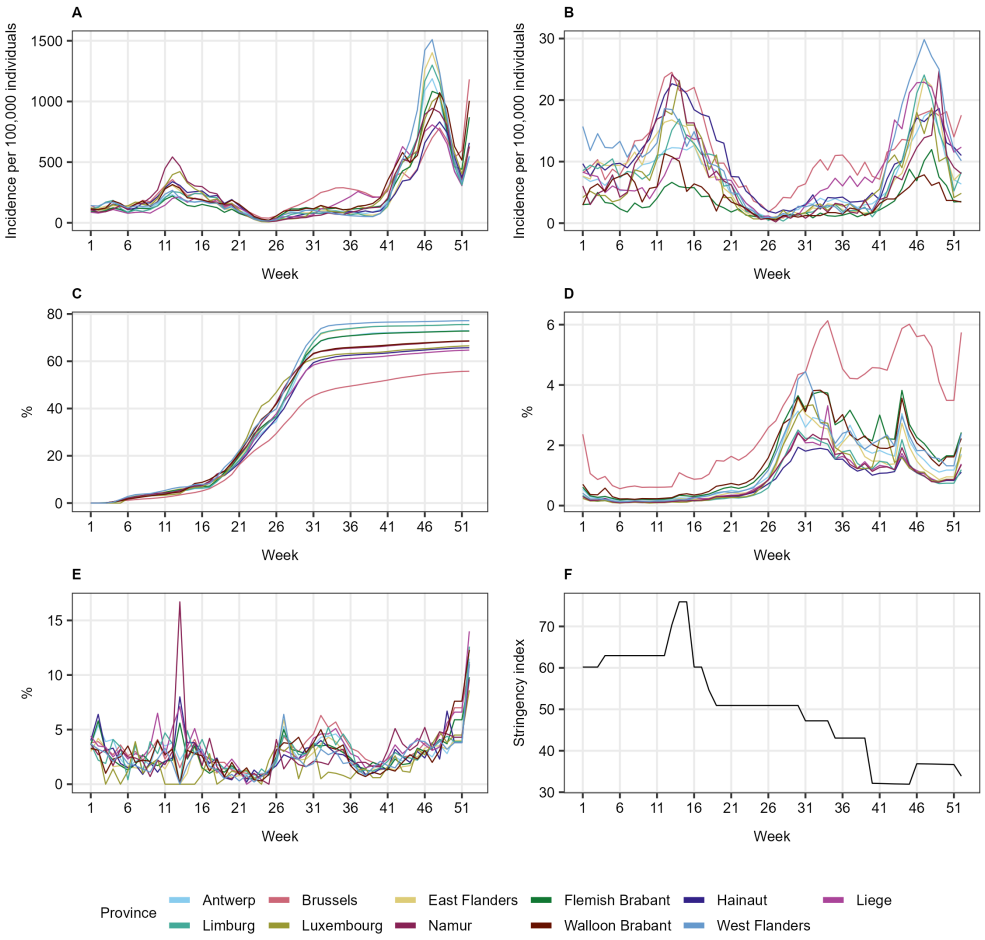


Figure 3.1. Weekly time trend of variables used in modeling COVID-19 incidences. (A) Confirmed cases, (B) hospitalizations, (C) full primary vaccination coverage, (D) incoming travel rate, (E) positivity rate among the incoming travelers, and (F) stringency index.

3.3.1.1 Linear mixed models for confirmed cases

The model selection process is outlined in Appendix B.2. Among the 2,900 candidate models evaluated for estimating $\log(Y_{ij1})$, one model achieved

the lowest AIC value of 965.658, which is given by:

$$\begin{aligned}\log Y_{ij1} = (\beta_{01} + b_{0j1}) + \beta_{11}\text{vfull}_{i-14,j1} + \beta_{21}\text{travel}_{ij1} + \beta_{31}\text{pos}_{ij1} + \beta_{41}\text{SI}_{i1} \\ +\beta_{51}\text{vfull}_{i-14,j1} \times \text{SI}_{i1} + \beta_{61}\text{travel} \times \text{pos}_{ij1} + \varepsilon_{ij1},\end{aligned} \tag{3.4}$$

with $\varepsilon_{ij1} \sim N(0, \sigma_1^2)$, $b_{0j1} \sim N(0, \sigma_{D1}^2)$, and ε_{ij1}, b_{0j1} mutually independent. This model highlights the important effect of lagged vaccination coverage ($l = 14$), incoming travel rates, positivity rates, and the stringency index on the incidence of confirmed cases. Furthermore, significant two-way interactions were identified between lagged vaccination coverage and the stringency index, as well as between travel rates and positivity rates.

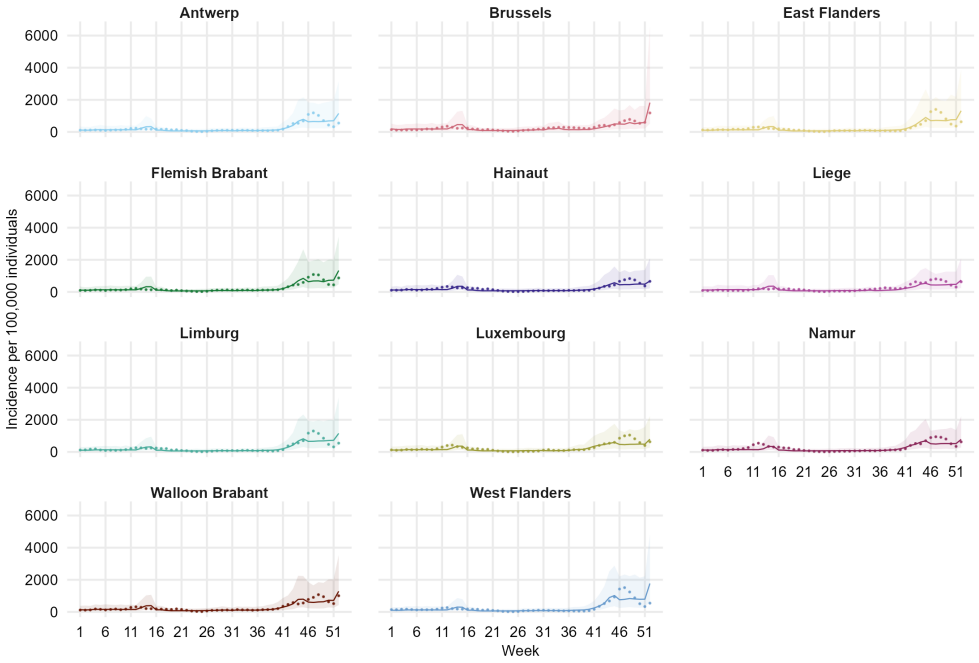


Figure 3.2. Fitted values of the weekly incidence of confirmed cases based on Equation 3.4. Observed values are depicted by points around the prediction line.

The complete parameter estimates of the model in Equation 3.4 are shown in Appendix B.3. The association between lagged vaccination coverage and confirmed cases was positive but strongly modified by the stringency index, with higher levels of stringency attenuating this relationship, as indicated by the negative interaction coefficient. A significant positive interaction between travel rate and positivity rate suggested a synergis-

tic pattern, i.e., a higher incidence of new cases could be expected with higher values of these explanatory variables. Using these estimates, the observed incidence of confirmed cases was compared with the 95% prediction intervals, as illustrated in Figure 3.2. The observed values consistently fall within the 95% prediction intervals, indicating a good overall fit of the model to the data. Notably, the width of the 95% prediction intervals increases starting from week 41 (October 11–17) in all provinces, suggesting greater uncertainty in the model's estimations during this period.

3.3.1.2 Linear mixed models for hospitalizations

Similar to the confirmed cases, a model with two interaction terms achieved the lowest AIC value of 1061.516 among the 2,900 candidate models. This model is denoted as:

$$\begin{aligned}\log Y_{ij2} = (\beta_{02} + b_{0j2}) + \beta_{12}\text{vfull}_{i-15,j2} + \beta_{22}\text{travel}_{ij2} + \beta_{32}\text{pos}_{ij2} + \beta_{42}\text{SI}_{i2} \\ +\beta_{52}\text{vfull}_{i-15,j2} \times \text{SI}_{i2} + \beta_{62}\text{travel}_{ij2} \times \text{SI}_{i2} + \varepsilon_{ij2},\end{aligned} \tag{3.5}$$

with $\varepsilon_{ij2} \sim N(0, \sigma_2^2)$, $b_{0j2} \sim N(0, \sigma_{D2}^2)$, and ε_{ij2}, b_{0j2} mutually independent. This model again highlights the importance of lagged vaccination coverage ($l = 15$), incoming travel rates, positivity rates, and the stringency index on the incidence of hospitalizations. Significant two-way interactions were found between lagged vaccination coverage and the stringency index, as well as between travel rates and the stringency index. The complete parameter estimates based on Equation 3.5 are shown in Appendix B.4. Consistent with the findings for confirmed cases, lagged vaccination coverage was positively associated with hospitalizations, although this relationship was strongly modified by the stringency index. In addition, travel rate was positively associated with hospitalizations. However, this association was significantly attenuated at higher levels of the stringency index, as indicated by the negative interaction term.

As shown in Figure 3.3, the observed values consistently fall within the 95% prediction intervals. However, a widening of the intervals could be observed during two distinct periods: prior to week 17 (April 26–May 2) and after week 41 (October 11–17), across all provinces.

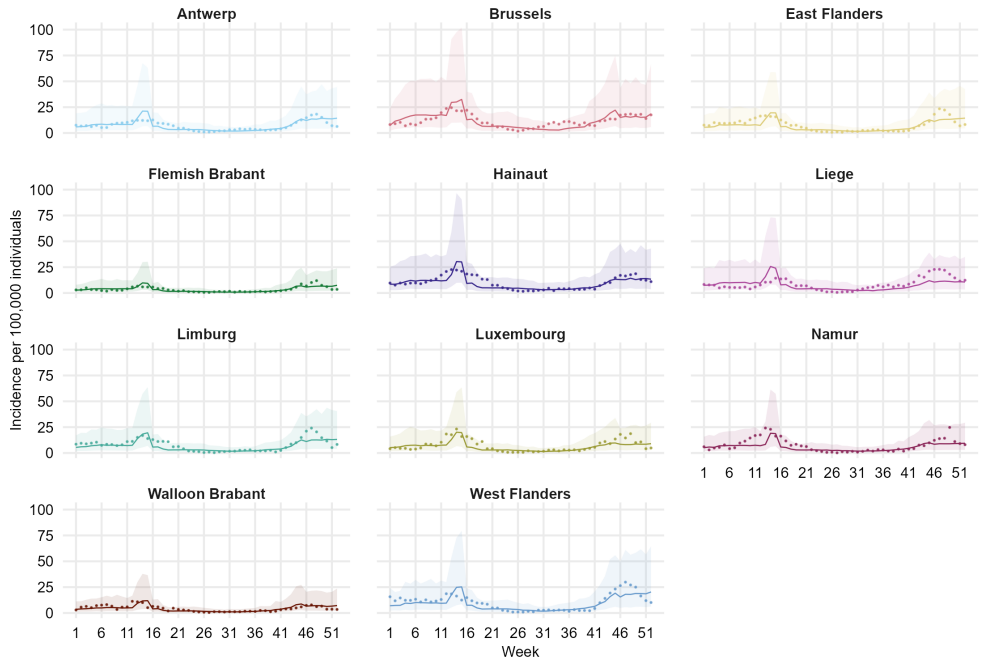


Figure 3.3. Fitted values of the weekly incidence of hospitalizations based on Equation 3.5. Observed values are depicted by points around the prediction line.

Incorporating the incidence of confirmed cases $\log(\text{IC7})_{ij}$ as an explanatory variable led to a considerable increase in the number of candidate models, with a total of 50,700 fitted models evaluated. Among these, the model achieving the lowest AIC value (432.708) is given by:

$$\begin{aligned}\log Y_{ij2} = {} & (\beta_{02} + b_{0j2}) + \beta_{12}\text{vfull}_{i-3,j2} + \beta_{22}\text{travel}_{ij2} + \beta_{32}\text{pos}_{ij2} + \beta_{42}\text{SI}_{i2} \\ & + \beta_{52}\log(\text{IC7})_{ij2} + \beta_{62}\log(\text{IC7}) \times \text{pos}_{ij2} + \beta_{72}\text{vfull}_{i-3,j2} \times \text{travel}_{ij2} \\ & + \varepsilon_{ij2}, \end{aligned} \tag{3.6}$$

with $\varepsilon_{ij2} \sim N(0, \sigma_2^2)$, $b_{0j2} \sim N(0, \sigma_{D2}^2)$, and ε_{ij2}, b_{0j2} mutually independent. A detailed summary of the parameter estimates for this model is provided in Appendix B.5. The incidence of confirmed cases was positively associated with hospitalizations and this association was modified by the positivity rate as indicated by the positive interaction term. Lagged vaccination coverage exhibited a modest association with lower hospitalizations. However, this relationship was attenuated at higher levels of mobility, as indicated by a positive interaction between these two explanatory variables.

This model also demonstrated a substantial improvement over the previously reported model in Equation 3.5, as evidenced by its notably lower AIC value. The better performance was also reflected in the predictive accuracy, with predictions aligning more closely with the observed data, as illustrated in Figure 3.4.

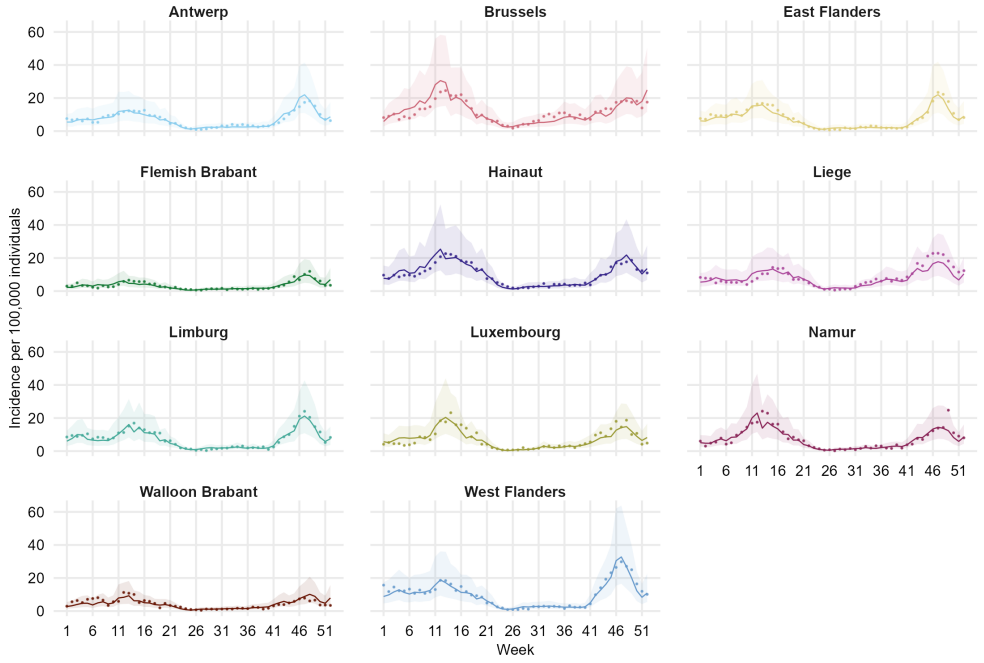


Figure 3.4. Fitted values of the weekly incidence of hospitalizations based on Equation 3.6. Observed values are depicted by points around the prediction line.

3.3.1.3 Sensitivity analysis

The model selection process using data from week 20 is summarized in Appendix B.6. Consistent with the results obtained from the complete dataset, models incorporating $\text{vfull}_{i-l,j}$ as predictor led to a substantially lower minimum AIC value compared to those including $\frac{10}{\text{vfull}_{i-l,j}}$, indicating a better model fit. An exception was observed, however, in the alternative model for hospitalizations that included $\log(\text{IC7})_{ij}$ as a predictor. In this case, the model using $\frac{10}{\text{vfull}_{i-l,j}}$ yielded the lowest AIC value, suggesting that the inverse transformation of vaccination coverage better captured the underlying dynamics of COVID-19 hospitalizations for this shorter period.

The formulations of the best-performing models are provided in Table 3.1 and the parameter estimates of these models are shown in Appendix B.7.

Table 3.1. Optimal models to estimate the incidence of confirmed cases and hospitalizations using data from week 20.

Model notation	AIC
$\log Y_{ij1} = (\beta_{01} + b_{0j1}) + \beta_{11}\text{vfull}_{i-15,j1} + \beta_{21}\text{travel}_{ij1} + \beta_{31}\text{pos}_{ij1} + \beta_{41}\text{SI}_{i1} + \varepsilon_{ij1},$ with $\varepsilon_{ij1} \sim N(0, \sigma_1^2)$, $b_{0j1} \sim N(0, \sigma_{D1}^2)$, and ε_{ij1}, b_{0j1} mutually independent.	621.231
$\log Y_{ij2} = (\beta_{02} + b_{0j2}) + \beta_{12}\text{vfull}_{i-23,j2} + \beta_{22}\text{travel}_{ij2} + \beta_{32}\text{pos}_{ij2} + \beta_{42}\text{SI}_{i2} + \beta_{52}\text{vfull}_{i-23,j2} \times \text{pos}_{ij2} + \beta_{62}\text{pos}_{ij2} \times \text{SI}_{i2} + \varepsilon_{ij2},$ with $\varepsilon_{ij2} \sim N(0, \sigma_2^2)$, $b_{0j2} \sim N(0, \sigma_{D2}^2)$, and ε_{ij2}, b_{0j2} mutually independent.	652.114
$\log Y_{ij2} = (\beta_{02} + b_{0j2}) + \beta_{12}\frac{10}{\text{vfull}_{i-11,j2}} + \beta_{22}\text{travel}_{ij2} + \beta_{32}\text{pos}_{ij2} + \beta_{42}\text{SI}_{i2} + \beta_{52}\text{log(IC7)}_{ij2} + \beta_{62}\text{log(IC7)} \times \text{pos}_{ij2} + \beta_{72}\text{log(IC7)} \times \text{travel}_{ij2} + \beta_{82}\frac{10}{\text{vfull}_{i-11,j2}} \times \text{pos}_{ij2} + \varepsilon_{ij2},$ with $\varepsilon_{ij2} \sim N(0, \sigma_2^2)$, $b_{0j2} \sim N(0, \sigma_{D2}^2)$, and ε_{ij2}, b_{0j2} mutually independent.	312.963

3.3.2 Excess mortality

The weekly observed all-cause mortality rate in 2009–2022 is shown in Figure 3.5. Similar mortality patterns were observed in the year 2020–2022 compared to the reference years (i.e., 2009–2019) in younger age groups (0–24 and 25–44). Starting from age 65, a seasonal pattern was notable with higher mortality rates in weeks 10–15, 31–33, and 42 onward of year 2020. Mortality rates in these age groups showed a clear peak in all-cause mortality, especially for year 2020. Compared to older age groups, age groups 0–24 and 25–44 had considerably lower weekly mortality.

Considering the distinct all-mortality patterns between the younger and older age groups, the model proposed in Equation 3.2 was also fitted separately for the younger age groups (0–24 and 25–44 years) and the older age groups (45+). All fitted models and their corresponding AIC values

are summarized in Appendix B.8.

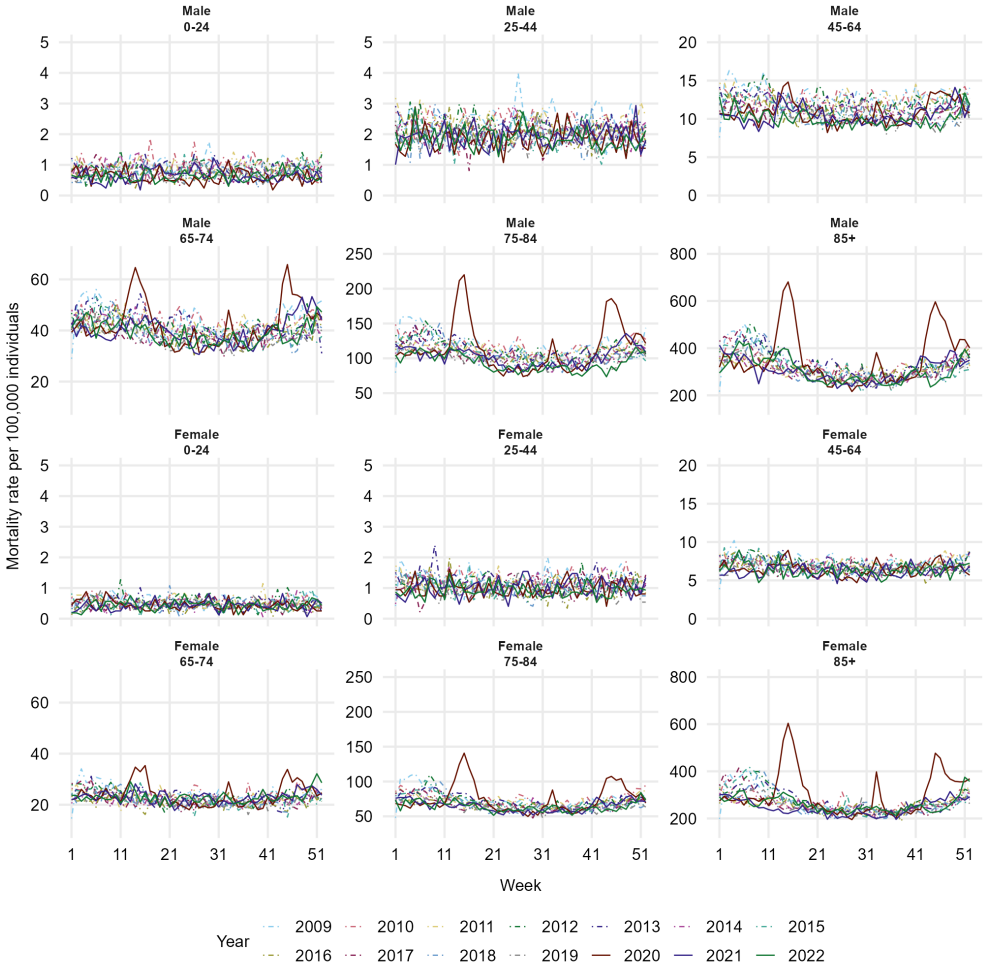


Figure 3.5. All-cause mortality rate in 2009–2022.

In the model that includes all age groups, significant associations of sex, age groups, and heatwave were found, together with the interaction between sex and age groups. This finding was consistent across all fitted models. Regarding cyclical parameters, there were initially four significant Fourier terms (i.e., $n = 2$ for both parameters α_m and γ_m), corresponding to a yearly sine wave (α_1), a yearly cosine wave (γ_1), a half-yearly sine wave (α_2), and a half-yearly cosine wave (γ_2). The full model is then reduced by excluding non-significant terms in each step until a reduced model with only significant parameters (i.e., Reduced model 2) was reached. This

selection is further supported by the lowest AIC value of Reduced model 2.

When fitting the model separately for the younger and older age groups, significant associations of sex and age groups with the mortality rate were identified in both age groups, together with an interaction between sex and age groups. Heatwave was significantly associated with the mortality rate only in the older age groups.

In age groups 0–24 and 25–44 years, a significant yearly sine wave α_1 was found. However, the seasonal pattern in these age groups was less pronounced compared to the older age groups (Figure 3.5). As shown in Appendix B.8, a model with a yearly sine wave has a higher AIC value compared to a model without any Fourier term (1363.9 and 1360.1, respectively). In view of this, the simpler model was used for the following model comparison.

Similar to the model with all age groups, four significant Fourier terms (yearly sine wave α_1, yearly cosine wave γ_1, half-yearly sine wave α_2, and half-yearly cosine wave γ_2) were identified for age groups ≥ 45 years. Separate models yielded a substantially lower AIC value (-6657.8) compared to the model that included all age groups (-2084.7). Therefore, these separate models were used as the final model to predict mortality rates for 2020–2022 as denoted in Table 3.2.

The parameter estimates for each model in Table 3.2 are presented in Appendix B.9. The observed mortality was compared with the 95% prediction interval derived from these estimates as shown in Figure 3.6. In general, higher predicted mortality rates were observed in males with wider prediction intervals. For age group 0–24 years, the observed mortality rate in 2020–2022 was lower relative to the predicted intervals. In the age group 25–44 years, the observed mortality rate in 2020–2022 followed the predicted trend. Higher mortality rates in 2020 were found from age 45 years and older. Higher values of observed mortality compared to the 95% predicted intervals in year 2020, particularly in age groups 65–74, 75–84, and 85+ years, indicated a higher excess mortality in these age groups. The mortality pattern changed in 2021–2022 where relatively lower observed mortality rates were found compared to the 95% predicted intervals, especially in the age group 45–64 years.

Table 3.2. Final separate linear mixed model for the younger and older age groups.

Age group k	**Model notation**
0-24, 25-44	$\log Y_{itk} = (\beta_0 + b_{0t}) + \beta_1 \text{female}_{itk} + \beta_{2,2} \text{age[25–44]}_{itk} + \beta_{3,2} \text{female} \times \text{age [25–44]}_{itk} + \varepsilon_{itk}$ with $\varepsilon_{itk} \sim N(0, \sigma^2)$, $b_{0j} \sim N(0, \sigma_D^2)$, and ε_{itk}, b_{0j} mutually independent.
45-64, 65-74, 75-84, 85+	$\log Y_{itk} = (\beta_0 + b_{0t}) + \beta_1 \text{female}_{itk} + \beta_{2,2} \text{age[65–74]}_{itk} + \beta_{2,3} \text{age[75–84]}_{itk} + \beta_{2,4} \text{age[85}_{+}\text{]}_{itk} + \beta_3 \text{heatwave}_{it} + \beta_{4,2} \text{female} \times \text{age[65–74]}_{itk} + \beta_{4,3} \text{female} \times \text{age[75–84]}_{itk} + \beta_{4,4} \text{female} \times \text{[age 85}_{+}\text{]}_{itk} + \alpha_1 \sin\left(\frac{2\pi i}{52}\right) + \alpha_2 \sin\left(\frac{4\pi i}{52}\right) + \gamma_1 \cos\left(\frac{2\pi i}{52}\right) + \gamma_2 \cos\left(\frac{4\pi i}{52}\right) + \varepsilon_{itk}$ with $\varepsilon_{itk} \sim N(0, \sigma^2)$, $b_{0j} \sim N(0, \sigma_D^2)$, and ε_{itk}, b_{0j} mutually independent.

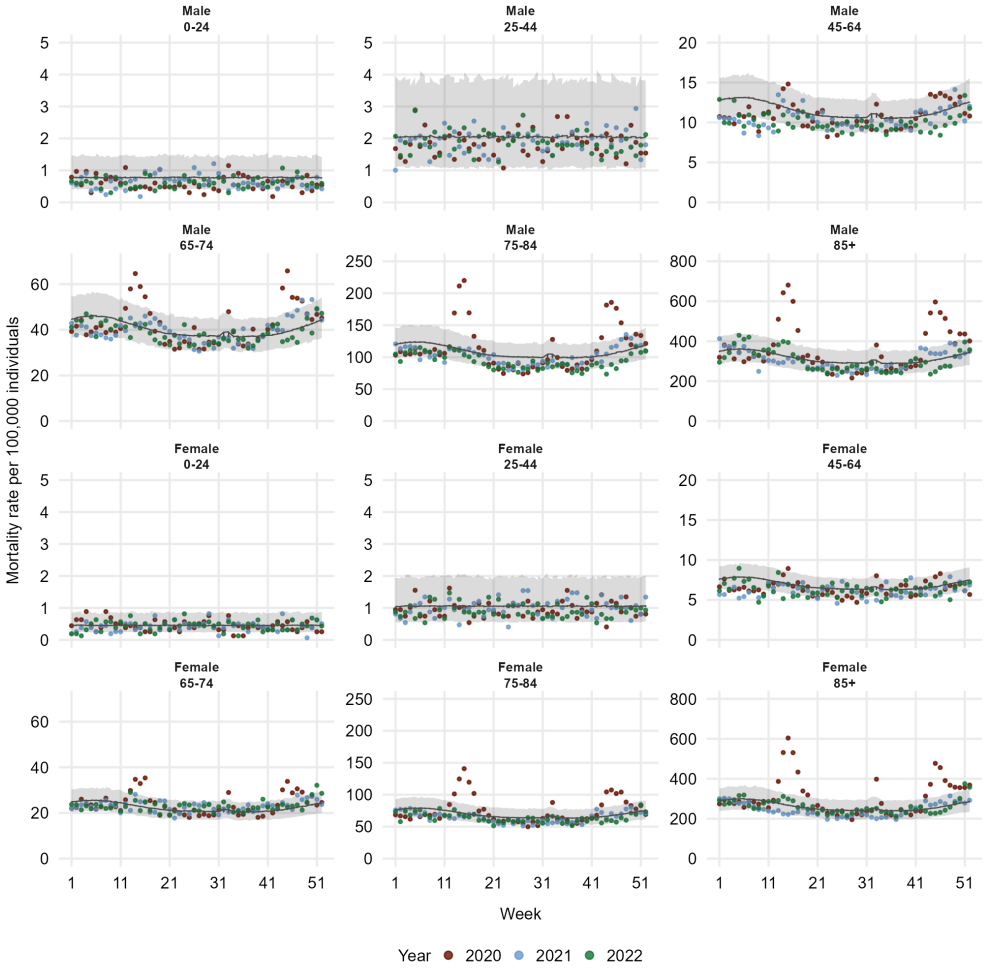


Figure 3.6. Excess mortality. Observed (coloured dots) and predicted (solid line) mortality rate in 2020–2022.

3.3.3 Cause-specific mortality

The monthly cause-specific mortality rates from 2009 to 2021 are shown in Figure 3.7. To ensure clarity in visualization, the causes of death were grouped into three broader categories: COVID-19, external causes, and internal causes (which include heart and vascular diseases, infectious diseases, mental and behavioral disorders, neoplasms, and other causes). A figure with seven groups is available in Appendix B.10.

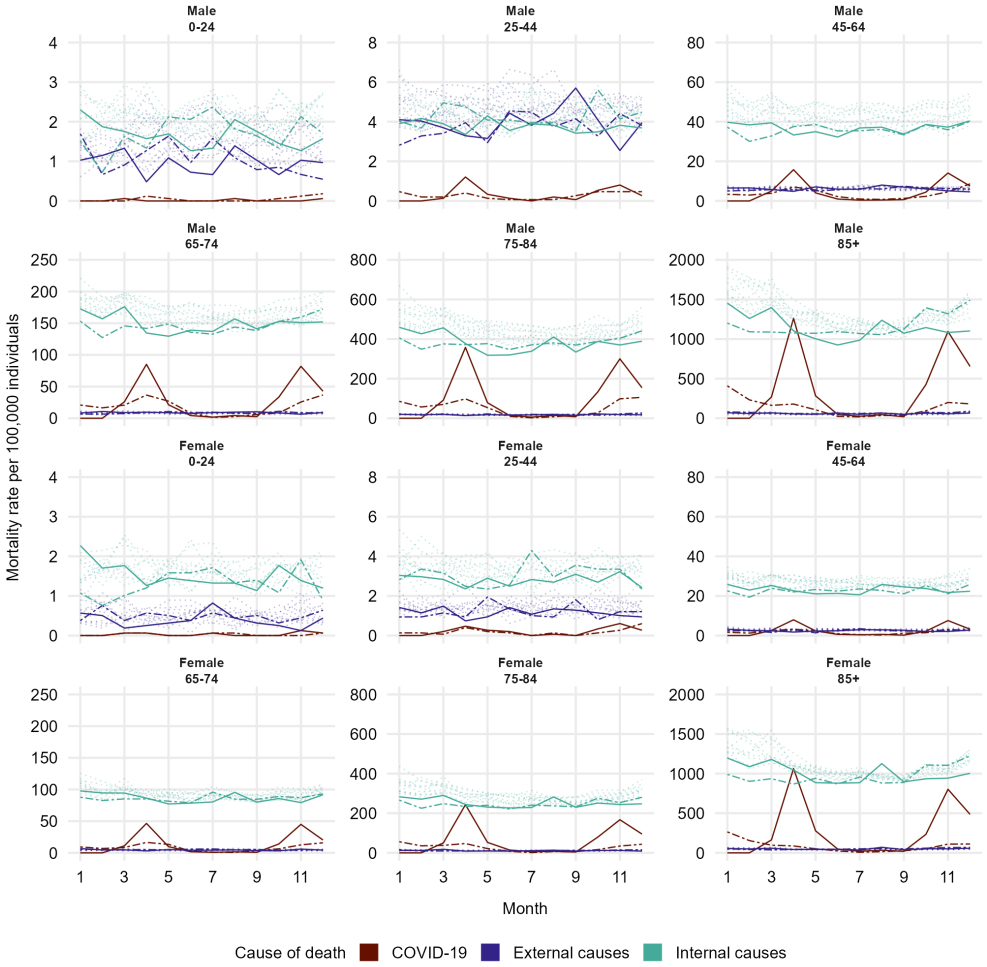


Figure 3.7. Monthly cause-specific mortality rate. Dotted lines: cause-specific rates in year 2009–2019. Solid lines: cause-specific rates in year 2020. Dashed lines: cause-specific rates in year 2021.

Varying trends in causes of death were found in different age groups. External and internal causes contributed almost equally to mortality in younger age groups, while internal causes were clearly more dominant in the older age groups (Figure 3.7). However, when the causes of death were further subdivided, mortality in younger age groups was predominantly attributed to external or other causes, particularly in males aged 25–44 years (Appendix B.10). From the age of 45 years onward, the primary causes shifted specifically towards neoplasms, heart and vascular diseases, or other causes. Interestingly, following the onset of the COVID-

19 pandemic in 2020, there were notable declines in all causes of death except for COVID-19 in all sex and age categories compared to the period 2009 to 2019, as shown by the relatively lower mortality rate in 2020 (Figure 3.7, solid lines) and 2021 (Figure 3.7, dashed lines) compared to the reference years (Figure 3.7, dotted lines). Notably, COVID-19 emerged as the predominant cause of death during the COVID-19 waves in 2020, especially in the older age groups. The trend changed again in 2021 where deaths caused by COVID-19 became considerably lower than in the year 2020.

All fitted models for each cause of death and their corresponding AIC values are presented in Appendix B.11. It should be noted that for COVID-19 mortality, the model with only fixed effects (Reduced model 8) exhibited a lower AIC value compared to Reduced model 5. However, the AIC values were relatively similar between Reduced models 5 and 8 (732.4 for Reduced model 5 and 725.1 for Reduced model 8). To maintain comparability across different causes of death, Reduced model 5 was chosen as the final model for all causes of death, as given by:

$$\begin{aligned}\log Y_{itkc} = {} & (\beta_0 + b_{0tc}) + \beta_1 \text{female}_{itkc} + \beta_{2,2} \text{age[25–44]}_{itkc} + \beta_{2,3} \text{age[45–64]}_{itkc} \\ & + \beta_{2,4} \text{age[65–74]}_{itkc} + \beta_{2,5} \text{age[75–84]}_{itkc} + \beta_{2,6} \text{age[85}_{+}\text{]}_{itkc} \\ & + \beta_3 \text{stringency}_{itc} + \beta_{4,2} \text{female} \times \text{age[25–44]}_{itkc} \\ & + \beta_{4,3} \text{female} \times \text{age[45–64]}_{itkc} + \beta_{4,4} \text{female} \times \text{age[65–74]}_{itkc} \\ & + \beta_{4,5} \text{female} \times \text{age[75–84]}_{itkc} + \beta_{4,6} \text{female} \times \text{age[85}_{+}\text{]}_{itkc} \\ & + \varepsilon_{itkc} \end{aligned} \tag{3.7}$$

with $\varepsilon_{itkc} \sim N(0, \sigma^2)$, $b_{0tc} \sim N(0, \sigma^2_{Dc})$, and ε_{itkc}, b_{0tc} mutually independent.

The complete estimates from the linear mixed models based on Equation 3.7 are provided in Appendix B.12. Significant associations of age group and sex could be observed in most models. However, the interaction between age group and sex was not statistically significant for COVID-19-related mortality. For non-COVID-19 causes, the stringency index demonstrated a significant negative relationship, i.e., a decrease in the mortality rate, for external causes, mental and behavioral disorders, neoplasms, and other causes.

Between 2009 and 2019, some causes of death exhibited a considerable decrease in YLL, while others remained relatively stable (Figure 3.8). When the pandemic started in 2020, the YLL attributed to COVID-19 im-

mediately reached a level similar to that of the main age-specific causes of death before the pandemic, especially in the oldest age groups. An interesting turning point was found in 2021, where the YLL attributed to COVID-19 decreased markedly, while other causes experienced an increase in YLL.

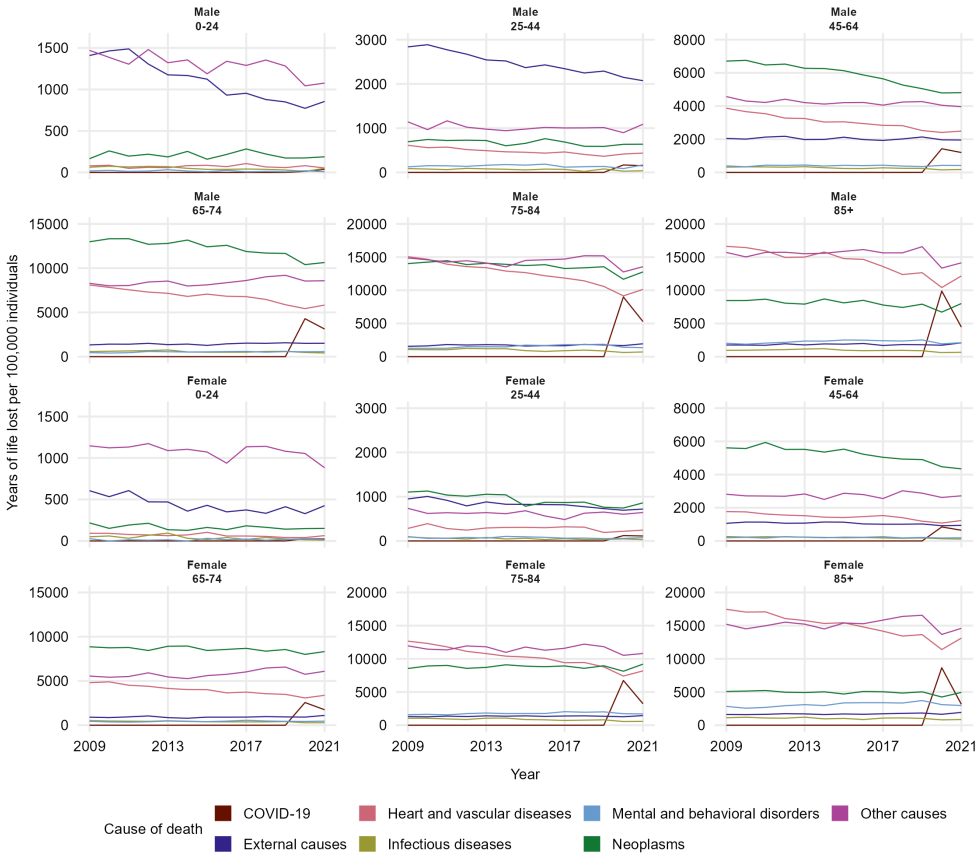


Figure 3.8. Cause-specific years of life lost.

3.4 Discussion

This chapter highlighted important associations of vaccination coverage, the rate of incoming international travel, the positivity rate among travelers, and the stringency index on the dynamics of COVID-19 confirmed cases and hospitalizations. Potential interactions among these variables were also identified, suggesting that their combined impact may be different than simply the sum of their individual contributions. Eventually, these factors would contribute to the observed mortality patterns during

the pandemic period.

Vaccination played a crucial role in reducing the incidence of COVID-19 cases and hospitalizations across various populations and geographic regions (Chen, Huang, et al., 2022; El-Shabasy et al., 2022). However, as indicated by the results in Section 3.3, there is no simple linear relationship between vaccination coverage and these outcomes, as the effect of vaccination was consistently modified by contextual factors, particularly the stringency index (Equations 3.4 and 3.5) and the positivity rate (Equation 3.6). This supports the notion of a non-uniform and context-dependent vaccination effect, rather than a single, constant association. Moreover, the inclusion of confirmed cases in Equation 3.6 highlighted that epidemic dynamics play a crucial role in shaping observed hospitalizations. Within this framework, the previously observed positive associations of vaccination coverage and stringency index as given in Equations 3.4 and 3.5 are certainly not causal, but instead reflect temporal alignment with epidemic phases. Specifically, vaccination rollout often coincided with periods of increasing incidence, while stricter policy measures were typically implemented in response to worsening epidemiological conditions. Together, these results underscore the importance of accounting for temporal dynamics and effect modification when interpreting associations in observational epidemic data.

For most optimal models, a similar lag l of approximately 14–15 weeks was identified. The importance of considering lagged effects when evaluating the impact of vaccination on COVID-19 dynamics has been highlighted due to the biological and immunological timeline of vaccine-induced protection. Polack Fernando et al. (2020) and Dagan et al. (2021) reported a measurable level of protection against SARS-CoV-2 infection typically begins to emerge approximately 14–20 days after one dose of BNT162b2 mRNA Covid-19 Vaccine and 21–27 days with gradual shifting between the first and second vaccine doses. H. Li et al. (2022) reported that the lag time for the protective effects of vaccination was approximately 40 days after the administration of the first dose of primary vaccination, with the potential for a rebound in the epidemic intensity afterward. In contrast, Lokonon et al. (2023) observed a shorter lag of 15–20 days among hospitalized patients and those in intensive care units in Germany. Notably, the shorter lag times reported by these studies compared to optimal models in this thesis may stem from differences in methodological approaches, par-

ticularly their use of daily data, which allows for finer temporal resolution and capturing more immediate effects.

While vaccination has been instrumental in providing individual immunity against COVID-19, it does not entirely eliminate the risk of infection (Tenforde et al., 2021). Increased human mobility within a specific time frame, particularly during periods of high SARS-CoV-2 circulation, increases the likelihood of contact with an infected individual, thereby amplifying transmission risks. As reported by Brackx et al. (2022), COVID-19 incidence was strongly influenced by epidemiological conditions in neighboring countries or regions, particularly during periods when cross-border mobility remained feasible. Differences in the timing, stringency, and enforcement of preventive measures across adjacent jurisdictions allowed continued population movement, facilitating the spread of infections beyond administrative boundaries. In line with this result, a consistent trend was found starting from week 25 (June 21–27) where increases in incoming travel rates and positivity rates among travelers were accompanied by increases in COVID-19 confirmed cases and hospitalizations (Figure 3.1). This pattern reinforces the findings from a previous study in Belgium, which demonstrated that international travel exacerbates COVID-19 incidence and suggested that restricting such travel could significantly mitigate epidemic growth (Nguyen et al., 2023). Similarly, a research from Ukraine reported a surge in COVID-19 cases during the summer of 2021, attributed to increased travel and tourism in the absence of travel restrictions (Yakovleva et al., 2022). In southern Taiwan, an outbreak in the summer of 2021 was linked to returning travelers from abroad, further underscoring the role of imported cases in local outbreaks (Shy et al., 2022).

In addition to vaccination and international travel, several other factors also changed during the investigation period. At the beginning of 2021, the stringency index remained relatively high in response to the resurgence of infections during the fall of 2020 (Natalia, Faes, et al., 2022). The restrictions were gradually eased from mid-April 2021 as part of a phased relaxation strategy (Figure 3.1F) following the nationwide vaccination rollout. The observed decline in stringency measures aligns with broader trends reported in European countries, where increased vaccination coverage facilitated policy shifts towards more lenient restrictions (Ge et al., 2022). The strong association between vaccination coverage and the stringency index

was further supported by the current results (Equations 3.4 and 3.5, see also Appendices B.3 and B.4). However, a widening of the 95% prediction intervals in the final weeks of 2021 could be observed in Figures 3.2–3.4. This increased variability likely reflects the growing complexity of factors affecting COVID-19 incidences, including changes in the dominant SARS-CoV-2 variants. In Belgium, the Alpha variant was predominant during the first half of 2021. However, the Delta variant began its ascent in May 2021, becoming dominant by July 1, 2021, and accounting for nearly 100% of cases by August 1, 2021. Subsequently, the Omicron variant emerged and began circulating in December 2021. These shifts in variants are critical, as each variant exhibits unique transmissibility and immune evasion characteristics (World Health Organization, 2021; Carabelli et al., 2023), which introduced additional complexity to the epidemiological landscape, influencing transmission dynamics and the effectiveness of NPIs.

The mortality trends in Belgium changed considerably during the three years of the COVID-19 pandemic, with distinct patterns between younger age groups (below 45 years) and older age groups (45 years and above). In particular, higher excess mortality could be observed in 2020–2022 starting from age 65 years. Similar age-related trends of excess mortality have been reported in other settings (Vestergaard et al., 2020; Islam et al., 2021). Higher excess mortality in the older age categories might be directly related to the number of COVID-19 infections in these groups (although transmission shifted to younger age groups in September 2020 as reported by Natalia, Faes, et al. (2022)) or caused by other factors such as the presence of comorbidities in older age groups or changes in other causes of death due to COVID-19 mitigation measures. On top of this, a decrease in emergency admissions for acute life-threatening conditions such as stroke or myocardial infarctions would eventually increase mortality (Lange et al., 2020; Khan et al., 2023).

Cause-specific mortality in 2020 and 2021 was generally lower compared to the reference period (2009–2019), with the exception of deaths attributed to COVID-19. Bahk and Jung-Choi (2022) reported similar results in Korea where mortality in 2020 decreased for most causes of death. In Belgium, a notable reduction was particularly evident in deaths due to external causes, such as road accidents or suicides, especially in males aged 25–44 years. This decline may be associated with the implementation of COVID-19 mitigation measures that prohibited travel or recreational

activities. The stratified linear mixed model analysis further supports the positive impact of NPIs in reducing mortality rates associated with external causes, mental and behavioral disorders, neoplasms, and other causes. In contrast to our results, Lee et al. (2023) found an increase in mortality due to external causes among individuals under 44 years, especially during periods of stricter interventions, with a major contribution from unintentional injuries, assaults, homicides, and drug overdoses. This contrast may be attributable to underlying sociodemographic differences related to the sublevel of external causes; however, further investigation was not possible as the available data did not permit disaggregation of external causes.

In older age groups, a more pronounced reduction in cause-specific mortality was observed in chronic diseases such as neoplasms or heart and vascular diseases. This phenomenon could be directly linked to COVID-19 infections, where individuals with pre-existing conditions might succumb to the infection rather than the underlying disease. In addition, delays in diagnosing these conditions may have occurred due to reduced ambulatory care for non-COVID-19 illnesses during the initial year of the pandemic, which could lead to incomplete or inaccurate registration in healthcare databases (Weber et al., 2022; Quaglio et al., 2022; Mafi et al., 2022).

In 2020, a substantial increase in YLL was observed attributable to COVID-19. However, this increase was accompanied by a decrease in YLL from other causes. This combination suggests a compensatory reduction in mortality, also known as the harvesting effect, due to the COVID-19 pandemic (Astengo et al., 2021). In 2021, a reversal in the YLL trends was evident. While the decline in COVID-19-related YLL is likely attributable to the vaccination campaign, which effectively reduced the number of severe cases and associated fatalities, a potential reverse harvesting effect should also be considered. This phenomenon refers to a subsequent increase in mortality due to other causes, including infectious diseases that occur outside their typical seasonal patterns (Walkowiak et al., 2023). Although direct assessment is currently limited by the unavailability of cause-specific mortality data for 2022 in Belgium, indirect evidence supports this possibility. Sciensano reports documented an unusual increase in respiratory syncytial virus (RSV) infections outside of the traditional winter season in 2022 (Sciensano, 2023c) in addition to five waves of COVID-19 and two influenza epidemics that caused a peak of mortality in December 2022 (Sciensano, 2023a).

The combination of various spatiotemporal units and predictor variables provides a comprehensive assessment of the impact of the COVID-19 pandemic in different areas. Some limitations, however, should be noted. The models employed in this chapter are statistical and descriptive in nature, designed primarily for interpolation rather than extrapolation. Their application to predictions beyond the observed data ranges of the predictor variables should be approached with caution. When the emphasis is on mobility patterns across different areas, such as regions or provinces, a meta-population approach might be relevant (Kwok et al., 2021; Parino et al., 2021).

Further, it is important to acknowledge that some data used in this chapter were only available at a certain aggregation level, which limits the ability to capture finer-scale variation. For example, the stringency index was reported only at the national level and mortality data was aggregated by broad age groups. Consequently, these data do not reflect heterogeneity at more localized spatial or demographic levels. Moreover, the analysis assumed a uniform evolution of the SARS-CoV-2 variants in the entire country, a simplification that may not hold when examined at a finer spatial or temporal resolution. Future studies incorporating more granular data are essential to better disentangle these effects and refine our understanding of the interplay between these variables.

CHAPTER 4

APPLICATION OF FRACTAL DIMENSION TO CHARACTERIZE COVID-19 TIME SERIES PATTERNS

This chapter is based on the following publications:

- **Natalia, Y. A.**, Faes, C., Neyens, T., Chys, P., Hammami, N., Molenberghs, G. (2023). Fractal dimension based geographical clustering of COVID-19 time series data. *Sci Rep* 13, 4322. DOI: 10.1038/s41598-023-30948-7.
- **Natalia, Y. A.**, Faes, C., Neyens, T., Hammami, N., Molenberghs, G. (2023). Key risk factors associated with fractal dimension based geographical clustering of COVID-19 data in the Flemish and Brussels region, Belgium. *Front. Public Health* 11:1249141. DOI: 10.3389/fpubh.2023.1249141.

4.1 Introduction

The results presented in Chapters 2 and 3 highlight key insights derived from spatial and longitudinal analyses of COVID-19 incidence and mortality. Such models would provide more benefit at the highest possible spatiotemporal resolution, for example daily counts rather than weekly aggregates, or postal code-level data instead of municipality-level, as finer granularity enables a more detailed and localized understanding of transmission dynamics (Lawson, 2018). However, data on infectious diseases

at this level of detail are often characterized by considerable heterogeneity.

During relatively stable periods, fluctuations in new cases or hospitalizations may be minimal, resulting in sparse or static data that can challenge model estimation. On the other hand, during epidemic surges, incidence patterns may change rapidly, with sharp peaks and regional variation (Viboud et al., 2016). These periods of high transmission often coincide with increased strain on surveillance systems, leading to a greater risk of missing or delayed data reporting (Dureau et al., 2013; Gibbons et al., 2014). Eventually, this spatial and temporal variability introduces considerable complexity into the modeling process (Gostic et al., 2020).

Furthermore, as the granularity of the data increases, so does the computational burden required to fit the models, especially when employing Bayesian or hierarchical approaches (Rue et al., 2009). In the presence of both increased heterogeneity and computational demand, there is a trade-off between resolution and model stability. Consequently, while high-resolution data are ideal in theory, practical constraints, such as data completeness, quality, and processing capacity, must be carefully considered in the design and interpretation of spatiotemporal analyses in infectious disease research (Elliott and Wartenberg, 2004).

Recently, the fractal dimension has gained attention as a promising metric to characterize the complexity of spatiotemporal datasets, particularly in the context of epidemiological modeling (Bianchi and Frezza, 2017; Ruiz de Miras et al., 2023; Tan et al., 2025). Many epidemiological processes exhibit self-similarity and irregularity across different temporal and spatial scales, features that are hallmarks of fractal structures (Mandelbrot, 1982). The fractal dimension provides a quantitative measure of this complexity and has been successfully applied to analyze disease spread patterns. For instance, Păcurar and Necula (2020) demonstrated that fractal analysis could effectively capture key characteristics of epidemic outbreaks. Furthermore, Castillo and Melin (2021) proposed a hybrid approach that integrates fractal theory with fuzzy logic to forecast COVID-19 time series, offering a novel method to deal with uncertainty and nonlinearity in epidemic dynamics. Some studies combined mathematical modeling with fractal dimensions to assess transmission and control of COVID-19 cases (Arfan et al., 2021; Zhou et al., 2022). These findings suggest that the fractal dimension could be used as a valuable complementary tool

to summarize and compare epidemic curves across regions and time periods, especially in settings where traditional metrics may fail to capture the inherent variability and complexity of transmission patterns.

In this chapter, the COVID-19 incidence patterns were evaluated through the lens of fractal dimension analysis. The underlying assumption is that areas with more intricate and variable incidence trends reflect higher structural complexity, which eventually leads to higher estimates of fractal dimension. Based on the statistical characteristics captured through the fractal dimension measures, it becomes possible to identify clusters of areas that exhibit similar epidemic dynamics. The proposed approach was initially tested on simulated data to evaluate its performance before being applied to real-world COVID-19 data. On top of this, many studies highlighted the importance of socio-demographic factors in shaping the transmission of infectious diseases, including COVID-19 (McLaughlin et al., 2021; Sartorius et al., 2021; Lo et al., 2021). Given the multidimensional and dynamic nature of these factors, it is essential to explore their various combinations and associations with epidemiological indicators. In this context, the relationship between sociodemographic characteristics and the complexity of local epidemic curves captured through fractal dimension estimates was investigated at a fine spatial resolution. This analysis contributes to the methodological and applied public health literature by introducing a novel approach to characterize epidemic patterns in small geographic units using routinely available sociodemographic data. The findings may inform the development of more tailored and responsive interventions in future public health crises.

4.2 Materials and methods

4.2.1 Study area

As shown in Figure 2.2, Belgium is administratively divided into 581 municipalities. These municipalities are further subdivided into statistical sectors, the smallest administrative units, defined based on structural characteristics such as social composition, economic activity, urban planning, and morphological features (Jamagne et al., 2020). In 2020, Belgium comprised a total of 19,794 statistical sectors: 9194 in the Flemish Region, 724 in the Brussels-Capital Region, and the remaining sectors in the Walloon Region, as illustrated in Figure 4.1.

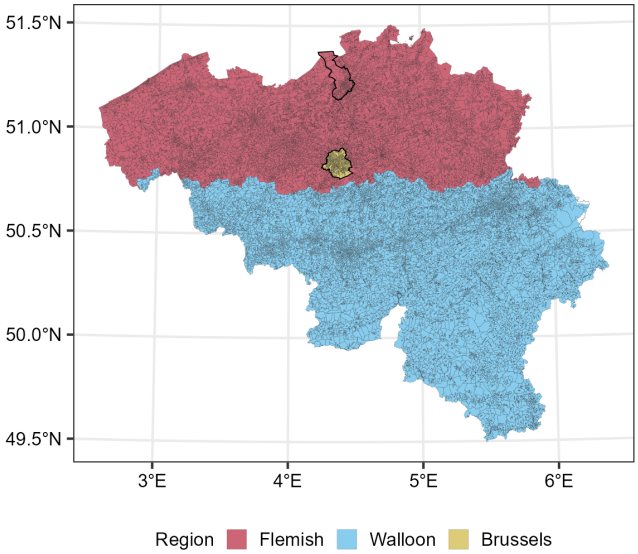


Figure 4.1. Statistical sectors in Belgium. The location of selected study area described in Subsection 4.2.3.1 is marked with thick black lines.

4.2.2 Data

4.2.2.1 Simulation

To obtain extra background information on the proposed method, possible daily incidences of COVID-19 cases at the statistical sector level were simulated using white noise from an ARIMA model, defined as:

$$y_t = \mu + \omega_t, \tag{4.1}$$

where y_t is the number of cases at a certain time point, μ is a constant, and error term $\omega_t \sim Po(\lambda)$. The Poisson distribution was selected to generate ω_t since the simulated outcome represents daily counts of COVID-19 cases, which are discrete, non-negative, and typically modeled using count-based distributions. Different values of the λ parameter in the Poisson distribution defined the expected number of events per time unit and were used to increase the complexity of the daily incidence curve. It should be noted that the λ parameter is constant over time within each simulated series, while different values of λ are used across simulations to represent different incidence levels. Specifically, higher values of λ correspond to higher expected counts and greater variability in the simulated series, which in

turn results in more irregular and complex incidence curves. Thus, the term ω_t refers to the stochastic realization of error at each time point, not to the changes in the underlying λ parameter.

4.2.2.2 Real world

Individual data of daily confirmed COVID-19 cases at the statistical sector level were retrieved via the ZorgAtlas platform, managed by the Department of Care of the Flemish Region (Departement Zorg, 2023). Although the agency primarily collects data for the Flemish Region, it also integrates data from the Brussels-Capital Region. Consequently, the analyses in this study are focused on these two regions.

The study period was defined from January 1 to December 31, 2021, corresponding to the start of the COVID-19 vaccination campaign in Belgium. Arguably, the vaccination coverage has an impact on COVID-19 transmission, especially in the short term. Therefore, vaccination data for the Flemish and the Brussels-Capital Region were also retrieved in collaboration with the Joint Community Commission of Brussels (De Gemeenschappelijke Gemeenschapscommissie, 2023). Vaccination coverage was calculated as the percentage of fully vaccinated residents (i.e., with primary doses) per statistical sector. Additional population data and statistical sector shapefiles were obtained from Statbel (Statbel, 2025).

Key socio-demographic factors used in this study were ethnic composition, yearly median income as a proxy for socioeconomic status, number of inhabitants, population density per km^2, and the proportion of older age group ($\geq$ 50 years). To summarize ethnic composition within each statistical sector, the Shannon diversity index was calculated, with higher values indicating greater ethnic diversity (Spellerberg and Fedor, 2003). Specific for Flemish Region, additional indicators were considered, including the proportion of the population reporting high trust in federal and regional governments and satisfaction with healthcare services, based on a municipality-level survey conducted in 2020 (Agentschap Binnenlands Bestuur, 2023).

4.2.3 Statistical methods

The following methods were performed in R through packages fractaldim (Gneiting et al., 2012) and candisc (Friendly and Fox, 2021).

4.2.3.1 Fractal dimension analysis

Time series data often exhibit (1) self-similarity across scales so that repeated patterns could be observed in short and long horizons, (2) scaling laws that follow power-law decay, and (3) complexities that are not well described by linear models. When a real-valued time series $X(t)$ is viewed as the graph $t \mapsto X(t)$ in the plane, its geometric roughness can be quantified by a fractal dimension D of that graph. For continuous-time processes observed on a fine grid, typical values satisfy $1 \leq D \leq 2$ where $D = 1$ corresponds to a structure that is similar to a straight line, while $D = 2$ indicate a more 'space-filling' structure (Figure 4.2). Thus, a higher fractal dimension value indicates a higher complexity of the daily COVID-19 incidence curve.

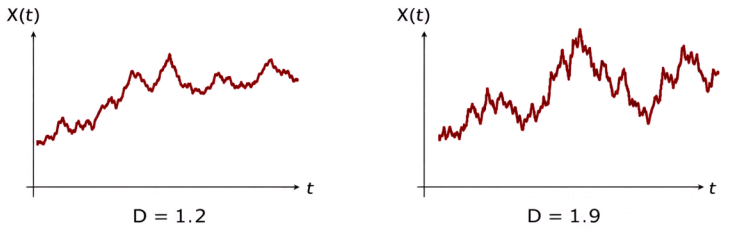


Figure 4.2. Comparison of two time series with different geometric complexities. Left panel: a smoother trajectory with lower fractal dimension. Right panel: a more irregular, highly fluctuating trajectory with higher fractal dimension.

The fractal dimension of a time series can be estimated using a variety of techniques, each with different theoretical underpinnings and practical advantages. In this thesis, four fundamental estimators were employed: the box-count, Hall–Wood, variogram, and madogram methods.

4.2.3.1.1 Boxcount

The box-count estimator quantifies the rate at which the time series, develop at increasingly fine scales (Falconer, 2013). In practice, this means that the time series data are first covered by the largest binding box, which is subsequently divided into four quadrants. Each quadrant is iteratively further divided into four sub-quadrants until the resulting box widths are equal to the resolution of the data. Afterwards, the number of boxes required to cover the time series curve is counted for each box

width and the box-count estimator takes the following form:

$$\mathsf{D}_{BC} = \lim_{\delta \to 0} \frac{\log N_\delta}{-\log \delta}, \tag{4.2}$$

where N_δ is the number of required boxes to cover the time series data and δ is the width of the box. The fractal dimension is equal to the slope in an ordinary least squares regression fit of $\log N_\delta$ on $-\log \delta$ (Gneiting et al., 2012). This estimator is widely used due to its simple and intuitive formulation. However, Dubuc et al. (1989), Liebovitch and Toth (1989), and Hall and Wood (1993) reported increasing mean squared errors and asymptotic bias with increasing number of points in the regression.

4.2.3.1.2 Hall-Wood

Hall and Wood (1993) introduced a modification of the box-count estimator that operates on the smallest observable scale. This method uses, instead of the number of boxes, the area of the boxes that cover the time series line. The total area of the boxes of scale δ that intersect with the time series data is denoted as A_δ, which is proportional to $N_\delta \times \delta^2$. This leads to a reformulation of definition (4.2) into:

$$\mathsf{D}_{HW} = 2 - \lim_{\delta \to 0} \frac{\log A_\delta}{\log \delta} \tag{4.3}$$

A natural variant of the box-count estimator uses scale $\delta = l/n$, where n is the total number of data series and $l = 1, 2, \ldots, n$. On this scale, an estimator of $A(l/n)$ of time series data X_t is given by :

$$A(l/n) = \frac{l}{n} \sum_{i=1}^{\lfloor n/l \rfloor} |X_{(i)l/n} - X_{(i-1)l/n}| \tag{4.4}$$

where $\lfloor n/l \rfloor$ denotes the number of points in the regression, which is the greatest integer less than or equal to n/l. The Hall–Wood estimator is based on an ordinary least squares regression fit of $\log A(l/n)$ on $\log l/n$,

$$\mathsf{D}_{HW} = 2 - \frac{\sum_{l=1}^{L} (s_l - \bar{s}) \log A(l/n)}{\sum_{l=1}^{L} (s_l - \bar{s})^2}, \tag{4.5}$$

with $L \leq 2$, $s_l = \log(l/n)$ and $\bar{s} = \frac{1}{L}\sum_{l=1}^{L} s_l$. To minimize the bias, Hall-Wood recommended to use a L equal to two, which leads to a simplified implementation as follows :

$$\mathsf{D}_{HW} = 2 - \frac{\log\ A(2/n) - \log\ A(1/n)}{\log\ 2} \tag{4.6}$$

Hall-Wood estimator is more robust than a simple box-counting, however it can only be applied to a stationary Gaussian process with equally spaced points (Hall and Wood, 1993).

4.2.3.1.3 Variogram and madogram

The variogram is a function that characterizes temporal dependence by quantifying the average variability between data points as a function of distance or lag t (Burrough, 1981; Constantine and Hall, 1994; Kent and Andrew, 1997). For a stochastic process with stationary increments, the p-th order variogram is defined as

$$\gamma_p(t) = \frac{1}{2}\mathbb{E}[(X_u + X_{u+t})^p] \tag{4.7}$$

with $p > 0$ denoting the order of the moment. For discrete data with lag $t = l/n$, this expression can be rewritten as:

$$V_p(l/n) = \frac{1}{2(n-l)}\sum_{i=l}^{n} |X_{i/n} - X_{(i-l)/n}|^p. \tag{4.8}$$

The fractal dimension can then be estimated by regressing $\log\ V_p(l/n)$ on $\log\,(l/n)$, yielding the simplified estimator:

$$\mathsf{D}_{V;p} = 2 - \frac{1}{p}\frac{\log\ V_p(2/n) - \log\ V_p(1/n)}{\log\ 2}. \tag{4.9}$$

Here, the classical variogram estimator corresponds to $p = 2$, whereas the madogram estimator corresponds to $p = 1$. Both are generally more efficient than the Hall–Wood estimator. In addition, the madogram provides greater robustness to outliers, making it particularly advantageous in the presence of heavy-tailed distributions or irregular fluctuations. However, an important limitation is that these methods, especially the variogram estimator, may perform poorly for certain non-Gaussian processes, where

the assumed scaling relationships break down (Gneiting et al., 2012).

4.2.3.1.4 Application to the available data

The local fractal dimension was estimated for each statistical sector within a certain period to capture the evolution of daily COVID-19 incidence rates. To account for short-term fluctuations in reporting, sliding windows of 7, 14, and 21 days were applied to estimate the fractal dimensions over time. These estimates were plotted to generate a fractal dimension curve, where each point reflects the estimate for the sliding window ending on a specific date. Subsequently, fractal dimension indicators, i.e., the mean, variance, and autocorrelation of the fractal dimension curves, were computed for each statistical sector. These indicators were visualized using choropleth maps to examine spatial patterns in the complexity of COVID-19 incidence dynamics.

Qualitative combinations of the mean, variance, and autocorrelation values can be used to characterize statistical sectors in terms of incidence pattern complexity. However, as noted by Gneiting et al. (2012), different methods for estimating fractal dimensions may yield varying results, which in turn influence the calculated mean, variance, and autocorrelation values. To address this and identify groups of statistical sectors with similar local fractal dimension indicators, k-means clustering was applied. This method, representing the simplest form of finite mixture models, offers computational efficiency while effectively detecting clusters of comparable patterns (Steinley, 2006). The optimal number of clusters was first determined using the elbow method, which is heuristic but simple to implement. This method optimizes the number of clusters based on the sum of squares of the Euclidean distances between each point and its corresponding centroid. The relationship between the sum of the square and the possible number of clusters k is plotted in a curve. The curve will be flattened out when the value of k increases and the optimal number of clusters lies in the highest curvature of elbow; i.e., adding another cluster will not give additional benefit to classifying the data (Sammouda and El-Zaart, 2021). The value of each centroid cluster was obtained and compared with the mean of local fractal dimension indicators. For each of these, the label 'low' indicates that the centroid values are lower than the mean of the respective indicator and the label 'high' indicates that the centroid values are higher than that.

4.2.3.2 Canonical correlation analysis

For each statistical sector, the various socio-demographic factors and vaccination coverage were collectively referred to as population indicators in subsequent sections. Given the presence of multiple, inter-correlated local fractal dimension indicators (mean, variance, and autocorrelation) alongside multiple population indicators, canonical correlation analysis (CCA) was employed to assess the relationships between these two sets of variables.

Given two sets of variables $X = X_1, X_2, \ldots, X_m$ and $Y = Y_1, Y_2, \ldots, Y_n$, CCA seeks the orthogonal linear combinations of the variables within each set of indicators based on a weighted average, such that the linear combination of the X variables (i.e. the population indicators that include ethnic diversity index, median income, population size, population density, the proportion of older age group, vaccination coverage, trust, and satisfaction in the government), denoted as U, given by:

$$U = a_1 X_1 + a_2 X_2 + \cdots + a_m X_m \tag{4.10}$$

and the combination of Y variables (i.e., the mean, variance, and autocorrelation value of the local fractal dimension), denoted as V, given by:

$$V = b_1 Y_1 + b_2 Y_2 + \cdots + b_n Y_n \tag{4.11}$$

has a maximum correlation. U and V are the so-called canonical variates that will be used to explain the correlation both within and between sets with constraints that Cov(U_j, U_k), Cov(V_j, V_k), and Cov(U_j, V_k) are equal to 0 for all $j \neq k$, j & $k \in 1, \ldots, i$ (ter Braak, 1986). The number of canonical variates i is equal to the smallest set of variables so that $i = 3$ in this study. The association between X and Y variables is evaluated using canonical loading values, which signify the degree of correlation between these variables and their canonical variate. Higher canonical loadings serve as an indicator of a stronger contribution to the latent dimension represented by its own canonical variate. In addition, the sign of a canonical loading determines the direction of their correlation. A positive loading indicates a positive contribution to the canonical correlation, thus establishing a positive association with other variables exhibiting positive loadings on the same canonical variate. This interpretation is equally applicable to nega-

tive canonical loadings, which denote a contrary orientation of the association. The redundancy analysis was used to assess the amount of variability in the fractal dimension indicators that can be explained by the population indicators.

4.3 Results

4.3.1 Simulation study

An example of the simulated curves and their fractal dimension are shown in Appendix C.1. Daily incidence rate curves with smaller values of λ (ranging from 0.01 to 0.11) reflect patterns of mild to heavy sporadic transmission, whereas curves with λ values between 0.16 and 0.26 mimic a mild community transmission. Higher λ values correspond to more intense community transmission dynamics. The fractal dimension curves, derived using various estimation methods, are displayed in the upper section of each panel. Notably, different estimators yielded varying estimates, with the box-counting method consistently producing the lowest fractal dimension values. Fractal dimension curves based on 7-day sliding windows exhibited more pronounced fluctuations compared to those using 14- or 21-day windows. For each curve, fractal dimension indicators (i.e., mean, variance, and autocorrelation) were computed to further characterize their behavior.

Following 1,000 simulation replications, a consistent increase in the mean value of the fractal dimension curve could be observed with increasing λ values (Appendix C.2, top panels), reinforcing the interpretation that higher mean values reflect greater complexity in the incidence curves, which is a likely characteristic of community transmission, where new cases occur more consistently at successive time points. The variance value was higher at lower λ values (middle panels), suggesting greater local fluctuations in transmission during sporadic and early community transmission. In contrast, a lower variance at higher λ values indicates more stable, sustained transmission dynamics. The boxplots of autocorrelation values (bottom panels) further revealed a high temporal dependency across all λ values, especially when using the variogram and madogram estimators, underscoring the persistence of transmission patterns captured by these methods.

For each estimator and sliding window, the optimal number of clusters

was identified at $k = 4$, corresponding to the maximum curvature point in the elbow plot. Beyond this value, the rate of change in curvature diminished, justifying the choice of $k = 4$ for subsequent classification using k-means clustering as shown in Appendix C.3. The X- and Y-axes indicate the dimension with the largest explained variance after dimensionality reduction using principal component analysis. The resulting clusters exhibited a clear separation between lower and higher λ values, reflecting distinct transmission patterns in the simulated data. Cluster centroids were compared against the mean values of local fractal dimension statistics, as indicated in the accompanying legends. For clarity, the classification scheme is summarized in Table 4.1.

Table 4.1. Proposed classification of local fractal dimension based on mean, variance, and autocorrelation value.

Mean	Variance	Autocorrelation	Transmission type
Low	Low	Low	Mild sporadic
Low	Low	High	Mild sporadic
Low	High	Low	Heavy sporadic
Low	High	High	Heavy sporadic
High	High	Low	Mild community
High	High	High	Mild community
High	Low	Low	Heavy community
High	Low	High	Heavy community

For comparison purposes, the simulated data were also analyzed using a previously known approach called dynamic time warping as implemented in the package `dtwclust`. This method utilizes the dynamic time warping distance as a dissimilarity measure to find the optimum warping path between two series under certain constraints (Sardá-Espinosa, 2019). The results showed that the proposed fractal dimension method can detect clusters with similar features as dynamic time warping, i.e., incidence rate curves with similar shapes are clustered together. The time required to complete an analysis is relatively short for both methods (2.7 seconds for local fractal dimension and 1.3 seconds for k-means dynamic time warping). An advantage of the fractal dimension method is the ability to retrieve the characteristics of each cluster via the mean, variance, and autocorrelation values classification to directly compare different curves. An analog of this feature is not readily available for dynamic time warping, given the necessity to manually compare each centroid curve to retrieve the characteristics of each cluster.

4.3.2 Real world data analysis

Between January 1 and December 31, 2021, a total of 917,565 COVID-19 cases with a known residential statistical sector were reported in the Flemish Region and the Brussels-Capital Region. As illustrated in Figure 4.4, the COVID-19 incidence rate in the three study areas (i.e., the Flemish Region, the Brussels-Capital Region, and the municipality of Antwerp) followed broadly similar temporal patterns, with slightly higher incidence observed in the municipality of Antwerp during the first half of 2021. Incidence declined in June 2021 before rising again in July, leading to a pronounced peak in early December 2021 in the Flemish Region and in late December 2021 in the Brussels-Capital Region. Between August and October 2021, incidence in the municipality of Antwerp remained between the Flemish and Brussels-Capital Regions, after which it stayed lower than both regions through the end of 2021.

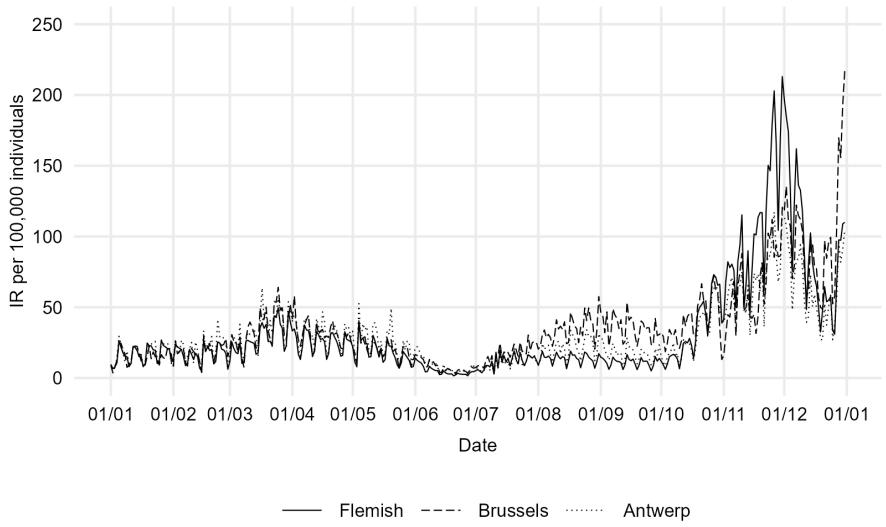


Figure 4.3. Daily incidence rate (IR) per 100,000 individuals in 2021.

4.3.2.1 Local fractal dimension and cluster detection

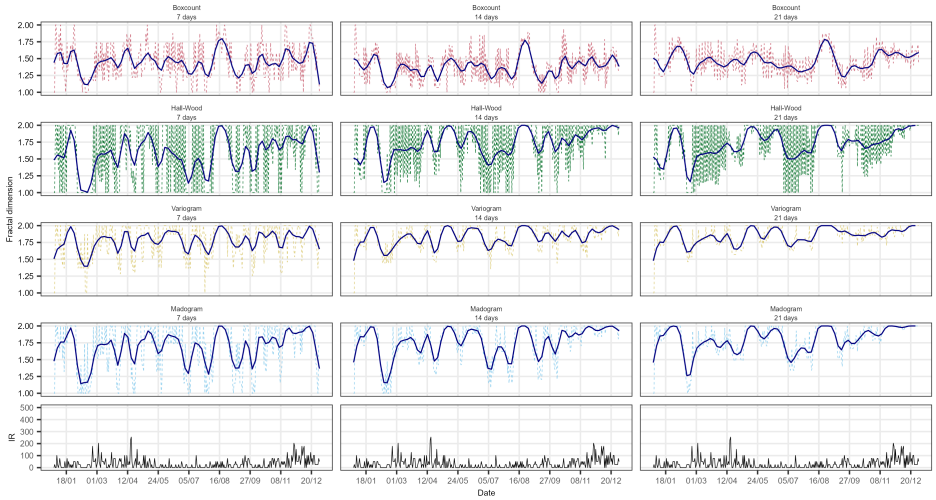
Given the extensive number of statistical sectors in the dataset, the results in this subsection are limited to the municipality of Antwerp and the Brussels-Capital Region (Figure 4.1, black thick lines). Antwerp, located in the northern part of Flemish Region, comprises 299 statistical sectors,

while the Brussels-Capital Region, situated centrally within Belgium, includes 724 sectors. Around 529,247 inhabitants were registered in the municipality of Antwerp while the Brussels-Capital Region was home to 1,215,012 individuals. These two areas represent typical urban environments in Belgium, characterized by middle socioeconomic status, high population density, and a heterogeneous composition of national and cultural backgrounds (Statbel, 2025).

To illustrate the difference of local fractal dimension curves in a real world setting, two statistical sectors in the municipality of Antwerp were selected: *De Peperbus* and *Prinshoeveland*. COVID-19 cases have been continually reported in *De Peperbus* since January 2021 with higher daily incidence rates in April and December 2021 (Figure 4.4A). In contrast to this sector, *Prinshoeveland* had relatively lower incidence rates with a longer period of no reported cases (Figure 4.4B). Similar to Figure 4.2 (right panel), the incidence curve in *De Peperbus* exhibited a more 'space-filling' pattern compared to *Prinshoeveland*. This notion is further confirmed by relatively higher fractal dimension estimates with lower variability of the fractal dimension curve, which indicates a higher complexity of the COVID-19 daily incidence curve in this sector.

The fractal dimension indicators of each fractal dimension curve were computed and visualized as shown in Appendices C.4 and C.5. The combination of these indicators revealed similarities among statistical sectors despite having different incidence rate curves. Based on these metrics, clusters representing transmission type as proposed in Table 4.1 were subsequently identified and are illustrated in Figures 4.5 and 4.6. All four estimators identified similar transmission types when applying a 7-day sliding window in both study areas. However, the use of longer sliding windows, particularly with the variogram and madogram estimators, revealed a greater tendency to detect heavy sporadic and heavy community transmission. Despite minor discrepancies between estimators, the results consistently indicated clusters of mild to heavy community transmission concentrated around the city center and, to a lesser extent, the northern part of the Antwerp municipality. In contrast, the Brussels-Capital Region exhibited a more widespread distribution of such clusters across the entire region.

(A) *De Peperbus*



(B) *Prinshoeveland*

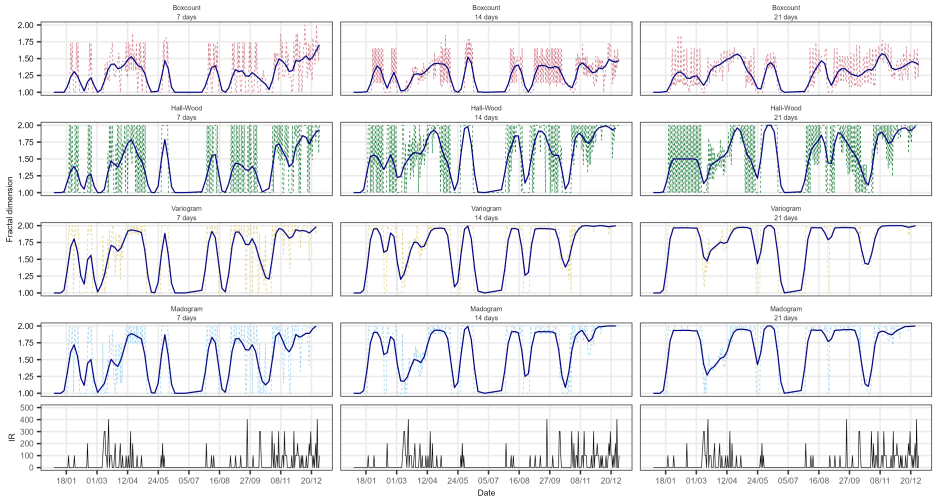


Figure 4.4. Local fractal dimension of COVID-19 incidence rate in selected areas. For ease of interpretation, the fractal dimension curve was smoothed using locally estimated scatterplot smoothing (LOESS).

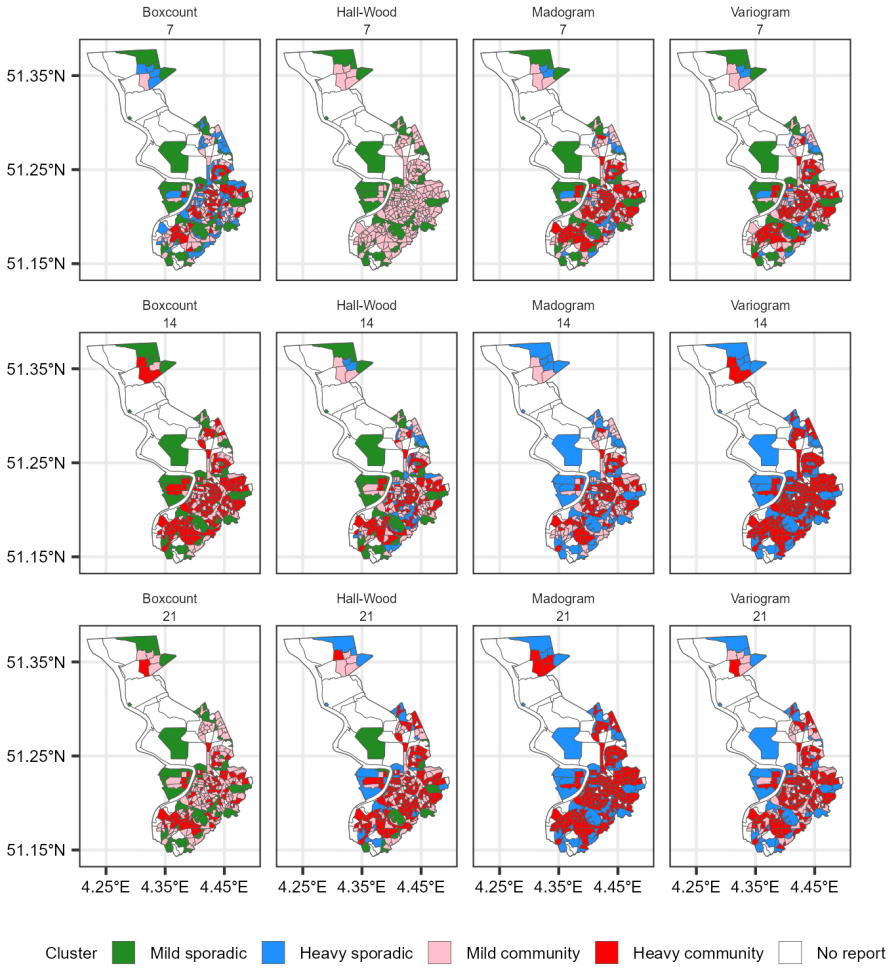


Figure 4.5. Transmission types detected in the municipality of Antwerp.

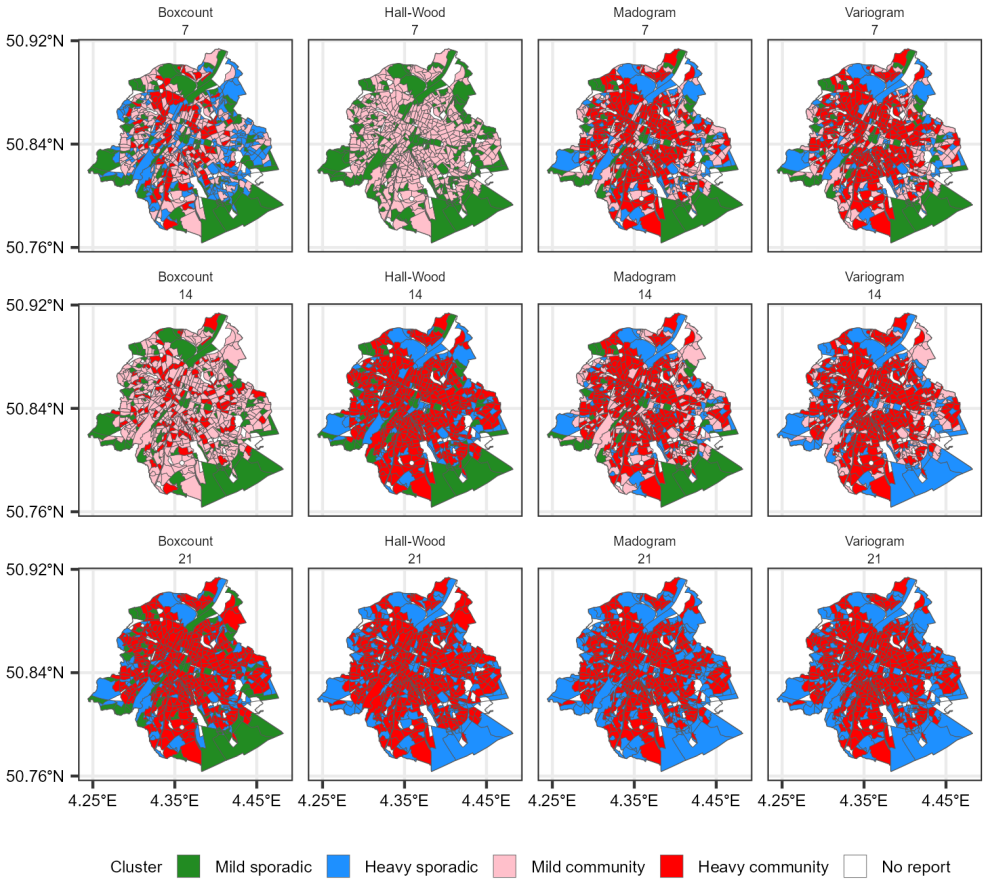


Figure 4.6. Transmission types detected in the Brussels-Capital region.

4.3.2.2 Association between population indicators and fractal dimension curve

The fractal dimension indicators per statistical sector are presented in Appendix C.6. Higher mean values were observed in numerous statistical sectors, particularly in urban areas such as Brussels, Antwerp, and Ghent, indicating greater complexity in the local COVID-19 incidence curves. Areas exhibiting high mean and low variance values were indicative of sustained community transmission, whereas higher variance values suggested more sporadic patterns.

Based on country or region of origin, the population was categorized into 14 ethnic groups: (i) Belgium, (ii) the Netherlands, (iii) France, (iv) North and other Western European countries, (v) Southern Europe, (vi) Eastern European members of the European Union, (vii) Eastern European non-members of the European Union, (viii) Organisation for Economic Co-operation and Development (OECD) countries, (ix) Maghreb countries, (x) other African countries, (xi) Asia, (xii) Turkey, (xiii) Central and South America, and (xiv) unknown origin. The Shannon diversity index ranged from 0 (no diversity) to 2.38 (very high diversity). Higher index values were observed in statistical sectors located in larger cities, along the Dutch border, and in former coal-mining municipalities in eastern Flemish Region (Figure 4.7A).

Yearly median income across statistical sectors varied widely, ranging from €2,213 to €55,949 (Figure 4.7B). In 2021, around 68% of Belgium's population resided in the Flemish and Brussels-Capital regions with highest number of inhabitants (Figure 4.7C) and density (Figure 4.7D) in major urban centers, particularly in the municipality of Antwerp and in the Brussels-Capital Region. Individuals aged 50 years and older comprised slightly more than 39% of the total population, with 43% of them residing in the Flemish Region. Higher proportions of older age residents were found in the northeast and coastal areas (Figure 4.7E). Vaccination coverage was notably high across Flemish Region, with most statistical sectors achieving 75–100% primary full vaccination (Figure 4.7F). In contrast, coverage in the Brussels-Capital Region generally ranged between 50–75%.

In the Flemish Region, higher proportions of people with trust in both the federal and regional governments were predominantly observed in larger municipalities around the capital of each province (i.e., Antwerp, Ghent, Leuven, Hasselt, Bruges) and to some extent in the coastal areas

(Figures 4.7G-H). Most municipalities also reported a high proportion of residents expressing satisfaction with the healthcare system (Figure 4.7I).

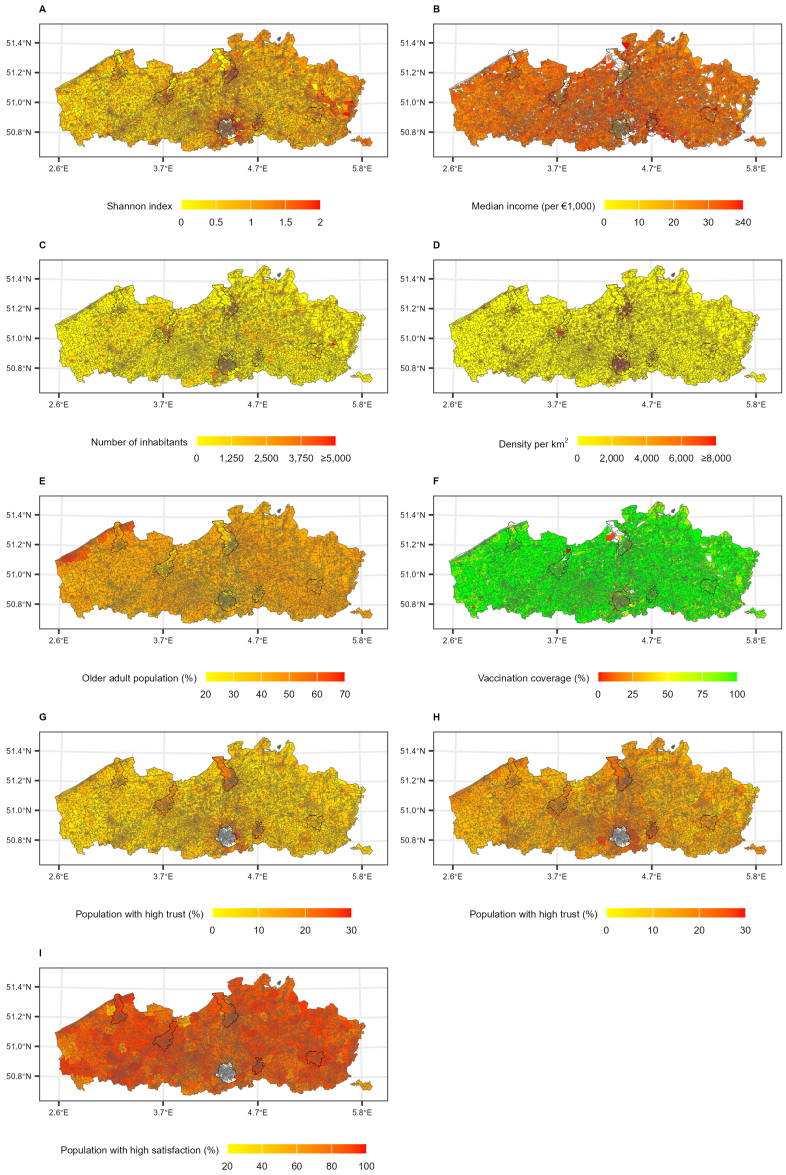


Figure 4.7. Population indicators per statistical sector: (A) Shannon index, (B) yearly median income, (C) population size, (D) population density, (E) proportion of older adult population, (F) vaccination coverage, (G) trust to the federal government, (H) trust to the regional (Flemish) government, and (I) satisfaction with healthcare services. The capital of each province is marked with a black line. The white color indicates no values.

Statistical sectors with missing autocorrelation values (due to an absence of reported COVID-19 cases during the study period) were excluded from the CCA, resulting in a final sample of 9,517 sectors. Considering different sets of population indicators between Flemish and Brussels-Capital regions, the CCA was performed separately for each region. There were at least two significant correlations among three canonical variates in both regions, as shown in Tables 4.2 and 4.3. For each method, the first canonical variate showed a very strong correlation and explained more than 93% of the correlation between the two sets of indicators.

The canonical loading between each set of indicators and the first canonical variate based on a 7-day sliding window are presented in Table 4.4 (results based on longer sliding windows are shown in Appendix C.7). Population indicators with an absolute canonical loading above 0.5, such as population size and population density, demonstrated the strongest explanatory power. For fractal dimension indicators, the first canonical variate was strongly represented by the mean, variance, and autocorrelation values. In the Flemish region, the canonical loadings for population size, population density, Shannon index, satisfaction with healthcare, and trust in the government shared the same sign as the mean and variance of fractal dimension indicators, indicating that higher values of these population indicators were associated with higher mean and variance values of the fractal dimension. In contrast, elderly proportion, median income, and vaccination coverage exhibited opposite signs, suggesting that higher values of these indicators were linked to lower mean and variance values. Similar trends were observed in the Brussels-Capital Region, except for predominantly positive canonical loading of the fractal dimension variance.

Redundancy analysis indicated that a substantial proportion of the variability in fractal dimension indicators could be attributed to population characteristics, although this relationship varied by region and estimator. In the Flemish Region, between 27% and 53% of the total variability was explained, with the box-count estimator providing the highest explanatory proportion. In contrast, the Brussels-Capital Region exhibited a lower explanatory range (14%–45%), with the variogram estimator accounting for the largest share. These findings suggest that the strength and nature of the association between spatial complexity and population factors are context-dependent, and may be influenced by both regional characteristics

and the methodological approach used to quantify fractal properties.

Table 4.2. Canonical correlation between population indicators and fractal dimension indicators in the Flemish region.

Estimator	Sliding window	Canonical variate	R_c	%	p-value
Boxcount	7	1	0.934	99.401	< 0.001
		2	0.17	0.432	< 0.001
		3	0.106	0.166	< 0.001
	14	1	0.918	99.236	< 0.001
		2	0.158	0.469	< 0.001
		3	0.126	0.295	< 0.001
	21	1	0.91	99.037	< 0.001
		2	0.174	0.634	< 0.001
		3	0.126	0.33	< 0.001
Hall-Wood	7	1	0.932	99.303	< 0.001
		2	0.17	0.45	< 0.001
		3	0.127	0.248	< 0.001
	14	1	0.909	99.118	< 0.001
		2	0.172	0.634	< 0.001
		3	0.109	0.247	< 0.001
	21	1	0.888	99.059	< 0.001
		2	0.157	0.677	< 0.001
		3	0.099	0.263	< 0.001
Variogram	7	1	0.918	99.235	< 0.001
		2	0.181	0.628	< 0.001
		3	0.086	0.137	< 0.001
	14	1	0.874	99.04	< 0.001
		2	0.16	0.811	< 0.001
		3	0.07	0.149	< 0.001
	21	1	0.833	98.944	< 0.001
		2	0.146	0.949	< 0.001
		3	0.05	0.108	0.003
Madogram	7	1	0.913	99.302	< 0.001
		2	0.175	0.625	< 0.001
		3	0.06	0.072	< 0.001
	14	1	0.87	99.108	< 0.001
		2	0.158	0.815	< 0.001
		3	0.049	0.077	0.003
	21	1	0.832	99.001	< 0.001
		2	0.144	0.937	< 0.001
		3	0.037	0.062	0.089

$R_c =$ canonical correlation value.
% = percentage of the correlation explained by this canonical variate.

Table 4.3. Canonical correlation between population indicators and fractal dimension indicators in the Brussels-Capital region.

Estimator	Sliding window	Canonical variate	R_c	%	p-value
Boxcount	7	1	0.894	99.255	< 0.001
		2	0.17	0.737	0.029
		3	0.018	0.008	0.994
	14	1	0.865	98.441	< 0.001
		2	0.172	1.018	< 0.001
		3	0.126	0.538	0.028
	21	1	0.84	97.876	< 0.001
		2	0.196	1.628	< 0.001
		3	0.11	0.495	0.085
Hall-Wood	7	1	0.913	99.171	< 0.001
		2	0.176	0.628	0.002
		3	0.101	0.202	0.142
	14	1	0.858	98.549	< 0.001
		2	0.178	1.151	0.002
		3	0.092	0.3	0.221
	21	1	0.815	97.591	< 0.001
		2	0.189	1.833	< 0.001
		3	0.107	0.576	0.098
Variogram	7	1	0.851	98.593	< 0.001
		2	0.183	1.306	0.006
		3	0.052	0.1	0.773
	14	1	0.751	96.814	< 0.001
		2	0.193	2.909	< 0.001
		3	0.061	0.278	< 0.001
	21	1	0.66	93.096	< 0.001
		2	0.222	6.248	< 0.001
		3	0.073	0.656	0.45
Madogram	7	1	0.858	98.818	< 0.001
		2	0.166	1.007	0.014
		3	0.07	0.174	0.506
	14	1	0.794	97.85	< 0.001
		2	0.167	1.649	0.005
		3	0.093	0.501	0.208
	21	1	0.706	95.876	< 0.001
		2	0.174	2.995	0.001
		3	0.108	1.129	0.095

$R_c =$ canonical correlation value.
% = percentage of the correlation explained by this canonical variate.

Table 4.4. Canonical loading between each set of indicators and their first canonical variate.

Region	Variable	Boxcount	Hall-Wood	Variogram	Madogram
Flemish	Population size	-0.991	-0.992	-0.99	-0.99
	Population density	-0.688	-0.684	-0.691	-0.69
	Shannon index	-0.368	-0.365	-0.376	-0.376
	Older age population	0.281	0.282	0.28	0.28
	Median income	0.295	0.289	0.302	0.299
	Vaccination coverage	0.156	0.154	0.154	0.153
	Satisfaction	-0.07	-0.066	-0.074	-0.073
	Trust in the federal government	-0.333	-0.328	-0.342	-0.341
	Trust in the regional government	-0.262	-0.256	-0.276	-0.275
	Mean FD	-0.939	-0.946	-0.917	-0.924
	Variance FD	-0.636	-0.73	-0.146	-0.427
	ACF FD	0.709	0.104	0.828	0.68
Brussels-Capital	Population size	-0.99	-0.995	-0.987	-0.991
	Population density	-0.611	-0.588	-0.625	-0.618
	Shannon index	-0.258	-0.218	-0.266	-0.243
	Older age population	0.329	0.31	0.316	0.305
	Median income	0.329	0.321	0.33	0.329
	Vaccination coverage	0.273	0.256	0.154	0.273
	Mean FD	-0.811	-0.79	-0.54	-0.622
	Variance FD	0.139	0.042	0.97	0.723
	ACF FD	0.918	0.269	0.766	0.592

ACF = autocorrelation value.
FD = fractal dimension.

4.4 Discussion

The results illustrated that structural elements of the fractal dimension curve derived from a statistical sector's COVID-19 incidence curve can be used to characterize how the epidemic behaves in a given area. Castillo and Melin (2021) reported that the complex behavior of time series data could be measured and compared between different periods and countries to forecast the evolution and decide on possible NPIs based on the current situation. Other studies also reported similar applications of fractal dimension to characterize the complexity of time series data in different fields (Krstacic et al., 2007; Swapna et al., 2020; Sarkar and Leong, 2003).

Different estimators produced varying fractal dimension estimates, with

the box-count method consistently yielding the lowest values. This aligns with the findings of Gneiting et al. (2012), who reported a downward bias in the box-count estimator, a lower mean square error for the variogram estimator, and greater robustness of the madogram estimator to outliers. Based on the results of this chapter, the variogram and madogram estimators are preferred.

The time series length also influenced the fractal dimension estimation, as different sliding windows produced different estimates and curve shapes. The box-count estimator is particularly sensitive to the length of the data window and small amounts of data can have detrimental effects on its estimation, particularly when trends are presented (Falconer, 2013; Guo, 2017). This sensitivity can affect the ability to distinguish between clusters. Simulation and empirical results suggest that for shorter time series, madogram or variogram estimators combined with shorter sliding windows (e.g., 7 or 14 days) are recommended. For longer time series, box-count or Hall–Wood estimators with longer sliding windows may also be appropriate.

It is important to note that the proposed method serves a different purpose than scan-statistics tools (e.g., SaTScan) or model-based disease mapping approaches. Those methods typically assess whether events occur randomly, and when they do not, they identify clusters based on spatial and/or temporal anomalies. In contrast, the fractal dimension approach focuses on quantifying the complexity of multivariate outcomes over time and classifying them according to characteristics of this complexity. Crucially, it avoids imposing any spatial mechanism as a data-generating process. This allows for the identification of similarities and differences in complexity patterns independent of their geographic location.

Based on the CCA results, a strong association was identified between various sociodemographic factors and COVID-19 incidence. Consistent with this finding, COVID-19 cases and/or deaths have been positively correlated with population size (Groppo et al., 2022), population density (Hu et al., 2020), and urban characteristics (McLaughlin et al., 2021). Oh et al. (2022) further demonstrated that racial and ethnic composition, age, income, household size, and population density significantly influence COVID-19 incidence. In contrast, Cifuentes-Faura (2022) reported that countries a with higher population density exhibited lower death rates, suggesting that disparities in demographic and socioeconomic conditions

may explain these differences. Supporting this view, Mollalo et al. (2020) found that higher income inequality (measured as the ratio of household income at the 80^{th} percentile to that at the 20^{th} percentile) was a significant predictor of increased COVID-19 incidence, particularly in the tri-state area of the United States. Additionally, inequities in access and quality of healthcare have been shown to exacerbate the impact of the pandemic (The Lancet Public Health, 2019).

While not explicitly included in the CCA analysis, social contact patterns remain a fundamental component in the modeling of infectious diseases. This dimension may be partially captured through the inclusion of the Shannon index, which reflects population heterogeneity. Although Mossong et al. (2008) reported broadly similar age-specific contact patterns and intensities across different European countries, this assumption may not hold at finer spatial scales, especially those characterized by high population diversity. For instance, in the country of Luxembourg where nearly half of the population is of foreign nationality, residents with Portuguese nationality were found to have a higher number of social contacts compared to Luxembourgish residents during the lockdown period (Latsuzbaia et al., 2020). While in theory residents of homogeneous areas (Shannon index closer to zero) are expected to have more contacts in their local community, this would be less relevant in cases where these homogeneous areas are predominantly populated by people from foreign origins, which increases the possibility of contact with their home country. On the other hand, different social contact patterns in individuals of foreign origin may be driven by structural determinants rather than migration background itself. For example, Dorélien et al. (2021) found that Hispanic individuals have the highest number of household contacts, whereas non-Hispanic Black individuals have the lowest number and shortest duration of household contacts across nearly all age groups. At the same time, non-Hispanic Black populations tend to have occupations requiring high levels of physical proximity, thereby increasing the risk of COVID-19 transmission outside the household.

A relatively low explanatory power of the vaccination coverage was found in the first canonical variate even in the presence of high coverage. Vaccination showed a favorable effect on alleviating the burden of COVID-19, for example by reducing COVID-19 infection, severity, hospitalization, and mortality in the first period of the pandemic (Tadesse et al., 2022;

Mohammed et al., 2022). However, the protection wanes over time, and at some point the vaccination aids in reducing the severity or mortality but less on the transmission, especially with a new variant of concern (Braeye et al., 2022). Moreover, Wong et al. (2023) reported an increasing daily mean number of contacts following summer vacation in 2021 in Belgium. Hence, it is possible to observe many COVID-19 cases in areas with high vaccination coverage during the study period. On top of this, there were still some areas where people were very hesitant to get vaccinated. Areas with a more diverse population or lower socioeconomic status had a lower vaccination coverage despite the free COVID-19 vaccination (Faes et al., 2022), which further created an imbalance in the overall vaccination coverage observed in the two regions. The implementation of the COVID-19 vaccination certificate in October 2021 also contributed to the improvement of vaccination coverage, particularly in the younger population aged 18–34 years, with an increase in vaccination uptake between 0.5–1.2 percentage points and 1–1.6 percentage points for the first and second doses, respectively (Sciensano, 2022). However, its effectiveness was likely heterogeneous across population groups. In particular, the policy may have had limited impact among individuals or communities with strong pre-existing anti-vaccination beliefs, or in contexts where the perceived consequences of remaining unvaccinated were minimal (Natalia, Delporte, et al., 2023).

In the Flemish region, higher levels of satisfaction and trust in the government were associated with higher mean and variance values, reflecting greater complexity in COVID-19 incidence patterns. Greater satisfaction and trust are generally expected to improve compliance with government policies, particularly during crises. Previous studies have shown that public trust plays a key role in vaccination uptake (Lamot et al., 2022; Y. Wang et al., 2022). However, in this analysis, higher satisfaction and trust were observed in larger municipalities, which also exhibited more complex COVID-19 incidence curves. In contrast, lower satisfaction and trust may contribute to underreporting of COVID-19 cases. It should also be noted that the underlying survey was conducted online using a self-administered questionnaire among residents aged 17–85, which introduces the potential for selection bias.

This chapter highlights the potential of the fractal dimension as an alternative approach to analyze data at very fine-scale administrative units.

Conventional surveillance systems rely primarily on burden indicators such as incidence, prevalence, mortality, or reproduction number of infectious diseases. These measures quantify the magnitude of transmission, yet they provide limited insight into the structural organization and temporal irregularity of epidemic spread. In contrast, the fractal dimension characterizes the geometric roughness or scaling behavior of epidemic time series and therefore introduces a complementary indicator that captures epidemic complexity. By quantifying how irregularly case counts evolve over time, fractal dimension provides information about the underlying transmission structure, including heterogeneity, clustering, and intermittency, which can be beneficial in terms of early outbreak detection.

Moreover, conventional epidemiological indicators are typically threshold-based and depend on observable increases in counts relative to historical baselines. As a result, they may lag behind structural changes in transmission dynamics. In theory, changes in fractal dimension could precede visible rises in incidence, as increasing complexity in the time series may reflect early fragmentation of transmission chains, emerging local clusters, or heightened variability before sustained growth becomes apparent. Although formal validation in population-based disease surveillance remains limited, fractal analysis has demonstrated sensitivity to early dynamical alterations in other biomedical contexts. For example, increased physiological complexity has been shown to precede clinical deterioration in patients with heart failure (Goldberger et al., 2002), while alterations in fractal properties have been observed prior to neurodegenerative disorders such as Alzheimer's disease (Lau et al., 2023) or in seizure dynamics in epilepsy (Tan et al., 2025).

Additionally, fractal properties are theoretically linked to long-range dependence and persistence. Lower values of the fractal dimension indicate smoother and more trend-reinforcing dynamics, suggesting stability of the short-term forecasts. Conversely, higher fractal dimension values reflect greater irregularity and anti-persistence, signaling increased uncertainty and susceptibility to abrupt local outbreaks. This distinction may be particularly valuable at fine spatial scales, where surveillance data are often affected by small population sizes or reporting noise. In such contexts, fractal dimension offers a measure of structural complexity that is less dependent on absolute case counts.

Fractal dimension provides a relative complexity ranking between ar-

eas, which can enhance spatial prioritization strategies. By summarizing epidemic structure through local fractal dimension curves, areas transitioning toward an outbreak behavior can be detected, so that public health interventions can be prioritized. On top of this, fractal dimension method can be combined with an existing method such as CCA, facilitating the joint assessment of epidemic complexity alongside sociodemographic or environmental factors. Differences in results across geographical units should be expected, reflecting variations in both population characteristics and fractal dimension indicators. In this sense, the fractal dimension approach enriches surveillance by adding a complexity-based perspective to conventional burden-based monitoring. Together, these perspectives may provide a more comprehensive framework for early detection, uncertainty assessment, and spatial prioritization in epidemic monitoring.

Despite these strengths, several limitations should be acknowledged. In the CCA, vaccination coverage was represented by a single value at the end of the study period (31 December 2021). Capturing temporal changes in vaccination coverage could provide additional explanatory value. In Belgium, vaccination uptake followed a dynamic and non-linear trajectory, particularly during the implementation of certain NPIs such as the COVID-19 vaccination certificate. This temporal heterogeneity is not adequately captured by a static end-point measure and may partly explain the relatively modest contribution of vaccination coverage observed in the CCA. Moreover, the dataset may contain discrepancies in recorded vaccination data. For instance, some individuals could receive vaccination outside their registered residence or some statistical sectors reporting more vaccinated individuals than their total registered population, particularly when the latter was zero. To address this, statistical sectors with zero population were excluded from the CCA.

While CCA is widely used in infectious disease and public health research, the application of fractal dimensions in this field remains rare. Even though the general idea of fractal dimensions might be intuitive, understanding its mathematical intricacies can become more challenging. Encouraging regular use of fractal dimensions, particularly as part of a surveillance system, will require clear communication of its added value and practical benefits.

CHAPTER 5

GENERAL DISCUSSION AND CONCLUDING REMARKS

All models are wrong, but some are useful
— George E. P. Box

Over the years, a wide range of statistical methods have been developed to address the challenges of analyzing complex spatiotemporal data. Each method is rooted in its own theoretical framework and designed to serve specific analytical purposes, whether to detect clusters, model spatial dependence, identify temporal trends, or quantify structural complexity. However, no single approach can fully capture all aspects of such multifaceted phenomena. Infectious disease transmission, in particular, is influenced by a combination of demographic, spatial, behavioral, and other processes that operate simultaneously across different scales. As a result, relying on a single methodological framework may provide only a partial representation of the underlying dynamics. Integrating complementary methods can therefore offer a more comprehensive perspective, enabling researchers to explore different facets of the data while strengthening the robustness of the conclusions.

In this thesis, the analytical framework begins with the application of Bayesian spatial models in Chapter 2. These models are particularly well suited for disease mapping and spatial epidemiology because they explic-

itly account for spatial dependence between neighboring areas and allow the estimation of smoothed outcomes across geographical units. By incorporating spatial random effects within a hierarchical Bayesian structure, these models help stabilize estimates in areas with small populations and limited case counts, while simultaneously quantifying uncertainty in the estimates. Such approaches are widely used in epidemiological research to identify spatial patterns of disease occurrence and to explore associations between sociodemographic characteristics and health outcomes at the population level (Lawson, 2018; Blangiardo and Cameletti, 2015).

Building upon this spatial framework, Chapter 3 employs linear mixed models to investigate the relationship between COVID-19 incidence and excess mortality with sociodemographic determinants over time. Linear mixed models are particularly suitable for longitudinal and hierarchical data structures, where repeated observations are recorded for the same spatial units across multiple time points. By incorporating both fixed effects and random effects, these models account for correlations within municipalities and allow the estimation of population-level associations while controlling for unobserved heterogeneity between areas (Verbeke and Molenberghs, 2009). Compared to fully Bayesian hierarchical frameworks, linear mixed models offer practical advantages in terms of computational efficiency and interpretability when analyzing datasets with numerous covariates and repeated measurements.

Finally, Chapter 4 introduces fractal dimension analysis as a complementary methodological perspective to characterize the structural complexity of the epidemic time series. Whereas regression-based approaches primarily aim to estimate associations between explanatory variables and disease outcomes, fractal analysis focuses on the intrinsic geometric properties of time series, such as irregularity, scaling behavior, and persistence. In this sense, fractal dimension analysis addresses a different perspective. Instead of explaining why incidence varies across space and time, it aims to describe how epidemic trajectories evolve in terms of their temporal complexity.

Although the methods applied in Chapters 3 and 4 could theoretically be embedded within a Bayesian hierarchical framework as implemented in Chapter 2, the use of distinct analytical approaches was a deliberate methodological choice. Each method highlights different characteristics of the data and therefore combining these approaches provides complemen-

tary insights into the same epidemiological phenomenon and ultimately enhance the robustness of our conclusions.

Age and sex remain fundamental explanatory variables for many diseases, including COVID-19. The findings presented in Chapters 2 and 3 reinforce this notion, as clear distinctions were observed between younger and older populations in both COVID-19 incidence and excess mortality patterns. The influence of these demographic effects is consistent across the pandemic waves, as further demonstrated by the longitudinal models in Chapter 3. This aligns with international evidence showing that age and sex are among the strongest predictors of COVID-19 outcomes, and areas with higher population of older adults and males experiencing higher risks of severe disease and mortality (Sundaram et al., 2022).

Before the widespread availability of COVID-19 vaccines, additional factors such as income, population density, and the size of the student population played an important role in shaping COVID-19 dynamics. While other important factors such as mobility and social contact patterns are not modeled explicitly, student population could be considered as a proxy for local mobility through their commuting patterns. Moreover, higher education institutions have been identified as transmission hubs, where dense social interactions can amplify the spread of COVID-19 (Wilson et al., 2020; Vang et al., 2021; Leidner et al., 2021). However, this association should also be viewed in the background of changes in testing capacity, eligibility criteria, and testing intensity over the course of the pandemic. In Belgium, as in many countries, testing strategies evolved substantially between 2020 and 2022 in response to laboratory capacity and public health priorities. Consequently, fluctuations in reported incidence partly reflect changes in surveillance practices in addition to underlying transmission dynamics.

During the early phase of the pandemic, testing capacity was limited and largely restricted to symptomatic cases. At that time, PCR testing was initially performed only by the national reference laboratory before additional clinical laboratories were gradually authorized to conduct SARS-CoV-2 diagnostics (Janssen et al., 2024). As testing capacity improved, eligibility criteria were gradually expanded to include more individuals in the community, high-risk contacts, and travelers returning from high-incidence regions. However, during periods of intense transmission, particularly during the second wave in autumn 2020, testing demand temporarily exceeded

capacity. In response, the Belgian authorities adjusted the testing strategy by prioritizing symptomatic individuals and healthcare personnel while temporarily limiting testing among asymptomatic contacts (European Observatory on Health Systems and Policies, 2024). Such changes may have influenced the number of cases detected, potentially leading to underestimation of true incidence during periods when testing eligibility was restricted. Therefore, the relationship between reported incidence and actual infection levels likely varied across time and the temporal changes in reported incidence should be interpreted cautiously, particularly when comparing different phases of the pandemic. Observed increases may partly reflect expanded testing capacity or broader eligibility criteria, while apparent declines may coincide with periods when testing was restricted or when testing demand decreased.

Following the rollout of vaccination campaigns, the role of mobility, particularly international and inter-regional travel, became more pronounced as restrictions eased. Modeling mobility explicitly during this period allowed for a more accurate representation of disease dynamics, especially given the resurgence of cross-border travel and its contribution to renewed transmission risks (Ensoy-Musoro et al., 2023; Nguyen et al., 2023). The results from Chapter 3 further supported the importance of explicitly modeling mobility alongside vaccination and overall preventive measures.

Hierarchical models, as applied in Chapters 2 and 3, represent a well-established and powerful statistical framework to investigate the influence of multiple variables with different structures, making them highly suitable for epidemiological applications, such as modeling the incidence and mortality of COVID-19. In particular, hierarchical frameworks are able to disentangle effects operating at different levels (e.g., individual, community, or regional) and account for spatial or temporal dependence that would otherwise bias estimates if ignored (Gelman and Hill, 2006; Lawson, 2018). This ability to "borrow strength" across groups or areas is especially relevant when dealing with small-area estimation or sparse data, which are common challenges in public health surveillance.

However, this flexibility comes at a price. Building complex hierarchical models requires careful evaluation of various assumptions, including distributional properties, model fit, or the specification of appropriate priors in the Bayesian context. Computational demands also increase substantially with model complexity. Approximate methods like INLA have

greatly improved computational feasibility in various settings, but fitting large spatiotemporal models might still require substantial time and resources. For instance, in Chapter 2, the average time to fit a single model using INLA was approximately 4.31 seconds, which would have been considerably longer with traditional MCMC methods. Although the longitudinal models in Chapter 3 were generally very fast to estimate (around 0.1 seconds on average), some models required up to 47 seconds, particularly when the complexity of the model increased. Moreover, overly complex models risk issues such as identifiability, convergence problems, or singularities, highlighting the need for careful model specification and validation. Similar challenges have been reported in hierarchical Bayesian modeling of infectious disease dynamics, where balancing computational feasibility with model realism remains an ongoing methodological concern (Lawson, 2018; Vranckx et al., 2023).

The results in Chapter 4 underscore the potential of fractal dimension as a complementary approach to analyze spatiotemporal data. Unlike conventional methods that often require explicit modeling of incidence dynamics or assumptions about spatial dependence, the fractal dimension approach rests on the relatively simple assumption that time-series data exhibit fractal properties. By summarizing the complexity of daily COVID-19 incidence through the mean, variance, and autocorrelation of the local fractal dimension curve, this approach offers a comprehensive characterization of epidemic dynamics without the need to directly examine each individual incidence curve at the statistical sector level. This methodological efficiency will be beneficial when time is crucial, for example during epidemic surges, and rapid assessment of multiple areas is required. Moreover, the proposed cluster classification framework offers a practical tool for detecting areas at increased risk of sustained transmission. Such classifications could inform local authorities on the tailoring of preventive measures, facilitating more agile, targeted, and resource-efficient interventions.

The potential of this approach lies not only in its ability to independently characterize epidemic dynamics but also in its compatibility with existing data and methods. As demonstrated in Chapter 4, routinely collected sociodemographic and epidemiological indicators were integrated into the analysis to better interpret epidemic patterns and identify the underlying characteristics that exhibit strong associations with epidemic

propagation. Additional variables such as temperature or heatwave as modeled in Chapter 3, as well as other approaches such as mathematical modeling, could be integrated with the fractal dimension framework to provide a more comprehensive understanding of epidemic dynamics. For instance, compartmental models can help simulate transmission dynamics under varying assumptions of susceptibility, exposure, and immunity to detect a higher risk of epidemic transmission (Nian and F. Li, 2024). Such adaptability improves the practical utility of the fractal dimension approach, making it suitable for dynamic public health surveillance systems. Ultimately, this framework provides a scalable and computationally efficient alternative to more traditional spatiotemporal models, while still allowing meaningful interpretation and application in decision-making.

Approaches using modern ML methods for spatiotemporal pattern recognition, such as random forests, gradient boosting, or deep learning architectures, also demonstrated strong predictive performance in many fields. However, they are typically designed to optimize prediction accuracy and may require large volumes of high-quality data, extensive tuning, and substantial computational resources (Bzdok et al., 2018; Beam and Kohane, 2018). When applied to small-area data with limited observations, which could be frequently encountered in disease surveillance at granular administrative levels, ML approaches may suffer from overfitting or reduced generalizability. Moreover, many of the ML models are "black boxes", in the sense that the way they represent and combine information becomes extremely hard to interpret in human terms, especially compared to widely used statistical models (Rudin, 2019). Fractal dimension analysis, on the other hand, can be applied directly to individual time series without the need for extensive training data, and has been shown to be relatively robust to noise and irregular sampling when appropriate estimators (e.g., madogram) are used (Gneiting et al., 2012). Furthermore, fractal analysis captures aspects of long-range dependence and temporal persistence, which are not always explicitly modeled in standard ML approaches unless specifically designed. By summarizing these properties into a small set of interpretable indicators, the approach enables comparative analysis across regions and facilitates the identification of areas exhibiting atypical or unstable epidemic dynamics. Although it is outside the focus of this thesis, fractal dimension has a potential to be used as a feature extraction or representation tool, which could also be integrated into ML pipelines to

enhance predictive performance.

An explicit assessment of model adequacy and goodness-of-fit is essential given the central role of the modeling framework in this thesis. Bayesian hierarchical models are commonly evaluated using the Deviance Information Criterion (DIC) and the Watanabe-Akaike Information Criterion (WAIC) which balance model fit and complexity (Spiegelhalter et al., 2002; Watanabe, 2010). Lower values of these criteria indicate improved model performance and were used to guide model selection and prior specification, including the comparison between default and PC priors. Since only a single model specification was fitted for each time period in Section 2.2.3.1, the model selection criteria were not explicitly applied. For models in Section 2.2.3.2, the differences in DIC and WAIC between the default and PC priors were negligible (approximately one point), suggesting a comparable model fit. For the linear mixed models applied in Chapter 3, the goodness-of-fit was assessed through standard diagnostics, including residual inspection and AIC values. For the fractal dimension approach introduced in Chapter 4, the model adequacy is inherently different, as the method does not rely on a probabilistic likelihood framework. Instead, the evaluation focused on the stability and consistency of the estimated fractal dimension curves, the sensitivity to sliding window choices, and the agreement between different estimators, which serve as an indirect validation of the robustness of the approach (Gneiting et al., 2012).

An important limitation of applying fractal dimension methods in epidemiological research is that the fractal characteristics of time-series data may not hold when aggregated at larger temporal or spatial scales. For example, aggregating incidence data from daily to weekly or monthly intervals could change a specific seasonal pattern (Alarcon Falconi et al., 2020). In terms of fractal processes, this can lead to a loss of fine-scale variability. Similarly, when data are aggregated to higher administrative levels such as municipalities or provinces, the heterogeneity observed at the neighborhood or statistical sector level might be smoothed out, masking critical local fluctuations in epidemic dynamics (Jeffery et al., 2014). This limitation highlights the importance of carefully selecting the appropriate scale of analysis, especially when the goal is to identify early signals of transmission clusters or localized outbreaks.

Another challenge lies in capturing the real effect of variables with

strong temporal dynamics, such as vaccination coverage. Vaccination uptake changes over time and across population subgroups (H. Wang et al., 2024), therefore, using a single cumulative value at the end of the study period risks underestimating or misrepresenting its impact on transmission patterns. Temporal differences between vaccination rollout and epidemic waves further complicate interpretation, as protection builds progressively and interacts with changing mobility, policy measures, and circulating variants.

The analyses presented in this thesis were conducted at several administrative levels (national, provincial, municipal, and statistical sector), reflecting an ecological study design in which the unit of observation is a group or population, rather than an individual. Aggregated data are particularly relevant for surveillance-based research, where routinely collected information such as case notifications, demographic characteristics, or mobility indicators is often available only at certain administrative levels. On the other hand, associations observed using aggregated data may arise from contextual influences, compositional effects, or statistical artifacts introduced by aggregation, a phenomenon widely referred to as the ecological fallacy (Roumeliotis et al., 2021). Consequently, results derived from administrative unit analyses should be interpreted as area-level correlations rather than individual-level causal relationships. For example, in Chapter 2, the proportion of higher education students was found to be associated with COVID-19 incidence. This result should be interpreted as an indication that areas with a higher proportion of higher education students tended to exhibit higher COVID-19 incidence rates during the study period, rather than implying that being a higher education student directly increases (or decreases) an individual's risk of COVID-19 infection.

Another limitation concerns within-area heterogeneity. Administrative units often encompass populations with diverse socioeconomic conditions, mobility patterns, and contact structures. Aggregation may mask substantial variation within municipalities, potentially obscuring localized transmission clusters or neighborhood-level inequalities. Statistical results may vary depending on the spatial scale or boundaries used for aggregation (Lawson, 2018). While higher administrative units provide a practical and policy-relevant scale for public health decision making, it should be acknowledged that a finer spatial resolution could reveal additional heterogeneity.

Another consideration relates to the external validity and generalizability of the findings beyond Belgium. The epidemiological patterns observed in one context may not fully translate to other settings due to differences in demographic structures, health system organization, and public health policies. Variations in testing strategies, reporting completeness, and health system capacity may affect the reliability of incidence data and the detectability of spatial patterns. Differences in population age structure, household composition, urbanization, and mobility networks can also shape transmission dynamics in ways that may alter the applicability of results obtained in another context. Furthermore, policy responses such as lockdown measures, vaccination strategies, or regional containment policies, varied widely across countries, which eventually altered epidemic trajectories.

Consequently, while the methodological framework presented in this thesis may be broadly applicable, the specific empirical findings should be interpreted within the Belgian epidemiological and institutional context. Future research could assess the transferability of the approach by applying similar analyses to surveillance data from other countries or regions with different demographic and health system characteristics. Comparative studies in various settings would help determine whether observed patterns represent general features of epidemic dynamics or context-specific phenomena. Moreover, the application of the fractal dimension approach in this thesis demonstrates its potential as a complementary framework for analyzing complex spatiotemporal data. Unlike many conventional methods, fractal-based measures capture the inherent variability and irregularity of epidemic processes in a compact yet informative way. Further refinement of this approach should be explored, for instance, by applying it to longer time series to assess stability over extended periods, or by integrating it with regression-based frameworks and other statistical models to strengthen its explanatory and predictive capacity.

Although the mathematical foundations of fractal theory may appear abstract, it is important to recognize that fractal-like structures are deeply embedded in natural and human systems, including neural encoding processes in the human brain. Fractal organization is central to neuronal dynamics and cognitive processing, suggesting that the human brain itself relies on fractal-like encoding schemes to handle complex informa-

tion (Werner, 2010). This implies that interpreting epidemic dynamics through fractal measures is not only mathematically sound but also conceptually aligned with how humans intuitively process complexity. As the well-known notion implies, there is "order within chaos and chaos within order", a principle that resonates strongly with epidemic behavior.

To strengthen the translation of methodological findings into public health practice, the fractal dimension approach can be framed as a complementary analytical tool that supports the surveillance, preparedness, and decision making process. From a public health perspective, fractal analysis could contribute in several practical ways:

1. A structural monitoring tool that complements burden indicators.
   Monitoring magnitude (via incidence) and structural complexity (via fractal dimension indicators) simultaneously may allow public health authorities to detect shifts in transmission dynamics that are not yet visible in case counts alone. This could improve the sensitivity of surveillance systems and support the development of early warning signals for emerging outbreaks.

2. An indicator of epidemic stability versus instability.
   Changes in the fractal dimension may reflect transitions between relatively stable transmission patterns and more irregular, unstable dynamics. Periods characterized by increasing complexity may indicate increased uncertainty or fragmentation of transmission chains, suggesting an increased risk of localized outbreaks. This information could help epidemiological teams interpret short-term forecasts and assess the robustness of observed epidemic trends.

3. A spatial prioritization metric at granular administrative levels.
   At fine spatial scales such as municipalities or statistical sectors, case counts are often small and subject to considerable variability. The fractal dimension provides a relative measure of epidemic complexity that can be compared between areas, enabling the identification of locations where transmission patterns are becoming increasingly irregular. These signals could help support targeted interventions in higher-risk areas, for example by guiding intensified testing, localized communication campaigns, or enhanced contact tracing capacity.

4. A tool for assessing surveillance sensitivity and data irregularity.
   Fractal analysis may also serve as a diagnostic tool to evaluate the

quality and stability of surveillance data. The high complexity of time series data may reflect genuine epidemiological variability, but it may also indicate reporting delays, inconsistent testing practices, or other surveillance artifacts. Identifying such irregularities can help public health authorities interpret epidemiological indicators more cautiously and improve data quality monitoring.

A promising course for future research is the integration of fractal dimension indicators into real-time surveillance systems. By providing a rapid summary of epidemic complexity in fine-scale units, fractal-based monitoring could enable early detection of anomalies and localized outbreaks, thus providing valuable support for public health decision-making and targeted interventions. In this way, fractal analysis can complement existing surveillance tools and contribute to more adaptive and resilient epidemic monitoring frameworks.

BIBLIOGRAPHY

Agentschap Binnenlands Bestuur (2023). *Gemeente-Stadsmonitor*. https://gemeente-stadsmonitor.vlaanderen.be/download-alle-cijfers.

Alarcon Falconi, T. M., Estrella, B., Sempértegui, F., and Naumova, E. N. (2020). "Effects of Data Aggregation on Time Series Analysis of Seasonal Infections". *International Journal of Environmental Research and Public Health* 17.16, p. 5887.

Alsan, M., Stantcheva, S., Yang, D., and Cutler, D. (2020). "Disparities in Coronavirus 2019 Reported Incidence, Knowledge, and Behavior Among US Adults". *JAMA Netw Open* 3.6, e2012403. DOI: 10.1001/jamanetworkopen.2020.12403.

Arfan, M. et al. (2021). "Investigation of fractal-fractional order model of COVID-19 in Pakistan under Atangana-Baleanu Caputo (ABC) derivative". *Results Phys* 24, p. 104046. DOI: 10.1016/j.rinp.2021.104046.

Aron, J., Muellbauer, J., Giattino, C., and Ritchie, H. (2020). *A pandemic primer on excess mortality statistics and their comparability across countries*. https://ourworldindata.org/covid-excess-mortality, access date 2023-03-22.

Astengo, M. et al. (2021). "Weight of risk factors for mortality and short-term mortality displacement during the COVID-19 pandemic". *J Prev Med Hyg* 62.4, E864–e870. DOI: 10.15167/2421-4248/jpmh2021.62.4.2269.

Avnir, D., Biham, O., Lidar, D., and Malcai, O. (1998). "Is the Geometry of Nature Fractal?" *Science* 279.5347, pp. 39–40.

Bae, S. et al. (2020). "Epidemiological characteristics of COVID-19 outbreak at fitness centers in Cheonan, Korea". *Journal of Korean medical science* 35.31, e288–e288. DOI: 10.3346/jkms.2020.35.e288.

Bahk, J. and Jung-Choi, K. (2022). "Cause-specific mortality in Korea during the first year of the COVID-19 pandemic". *Epidemiol Health* 44, e2022110. DOI: 10.4178/epih.e2022110.

Basso, C. et al. (2023). "Veneto Region dementia-related mortality during the COVID-19 pandemic: Multiple causes of death and time series analysis". *Eur J Public Health*. DOI: 10.1093/eurpub/ckad005.

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). "Fitting Linear Mixed-Effects Models Using lme4". *Journal of Statistical Software* 67.1, pp. 1–48. DOI: 10.18637/jss.v067.i01.

Beam, A. L. and Kohane, I. S. (2018). "Big Data and Machine Learning in Health Care". *JAMA* 319.13, pp. 1317–1318. DOI: 10.1001/jama.2017.18391.

Besag, J., York, J., and Mollié, A. (1991). "Bayesian image restoration, with two applications in spatial statistics". *Annals of the Institute of Statistical Mathematics* 43.1, pp. 1–20. DOI: 10.1007/BF00116466.

Bhadra, A., Mukherjee, A., and Sarkar, K. (2020). "Impact of population density on Covid-19 infected and mortality rate in India". *Model Earth Syst Environ*, pp. 1–7. DOI: 10.1007/s40808-020-00984-7.

Bianchi, S. and Frezza, M. (2017). "Fractal stock markets: International evidence of dynamical (in)efficiency". *Chaos* 27.7, p. 071102. DOI: 10.1063/1.4987150.

Bisanzio, D. et al. (2022). "Estimating the effect of non-pharmaceutical interventions to mitigate COVID-19 spread in Saudi Arabia". *BMC Med* 20.1, p. 51. DOI: 10.1186/s12916-022-02232-4.

Blangiardo, M. and Cameletti, M. (2015). "Spatial and Spatio□temporal Bayesian Models with R□INLA". *Wiley*. DOI: 10.1002/9781118950203.ch6.

Bobrovitz, N. et al. (2023). "Protective effectiveness of previous SARS-CoV-2 infection and hybrid immunity against the omicron variant and severe disease: A systematic review and meta-regression". *The Lancet Infectious Diseases* 23.5, pp. 556–567. DOI: 10.1016/S1473-3099(22)00801-5.

Boehmer, T. K. et al. (2020). "Changing age distribution of the COVID-19 pandemic - United States, May-August 2020". *MMWR. Morbidity and*

mortality weekly report 69.39, pp. 1404–1409. DOI: 10.15585/mmwr.mm6939e1.

Bonneux, L. (2002). "How to measure the burden of mortality?" *J Epidemiol Community Health* 56.2, pp. 128–31. DOI: 10.1136/jech.56.2.128.

Box, G., Jenkins, G., and Reinsel, G. (2015). *Time series analysis: Forecasting and control*. 5th ed. Wiley Series in Probability and Statistics. Wiley, p. 720.

Brackx, F., Vanongeval, F., Natalia, Y. A., Molenberghs, G., and Steenberghen, T. (2022). "The Effect of Transborder Mobility on COVID-19 Incidences in Belgium". *International Journal of Environmental Research and Public Health* 19.16, p. 9968.

Braeye, T. et al. (2022). "Vaccine effectiveness against onward transmission of SARS-CoV2-infection by variant of concern and time since vaccination, Belgian contact tracing, 2021". *Vaccine* 40.22, pp. 3027–3037. DOI: 10.1016/j.vaccine.2022.04.025.

British Broadcasting Corporation (2020). *Coronavirus: The world in lockdown in maps and charts*. https://www.bbc.com/news/world-52103747.

Burrough, P. A. (1981). "Fractal dimensions of landscapes and other environmental data". *Nature* 294.5838, pp. 240–242. DOI: 10.1038/294240a0.

Bzdok, D., Altman, N., and Krzywinski, M. (2018). "Statistics versus machine learning". *Nature Methods* 15.4, pp. 233–234. DOI: 10.1038/nmeth.4642.

Carabelli, A. M. et al. (2023). "SARS-CoV-2 variant biology: Immune escape, transmission and fitness". *Nat Rev Microbiol* 21.3, pp. 162–177. DOI: 10.1038/s41579-022-00841-7.

Castillo, O. and Melin, P. (2021). "A novel method for a COVID-19 classification of countries based on an intelligent fuzzy fractal approach". *Healthcare (Basel)* 9.2. DOI: 10.3390/healthcare9020196.

Checchi, F. and Roberts, L. (2005). *Interpreting and using mortality data in humanitarian emergencies: A primer for non-epidemiologists*. https://odihpn.org/wp-content/uploads/2005/09/networkpaper052.pdf, access date 2023-11-10.

Chen, X. and Fu, F. (2022). "Highly coordinated nationwide massive travel restrictions are central to effective mitigation and control of COVID-19

outbreaks in China". *Proc Math Phys Eng Sci* 478.2260, p. 20220040. DOI: 10.1098/rspa.2022.0040.

Chen, X., Huang, H., Ju, J., Sun, R., and Zhang, J. (2022). "Impact of vaccination on the COVID-19 pandemic in U.S. states". *Scientific Reports* 12.1, p. 1554. DOI: 10.1038/s41598-022-05498-z.

Cifuentes-Faura, J. (2022). "Effect of Population Density and Economic Indicators on COVID-19 Death Rates in the Community of Latin American and Caribbean States". *P R Health Sci J* 41.4, pp. 192–196.

Constantine, A. G. and Hall, P. (1994). "Characterizing Surface Smoothness via Estimation of Effective Fractal Dimension". *Journal of the Royal Statistical Society. Series B (Methodological)* 56.1, pp. 97–113.

COVID-19 Forecasting Team (2023). "Past SARS-CoV-2 infection protection against re-infection: A systematic review and meta-analysis". *Lancet* 401.10379, pp. 833–842. DOI: 10.1016/s0140-6736(22)02465-5.

Cronin, C. J. and Evans, W. N. (2022). "Nursing home quality, COVID-19 deaths, and excess mortality". *Journal of Health Economics* 82, p. 102592. DOI: https://doi.org/10.1016/j.jhealeco.2022.102592.

D'Angelo, S., Bevilacqua, G., Bloom, I., Ntani, G., and Walker-Bone, K. (2022). "Predictors and consequences of not seeking healthcare during the COVID-19 pandemic: Findings from the HEAF cohort". *Int J Environ Res Public Health* 19.20. DOI: 10.3390/ijerph192013271.

Dagan, N. et al. (2021). "BNT162b2 mRNA Covid-19 Vaccine in a Nationwide Mass Vaccination Setting". *New England Journal of Medicine* 384.15, pp. 1412–1423. DOI: 10.1056/NEJMoa2101765.

De Gemeenschappelijke Gemeenschapscommissie (2023). *De Gemeenschappelijke Gemeenschapscommissie*. https://www.ccc-ggc.brussels/nl.

Departement Zorg (2023). *Department of Care*. https://www.departementzorg.be/en.

Di Angelantonio, E. et al. (2015). "Association of cardiometabolic multimorbidity with mortality". *Jama* 314.1, pp. 52–60. DOI: 10.1001/jama.2015.7008.

Dorélien, A. M., Ramen, A., Swanson, I., and Hill, R. (2021). "Analyzing the demographic, spatial, and temporal factors influencing social contact patterns in U.S. and implications for infectious disease spread". *BMC infectious diseases* 21.1, pp. 1009–1009. DOI: 10.1186/s12879-021-06610-w.

Dubuc, B., Quiniou, J. F., Roques-Carmes, C., Tricot, C., and Zucker, S. W. (1989). "Evaluating the fractal dimension of profiles". *Physical Review A* 39.3. PRA, pp. 1500–1512. DOI: 10.1103/PhysRevA.39.1500.

Dureau, J., Kalogeropoulos, K., and Baguelin, M. (2013). "Capturing the time-varying drivers of an epidemic using stochastic dynamical systems". *Biostatistics* 14.3, pp. 541–55. DOI: 10.1093/biostatistics/kxs052.

Elliott, P. and Wartenberg, D. (2004). "Spatial epidemiology: Current approaches and future challenges". *Environmental Health Perspectives* 112.9, pp. 998–1006. DOI: doi:10.1289/ehp.6735.

Ensoy-Musoro, C., Nguyen, M. H., Hens, N., Molenberghs, G., and Faes, C. (2023). "Spatio-temporal model to investigate COVID-19 spread accounting for the mobility amongst municipalities". *Spatial and Spatio-Temporal Epidemiology* 45, pp. 100568–100568.

European Centre for Disease Prevention and Control (2023a). *COVID-19 Vaccine Tracker*. https://vaccinetracker.ecdc.europa.eu/public/extensions/Covid-19/vaccine-tracker.html#uptake-tab.

European Centre for Disease Prevention and Control (2023b). *Overview of the implementation of COVID-19 vaccination strategies and deployment plans in the EU/EEA*. https://www.ecdc.europa.eu/en/publications-data/overview-implementation-covid-19-vaccination-strategies-and-deployment-plans.

European Commission (2020). *Council Recommendation (EU) 2020/1475 of 13 October 2020 on a coordinated approach to the restriction of free movement in response to the COVID-19 pandemic (Text with EEA relevance)*. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32020H1475.

European Observatory on Health Systems and Policies (2024). *COVID-19 Health System Response Monitor (HSRM)*. https://eurohealthobservatory.who.int/monitors/hsrm/hsrm-countries/hsrm/hsrm-belgium/hsrm-bel-preventing-transmission/hsrm-bel-testing.

Faes, C. et al. (2022). "Geographical variation of COVID-19 vaccination coverage, ethnic diversity and population composition in Flanders". *Vaccine X* 11, p. 100194.

Falconer, K. (2013). "Fractal geometry: Mathematical foundations and applications". *Wiley*.

Fang, F., Clemens, J. D., Zhang, Z. F., and Brewer, T. F. (2024). "Impact of SARS-CoV-2 vaccines on COVID-19 incidence and mortality in the United States". *PLoS One* 19.4, e0301830. DOI: 10.1371/journal.pone.0301830.

FPS Health, Food Chain Safety and Environment (2020). *One repatriated Belgian has tested positive for the novel coronavirus*. https://web.archive.org/web/20200406101412/https://www.info-coronavirus.be/en/news/one-repatriated-belgian-has-tested-positive-for-the-novel-coronavirus/.

Friendly, M. and Fox, J. (2021). *candisc: Visualizing Generalized Canonical Discriminant and Canonical Correlation Analysis*. R package version 0.8-6.

Gandhi, M., Yokoe, D. S., and Havlir, D. V. (2020). "Asymptomatic transmission, the achilles' heel of current strategies to control COVID-19". *The New England journal of medicine* 382.22, pp. 2158–2160. DOI: 10.1056/NEJMe2009758.

Garland, P., Babbitt, D., Bondarenko, M., Sorichetta, A., Tatem, A. J., and Johnson, O. (2020). *The COVID-19 pandemic as experienced by the individual*.

Gausman, J. and Langer, A. (2020). "Sex and Gender Disparities in the COVID-19 Pandemic". *J Womens Health (Larchmt)* 29.4, pp. 465–466. DOI: 10.1089/jwh.2020.8472.

Ge, Y. et al. (2022). "Untangling the changing impact of non-pharmaceutical interventions and vaccination on European COVID-19 trajectories". *Nat Commun* 13.1, p. 3106. DOI: 10.1038/s41467-022-30897-1.

Gelman, A. and Hill, J. (2006). "Data Analysis Using Regression and Multilevel/Hierarchical Models". Analytical Methods for Social Research. DOI: 10.1017/CBO9780511790942.

Gibbons, C. L. et al. (2014). "Measuring underreporting and under-ascertainment in infectious disease datasets: A comparison of methods". *BMC Public Health* 14.1, p. 147. DOI: 10.1186/1471-2458-14-147.

Gneiting, T., Ševčíková, H., and Percival, D. B. (2012). "Estimators of fractal dimension: Assessing the roughness of time series and spatial data". *Statistical Science* 27.2, pp. 247–277, 31. DOI: 10.1214/11-STS370.

Goldberger, A. L., Amaral, L. A. N., Hausdorff, J. M., Ivanov, P. C., Peng, C. K., and Stanley, H. E. (2002). "Fractal dynamics in physiology: Al-

terations with disease and aging". *Proceedings of the National Academy of Sciences* 99, pp. 2466–2472. DOI: 10.1073/pnas.012579499.

Good, I. J. (1980). "Some history of the hierarchical Bayesian methodology". *Trabajos de Estadistica Y de Investigacion Operativa* 31.1, pp. 489–519. DOI: 10.1007/BF02888365.

Gostic, K. M. et al. (2020). "Practical considerations for measuring the effective reproductive number, Rt". *PLOS Computational Biology* 16.12, e1008409. DOI: 10.1371/journal.pcbi.1008409.

Greene, S. K. et al. (2022). "Reduced COVID-19 hospitalizations among New York City residents following age-based SARS-CoV-2 vaccine eligibility: Evidence from a regression discontinuity design". *Vaccine X* 10, p. 100134. DOI: 10.1016/j.jvacx.2021.100134.

Groppo, M. F., Groppo, F. C., Figueroba, S. R., and Pereira, A. C. (2022). "Influence of population size, the Human Development Index and the Gross Domestic Product on mortality by COVID-19 in the Southeast Region of Brazil". *Int J Environ Res Public Health* 19.21. DOI: 10.3390/ijerph192114459.

Guo, H. (2017). "Exploring online learning data using fractal dimensions". *ETS Research Report Series* 2017.1, pp. 1–14. DOI: 10.1002/ets2.12143.

Haitao, T. et al. (2020). "COVID-19 and sex differences: Mechanisms and biomarkers". In: *Mayo Clinic Proceedings*. Vol. 95. 10. Elsevier, pp. 2189–2203.

Hale, T. et al. (2021). "A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker)". *Nat Hum Behav* 5.4, pp. 529–538. DOI: 10.1038/s41562-021-01079-8.

Hall, P. and Wood, A. (1993). "On the performance of box-counting estimators of fractal dimension". *Biometrika* 80.1, pp. 246–251. DOI: 10.1093/biomet/80.1.246.

Hatlee, M. D. and Kozak, J. J. (1981). "Stochastic flows in integral and fractal dimensions and morphogenesis". *Proc Natl Acad Sci U S A* 78.2, pp. 972–5. DOI: 10.1073/pnas.78.2.972.

Hawkins, R. B., Charles, E. J., and Mehaffey, J. H. (2020). "Socio-economic status and COVID-19-related cases and fatalities". *Public Health* 189, pp. 129–134. DOI: 10.1016/j.puhe.2020.09.016.

Het Koninklijk Meteorologisch Instituut van België (2025). *Klimatologische overzichten*. https://www.meteo.be/nl/klimaat/klimaat-van-belgie/klimatologisch-overzicht.

Hu, Z. et al. (2020). "Clinical characteristics of 24 asymptomatic infections with COVID-19 screened among close contacts in Nanjing, China". *Science China. Life sciences* 63.5, pp. 706–711. DOI: 10.1007/s11427-020-1661-4.

Islam, N. et al. (2021). "Excess deaths associated with covid-19 pandemic in 2020: Age and sex disaggregated time series analysis in 29 high income countries". *Bmj* 373, n1137. DOI: 10.1136/bmj.n1137.

Jamagne, P., Lebrun, L., and Sajotte, C. (2020). *Vademecum statistische sectoren*. https://statbel.fgov.be/sites/default/files/files/opendata/Statistische%20sectoren/Secteurs%20stat-NL_tcm325-174181.pdf.

Janssen, R. et al. (2024). "Nationwide quality assurance of high-throughput diagnostic molecular testing during the SARS-CoV-2 pandemic: Role of the Belgian National Reference Centre". *Virology Journal* 21.1, p. 40. DOI: 10.1186/s12985-024-02308-y.

Jefferies, S. et al. (2020). "COVID-19 in New Zealand and the impact of the national response: A descriptive epidemiological study". *Lancet Public Health* 5.11, e612–e623. DOI: 10.1016/s2468-2667(20)30225-5.

Jeffery, C., Ozonoff, A., and Pagano, M. (2014). "The effect of spatial aggregation on performance when mapping a risk of disease". *International Journal of Health Geographics* 13.1, p. 9. DOI: 10.1186/1476-072X-13-9.

Kadi, N. and Khelfaoui, M. (2020). "Population density, a factor in the spread of COVID-19 in Algeria: Statistic study". *Bull Natl Res Cent* 44.1, p. 138. DOI: 10.1186/s42269-020-00393-x.

Kent, J. T. and Andrew, T. A. W. (1997). "Estimating the Fractal Dimension of a Locally Self-Similar Gaussian Process by Using Increments". *Journal of the Royal Statistical Society. Series B (Methodological)* 59.3, pp. 679–699.

Kesić, S. and Spasić, S. Z. (2016). "Application of Higuchi's fractal dimension from basic to clinical neurophysiology: A review". *Comput Methods Programs Biomed* 133, pp. 55–70. DOI: 10.1016/j.cmpb.2016.05.014.

Khan, Y. et al. (2023). "Impact of the COVID-19 pandemic on delayed care of cardiovascular diseases in Europe: A systematic review". *The Lancet* 402, S61. DOI: 10.1016/S0140-6736(23)02117-7.

Klein, S. L. and Flanagan, K. L. (2016). "Sex differences in immune responses". *Nat Rev Immunol* 16.10, pp. 626–38. DOI: 10.1038/nri.2016.90.

Kontis, V. et al. (2020). "Magnitude, demographics and dynamics of the effect of the first wave of the COVID-19 pandemic on all-cause mortality in 21 industrialized countries". *Nat Med* 26.12, pp. 1919–1928. DOI: 10.1038/s41591-020-1112-0.

Krstacic, G., Krstacic, A., Smalcelj, A., Milicic, D., and Jembrek-Gostovic, M. (2007). "The "Chaos Theory" and nonlinear dynamics in heart rate variability analysis: Does it work in short-time series in patients with coronary heart disease?" *Annals of noninvasive electrocardiology: The official journal of the International Society for Holter and Noninvasive Electrocardiology, Inc* 12.2, pp. 130–136. DOI: 10.1111/j.1542-474X.2007.00151.x.

Kwok, W. C. et al. (2021). "Modelling the impact of travel restrictions on COVID-19 cases in Hong Kong in early 2020". *BMC Public Health* 21.1, p. 1878. DOI: 10.1186/s12889-021-11889-0.

Lamot, M., Kerman, K., and Kirbiš, A. (2022). "Distrustful, dissatisfied, and conspiratorial: A latent profile analysis of COVID-19 vaccination rejection". *Int J Environ Res Public Health* 19.16. DOI: 10.3390/ijerph191610096.

Lampl, B. M. J. and Salzberger, B. (2020). "Changing epidemiology of COVID-19". *GMS hygiene and infection control* 15, Doc27–Doc27. DOI: 10.3205/dgkh000362.

Lange, S. J. et al. (2020). "Potential indirect effects of the COVID-19 pandemic on use of emergency departments for acute life-threatening conditions - United States, January-May 2020". *MMWR Morb Mortal Wkly Rep* 69.25, pp. 795–800. DOI: 10.15585/mmwr.mm6925e2.

Latsuzbaia, A., Herold, M., Bertemes, J., and Mossong, J. (2020). "Evolving social contact patterns during the COVID-19 crisis in Luxembourg". *PloS one* 15.8, e0237128–e0237128. DOI: 10.1371/journal.pone.0237128.

Lau, C. I., Yeh, J. H., Tsai, Y. F., Hsiao, C. Y., Wu, Y. T., and Jao, C. W. (2023). "Decreased brain structural network connectivity in patients with mild cognitive impairment: A novel fractal dimension analysis". *Brain Sci* 13.1. DOI: 10.3390/brainsci13010093.

Lawson, A. B. (2018). "Bayesian disease mapping: Hierarchical modeling in spatial epidemiology (3rd ed.)" *Taylor & Francis*. DOI: 10.1201/9781351271769.

Lee, W.-E. et al. (2023). "Direct and indirect mortality impacts of the COVID-19 pandemic in the United States, March 1, 2020 to January 1, 2022". *eLife* 12, e77562. DOI: 10.7554/eLife.77562.

Leidner, A. et al. (2021). "Opening of large institutions of higher education and county-level COVID-19 incidence — United States, July 6–September 1, 2020". *MMWR Morb Mortal Wkly Rep* 70.1, pp. 14–19. DOI: 10.15585/mmwr.mm7001a4.

Lemmens, S., Devulder, A., Van Keer, K., Bierkens, J., De Boever, P., and Stalmans, I. (2020). "Systematic review on fractal dimension of the retinal vasculature in neurodegeneration and stroke: Assessment of a Potential Biomarker". *Front Neurosci* 14, p. 16. DOI: 10.3389/fnins.2020.00016.

Lesaffre, E. and Lawson, A. B. (2012). "Bayesian biostatistics". *Wiley*. DOI: 10.1002/9781119942412.

Li, H., Wang, L., Zhang, M., Lu, Y., and Wang, W. (2022). "Effects of vaccination and non-pharmaceutical interventions and their lag times on the COVID-19 pandemic: Comparison of eight countries". *PLOS Neglected Tropical Diseases* 16.1, e0010101. DOI: 10.1371/journal.pntd.0010101.

Liebovitch, L. S. and Toth, T. (1989). "A fast algorithm to determine fractal dimensions by box counting". *Physics Letters A* 141.8, pp. 386–390. DOI: 10.1016/0375-9601(89)90854-2.

Lindgren, F. and Rue, H. (2015). "Bayesian Spatial Modelling with R-INLA". *Journal of Statistical Software* 63.19, pp. 1–25. DOI: 10.18637/jss.v063.i19.

Liu, X. (2012). "Appendix A: The Delta Method". In: *Survival Analysis*. Wiley Online Books, pp. 405–406. DOI: 10.1002/9781118307656.app1.

Lo, C.-H. et al. (2021). "Race, ethnicity, community-level socioeconomic factors, and risk of COVID-19 in the United States and the United Kingdom". *EClinicalMedicine* 38, pp. 101029–101029. DOI: 10.1016/j.eclinm.2021.101029.

Lokonon, B. E., Montcho, Y., Klingler, P., Tovissodé, C. F., Glèlè Kakaï, R., and Wolkewitz, M. (2023). "Lag-time effects of vaccination on SARS-CoV-2 dynamics in German hospitals and intensive-care units". *Front Public Health* 11, p. 1085991. DOI: 10.3389/fpubh.2023.1085991.

Lovejoy, S. (1982). "Area-perimeter relation for rain and cloud areas". *Science* 216.4542, pp. 185–7. DOI: 10.1126/science.216.4542.185.

Mafi, J. N. et al. (2022). "Trends in US ambulatory care patterns during the COVID-19 pandemic, 2019-2021". *Jama* 327.3, pp. 237–247. DOI: 10.1001/jama.2021.24294.

Mandelbrot, B. (1975). "Stochastic models for the Earth's relief, the shape and the fractal dimension of the coastlines, and the number-area rule for islands". *Proc Natl Acad Sci U S A* 72.10, pp. 3825–8. DOI: 10.1073/pnas.72.10.3825.

Mandelbrot, B. (1982). *The fractal geometry of nature*. Vol. 2. WH freeman New York.

Matusiak, Ł., Szepietowska, M., Krajewski, P. K., Białynicki-Birula, R., and Szepietowski, J. C. (2020). "The use of face masks during the COVID-19 pandemic in Poland: A survey study of 2315 young adults". *Dermatol Ther* 33.6, e13909. DOI: 10.1111/dth.13909.

McLaughlin, J. M. et al. (2021). "County-level predictors of Coronavirus Disease 2019 (COVID-19) cases and deaths in the United States: What happened, and where do we go from here?" *Clin Infect Dis* 73.7, e1814–e1821. DOI: 10.1093/cid/ciaa1729.

McNeil, N., Odton, P., and Ueranantasun, A. (2011). "Spline interpolation of demographic data revisited". *Sonklanakarin Journal of Science and Technology* 33.1, p. 117.

Mohammed, I. et al. (2022). "The efficacy and effectiveness of the COVID-19 vaccines in reducing infection, severity, hospitalization, and mortality: A systematic review". *Hum Vaccin Immunother* 18.1, p. 2027160. DOI: 10.1080/21645515.2022.2027160.

Molenberghs, G. et al. (2022). "COVID-19 mortality, excess mortality, deaths per million and infection fatality ratio, Belgium, 9 March 2020 to 28 June 2020". *Euro Surveill* 27.7. DOI: 10.2807/1560-7917.Es.2022.27.7.2002060.

Mollalo, A., Vahedi, B., and Rivera, K. M. (2020). "GIS-based spatial modeling of COVID-19 incidence rate in the continental United States". *Science of The Total Environment* 728, p. 138884. DOI: 10.1016/j.scitotenv.2020.138884.

Moran, P. A. P. (1950). "Notes on continuous stochastic phenomena". *Biometrika* 37.1/2, pp. 17–23. DOI: 10.2307/2332142.

Mossong, J. et al. (2008). "Social contacts and mixing patterns relevant to the spread of infectious diseases". *PLoS medicine* 5.3, e74–e74. DOI: 10.1371/journal.pmed.0050074.

Msemburi, W., Karlinsky, A., Knutson, V., Aleshin-Guendel, S., Chatterji, S., and Wakefield, J. (2023). "The WHO estimates of excess mortality associated with the COVID-19 pandemic". *Nature* 613.7942, pp. 130–137. DOI: 10.1038/s41586-022-05522-2.

Nagata, J. M. (2020). "Supporting young adults to rise to the challenge of COVID-19". *The Journal of adolescent health: Official publication of the Society for Adolescent Medicine* 67.2, pp. 297–298. DOI: 10.1016/j.jadohealth.2020.04.020.

Natalia, Y. A., Delporte, M., De Witte, D., Beutels, P., Dewatripont, M., and Molenberghs, G. (2023). "Assessing the impact of COVID-19 passes and mandates on disease transmission, vaccination intention, and uptake: A scoping review". *BMC Public Health* 23.1, p. 2279. DOI: 10.1186/s12889-023-17203-4.

Natalia, Y. A., Faes, C., Neyens, T., and Molenberghs, G. (2022). "The COVID-19 wave in Belgium during the Fall of 2020 and its association with higher education". *PLOS ONE* 17.2, e0264516. DOI: 10.1371/journal.pone.0264516.

Nguyen, M. H., Nguyen, T. H. T., Molenberghs, G., Abrams, S., Hens, N., and Faes, C. (2023). "The impact of national and international travel on spatio-temporal transmission of SARS-CoV-2 in Belgium in 2021". *BMC Infect Dis* 23.1, p. 428. DOI: 10.1186/s12879-023-08368-9.

Nian, F. and Li, F. (2024). "Epidemic propagation risk study with effective fractal dimension". *Risk Analysis* 44.3, pp. 612–622. DOI: 10.1111/risa.14192.

Nielsen, J., Mazick, A., et al. (2013). "Pooling European all-cause mortality: Methodology and findings for the seasons 2008/2009 to 2010/2011". *Epidemiology and Infection* 141.9, pp. 1996–2010. DOI: 10.1017/S0950268812002580.

Nielsen, J., Nørgaard, S. K., Lanzieri, G., Vestergaard, L. S., and Moelbak, K. (2021). "Sex-differences in COVID-19 associated excess mortality is not exceptional for the COVID-19 pandemic". *Sci Rep* 11.1, p. 20815. DOI: 10.1038/s41598-021-00213-w.

Oh, D. L. et al. (2022). "Neighborhood factors associated with COVID-19 cases in California". *J Racial Ethn Health Disparities*, pp. 1–10. DOI: 10.1007/s40615-022-01443-y.

Our World in Data (2025). *COVID-19 Data Explorer*. https://ourworldindata.org/explorers/covid.

Păcurar, C. M. and Necula, B. R. (2020). "An analysis of COVID-19 spread based on fractal interpolation and fractal dimension". *Chaos Solitons Fractals* 139, p. 110073. DOI: 10.1016/j.chaos.2020.110073.

Paelinck, J. and Klaassen, L. (1979). *Spatial econometrics*. Saxon House.

Parino, F., Zino, L., Porfiri, M., and Rizzo, A. (2021). "Modelling and predicting the effect of social distancing and travel restrictions on COVID-19 spreading". *J R Soc Interface* 18.175, p. 20200875. DOI: 10.1098/rsif.2020.0875.

Paumgartner, D., Losa, G., and Weibel, E. R. (1981). "Resolution effect on the stereological estimation of surface and volume and its interpretation in terms of fractal dimensions". *J Microsc* 121.Pt 1, pp. 51–63. DOI: 10.1111/j.1365-2818.1981.tb01198.x.

Pedersen, M. J. and Favero, N. (2020). "Social distancing during the COVID-19 pandemic: Who are the present and future non-compliers?" *Public Adm Rev*. DOI: 10.1111/puar.13240.

Peitgen, H.-O., Jürgens, H., and Saupe, D. (2004). *Chaos and fractals: New frontiers of science*. Ed. by H.-O. Peitgen, H. Jürgens, and D. Saupe. New York, NY: Springer New York, pp. 173–214. DOI: 10.1007/0-387-21823-8_5.

Polack Fernando, P. et al. (2020). "Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine". *New England Journal of Medicine* 383.27, pp. 2603–2615. DOI: 10.1056/NEJMoa2034577.

Pouwels, K. B. et al. (2021). "Community prevalence of SARS-CoV-2 in England from April to November, 2020: Results from the ONS Coronavirus Infection Survey". *Lancet Public Health* 6.1, e30–e38. DOI: 10.1016/s2468-2667(20)30282-6.

Quaglio, G. et al. (2022). "The impact of the COVID-19 pandemic on health service use in sub-Saharan Africa". *Public Health Action* 12.1, pp. 34–39. DOI: 10.5588/pha.21.0073.

Roumeliotis, S. et al. (2021). "Be careful with ecological associations". *Nephrology* 26.6, pp. 501–505. DOI: 10.1111/nep.13861.

Rudin, C. (2019). "Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead". *Nature Machine Intelligence* 1.5, pp. 206–215. DOI: 10.1038/s42256-019-0048-x.

Rue, H., Martino, S., and Chopin, N. (2009). "Approximate Bayesian inference for latent Gaussian Models using integrated nested Laplace approximations (with discussion)." *Journal of the Royal Statistical Society B* 71, pp. 319–392.

Ruiz de Miras, J., Ibáñez-Molina, A. J., Soriano, M. F., and Iglesias-Parro, S. (2023). "Fractal dimension analysis of resting state functional networks in schizophrenia from EEG signals". *Front Hum Neurosci* 17, p. 1236832. DOI: 10.3389/fnhum.2023.1236832.

Sammouda, R. and El-Zaart, A. (2021). "An Optimized Approach for Prostate Image Segmentation Using K-Means Clustering Algorithm with Elbow Method". *Computational intelligence and neuroscience* 2021, pp. 4553832–4553832. DOI: 10.1155/2021/4553832.

Sardá-Espinosa, A. (2019). "Time-Series Clustering in R Using the dtwclust Package". *The R Journal* 11.1, pp. 22–43. DOI: 10.32614/RJ-2019-023.

Sarkar, M. and Leong, T. Y. (2003). "Characterization of medical time series using fuzzy similarity-based fractal dimensions". *Artif Intell Med* 27.2, pp. 201–22. DOI: 10.1016/s0933-3657(02)00114-8.

Sartorius, B., Lawson, A. B., and Pullan, R. L. (2021). "Modelling and predicting the spatio-temporal spread of COVID-19, associated deaths and impact of key risk factors in England". *Sci Rep* 11.1, p. 5378. DOI: 10.1038/s41598-021-83780-2.

Sciensano (2022). *COVID-19 - epidemiologische situatie*. https://covid-19.sciensano.be/nl/covid-19-epidemiologische-situatie.

Sciensano (2023a). *Belgian mortality monitoring (BE-MOMO)*. https://epistat.sciensano.be/momo/.

Sciensano (2023b). *Epidemiology of COVID-19 mortality in Belgium from wave 1 to wave 7 (March 2020 – 11 September 2022)*. https://www.sciensano.be/en/biblio/epidemiology-covid-19-mortality-belgium-wave-1-wave-7-march-2020-11-september-2022.

Sciensano (2023c). *Infectious diseases dashboard*. https://epistat.sciensano.be/dashboard/.

Sciensano (2025a). *Belgium COVID-19 epidemiological situation*. https://covid-19.sciensano.be/nl/covid-19-epidemiologische-situatie.

Sciensano (2025b). *Sciensano*. https://www.sciensano.be/en.

Scobie, H. M. et al. (2021). "Monitoring Incidence of COVID-19 Cases, Hospitalizations, and Deaths, by Vaccination Status - 13 U.S. Jurisdictions, April 4-July 17, 2021". *MMWR Morb Mortal Wkly Rep* 70.37, pp. 1284–1290. DOI: 10.15585/mmwr.mm7037e1.

Scortichini, M. et al. (2021). "Excess mortality during the COVID-19 outbreak in Italy: A two-stage interrupted time-series analysis". *Int J Epidemiol* 49.6, pp. 1909–1917. DOI: 10.1093/ije/dyaa169.

Seemann, T. et al. (2020). "Tracking the COVID-19 pandemic in Australia using genomics". *Nat Commun* 11.1, p. 4376. DOI: 10.1038/s41467-020-18314-x.

El-Shabasy, R. M., Nayel, M. A., Taher, M. M., Abdelmonem, R., Shoueir, K. R., and Kenawy, E. R. (2022). "Three waves changes, new variant strains, and vaccination effect against COVID-19 pandemic". *International Journal of Biological Macromolecules* 204, pp. 161–168. DOI: 10.1016/j.ijbiomac.2022.01.118.

Shy, C.-G. et al. (2022). "Rapid control of a SARS-CoV-2 B.1.617.2 (Delta) variant COVID-19 community outbreak: The successful experience in Pingtung County of Taiwan". *International journal of environmental research and public health* 19.3, p. 1421. DOI: 10.3390/ijerph19031421.

Simpson, D., Rue, H., Riebler, A., Martins, T. G., and Sørbye, S. H. (2017). "Penalising model component complexity: A principled, practical approach to constructing priors". *Statistical Science* 32.1, pp. 1–28. DOI: 10.1214/16-STS576.

Siu, J. Y. (2020). "Health inequality experienced by the socially disadvantaged populations during the outbreak of COVID-19 in Hong Kong: An interaction with social inequality". *Health Soc Care Community*. DOI: 10.1111/hsc.13214.

Sørbye, S. H. and Rue, H. (2014). "Scaling intrinsic Gaussian Markov random field priors in spatial modelling". *Spatial Statistics* 8, pp. 39–51. DOI: 10.1016/j.spasta.2013.06.004.

Spellerberg, I. F. and Fedor, P. J. (2003). "A tribute to Claude Shannon (1916–2001) and a plea for more rigorous use of species richness, species diversity and the 'Shannon–Wiener' Index". *Global Ecology and Biogeography* 12.3, pp. 177–179. DOI: 10.1046/j.1466-822X.2003.00015.x.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., and Van Der Linde, A. (2002). "Bayesian measures of model complexity and fit". *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 64.4, pp. 583–639. DOI: 10.1111/1467-9868.00353.

Statbel (2025). *Statbel: Open Data*. https://https://statbel.fgov.be/en/open-data.

Steinley, D. (2006). "K-means clustering: A half-century synthesis". *Br J Math Stat Psychol* 59.Pt 1, pp. 1–34. DOI: 10.1348/000711005x48266.

Sundaram, S. S. et al. (2022). "Impact of age, sex, race, and regionality on major clinical outcomes of COVID-19 in hospitalized patients in the United States". *BMC Infectious Diseases* 22.1, p. 659. DOI: 10.1186/s12879-022-07611-z.

Swapna, M. S., Renjini, A., Raj, V., Sreejyothi, S., and Sankararaman, S. (2020). "Time series and fractal analyses of wheezing: A novel approach". *Physical and engineering sciences in medicine* 43.4, pp. 1339–1347. DOI: 10.1007/s13246-020-00937-5.

Tadesse, B. T. et al. (2022). "Impact of vaccination with SCB-2019 COVID-19 vaccine on transmission of SARS-CoV-2 infection: A household contact study in the Philippines". *Clin Infect Dis*. DOI: 10.1093/cid/ciac914.

Tan, X. et al. (2025). "The spatiotemporal scaling laws of urban population dynamics". *Nat Commun* 16.1, p. 2881. DOI: 10.1038/s41467-025-58286-4.

Tenforde, M. W. et al. (2021). "Association between mRNA vaccination and COVID-19 hospitalization and disease severity". *JAMA* 326.20, pp. 2043–2054. DOI: 10.1001/jama.2021.19499.

ter Braak, C. J. F. (1986). "Canonical correspondence analysis: A new eigenvector technique for multivariate direct gradient analysis". *Ecology* 67.5, pp. 1167–1179. DOI: 10.2307/1938672.

The Lancet Public Health (2019). "Achieving health equity in the European region". *The Lancet Public Health* 4.10. doi: 10.1016/S2468-2667(19)30181-1, e482. DOI: 10.1016/S2468-2667(19)30181-1.

Thompson, A. E., Anisimowicz, Y., Miedema, B., Hogg, W., Wodchis, W. P., and Aubrey-Bassler, K. (2016). "The influence of gender and other patient characteristics on health care-seeking behaviour: A QUALICOPC study". *BMC Fam Pract* 17, p. 38. DOI: 10.1186/s12875-016-0440-0.

Tobler, W. R. (1959). "Automation and cartography". *Geographical Review* 49.4, pp. 526–534. DOI: 10.2307/212211.

Torres, M. J., Coste, J., Canouï-Poitrine, F., Pouchot, J., Rachas, A., and Carcaillon-Bentata, L. (2023). "Impact of the first COVID-19 pandemic wave on hospitalizations and deaths caused by geriatric syndromes in France: A nationwide study". *J Gerontol A Biol Sci Med Sci*. DOI: 10.1093/gerona/glad032.

Tulchinsky, T. H. (2018). "John Snow, Cholera, the Broad Street Pump; Waterborne Diseases Then and Now". *Case Studies in Public Health*, pp. 77–99.

Vang, K. E. et al. (2021). "Participation in fraternity and sorority activities and the spread of COVID-19 among residential university communities - Arkansas, August 21-September 5, 2020". *MMWR Morb Mortal Wkly Rep* 70.1, pp. 20–23. DOI: 10.15585/mmwr.mm7001a5.

Venkatesan, P. (2020). "The changing demographics of COVID-19". *Lancet Respir Med* 8.12, e95. DOI: 10.1016/s2213-2600(20)30461-6.

Verbeeck, J. et al. (2021). "A linear mixed model to estimate COVID-19-induced excess mortality". *Biometrics*. DOI: 10.1111/biom.13578.

Verbeke, G. and Molenberghs, G. (2009). "Linear Mixed Models for Longitudinal Data". *Springer New York*.

Vestergaard, L. S. et al. (2020). "Excess all-cause mortality during the COVID-19 pandemic in Europe - preliminary pooled estimates from the EuroMOMO network, March to April 2020". *Euro Surveill* 25.26. DOI: 10.2807/1560-7917.Es.2020.25.26.2001214.

Viboud, C., Simonsen, L., and Chowell, G. (2016). "A generalized-growth model to characterize the early ascending phase of infectious disease outbreaks". *Epidemics* 15, pp. 27–37. DOI: 10.1016/j.epidem.2016.01.002.

Vila-Corcoles, A. et al. (2021). "COVID19-related and all-cause mortality risk among middle-aged and older adults across the first epidemic wave of SARS-COV-2 infection: A population-based cohort stuJune 2020.dy in Southern Catalonia, Spain, March". *BMC Public Health* 21.1, p. 1795. DOI: 10.1186/s12889-021-11879-2.

Vranckx, M. et al. (2023). "A spatial model to jointly analyze self-reported survey data of COVID-19 symptoms and official COVID-19 incidence data". *Biometrical Journal* 65.1, p. 2100186. DOI: 10.1002/bimj.202100186.

Walkowiak, M. P., Domaradzki, J., and Walkowiak, D. (2023). "Unmasking the COVID-19 pandemic prevention gains: Excess mortality reversal in

2022". *Public Health* 223, pp. 193–201. DOI: 10.1016/j.puhe.2023.08.004.

Wang, H., Varol, T., Gültzow, T., Zimmermann, H., Ruiter, R., and Jonas, K. (2024). "Spatio-temporal distributions of COVID-19 vaccine doses uptake in the Netherlands: A Bayesian ecological modelling analysis". *Epidemiology and Infection* 152, e119. DOI: 10.1017/S0950268824001249.

Wang, Y., Huang, Y. C., and Cai, Q. (2022). "Exploring the mediating role of government-public relationships during the COVID-19 pandemic: A model comparison approach". *Public Relat Rev* 48.4, p. 102231. DOI: 10.1016/j.pubrev.

Watanabe, S. (2010). "Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory". *Journal of Machine Learning Research* 11. Export Date: 22 August 2024; Cited By: 1784, pp. 3571–3594.

Weber, T. et al. (2022). "COVID-19 associated reduction in hypertension-related diagnostic and therapeutic procedures in Excellence Centers of the European Society of Hypertension". *Blood Press* 31.1, pp. 71–79. DOI: 10.1080/08037051.2022.2060182.

Weitz, D. A., Huang, J. S., Lin, M. Y., and Sung, J. (1985). "Limits of the fractal dimension for irreversible kinetic aggregation of gold colloids". *Phys Rev Lett* 54.13, pp. 1416–1419. DOI: 10.1103/PhysRevLett.54.1416.

Werner, G. (2010). "Fractals in the nervous system: Conceptual implications for theoretical neuroscience". *Frontiers in Physiology* Volume 1 - 2010. DOI: 10.3389/fphys.2010.00015.

Wilson, E. et al. (2020). "Multiple COVID-19 clusters on a university campus - North Carolina, August 2020". *MMWR Morb Mortal Wkly Rep* 69.39, pp. 1416–1418. DOI: 10.15585/mmwr.mm6939e3.

Wong, K. L. M. et al. (2023). "Social contact patterns during the COVID-19 pandemic in 21 European countries - evidence from a two-year study". *BMC Infect Dis* 23.1, p. 268. DOI: 10.1186/s12879-023-08214-y.

World Health Organization (2020a). *Novel Coronavirus (2019-nCoV) situation report - 5, 25 January 2020*. https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200125-sitrep-5-2019-ncov.pdf?sfvrsn=429b143d_8.

World Health Organization (2020b). *WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020*. https://

www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020.

World Health Organization (2021). *Coronavirus disease (COVID-19) weekly epidemiological update and weekly operational update*. https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports/situation-reports-archive.

World Health Organization (2024). *International Statistical Classification of Diseases and Related Health Problems (ICD) - 11*. https://www.who.int/classifications/classification-of-diseases.

Wu, F. et al. (2020). "A new coronavirus associated with human respiratory disease in China". *Nature* 579.7798, pp. 265–269. DOI: 10.1038/s41586-020-2008-3.

Yakovleva, A. et al. (2022). "Tracking SARS-COV-2 variants using Nanopore sequencing in Ukraine in 2021". *Sci Rep* 12.1, p. 15749. DOI: 10.1038/s41598-022-19414-y.

Yamana, T. K. et al. (2023). "The impact of COVID-19 vaccination in the US: Averted burden of SARS-COV-2-related cases, hospitalizations and deaths". *PLoS One* 18.4, e0275699. DOI: 10.1371/journal.pone.0275699.

Yasri, S. and Wiwanitkit, V. (2020). "Sharing Cigarette Smoking and COVID-19 Outbreak in a Party Group". *International journal of preventive medicine* 11, pp. 50–50. DOI: 10.4103/ijpvm.IJPVM_121_20.

Zhou, J. C., Salahshour, S., Ahmadian, A., and Senu, N. (2022). "Modeling the dynamics of COVID-19 using fractal-fractional operator with a case study". *Results Phys* 33, p. 105103. DOI: 10.1016/j.rinp.2021.105103.

Zou, Y., Yang, W., Lai, J., Hou, J., and Lin, W. (2022). "Vaccination and quarantine effect on COVID-19 transmission dynamics incorporating Chinese-Spring-Festival travel rush: Modeling and simulations". *Bulletin of mathematical biology* 84.2, pp. 30–30. DOI: 10.1007/s11538-021-00958-5.

APPENDIX A

APPENDIX FOR CHAPTER 2

A.1 Parameter estimates for the spatial model based on Equation 2.2 in Flemish Region.

Period	Effect	Parameter	Mean	SD	95% credible interval	
					Lower	Upper
June 1-15	Age ≤17	α_{11}	-10,140	0,164	-10,463	-9,821
	Age 18-29	α_{12}	-9,249	0,149	-9,545	-8,960
	Age 30-39	α_{13}	-9,359	0,154	-9,663	-9,061
	Age 40-49	α_{14}	-9,446	0,155	-9,753	-9,144
	Age 50-59	α_{15}	-9,468	0,153	-9,771	-9,171
	Age 60-69	α_{16}	-9,833	0,170	-10,168	-9,503
	Age 70-79	α_{17}	-9,429	0,169	-9,762	-9,101
	Age ≥80	α_{18}	-8,611	0,154	-8,916	-8,310
	Female	α_{22}	0,250	0,067	0,119	0,381
	Income	β_1	-0,085	0,046	-0,176	0,006
	Log(popdens)	β_2	0,060	0,057	-0,051	0,173
	Age ≤17 × sturatio	γ_{11}	0,226	0,142	-0,052	0,504
	Age 18-29 × sturatio	γ_{12}	0,128	0,126	-0,119	0,377
	Age 30-39 × sturatio	γ_{13}	0,001	0,144	-0,281	0,284
	Age 40-49 × sturatio	γ_{14}	0,129	0,140	-0,145	0,405
	Age 50-59 × sturatio	γ_{15}	-0,001	0,150	-0,296	0,294
	Age 60-69 × sturatio	γ_{16}	-0,047	0,178	-0,397	0,303
	Age 70-79 × sturatio	γ_{17}	-0,151	0,187	-0,518	0,217
	Age ≥80 × sturatio	γ_{18}	-0,109	0,154	-0,411	0,194
	Precision SS	τ_v	4,502	4,436	0,852	16,217
	Precision SU	τ_u	2,669	0,790	1,397	4,479

Period	Effect	Parameter	Mean	SD	95% credible interval	
					Lower	Upper
June 16-30	Age ≤17	α_{11}	-9,726	0,159	-10,040	-9,417
	Age 18-29	α_{12}	-8,869	0,145	-9,157	-8,587
	Age 30-39	α_{13}	-9,026	0,151	-9,325	-8,731
	Age 40-49	α_{14}	-9,259	0,157	-9,570	-8,953
	Age 50-59	α_{15}	-9,519	0,163	-9,842	-9,200
	Age 60-69	α_{16}	-9,617	0,173	-9,959	-9,279
	Age 70-79	α_{17}	-9,593	0,192	-9,971	-9,218
	Age ≥80	α_{18}	-8,879	0,172	-9,218	-8,544
	Female	α_{22}	0,133	0,073	-0,009	0,275
	Income	β_1	-0,160	0,048	-0,255	-0,066
	Log(popdens)	β_2	-0,021	0,057	-0,132	0,093
	Age ≤17 × sturatio	γ_{11}	0,103	0,175	-0,240	0,446
	Age 18-29 × sturatio	γ_{12}	0,057	0,143	-0,224	0,338
	Age 30-39 × sturatio	γ_{13}	0,299	0,136	0,031	0,567
	Age 40-49 × sturatio	γ_{14}	0,176	0,163	-0,144	0,495
	Age 50-59 × sturatio	γ_{15}	-0,066	0,201	-0,460	0,329
	Age 60-69 × sturatio	γ_{16}	0,157	0,193	-0,221	0,535
	Age 70-79 × sturatio	γ_{17}	-0,082	0,253	-0,577	0,414
	Age ≥80 × sturatio	γ_{18}	0,122	0,189	-0,248	0,491
	Precision SS	τ_v	1,310	0,762	0,337	3,231
	Precision SU	τ_u	7,371	5,961	1,756	23,269
July 1-15	Age ≤17	α_{11}	-9,847	0,153	-10,149	-9,551
	Age 18-29	α_{12}	-8,796	0,139	-9,073	-8,526
	Age 30-39	α_{13}	-9,044	0,145	-9,331	-8,763
	Age 40-49	α_{14}	-9,172	0,147	-9,464	-8,887
	Age 50-59	α_{15}	-9,259	0,147	-9,551	-8,973
	Age 60-69	α_{16}	-9,658	0,163	-9,981	-9,341
	Age 70-79	α_{17}	-9,641	0,177	-9,991	-9,296
	Age ≥80	α_{18}	-9,220	0,175	-9,565	-8,879
	Female	α_{22}	-0,103	0,059	-0,220	0,013
	Income	β_1	-0,181	0,049	-0,277	-0,084
	Log(popdens)	β_2	0,055	0,056	-0,054	0,167
	Age ≤17 × sturatio	γ_{11}	0,199	0,170	-0,133	0,532
	Age 18-29 × sturatio	γ_{12}	0,355	0,129	0,104	0,608
	Age 30-39 × sturatio	γ_{13}	0,353	0,140	0,079	0,628
	Age 40-49 × sturatio	γ_{14}	0,168	0,161	-0,147	0,484
	Age 50-59 × sturatio	γ_{15}	0,340	0,150	0,047	0,634
	Age 60-69 × sturatio	γ_{16}	0,274	0,179	-0,077	0,626
	Age 70-79 × sturatio	γ_{17}	0,031	0,229	-0,417	0,480
	Age ≥80 × sturatio	γ_{18}	-0,176	0,235	-0,637	0,284
	Precision SS	τ_v	0,468	0,077	0,334	0,637
	Precision SU	τ_u	2211,629	2419,826	149,109	8636,667
July 16-31	Age ≤17	α_{11}	-8,332	0,092	-8,515	-8,153
	Age 18-29	α_{12}	-7,542	0,089	-7,718	-7,368
	Age 30-39	α_{13}	-7,755	0,091	-7,936	-7,578
	Age 40-49	α_{14}	-7,856	0,092	-8,038	-7,677
	Age 50-59	α_{15}	-8,135	0,094	-8,321	-7,952

Period	Effect	Parameter	Mean	SD	95% credible interval	
					Lower	Upper
	Age 60-69	α_{16}	-8,780	0,107	-8,992	-8,571
	Age 70-79	α_{17}	-8,897	0,118	-9,129	-8,667
	Age ≥80	α_{18}	-8,503	0,115	-8,730	-8,277
	Female	α_{22}	-0,018	0,031	-0,077	0,042
	Income	β_1	-0,109	0,036	-0,179	-0,038
	Log(popdens)	β_2	0,025	0,041	-0,054	0,105
	Age ≤17 × sturatio	γ_{11}	-0,016	0,104	-0,219	0,188
	Age 18-29 × sturatio	γ_{12}	0,046	0,094	-0,137	0,230
	Age 30-39 × sturatio	γ_{13}	-0,001	0,100	-0,197	0,195
	Age 40-49 × sturatio	γ_{14}	0,041	0,101	-0,158	0,240
	Age 50-59 × sturatio	γ_{15}	0,032	0,105	-0,174	0,238
	Age 60-69 × sturatio	γ_{16}	-0,223	0,138	-0,493	0,047
	Age 70-79 × sturatio	γ_{17}	0,187	0,134	-0,075	0,450
	Age ≥80 × sturatio	γ_{18}	0,083	0,136	-0,184	0,350
	Precision SS	τ_v	0,621	0,083	0,473	0,800
	Precision SU	τ_u	2161,983	2297,688	156,350	8265,198
August 1-15	Age ≤17	α_{11}	-8,209	0,091	-8,389	-8,032
	Age 18-29	α_{12}	-7,407	0,088	-7,582	-7,235
	Age 30-39	α_{13}	-7,625	0,090	-7,803	-7,450
	Age 40-49	α_{14}	-7,784	0,091	-7,964	-7,606
	Age 50-59	α_{15}	-8,103	0,093	-8,288	-7,921
	Age 60-69	α_{16}	-8,712	0,105	-8,919	-8,506
	Age 70-79	α_{17}	-8,977	0,122	-9,216	-8,739
	Age ≥80	α_{18}	-8,342	0,112	-8,563	-8,124
	Female	α_{22}	0,039	0,030	-0,020	0,097
	Income	β_1	-0,128	0,035	-0,197	-0,058
	Log(popdens)	β_2	0,077	0,041	-0,003	0,157
	Age ≤17 × sturatio	γ_{11}	0,022	0,101	-0,176	0,220
	Age 18-29 × sturatio	γ_{12}	0,202	0,089	0,027	0,379
	Age 30-39 × sturatio	γ_{13}	0,202	0,093	0,019	0,386
	Age 40-49 × sturatio	γ_{14}	0,156	0,098	-0,035	0,348
	Age 50-59 × sturatio	γ_{15}	0,206	0,100	0,009	0,403
	Age 60-69 × sturatio	γ_{16}	0,097	0,124	-0,146	0,340
	Age 70-79 × sturatio	γ_{17}	0,043	0,155	-0,261	0,348
	Age ≥80 × sturatio	γ_{18}	0,369	0,115	0,143	0,595
	Precision SS	τ_v	0,956	0,312	0,443	1,650
	Precision SU	τ_u	13,464	10,843	3,622	42,486
August 16-31	Age ≤17	α_{11}	-8,317	0,090	-8,495	-8,141
	Age 18-29	α_{12}	-7,631	0,088	-7,804	-7,459
	Age 30-39	α_{13}	-7,785	0,090	-7,963	-7,610
	Age 40-49	α_{14}	-7,881	0,090	-8,059	-7,704
	Age 50-59	α_{15}	-8,373	0,095	-8,562	-8,187
	Age 60-69	α_{16}	-8,911	0,109	-9,126	-8,698
	Age 70-79	α_{17}	-9,475	0,142	-9,755	-9,197
	Age ≥80	α_{18}	-9,056	0,138	-9,326	-8,786
	Female	α_{22}	-0,020	0,034	-0,088	0,047
	Income	β_1	-0,082	0,033	-0,146	-0,016

Period	Effect	Parameter	Mean	SD	95% credible interval	
					Lower	Upper
	Log(popdens)	β_2	0,127	0,038	0,053	0,202
	Age ≤17 × sturatio	γ_{11}	-0,026	0,092	-0,207	0,154
	Age 18-29 × sturatio	γ_{12}	-0,008	0,085	-0,175	0,159
	Age 30-39 × sturatio	γ_{13}	-0,063	0,090	-0,241	0,114
	Age 40-49 × sturatio	γ_{14}	-0,141	0,095	-0,328	0,045
	Age 50-59 × sturatio	γ_{15}	0,021	0,098	-0,172	0,213
	Age 60-69 × sturatio	γ_{16}	-0,113	0,124	-0,357	0,131
	Age 70-79 × sturatio	γ_{17}	-0,041	0,167	-0,368	0,286
	Age ≥80 × sturatio	γ_{18}	0,180	0,132	-0,080	0,439
	Precision SS	τ_v	0,797	0,124	0,581	1,069
	Precision SU	τ_u	2186,305	2386,827	146,828	8524,154
September 1-15	Age ≤17	α_{11}	-7,644	0,079	-7,800	-7,491
	Age 18-29	α_{12}	-7,023	0,077	-7,176	-6,873
	Age 30-39	α_{13}	-7,278	0,079	-7,434	-7,124
	Age 40-49	α_{14}	-7,428	0,080	-7,586	-7,272
	Age 50-59	α_{15}	-7,775	0,082	-7,938	-7,615
	Age 60-69	α_{16}	-8,271	0,091	-8,451	-8,092
	Age 70-79	α_{17}	-8,510	0,104	-8,715	-8,306
	Age ≥80	α_{18}	-8,214	0,104	-8,420	-8,010
	Female	α_{22}	0,014	0,027	-0,038	0,066
	Income	β_1	-0,090	0,032	-0,152	-0,027
	Log(popdens)	β_2	0,077	0,037	0,006	0,149
	Age ≤17 × sturatio	γ_{11}	-0,018	0,089	-0,193	0,156
	Age 18-29 × sturatio	γ_{12}	0,113	0,083	-0,051	0,276
	Age 30-39 × sturatio	γ_{13}	0,080	0,087	-0,091	0,251
	Age 40-49 × sturatio	γ_{14}	0,023	0,090	-0,154	0,201
	Age 50-59 × sturatio	γ_{15}	0,063	0,094	-0,121	0,246
	Age 60-69 × sturatio	γ_{16}	0,148	0,103	-0,054	0,349
	Age 70-79 × sturatio	γ_{17}	0,144	0,120	-0,092	0,379
	Age ≥80 × sturatio	γ_{18}	0,006	0,126	-0,242	0,254
	Precision SS	τ_v	0,646	0,079	0,504	0,815
	Precision SU	τ_u	2163,268	2434,018	135,045	8620,921
September 16-30	Age ≤17	α_{11}	-7,072	0,055	-7,181	-6,964
	Age 18-29	α_{12}	-6,282	0,054	-6,388	-6,178
	Age 30-39	α_{13}	-6,682	0,056	-6,791	-6,573
	Age 40-49	α_{14}	-6,724	0,056	-6,834	-6,615
	Age 50-59	α_{15}	-6,888	0,056	-7,000	-6,778
	Age 60-69	α_{16}	-7,513	0,064	-7,638	-7,389
	Age 70-79	α_{17}	-7,623	0,071	-7,762	-7,485
	Age ≥80	α_{18}	-7,463	0,073	-7,607	-7,319
	Female	α_{22}	0,021	0,019	-0,017	0,059
	Income	β_1	-0,049	0,022	-0,092	-0,005
	Log(popdens)	β_2	0,066	0,025	0,016	0,116
	Age ≤17 × sturatio	γ_{11}	-0,041	0,063	-0,165	0,083
	Age 18-29 × sturatio	γ_{12}	0,129	0,058	0,016	0,242
	Age 30-39 × sturatio	γ_{13}	0,049	0,062	-0,072	0,171
	Age 40-49 × sturatio	γ_{14}	0,019	0,063	-0,105	0,143

Period	Effect	Parameter	Mean	SD	95% credible interval	
					Lower	Upper
	Age 50-59 × sturatio	γ_{15}	0,013	0,064	-0,113	0,139
	Age 60-69 × sturatio	γ_{16}	-0,028	0,075	-0,175	0,120
	Age 70-79 × sturatio	γ_{17}	-0,063	0,087	-0,233	0,107
	Age ≥80 × sturatio	γ_{18}	0,012	0,087	-0,159	0,182
	Precision SS	τ_v	1,328	0,158	1,044	1,665
	Precision SU	τ_u	2161,438	2352,218	145,389	8407,772
October 1-15	Age ≤17	α_{11}	-6,071	0,037	-6,143	-5,999
	Age 18-29	α_{12}	-5,023	0,035	-5,091	-4,955
	Age 30-39	α_{13}	-5,372	0,036	-5,442	-5,301
	Age 40-49	α_{14}	-5,388	0,036	-5,458	-5,317
	Age 50-59	α_{15}	-5,555	0,036	-5,626	-5,485
	Age 60-69	α_{16}	-6,057	0,039	-6,134	-5,981
	Age 70-79	α_{17}	-6,229	0,043	-6,313	-6,145
	Age ≥80	α_{18}	-5,886	0,043	-5,970	-5,803
	Female	α_{22}	0,055	0,011	0,033	0,078
	Income	β_1	-0,011	0,016	-0,042	0,020
	Log(popdens)	β_2	0,012	0,018	-0,023	0,046
	Age ≤17 × sturatio	γ_{11}	0,046	0,043	-0,040	0,131
	Age 18-29 × sturatio	γ_{12}	0,079	0,040	0,000	0,159
	Age 30-39 × sturatio	γ_{13}	0,025	0,042	-0,058	0,108
	Age 40-49 × sturatio	γ_{14}	0,027	0,042	-0,056	0,111
	Age 50-59 × sturatio	γ_{15}	-0,017	0,043	-0,101	0,068
	Age 60-69 × sturatio	γ_{16}	-0,034	0,047	-0,127	0,058
	Age 70-79 × sturatio	γ_{17}	-0,209	0,055	-0,317	-0,100
	Age ≥80 × sturatio	γ_{18}	-0,265	0,055	-0,374	-0,157
	Precision SS	τ_v	2,349	0,240	1,909	2,855
	Precision SU	τ_u	2168,917	2402,816	138,439	8547,238
October 16-31	Age ≤17	α_{11}	-5,253	0,025	-5,302	-5,204
	Age 18-29	α_{12}	-4,150	0,024	-4,196	-4,104
	Age 30-39	α_{13}	-4,262	0,024	-4,309	-4,216
	Age 40-49	α_{14}	-4,269	0,024	-4,316	-4,223
	Age 50-59	α_{15}	-4,353	0,024	-4,400	-4,307
	Age 60-69	α_{16}	-4,769	0,025	-4,818	-4,720
	Age 70-79	α_{17}	-4,895	0,027	-4,948	-4,843
	Age ≥80	α_{18}	-4,361	0,026	-4,412	-4,310
	Female	α_{22}	0,149	0,007	0,135	0,163
	Income	β_1	-0,029	0,011	-0,051	-0,007
	Log(popdens)	β_2	-0,019	0,012	-0,043	0,005
	Age ≤17 × sturatio	γ_{11}	0,022	0,031	-0,039	0,083
	Age 18-29 × sturatio	γ_{12}	0,029	0,029	-0,027	0,085
	Age 30-39 × sturatio	γ_{13}	0,004	0,029	-0,053	0,062
	Age 40-49 × sturatio	γ_{14}	0,014	0,029	-0,044	0,071
	Age 50-59 × sturatio	γ_{15}	-0,005	0,029	-0,063	0,053
	Age 60-69 × sturatio	γ_{16}	-0,024	0,031	-0,086	0,037
	Age 70-79 × sturatio	γ_{17}	-0,065	0,034	-0,131	0,001
	Age ≥80 × sturatio	γ_{18}	-0,093	0,033	-0,157	-0,029
	Precision SS	τ_v	4,550	0,437	3,747	5,466

Period	Effect	Parameter	Mean	SD	95% credible interval	
					Lower	Upper
	Precision SU	τ_u	2515,108	2573,438	229,561	9345,161
November 1-15	Age ≤17	α_{11}	-6,588	0,033	-6,654	-6,523
	Age 18-29	α_{12}	-5,095	0,028	-5,150	-5,041
	Age 30-39	α_{13}	-5,106	0,028	-5,162	-5,051
	Age 40-49	α_{14}	-5,093	0,028	-5,148	-5,038
	Age 50-59	α_{15}	-5,152	0,028	-5,207	-5,097
	Age 60-69	α_{16}	-5,570	0,030	-5,630	-5,511
	Age 70-79	α_{17}	-5,546	0,032	-5,609	-5,483
	Age ≥80	α_{18}	-4,603	0,029	-4,661	-4,546
	Female	α_{22}	0,243	0,010	0,222	0,263
	Income	β_1	-0,042	0,012	-0,065	-0,018
	Log(popdens)	β_2	-0,013	0,013	-0,039	0,013
	Age ≤17 × sturatio	γ_{11}	0,012	0,041	-0,068	0,093
	Age 18-29 × sturatio	γ_{12}	-0,065	0,033	-0,130	0,000
	Age 30-39 × sturatio	γ_{13}	-0,053	0,034	-0,120	0,013
	Age 40-49 × sturatio	γ_{14}	-0,052	0,034	-0,118	0,015
	Age 50-59 × sturatio	γ_{15}	-0,041	0,034	-0,107	0,026
	Age 60-69 × sturatio	γ_{16}	-0,038	0,037	-0,111	0,035
	Age 70-79 × sturatio	γ_{17}	-0,043	0,040	-0,121	0,035
	Age ≥80 × sturatio	γ_{18}	0,058	0,034	-0,009	0,125
	Precision SS	τ_v	4,384	0,471	3,528	5,381
	Precision SU	τ_u	2191,748	2426,474	142,170	8632,776
November 16-30	Age ≤17	α_{11}	-6,660	0,044	-6,746	-6,575
	Age 18-29	α_{12}	-5,672	0,041	-5,752	-5,592
	Age 30-39	α_{13}	-5,815	0,042	-5,897	-5,733
	Age 40-49	α_{14}	-5,787	0,041	-5,868	-5,706
	Age 50-59	α_{15}	-5,839	0,041	-5,920	-5,758
	Age 60-69	α_{16}	-6,265	0,044	-6,352	-6,178
	Age 70-79	α_{17}	-6,191	0,046	-6,281	-6,100
	Age ≥80	α_{18}	-4,957	0,041	-5,038	-4,876
	Female	α_{22}	0,287	0,014	0,260	0,315
	Income	β_1	-0,044	0,017	-0,078	-0,010
	Log(popdens)	β_2	-0,016	0,020	-0,055	0,022
	Age ≤17 × sturatio	γ_{11}	0,000	0,052	-0,101	0,101
	Age 18-29 × sturatio	γ_{12}	-0,016	0,046	-0,107	0,074
	Age 30-39 × sturatio	γ_{13}	-0,062	0,049	-0,157	0,033
	Age 40-49 × sturatio	γ_{14}	-0,039	0,048	-0,134	0,056
	Age 50-59 × sturatio	γ_{15}	0,007	0,048	-0,086	0,101
	Age 60-69 × sturatio	γ_{16}	-0,028	0,053	-0,132	0,075
	Age 70-79 × sturatio	γ_{17}	-0,133	0,058	-0,246	-0,021
	Age ≥80 × sturatio	γ_{18}	-0,023	0,047	-0,116	0,069
	Precision SS	τ_v	5,357	1,739	2,680	9,450
	Precision SU	τ_u	19,509	5,674	10,849	32,988
December 1-15	Age ≤17	α_{11}	-6,461	0,048	-6,555	-6,366
	Age 18-29	α_{12}	-5,767	0,047	-5,859	-5,675
	Age 30-39	α_{13}	-5,937	0,048	-6,031	-5,843
	Age 40-49	α_{14}	-5,920	0,048	-6,013	-5,826

Period	Effect	Parameter	Mean	SD	95% credible interval	
					Lower	Upper
	Age 50-59	α_{15}	-5,981	0,047	-6,074	-5,888
	Age 60-69	α_{16}	-6,384	0,050	-6,482	-6,286
	Age 70-79	α_{17}	-6,242	0,051	-6,342	-6,141
	Age ≥80	α_{18}	-5,104	0,048	-5,197	-5,010
	Female	α_{22}	0,244	0,014	0,217	0,271
	Income	β_1	-0,037	0,022	-0,080	0,006
	Log(popdens)	β_2	0,020	0,023	-0,026	0,066
	Age ≤17 × sturatio	γ_{11}	-0,017	0,059	-0,132	0,099
	Age 18-29 × sturatio	γ_{12}	-0,057	0,056	-0,168	0,054
	Age 30-39 × sturatio	γ_{13}	-0,004	0,058	-0,117	0,110
	Age 40-49 × sturatio	γ_{14}	-0,020	0,058	-0,134	0,094
	Age 50-59 × sturatio	γ_{15}	-0,001	0,058	-0,114	0,113
	Age 60-69 × sturatio	γ_{16}	0,045	0,061	-0,076	0,165
	Age 70-79 × sturatio	γ_{17}	0,010	0,064	-0,115	0,135
	Age ≥80 × sturatio	γ_{18}	0,094	0,056	-0,016	0,205
	Precision SS	τ_v	1,254	0,363	0,623	2,025
	Precision SU	τ_u	430,169	495,785	80,207	1710,756
December 16-31	Age ≤17	α_{11}	-6,446	0,040	-6,525	-6,368
	Age 18-29	α_{12}	-5,728	0,039	-5,804	-5,652
	Age 30-39	α_{13}	-5,829	0,039	-5,906	-5,752
	Age 40-49	α_{14}	-5,863	0,039	-5,940	-5,786
	Age 50-59	α_{15}	-5,879	0,039	-5,955	-5,803
	Age 60-69	α_{16}	-6,247	0,042	-6,329	-6,165
	Age 70-79	α_{17}	-6,239	0,045	-6,327	-6,151
	Age ≥80	α_{18}	-5,252	0,040	-5,331	-5,173
	Female	α_{22}	0,188	0,014	0,160	0,217
	Income	β_1	-0,036	0,017	-0,069	-0,003
	Log(popdens)	β_2	-0,013	0,018	-0,049	0,022
	Age ≤17 × sturatio	γ_{11}	-0,161	0,051	-0,262	-0,060
	Age 18-29 × sturatio	γ_{12}	-0,043	0,046	-0,133	0,047
	Age 30-39 × sturatio	γ_{13}	-0,116	0,049	-0,212	-0,021
	Age 40-49 × sturatio	γ_{14}	-0,106	0,049	-0,203	-0,010
	Age 50-59 × sturatio	γ_{15}	-0,050	0,048	-0,145	0,044
	Age 60-69 × sturatio	γ_{16}	-0,023	0,052	-0,126	0,079
	Age 70-79 × sturatio	γ_{17}	-0,053	0,057	-0,165	0,059
	Age ≥80 × sturatio	γ_{18}	-0,015	0,049	-0,110	0,081
	Precision SS	τ_v	2,294	0,247	1,846	2,818
	Precision SU	τ_u	1998,749	2314,436	102,610	8131,092

A.2 Parameter estimates for the spatial model based on Equation 2.2 in Wallo-Brux.

Period	Effect	Parameter	Mean	SD	95% credible interval	
					Lower	Upper
June 1-15	Age ≤17	α_{11}	-10,721	0,189	-11,094	-10,352
	Age 18-29	α_{12}	-9,761	0,157	-10,072	-9,456
	Age 30-39	α_{13}	-9,585	0,155	-9,893	-9,284
	Age 40-49	α_{14}	-9,786	0,162	-10,108	-9,471
	Age 50-59	α_{15}	-9,699	0,159	-10,013	-9,391
	Age 60-69	α_{16}	-9,649	0,165	-9,975	-9,329
	Age 70-79	α_{17}	-9,563	0,180	-9,919	-9,212
	Age ≥80	α_{18}	-8,602	0,160	-8,918	-8,292
	Female	α_{22}	0,135	0,079	-0,021	0,290
	Income	β_1	-0,113	0,032	-0,176	-0,051
	Log(popdens)	β_2	0,102	0,049	0,008	0,200
	Age ≤17 × sturatio	γ_{11}	-0,165	0,184	-0,525	0,197
	Age 18-29 × sturatio	γ_{12}	0,023	0,116	-0,204	0,251
	Age 30-39 × sturatio	γ_{13}	0,122	0,108	-0,089	0,334
	Age 40-49 × sturatio	γ_{14}	0,033	0,131	-0,223	0,290
	Age 50-59 × sturatio	γ_{15}	0,200	0,104	-0,003	0,405
	Age 60-69 × sturatio	γ_{16}	0,052	0,130	-0,201	0,307
	Age 70-79 × sturatio	γ_{17}	0,087	0,139	-0,184	0,359
	Age ≥80 × sturatio	γ_{18}	-0,021	0,125	-0,265	0,224
	Precision SS	τ_v	1,266	0,384	0,689	2,185
	Precision SU	τ_u	2162,901	2332,214	149,329	8356,947
June 16-30	Age ≤17	α_{11}	-10,453	0,174	-10,797	-10,115
	Age 18-29	α_{12}	-9,944	0,167	-10,275	-9,620
	Age 30-39	α_{13}	-9,803	0,168	-10,136	-9,476
	Age 40-49	α_{14}	-10,241	0,185	-10,607	-9,880
	Age 50-59	α_{15}	-10,096	0,177	-10,445	-9,751
	Age 60-69	α_{16}	-10,286	0,197	-10,675	-9,901
	Age 70-79	α_{17}	-10,296	0,243	-10,774	-9,822
	Age ≥80	α_{18}	-9,774	0,238	-10,243	-9,308
	Female	α_{22}	0,430	0,090	0,253	0,607
	Income	β_1	-0,083	0,036	-0,154	-0,012
	Log(popdens)	β_2	0,123	0,056	0,015	0,234
	Age ≤17 × sturatio	γ_{11}	0,062	0,134	-0,201	0,326
	Age 18-29 × sturatio	γ_{12}	0,116	0,108	-0,096	0,329
	Age 30-39 × sturatio	γ_{13}	-0,058	0,134	-0,320	0,205
	Age 40-49 × sturatio	γ_{14}	-0,062	0,162	-0,380	0,255
	Age 50-59 × sturatio	γ_{15}	0,123	0,126	-0,124	0,369
	Age 60-69 × sturatio	γ_{16}	-0,019	0,167	-0,346	0,310
	Age 70-79 × sturatio	γ_{17}	-0,480	0,242	-0,955	-0,005
	Age ≥80 × sturatio	γ_{18}	-0,769	0,227	-1,213	-0,324
	Precision SS	τ_v	0,832	0,232	0,471	1,376
	Precision SU	τ_u	2170,125	2313,787	155,576	8315,965

Period	Effect	Parameter	Mean	SD	95% credible interval	
					Lower	Upper
July 1-15	Age ≤17	α_{11}	-10,363	0,187	-10,731	-9,999
	Age 18-29	α_{12}	-9,267	0,162	-9,586	-8,952
	Age 30-39	α_{13}	-9,384	0,166	-9,712	-9,060
	Age 40-49	α_{14}	-9,564	0,171	-9,902	-9,230
	Age 50-59	α_{15}	-9,756	0,178	-10,107	-9,407
	Age 60-69	α_{16}	-10,049	0,203	-10,448	-9,653
	Age 70-79	α_{17}	-10,280	0,280	-10,830	-9,733
	Age ≥80	α_{18}	-9,597	0,243	-10,075	-9,120
	Female	α_{22}	-0,017	0,082	-0,177	0,143
	Income	β_1	-0,105	0,045	-0,192	-0,017
	Log(popdens)	β_2	0,074	0,069	-0,061	0,209
	Age ≤17 × sturatio	γ_{11}	-0,007	0,181	-0,362	0,348
	Age 18-29 × sturatio	γ_{12}	-0,029	0,139	-0,302	0,244
	Age 30-39 × sturatio	γ_{13}	0,286	0,120	0,051	0,523
	Age 40-49 × sturatio	γ_{14}	0,139	0,145	-0,146	0,423
	Age 50-59 × sturatio	γ_{15}	0,203	0,142	-0,075	0,482
	Age 60-69 × sturatio	γ_{16}	0,217	0,161	-0,100	0,534
	Age 70-79 × sturatio	γ_{17}	-0,571	0,296	-1,151	0,010
	Age ≥80 × sturatio	γ_{18}	-0,360	0,256	-0,862	0,142
	Precision SS	τ_v	1,189	0,747	0,326	3,135
	Precision SU	τ_u	2,471	1,181	0,910	5,444
July 16-31	Age ≤17	α_{11}	-9,274	0,103	-9,478	-9,073
	Age 18-29	α_{12}	-8,264	0,091	-8,443	-8,087
	Age 30-39	α_{13}	-8,330	0,093	-8,514	-8,148
	Age 40-49	α_{14}	-8,470	0,095	-8,658	-8,284
	Age 50-59	α_{15}	-8,665	0,099	-8,860	-8,473
	Age 60-69	α_{16}	-9,211	0,118	-9,444	-8,980
	Age 70-79	α_{17}	-9,304	0,140	-9,579	-9,030
	Age ≥80	α_{18}	-8,949	0,144	-9,232	-8,668
	Female	α_{22}	0,048	0,046	-0,042	0,137
	Income	β_1	-0,088	0,027	-0,142	-0,035
	Log(popdens)	β_2	0,146	0,041	0,066	0,228
	Age ≤17 × sturatio	γ_{11}	-0,053	0,103	-0,253	0,149
	Age 18-29 × sturatio	γ_{12}	0,204	0,071	0,065	0,345
	Age 30-39 × sturatio	γ_{13}	0,256	0,074	0,112	0,401
	Age 40-49 × sturatio	γ_{14}	0,111	0,085	-0,054	0,278
	Age 50-59 × sturatio	γ_{15}	0,112	0,089	-0,061	0,287
	Age 60-69 × sturatio	γ_{16}	0,190	0,104	-0,014	0,395
	Age 70-79 × sturatio	γ_{17}	0,060	0,139	-0,213	0,334
	Age ≥80 × sturatio	γ_{18}	-0,076	0,153	-0,375	0,223
	Precision SS	τ_v	0,706	0,124	0,494	0,981
	Precision SU	τ_u	2160,754	2393,383	137,234	8514,079
August 1-15	Age ≤17	α_{11}	-8,684	0,094	-8,870	-8,502
	Age 18-29	α_{12}	-7,810	0,090	-7,988	-7,637
	Age 30-39	α_{13}	-7,957	0,091	-8,139	-7,780
	Age 40-49	α_{14}	-8,036	0,092	-8,219	-7,857
	Age 50-59	α_{15}	-8,201	0,094	-8,387	-8,019

Period	Effect	Parameter	Mean	SD	95% credible interval	
					Lower	Upper
	Age 60-69	α_{16}	-8,686	0,104	-8,892	-8,482
	Age 70-79	α_{17}	-9,026	0,125	-9,273	-8,781
	Age ≥80	α_{18}	-8,580	0,123	-8,824	-8,339
	Female	α_{22}	-0,019	0,033	-0,084	0,046
	Income	β_1	-0,104	0,031	-0,164	-0,043
	Log(popdens)	β_2	0,181	0,046	0,092	0,273
	Age ≤17 × sturatio	γ_{11}	0,194	0,083	0,031	0,358
	Age 18-29 × sturatio	γ_{12}	0,135	0,077	-0,016	0,288
	Age 30-39 × sturatio	γ_{13}	0,134	0,081	-0,024	0,292
	Age 40-49 × sturatio	γ_{14}	0,107	0,083	-0,055	0,270
	Age 50-59 × sturatio	γ_{15}	0,128	0,084	-0,037	0,293
	Age 60-69 × sturatio	γ_{16}	0,198	0,092	0,018	0,378
	Age 70-79 × sturatio	γ_{17}	0,205	0,109	-0,008	0,418
	Age ≥80 × sturatio	γ_{18}	0,101	0,117	-0,128	0,329
	Precision SS	τ_v	5,257	3,575	1,514	14,756
	Precision SU	τ_u	2,612	0,490	1,773	3,695
August 16-31	Age ≤17	α_{11}	-8,733	0,090	-8,910	-8,558
	Age 18-29	α_{12}	-7,864	0,085	-8,034	-7,699
	Age 30-39	α_{13}	-8,104	0,088	-8,278	-7,933
	Age 40-49	α_{14}	-8,259	0,089	-8,436	-8,085
	Age 50-59	α_{15}	-8,402	0,091	-8,583	-8,225
	Age 60-69	α_{16}	-8,915	0,104	-9,119	-8,712
	Age 70-79	α_{17}	-9,081	0,121	-9,320	-8,844
	Age ≥80	α_{18}	-9,042	0,134	-9,305	-8,780
	Female	α_{22}	0,025	0,033	-0,040	0,090
	Income	β_1	-0,127	0,028	-0,181	-0,073
	Log(popdens)	β_2	0,235	0,042	0,154	0,317
	Age ≤17 × sturatio	γ_{11}	0,198	0,078	0,046	0,351
	Age 18-29 × sturatio	γ_{12}	0,247	0,068	0,113	0,382
	Age 30-39 × sturatio	γ_{13}	0,287	0,071	0,148	0,428
	Age 40-49 × sturatio	γ_{14}	0,262	0,075	0,116	0,409
	Age 50-59 × sturatio	γ_{15}	0,236	0,077	0,085	0,387
	Age 60-69 × sturatio	γ_{16}	0,240	0,090	0,064	0,416
	Age 70-79 × sturatio	γ_{17}	0,090	0,118	-0,142	0,321
	Age ≥80 × sturatio	γ_{18}	0,166	0,126	-0,081	0,414
	Precision SS	τ_v	6,008	4,434	1,651	17,842
	Precision SU	τ_u	3,854	0,836	2,441	5,714
September 1-15	Age ≤17	α_{11}	-7,535	0,062	-7,658	-7,414
	Age 18-29	α_{12}	-6,821	0,060	-6,939	-6,704
	Age 30-39	α_{13}	-7,124	0,062	-7,247	-7,002
	Age 40-49	α_{14}	-7,250	0,063	-7,374	-7,126
	Age 50-59	α_{15}	-7,407	0,065	-7,534	-7,280
	Age 60-69	α_{16}	-7,877	0,074	-8,022	-7,732
	Age 70-79	α_{17}	-8,237	0,092	-8,418	-8,056
	Age ≥80	α_{18}	-8,219	0,104	-8,423	-8,015
	Female	α_{22}	0,059	0,025	0,010	0,108
	Income	β_1	-0,058	0,022	-0,101	-0,014

Period	Effect	Parameter	Mean	SD	95% credible interval	
					Lower	Upper
	Log(popdens)	β_2	0,148	0,033	0,083	0,213
	Age ≤17 × sturatio	γ_{11}	0,167	0,061	0,048	0,287
	Age 18-29 × sturatio	γ_{12}	0,239	0,056	0,129	0,349
	Age 30-39 × sturatio	γ_{13}	0,253	0,059	0,137	0,368
	Age 40-49 × sturatio	γ_{14}	0,220	0,061	0,101	0,340
	Age 50-59 × sturatio	γ_{15}	0,141	0,064	0,015	0,266
	Age 60-69 × sturatio	γ_{16}	0,213	0,071	0,074	0,353
	Age 70-79 × sturatio	γ_{17}	0,070	0,096	-0,118	0,258
	Age ≥80 × sturatio	γ_{18}	0,061	0,109	-0,154	0,276
	Precision SS	τ_v	3,366	1,216	1,647	6,365
	Precision SU	τ_u	5,866	1,234	3,782	8,617
September 16-30	Age ≤17	α_{11}	-6,599	0,041	-6,679	-6,519
	Age 18-29	α_{12}	-5,554	0,038	-5,629	-5,479
	Age 30-39	α_{13}	-6,131	0,041	-6,212	-6,051
	Age 40-49	α_{14}	-6,166	0,041	-6,246	-6,085
	Age 50-59	α_{15}	-6,312	0,042	-6,395	-6,230
	Age 60-69	α_{16}	-6,782	0,048	-6,876	-6,688
	Age 70-79	α_{17}	-6,979	0,057	-7,091	-6,868
	Age ≥80	α_{18}	-6,656	0,058	-6,770	-6,543
	Female	α_{22}	0,065	0,016	0,033	0,096
	Income	β_1	-0,032	0,016	-0,063	0,000
	Log(popdens)	β_2	0,101	0,023	0,056	0,146
	Age ≤17 × sturatio	γ_{11}	0,152	0,043	0,068	0,237
	Age 18-29 × sturatio	γ_{12}	0,252	0,039	0,176	0,328
	Age 30-39 × sturatio	γ_{13}	0,167	0,042	0,084	0,250
	Age 40-49 × sturatio	γ_{14}	0,136	0,043	0,051	0,221
	Age 50-59 × sturatio	γ_{15}	0,144	0,044	0,059	0,230
	Age 60-69 × sturatio	γ_{16}	0,119	0,050	0,021	0,216
	Age 70-79 × sturatio	γ_{17}	0,074	0,059	-0,041	0,190
	Age ≥80 × sturatio	γ_{18}	0,042	0,061	-0,078	0,162
	Precision SS	τ_v	4,449	1,174	2,594	7,182
	Precision SU	τ_u	11,942	2,382	7,919	17,258
October 1-15	Age ≤17	α_{11}	-4,869	0,017	-4,902	-4,837
	Age 18-29	α_{12}	-3,841	0,015	-3,870	-3,812
	Age 30-39	α_{13}	-4,233	0,016	-4,265	-4,201
	Age 40-49	α_{14}	-4,267	0,016	-4,299	-4,235
	Age 50-59	α_{15}	-4,391	0,016	-4,424	-4,359
	Age 60-69	α_{16}	-4,880	0,020	-4,919	-4,842
	Age 70-79	α_{17}	-5,094	0,024	-5,142	-5,046
	Age ≥80	α_{18}	-4,863	0,026	-4,915	-4,812
	Female	α_{22}	0,100	0,008	0,084	0,115
	Income	β_1	0,026	0,008	0,009	0,043
	Log(popdens)	β_2	0,005	0,012	-0,018	0,028
	Age ≤17 × sturatio	γ_{11}	0,078	0,022	0,035	0,122
	Age 18-29 × sturatio	γ_{12}	0,119	0,021	0,078	0,159
	Age 30-39 × sturatio	γ_{13}	0,066	0,022	0,023	0,109
	Age 40-49 × sturatio	γ_{14}	0,032	0,022	-0,011	0,076

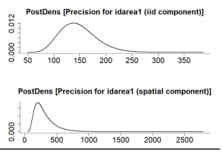
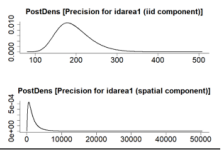
Period	Effect	Parameter	Mean	SD	95% credible interval	
					Lower	Upper
	Age 50-59 × sturatio	γ_{15}	0,061	0,022	0,018	0,105
	Age 60-69 × sturatio	γ_{16}	0,055	0,025	0,007	0,103
	Age 70-79 × sturatio	γ_{17}	-0,045	0,030	-0,103	0,013
	Age ≥80 × sturatio	γ_{18}	-0,183	0,033	-0,247	-0,119
	Precision SS	τ_v	4,031	0,423	3,260	4,923
	Precision SU	τ_u	2255,375	2479,871	155,856	8837,412
October 16-31	Age ≤17	α_{11}	-4,335	0,012	-4,359	-4,310
	Age 18-29	α_{12}	-3,251	0,011	-3,273	-3,231
	Age 30-39	α_{13}	-3,259	0,011	-3,281	-3,238
	Age 40-49	α_{14}	-3,267	0,011	-3,289	-3,246
	Age 50-59	α_{15}	-3,394	0,011	-3,416	-3,373
	Age 60-69	α_{16}	-3,785	0,013	-3,810	-3,760
	Age 70-79	α_{17}	-3,883	0,015	-3,912	-3,854
	Age ≥80	α_{18}	-3,403	0,014	-3,431	-3,375
	Female	α_{22}	0,171	0,005	0,161	0,182
	Income	β_1	0,033	0,006	0,021	0,046
	Log(popdens)	β_2	-0,050	0,009	-0,067	-0,033
	Age ≤17 × sturatio	γ_{11}	0,020	0,017	-0,013	0,054
	Age 18-29 × sturatio	γ_{12}	-0,016	0,016	-0,047	0,015
	Age 30-39 × sturatio	γ_{13}	-0,044	0,016	-0,075	-0,012
	Age 40-49 × sturatio	γ_{14}	-0,036	0,016	-0,067	-0,004
	Age 50-59 × sturatio	γ_{15}	-0,034	0,016	-0,066	-0,002
	Age 60-69 × sturatio	γ_{16}	-0,061	0,017	-0,095	-0,027
	Age 70-79 × sturatio	γ_{17}	-0,080	0,019	-0,117	-0,043
	Age ≥80 × sturatio	γ_{18}	-0,083	0,019	-0,119	-0,046
	Precision SS	τ_v	7,439	0,774	6,035	9,077
	Precision SU	τ_u	2138,571	2829,832	94,953	9527,038
November 1-15	Age ≤17	α_{11}	-6,033	0,028	-6,087	-5,979
	Age 18-29	α_{12}	-4,495	0,022	-4,538	-4,452
	Age 30-39	α_{13}	-4,355	0,022	-4,398	-4,313
	Age 40-49	α_{14}	-4,326	0,022	-4,369	-4,284
	Age 50-59	α_{15}	-4,361	0,022	-4,404	-4,318
	Age 60-69	α_{16}	-4,629	0,023	-4,675	-4,584
	Age 70-79	α_{17}	-4,587	0,025	-4,636	-4,538
	Age ≥80	α_{18}	-3,911	0,024	-3,958	-3,865
	Female	α_{22}	0,192	0,009	0,175	0,210
	Income	β_1	0,024	0,010	0,005	0,044
	Log(popdens)	β_2	-0,049	0,013	-0,074	-0,024
	Age ≤17 × sturatio	γ_{11}	-0,024	0,030	-0,082	0,034
	Age 18-29 × sturatio	γ_{12}	-0,124	0,025	-0,172	-0,075
	Age 30-39 × sturatio	γ_{13}	-0,111	0,025	-0,159	-0,062
	Age 40-49 × sturatio	γ_{14}	-0,117	0,025	-0,166	-0,068
	Age 50-59 × sturatio	γ_{15}	-0,084	0,025	-0,132	-0,036
	Age 60-69 × sturatio	γ_{16}	-0,127	0,026	-0,179	-0,075
	Age 70-79 × sturatio	γ_{17}	-0,112	0,028	-0,166	-0,057
	Age ≥80 × sturatio	γ_{18}	-0,054	0,026	-0,104	-0,003
	Precision SS	τ_v	10,548	2,377	6,624	15,932

Period	Effect	Parameter	Mean	SD	95% credible interval	
					Lower	Upper
	Precision SU	τ_u	39,132	8,384	25,237	58,090
November 16-30	Age ≤17	α_{11}	-6,496	0,032	-6,559	-6,434
	Age 18-29	α_{12}	-5,502	0,027	-5,554	-5,450
	Age 30-39	α_{13}	-5,497	0,027	-5,551	-5,444
	Age 40-49	α_{14}	-5,496	0,027	-5,550	-5,443
	Age 50-59	α_{15}	-5,509	0,027	-5,562	-5,455
	Age 60-69	α_{16}	-5,681	0,030	-5,740	-5,622
	Age 70-79	α_{17}	-5,571	0,034	-5,637	-5,506
	Age ≥80	α_{18}	-4,577	0,028	-4,632	-4,522
	Female	α_{22}	0,193	0,015	0,164	0,221
	Income	β_1	0,006	0,011	-0,015	0,027
	Log(popdens)	β_2	-0,052	0,015	-0,081	-0,023
	Age ≤17 × sturatio	γ_{11}	-0,085	0,037	-0,157	-0,013
	Age 18-29 × sturatio	γ_{12}	-0,137	0,031	-0,197	-0,077
	Age 30-39 × sturatio	γ_{13}	-0,099	0,032	-0,161	-0,037
	Age 40-49 × sturatio	γ_{14}	-0,119	0,032	-0,182	-0,056
	Age 50-59 × sturatio	γ_{15}	-0,122	0,032	-0,185	-0,059
	Age 60-69 × sturatio	γ_{16}	-0,113	0,035	-0,180	-0,045
	Age 70-79 × sturatio	γ_{17}	-0,188	0,039	-0,264	-0,112
	Age ≥80 × sturatio	γ_{18}	-0,071	0,031	-0,133	-0,010
	Precision SS	τ_v	3,321	0,416	2,578	4,215
	Precision SU	τ_u	1817,225	2185,313	72,397	7594,609
December 1-15	Age ≤17	α_{11}	-6,447	0,041	-6,528	-6,367
	Age 18-29	α_{12}	-5,879	0,039	-5,957	-5,803
	Age 30-39	α_{13}	-5,792	0,040	-5,871	-5,715
	Age 40-49	α_{14}	-5,860	0,040	-5,939	-5,782
	Age 50-59	α_{15}	-5,962	0,041	-6,042	-5,883
	Age 60-69	α_{16}	-6,088	0,043	-6,174	-6,004
	Age 70-79	α_{17}	-5,886	0,046	-5,976	-5,797
	Age ≥80	α_{18}	-4,916	0,041	-4,997	-4,836
	Female	α_{22}	0,153	0,017	0,119	0,187
	Income	β_1	-0,004	0,017	-0,036	0,029
	Log(popdens)	β_2	-0,071	0,022	-0,114	-0,027
	Age ≤17 × sturatio	γ_{11}	-0,045	0,047	-0,136	0,047
	Age 18-29 × sturatio	γ_{12}	-0,035	0,043	-0,120	0,051
	Age 30-39 × sturatio	γ_{13}	-0,063	0,045	-0,150	0,026
	Age 40-49 × sturatio	γ_{14}	-0,049	0,045	-0,138	0,040
	Age 50-59 × sturatio	γ_{15}	-0,050	0,046	-0,139	0,040
	Age 60-69 × sturatio	γ_{16}	-0,052	0,048	-0,146	0,043
	Age 70-79 × sturatio	γ_{17}	-0,039	0,050	-0,137	0,059
	Age ≥80 × sturatio	γ_{18}	-0,110	0,046	-0,200	-0,019
	Precision SS	τ_v	3,578	1,138	1,824	6,257
	Precision SU	τ_u	13,024	3,852	7,143	22,168
December 16-31	Age ≤17	α_{11}	-6,801	0,043	-6,885	-6,718
	Age 18-29	α_{12}	-5,988	0,039	-6,065	-5,913
	Age 30-39	α_{13}	-5,944	0,039	-6,021	-5,867
	Age 40-49	α_{14}	-5,945	0,039	-6,022	-5,869

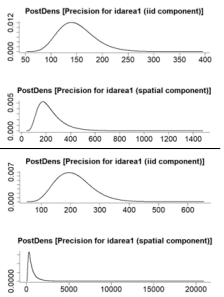
Period	Effect	Parameter	Mean	SD	95% credible interval Lower	Upper
	Age 50-59	α_{15}	-6,058	0,040	-6,137	-5,980
	Age 60-69	α_{16}	-6,218	0,043	-6,303	-6,134
	Age 70-79	α_{17}	-6,151	0,047	-6,244	-6,059
	Age ≥80	α_{18}	-5,257	0,042	-5,341	-5,175
	Female	α_{22}	0,161	0,019	0,124	0,197
	Income	β_1	0,030	0,015	0,001	0,058
	Log(popdens)	β_2	-0,012	0,021	-0,053	0,029
	Age ≤17 × sturatio	γ_{11}	-0,031	0,044	-0,117	0,054
	Age 18-29 × sturatio	γ_{12}	-0,083	0,040	-0,161	-0,005
	Age 30-39 × sturatio	γ_{13}	-0,058	0,041	-0,138	0,021
	Age 40-49 × sturatio	γ_{14}	-0,094	0,041	-0,176	-0,013
	Age 50-59 × sturatio	γ_{15}	-0,076	0,042	-0,157	0,006
	Age 60-69 × sturatio	γ_{16}	-0,129	0,045	-0,218	-0,040
	Age 70-79 × sturatio	γ_{17}	-0,080	0,048	-0,174	0,013
	Age ≥80 × sturatio	γ_{18}	-0,207	0,045	-0,295	-0,119
	Precision SS	τ_v	4,340	1,497	2,069	7,885
	Precision SU	τ_u	17,538	6,142	8,679	32,538

A.3 Prior sensitivity analysis for the spatial model based on Equation 2.5.

Default prior

Period	Effect	Parameter	Estimate	95% credible interval Lower	Upper	Posterior density of the random effect
January 1 – June 30	Intercept	α	-0.194	-0.244	-0.143	
	Age 65–84	β_{12}	0.079	0.053	0.106	
	Age $\geq$85	β_{13}	0.082	0.052	0.111	
	Female	β_{22}	-0.02	-0.036	-0.004	
	Log(covir)	β_3	0.036	0.029	0.036	
	Precision SS	τ_ν	310.53	106.88	769.02	
	Precision SU	τ_υ	147.51	88.2	227.62	
July 1 – December 31	Intercept	α	-0.972	-1.12	-0.826	
	Age 65–84	β_{12}	0.163	0.136	0.19	
	Age $\geq$85	β_{13}	0.127	0.1	0.155	
	Female	β_{22}	-0.053	-0.07	-0.037	
	Log(covir)	β_3	0.125	0.108	0.142	
	Precision SS	τ_ν	1909.05	205.11	7214.09	
	Precision SU	τ_υ	195.46	126.32	296.65	

PC prior

Period	Effect	Parameter	Estimate	95% credible interval Lower	Upper	Posterior density of the random effect
January 1 – June 30	Intercept	α	-0.192	-0.242	-0.142	
	Age 65–84	β_{12}	0.079	0.053	0.106	
	Age $\geq$85	β_{13}	0.082	0.053	0.111	
	Female	β_{22}	-0.02	-0.036	-0.004	
	Log(covir)	β_3	0.036	0.028	0.043	
	Precision SS	τ_ν	235.22	95.49	511.23	
	Precision SU	τ_υ	150.48	91.18	231.87	
July 1 – December 31	Intercept	α	-0.975	-1.128	-0.822	
	Age 65–84	β_{12}	0.163	0.136	0.19	
	Age $\geq$85	β_{13}	0.127	0.099	0.155	
	Female	β_{22}	-0.053	-0.07	-0.037	
	Log(covir)	β_3	0.125	0.108	0.143	
	Precision SS	τ_ν	732.75	150.82	2665.61	
	Precision SU	τ_υ	212.49	108.96	355.68	

APPENDIX B

APPENDIX FOR CHAPTER 3

B.1 ICD version 11 code translation.

ICD–11 code	Description	New group in the analysis
A00–B99	Certain infectious and parasitic diseases	Infectious diseases
C00–D48	Neoplasms	Neoplasms
D50–D89	Diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism	Other
E00–E90	Endocrine, nutritional and metabolic diseases	Other
F00–F99	Mental and behavioral disorders	Mental and behavioral disorders
G00–G99	Diseases of the nervous system	Other
H00–H59	Diseases of the eye and adnexa	Other
H60–H95	Diseases of the ear and mastoid process	Other
I00–I99	Diseases of the circulatory system	Other
J00–J997	Diseases of the respiratory system	Other
K00–K93	Diseases of the digestive system	Other
L00–L99	Diseases of the skin and subcutaneous tissue	Other
M00–M99	Diseases of the musculoskeletal system and connective tissue	Other
N00–N99	Diseases of the genitourinary system	Other
O00–O99	Pregnancy, childbirth and the puerperium	Other
P00–P96	Certain conditions originating in the perinatal period	Other
Q00–Q99	Congenital malformations, deformations and chromosomal abnormalities	Other
R00–R99	Symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified	Other
U071–U072	COVID-19	COVID-19
U00–U49	Codes for special purposes (Excl. COVID-19)	Other
V01–Y98	External causes of morbidity and mortality	External causes

B.2 Summary of the model selection process based on Equation 3.1 using the full dataset.

Outcome	Effect	Minimum n	Minimum n	Total fitted models	Minimum AIC	Maximum AIC
Confirmed cases ($\log Y_{ij1}$)	$\text{vfull}_{i-l,j}$ travel_{ij} pos_{ij} SI_i	4	10	1450	965.658	1471.538
	$\frac{10}{\text{vfull}_{i-l,j}}$ travel_{ij} pos_{ij} SI_i	4	10	1450	1318.882	1556.271
Hospitalizations ($\log Y_{ij2}$)	$\text{vfull}_{i-l,j}$ travel_{ij} pos_{ij} SI_i	4	10	1450	1061.516	1455.623
	$\frac{10}{\text{vfull}_{i-l,j}}$ travel_{ij} pos_{ij} SI_i	4	10	1450	1260.645	1538.833
	$\text{vfull}_{i-l,j}$ travel_{ij} pos_{ij} SI_i $\log \text{IC7}_{ij}$	5	15	25350	432.708	1455.623
	$\frac{10}{\text{vfull}_{i-l,j}}$ travel_{ij} pos_{ij} SI_i $\log \text{IC7}_{ij}$	5	15	25350	451.718	1538.833

B.3 Parameter estimates for model in Equation 3.4.

Effect	Parameter	Estimate	SE	p-value	σ_j^2	σ_{res}^2
Intercept	β_{01}	0.683	0.26	0.009	0.013	0.273
$\text{vfull}_{i-l,j}$	β_{11}	0.122	0.008	<0.001		
travel_{ij}	β_{21}	0.018	0.04	0.657		
pos_{ij}	β_{31}	-0.017	0.018	0.358		
SI_i	β_{41}	0.068	0.004	<0.001		
$\text{vfull}_{i-l,j} \times \text{SI}_i$	β_{51}	-0.002	0.0002	<0.001		
travel $\times$ pos_{ij}	β_{61}	0.028	0.009	0.002		

B.4 Parameter estimates for model in Equation 3.5.

Effect	Parameter	Estimate	SE	p-value	σ_j^2	σ_{res}^2
Intercept	β_{02}	-2.655	0.301	<0.001	0.186	0.309
vfull$_{i-l,j}$	β_{12}	0.096	0.012	<0.001		
travel$_{ij}$	β_{22}	0.593	0.129	<0.001		
pos$_{ij}$	β_{32}	-0.027	0.015	0.063		
SI$_i$	β_{42}	0.078	0.005	<0.001		
vfull$_{i-l,j}$ × SI$_i$	β_{52}	-0.002	0.0003	<0.001		
travel$_{ij}$ × SI$_i$	β_{62}	-0.016	0.003	<0.001		

B.5 Parameter estimates for model in Equation 3.6.

Effect	Parameter	Estimate	SE	p-value	σ_j^2	σ_{res}^2
Intercept	β_{02}	-2.209	0.232	<0.001	0.167	0.1
vfull$_{i-l,j}$	β_{12}	-0.011	0.001	<0.001		
travel$_{ij}$	β_{22}	-0.492	0.052	<0.001		
pos$_{ij}$	β_{32}	-0.149	0.046	0.001		
SI$_i$	β_{42}	0.01	0.003	<0.001		
log IC7$_{ij}$	β_{52}	0.796	0.024	<0.001		
log IC7 × pos$_{ij}$	β_{62}	0.019	0.008	0.014		
vfull$_{i-l,j}$ ×travel$_{ij}$	β_{72}	0.007	0.001	<0.001		

B.6 Summary of the model selection process based on Equation 3.1 using data from week 20.

Outcome	Effect	Minimum AIC	Maximum AIC
Confirmed cases ($\log Y_{ij1}$)	$\text{vfull}_{i-l,j}$ travel_{ij} pos_{ij} SI_i	621.231	834.721
	$\frac{10}{\text{vfull}_{i-l,j}}$ travel_{ij} pos_{ij} SI_i	711.745	875.875
Hospitalizations ($\log Y_{ij2}$)	$\text{vfull}_{i-l,j}$ travel_{ij} pos_{ij} SI_i	652.114	823.483
	$\frac{10}{\text{vfull}_{i-l,j}}$ travel_{ij} pos_{ij} SI_i	711.424	870.588
	$\text{vfull}_{i-l,j}$ travel_{ij} pos_{ij} SI_i $\log \text{IC7}_{ij}$	318.395	823.483
	$\frac{10}{\text{vfull}_{i-l,j}}$ travel_{ij} pos_{ij} SI_i $\log \text{IC7}_{ij}$	312.963	870.588

B.7 Parameter estimates for models in Table 3.1.

Outcome	Effect	Parameter	Estimate	SE	p-value	σ_j^2	σ_{res}^2
Confirmed cases ($\log Y_{ij1}$)	Intercept	β_{01}	3.48	0.454	<0.001	0.027	0.275
	$\text{vfull}_{i-l,j}$	β_{11}	0.04	0.003	<0.001		
	travel_{ij}	β_{21}	0.128	0.032	<0.001		
	pos_{ij}	β_{31}	0.024	0.016	0.141		
	SI_i	β_{41}	0.003	0.007	0.777		
Hospitalizations ($\log Y_{ij2}$)	Intercept	β_{02}	-1.352	0.607	0.027	0.172	0.268
	$\text{vfull}_{i-l,j}$	β_{12}	0.068	0.005	<0.001		
	travel_{ij}	β_{22}	-0.061	0.037	0.106		
	pos_{ij}	β_{32}	1.797	0.232	<0.001		
	SI_i	β_{42}	0.047	0.012	<0.001		
	$\text{vfull}_{i-l,j} \times \text{pos}_{ij}$	β_{52}	-0.012	0.001	<0.001		
	$\text{pos}_{ij} \times \text{SI}_i$	β_{62}	-0.038	0.005	<0.001		
	Intercept	β_{02}	-3.128	0.333	<0.001	0.16	0.096
	$\frac{10}{\text{vfull}_{i-l,j}}$	β_{12}	0.003	0.001	<0.001		
	travel_{ij}	β_{22}	-0.496	0.102	0.106		
	pos_{ij}	β_{32}	0.345	0.088	<0.001		
	SI_i	β_{42}	0.014	0.005	0.009		
	$\log \text{IC7}_{ij}$	β_{52}	0.816	0.036	<0.001		
	$\log \text{IC7} \times \text{pos}_{ij}$	β_{62}	-0.055	0.014	<0.001		
	$\log \text{IC7} \times \text{travel}_{ij}$	β_{72}	0.07	0.018	<0.001		
	$\frac{10}{\text{vfull}_{i-l,j}} \times \text{pos}_{ij}$	β_{82}	-0.002	0.0003	<0.001		

B.8 Summary of the model selection process based on Equation 3.2.

Age groups	Model	Effect	Parameter	p-value	AIC
All	Full	Female	β_1	< 0.001	-2061.9
		Age	β_{2k}	< 0.001	
		Heatwave	β_3	0.001	
		Female × Age	β_{4k}	< 0.001	
		Yearly sine wave	α_1	< 0.001	
		Half-yearly sine wave	α_2	< 0.001	
		Quarterly sine wave	α_3	0.934	
		Yearly cosine wave	γ_1	< 0.001	
		Half-yearly cosine wave	γ_2	0.015	
		Quarterly cosine wave	γ_3	0.719	

Age groups	Model	Effect	Parameter	p-value	AIC
	Reduced 1	Female	β_1	< 0.001	-2073.4
		Age	β_{2k}	< 0.001	
		Heatwave	β_3	0.001	
		Female × Age	β_{4k}	< 0.001	
		Yearly sine wave	α_1	< 0.001	
		Half-yearly sine wave	α_2	< 0.001	
		Yearly cosine wave	γ_1	< 0.001	
		Half-yearly cosine wave	γ_2	0.015	
		Quarterly cosine wave	γ_3	0.719	
	Reduced 2	Female	β_1	< 0.001	-2084.7
		Age	β_{2k}	< 0.001	
		Heatwave	β_3	0.001	
		Female × Age	β_{4k}	< 0.001	
		Yearly sine wave	α_1	< 0.001	
		Half-yearly sine wave	α_2	< 0.001	
		Yearly cosine wave	γ_1	< 0.001	
		Half-yearly cosine wave	γ_2	0.015	
Younger (0-24, 25-44)	Full	Female	β_1	< 0.001	1407.5
		Age	β_{2k}	< 0.001	
		Heatwave	β_3	0.054	
		Female × Age	β_{4k}	< 0.001	
		Yearly sine wave	α_1	0.012	
		Half-yearly sine wave	α_2	0.087	
		Quarterly sine wave	α_3	0.818	
		Yearly cosine wave	γ_1	0.135	
		Half-yearly cosine wave	γ_2	0.52	
		Quarterly cosine wave	γ_3	0.621	
	Reduced 1	Female	β_1	< 0.001	1398
		Age	β_{2k}	< 0.001	
		Heatwave	β_3	0.055	
		Female × Age	β_{4k}	< 0.001	
		Yearly sine wave	α_1	0.012	
		Half-yearly sine wave	α_2	0.088	
		Yearly cosine wave	γ_1	0.136	
		Half-yearly cosine wave	γ_2	0.52	
		Quarterly cosine wave	γ_3	0.621	
	Reduced 2	Female	β_1	< 0.001	1388.8
		Age	β_{2k}	< 0.001	
		Heatwave	β_3	0.057	
		Female × Age	β_{4k}	< 0.001	
		Yearly sine wave	α_1	0.012	
		Half-yearly sine wave	α_2	0.086	
		Yearly cosine wave	γ_1	0.136	
		Half-yearly cosine wave	γ_2	0.518	
	Reduced 3	Female	β_1	< 0.001	1379.75
		Age	β_{2k}	< 0.001	
		Heatwave	β_3	0.049	
		Female × Age	β_{4k}	< 0.001	
		Yearly sine wave	α_1	0.012	
		Half-yearly sine wave	α_2	0.087	
		Yearly cosine wave	γ_1	0.134	
	Reduced 4	Female	β_1	< 0.001	1372.5
		Age	β_{2k}	< 0.001	
		Heatwave	β_3	0.078	
		Female × Age	β_{4k}	< 0.001	
		Yearly sine wave	α_1	0.013	
		Half-yearly sine wave	α_2	0.084	

Age groups	Model	Effect	Parameter	p-value	AIC
	Reduced 5	Female	β_1	< 0.001	1366.1
		Age	β_{2k}	< 0.001	
		Heatwave	β_3	0.055	
		Female × Age	β_{4k}	< 0.001	
		Yearly sine wave	α_1	0.012	
	Reduced 6	Female	β_1	< 0.001	1363.9
		Age	β_{2k}	< 0.001	
		Female × Age	β_{3k}	< 0.001	
		Yearly sine wave	α_1	0.017	
	Reduced 7	Female	β_1	< 0.001	1360.1
		Age	β_{2k}	< 0.001	
		Female × Age	β_{3k}	< 0.001	
Older (45-64, 65-74, 75-84, 85+)	Full	Female	β_1	< 0.001	-7993.4
		Age	β_{2k}	< 0.001	
		Heatwave	β_3	< 0.001	
		Female × Age	β_{4k}	< 0.001	
		Yearly sine wave	α_1	< 0.001	
		Half-yearly sine wave	α_2	< 0.001	
		Quarterly sine wave	α_3	0.453	
		Yearly cosine wave	γ_1	< 0.001	
		Half-yearly cosine wave	γ_2	< 0.001	
		Quarterly cosine wave	γ_3	0.828	
	Reduced 1	Female	β_1	< 0.001	-8005.9
		Age	β_{2k}	< 0.001	
		Heatwave	β_3	< 0.001	
		Female × Age	β_{4k}	< 0.001	
		Yearly sine wave	α_1	< 0.001	
		Half-yearly sine wave	α_2	< 0.001	
		Quarterly sine wave	α_3	0.453	
		Yearly cosine wave	γ_1	< 0.001	
		Half-yearly cosine wave	γ_2	< 0.001	
	Reduced 2	Female	β_1	< 0.001	-8017.9
		Age	β_{2k}	< 0.001	
		Heatwave	β_3	< 0.001	
		Female × Age	β_{4k}	< 0.001	
		Yearly sine wave	α_1	< 0.001	
		Half-yearly sine wave	α_2	< 0.001	
		Yearly cosine wave	γ_1	< 0.001	
		Half-yearly cosine wave	γ_2	< 0.001	

B.9 Parameter estimates for models in Table 3.2.

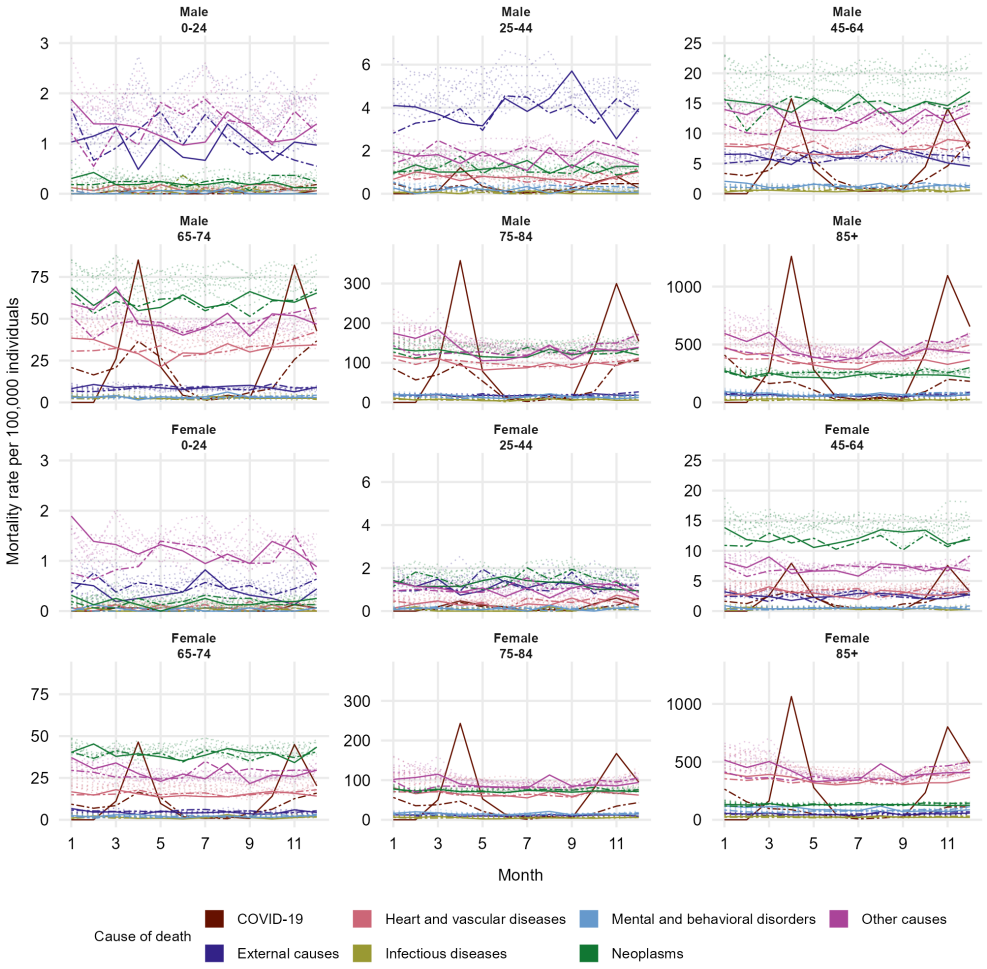
Younger age groups (0–24, 25–44)

Effect	Parameter	Estimate	SE	p-value	σ_j^2	σ_{res}^2
Intercept	β_0	-0.254	0.032	$<$0.001	0.009	0.103
Female	β_1	-0.537	0.019	$<$0.001		
Age 25–44	β_{22}	0.97	0.019	$<$0.001		
Female $\times$ Age 25–44	β_{32}	-0.127	0.027	$<$0.001		

Older age groups (45–64, 65–74, 75–84, 85+)

Effect	Parameter	Estimate	SE	p-value	σ_j^2	σ_{res}^2
Intercept	β_0	2.439	0.016	$<$0.001	0.003	0.01
Female	β_1	-0.516	0.006	$<$0.001		
Age 65–74	β_{22}	1.256	0.006	$<$0.001		
Age 75–84	β_{23}	2.243	0.006	$<$0.001		
Age 85+	β_{24}	3.31	0.006	$<$0.001		
Heatwave	β_3	0.05	0.012	$<$0.001		
Female $\times$ Age 65–74	β_{42}	-0.075	0.008	$<$0.001		
Female $\times$ Age 75–84	β_{43}	0.068	0.008	$<$0.001		
Female $\times$ Age 85+	β_{44}	0.33	0.008	$<$0.001		
Yearly sine wave	α_1	0.07	0.002	$<$0.001		
Half-yearly sine wave	α_2	0.082	0.002	$<$0.001		
Yearly cosine wave	γ_1	0.027	0.002	$<$0.001		
Half-yearly cosine wave	γ_2	0.01	0.002	$<$0.001		

B.10 Monthly cause-specific mortality rate with seven groups.



B.11 Summary of the model selection process based on Equation 3.3.

Cause of death	Model	Effect	Parameter	p-value	AIC
COVID-19	Full	Female	β_1	0.422	811
		Age	β_{2k}	< 0.001	
		SI	β_3	< 0.001	
		Female × Age	β_{4k}	0.994	
		Female × SI	β_5	0.907	
		Age × SI	β_{6k}	0.214	
		Female × SI	β_{7k}	0.998	
	Reduced 1	Female	β_1	0.426	776.3
		Age	β_{2k}	< 0.001	
		SI	β_3	< 0.001	
		Female × Age	β_{4k}	0.779	
		Female × SI	β_5	0.942	
		Age × SI	β_{6k}	0.204	
	Reduced 2	Female	β_1	0.377	741.6
		Age	β_{2k}	< 0.001	
		SI	β_3	< 0.001	
		Female × Age	β_{4k}	0.843	
		Female × SI	β_5	0.924	
	Reduced 3	Female	β_1	< 0.001	767.1
		Age	β_{2k}	< 0.001	
		SI	β_3	< 0.001	
		Female × Age	β_{4k}	0.777	
		Age × SI	β_{5k}	0.201	
	Reduced 4	Female	β_1	0.379	769.5
		Age	β_{2k}	< 0.001	
		SI	β_3	< 0.001	
		Female × SI	β_4	0.93	
		Age × SI	β_{5k}	0.228	
	Reduced 5	Female	β_1	0.001	732.4
		Age	β_{2k}	< 0.001	
		SI	β_3	< 0.001	
		Female × Age	β_{4k}	0.841	
	Reduced 6	Female	β_1	0.35	734.3
		Age	β_{2k}	< 0.001	
		SI	β_3	< 0.001	
		Female × SI	β_4	0.909	
	Reduced 7	Female	β_1	< 0.001	760.3
		Age	β_{2k}	< 0.001	
		SI	β_3	< 0.001	
		Age × SI	β_{4k}	0.225	
	Reduced 8	Female	β_1	< 0.001	725.1
		Age	β_{2k}	< 0.001	
		SI	β_3	< 0.001	

Cause of death	Model	Effect	Parameter	p-value	AIC
External causes	Full	Female	β_1	< 0.001	165.7
		Age	β_{2k}	< 0.001	
		SI	β_3	< 0.001	
		Female × Age	β_{4k}	0.01	
		Female × SI	β_5	0.174	
		Age × SI	β_{6k}	0.985	
		Female × SI	β_{7k}	0.857	
	Reduced 1	Female	β_1	< 0.001	112.3
		Age	β_{2k}	< 0.001	
		SI	β_3	< 0.001	
		Female × Age	β_{4k}	< 0.001	
		Female × SI	β_5	0.171	
		Age × SI	β_{6k}	0.985	
	Reduced 2	Female	β_1	< 0.001	50.6
		Age	β_{2k}	< 0.001	
		SI	β_3	< 0.001	
		Female × Age	β_{4k}	< 0.001	
		Female × SI	β_5	0.168	
	Reduced 3	Female	β_1	< 0.001	100.9
		Age	β_{2k}	< 0.001	
		SI	β_3	< 0.001	
		Female × Age	β_{4k}	< 0.001	
		Age × SI	β_{5k}	0.985	
	Reduced 4	Female	β_1	< 0.001	201.3
		Age	β_{2k}	< 0.001	
		SI	β_3	< 0.001	
		Female × SI	β_4	0.264	
		Age × SI	β_{5k}	0.994	
	Reduced 5	Female	β_1	< 0.001	39.2
		Age	β_{2k}	< 0.001	
		SI	β_3	< 0.001	
		Female × Age	β_{4k}	< 0.001	
	Reduced 6	Female	β_1	< 0.001	141.4
		Age	β_{2k}	< 0.001	
		SI	β_3	< 0.001	
		Female × SI	β_4	0.26	
	Reduced 7	Female	β_1	< 0.001	189.7
		Age	β_{2k}	< 0.001	
		SI	β_3	< 0.001	
		Age × SI	β_{4k}	0.994	
	Reduced 8	Female	β_1	< 0.001	129.8
		Age	β_{2k}	< 0.001	
		SI	β_3	< 0.001	
Heart and vascular diseases	Full	Female	β_1	< 0.001	261.4
		Age	β_{2k}	< 0.001	
		SI	β_3	0.224	
		Female × Age	β_{4k}	0.004	
		Female × SI	β_5	0.224	
		Age × SI	β_{6k}	0.5	
		Female × SI	β_{7k}	0.962	
	Reduced 1	Female	β_1	< 0.001	209.1
		Age	β_{2k}	< 0.001	
		SI	β_3	0.236	
		Female × Age	β_{4k}	< 0.001	
		Female × SI	β_5	0.186	
		Age × SI	β_{6k}	0.482	
	Reduced 2	Female	β_1	< 0.001	153.8
		Age	β_{2k}	< 0.001	
		SI	β_3	0.207	
		Female × Age	β_{4k}	< 0.001	

Cause of death	Model	Effect	Parameter	p-value	AIC
		Female × SI	β_5	0.194	
	Reduced 3	Female	β_1	< 0.001	198.4
		Age	β_{2k}	< 0.001	
		SI	β_3	0.216	
		Female × Age	β_{4k}	< 0.001	
		Age × SI	β_{5k}	0.492	
	Reduced 4	Female	β_1	< 0.001	281.9
		Age	β_{2k}	< 0.001	
		SI	β_3	0.34	
		Female × SI	β_4	0.302	
		Age × SI	β_{5k}	0.666	
	Reduced 5	Female	β_1	< 0.001	142.8
		Age	β_{2k}	< 0.001	
		SI	β_3	0.195	
		Female × Age	β_{4k}	< 0.001	
	Reduced 6	Female	β_1	< 0.001	226.8
		Age	β_{2k}	< 0.001	
		SI	β_3	0.302	
		Female × SI	β_4	0.314	
	Reduced 7	Female	β_1	< 0.001	270.5
		Age	β_{2k}	< 0.001	
		SI	β_3	0.32	
		Age × SI	β_{4k}	0.674	
	Reduced 8	Female	β_1	< 0.001	215.4
		Age	β_{2k}	< 0.001	
		SI	β_3	0.289	
Infectious diseases	Full	Female	β_1	0.048	349.2
		Age	β_{2k}	< 0.001	
		SI	β_3	0.095	
		Female × Age	β_{4k}	0.039	
		Female × SI	β_5	0.64	
		Age × SI	β_{6k}	0.165	
		Female × SI	β_{7k}	0.356	
	Reduced 1	Female	β_1	0.229	306.1
		Age	β_{2k}	< 0.001	
		SI	β_3	0.118	
		Female × Age	β_{4k}	< 0.001	
		Female × SI	β_5	0.369	
		Age × SI	β_{6k}	0.176	
	Reduced 2	Female	β_1	0.292	258.3
		Age	β_{2k}	< 0.001	
		SI	β_3	0.188	
		Female × Age	β_{4k}	< 0.001	
		Female × SI	β_5	0.382	
	Reduced 3	Female	β_1	< 0.001	294.8
		Age	β_{2k}	< 0.001	
		SI	β_3	0.114	
		Female × Age	β_{4k}	< 0.001	
		Age × SI	β_{5k}	0.177	
	Reduced 4	Female	β_1	0.372	330
		Age	β_{2k}	< 0.001	
		SI	β_3	0.264	
		Female × SI	β_4	0.263	
		Age × SI	β_{5k}	0.287	
	Reduced 5	Female	β_1	< 0.001	247
		Age	β_{2k}	< 0.001	
		SI	β_3	0.185	
		Female × Age	β_{4k}	< 0.001	
	Reduced 6	Female	β_1	0.39	281.3
		Age	β_{2k}	< 0.001	

Cause of death	Model	Effect	Parameter	p-value	AIC
		SI	β_3	0.262	
		Female × SI	β_4	0.265	
	Reduced 7	Female	β_1	< 0.001	319.4
		Age	β_{2k}	< 0.001	
		SI	β_3	0.249	
		Age × SI	β_{4k}	0.288	
	Reduced 8	Female	β_1	< 0.001	270.7
		Age	β_{2k}	< 0.001	
		SI	β_3	0.256	
Mental and behavioral disorders	Full	Female	β_1	0.054	311.3
		Age	β_{2k}	< 0.001	
		SI	β_3	< 0.001	
		Female × Age	β_{4k}	0.018	
		Female × SI	β_5	0.945	
		Age × SI	β_{6k}	0.737	
		Female × SI	β_{7k}	0.808	
	Reduced 1	Female	β_1	0.001	263.4
		Age	β_{2k}	< 0.001	
		SI	β_3	< 0.001	
		Female × Age	β_{4k}	< 0.001	
		Female × SI	β_5	0.533	
		Age × SI	β_{6k}	0.652	
	Reduced 2	Female	β_1	0.002	209.3
		Age	β_{2k}	< 0.001	
		SI	β_3	< 0.001	
		Female × Age	β_{4k}	< 0.001	
		Female × SI	β_5	0.561	
	Reduced 3	Female	β_1	< 0.001	251.4
		Age	β_{2k}	< 0.001	
		SI	β_3	< 0.001	
		Female × Age	β_{4k}	< 0.001	
		Age × SI	β_{5k}	0.659	
	Reduced 4	Female	β_1	0.004	347.1
		Age	β_{2k}	< 0.001	
		SI	β_3	< 0.001	
		Female × SI	β_4	0.571	
		Age × SI	β_{5k}	0.698	
	Reduced 5	Female	β_1	< 0.001	197.2
		Age	β_{2k}	< 0.001	
		SI	β_3	< 0.001	
		Female × Age	β_{4k}	< 0.001	
	Reduced 6	Female	β_1	0.005	294.6
		Age	β_{2k}	< 0.001	
		SI	β_3	< 0.001	
		Female × SI	β_4	0.601	
	Reduced 7	Female	β_1	< 0.001	335.5
		Age	β_{2k}	< 0.001	
		SI	β_3	< 0.001	
		Age × SI	β_{4k}	0.705	
	Reduced 8	Female	β_1	< 0.001	282.9
		Age	β_{2k}	< 0.001	
		SI	β_3	< 0.001	
Neoplasms	Full	Female	β_1	< 0.001	106.8
		Age	β_{2k}	< 0.001	
		SI	β_3	0.005	
		Female × Age	β_{4k}	0.003	
		Female × SI	β_5	0.712	
		Age × SI	β_{6k}	0.053	
		Female × SI	β_{7k}	0.744	
	Reduced 1	Female	β_1	< 0.001	53.1

Cause of death	Model	Effect	Parameter	p-value	AIC
		Age	β_{2k}	< 0.001	
		SI	β_3	0.005	
		Female × Age	β_{4k}	< 0.001	
		Female × SI	β_5	0.728	
		Age × SI	β_{6k}	0.045	
	Reduced 2	Female	β_1	< 0.001	1.1
		Age	β_{2k}	< 0.001	
		SI	β_3	0.006	
		Female × Age	β_{4k}	< 0.001	
		Female × SI	β_5	0.696	
	Reduced 3	Female	β_1	< 0.001	39.8
		Age	β_{2k}	< 0.001	
		SI	β_3	0.005	
		Female × Age	β_{4k}	< 0.001	
		Age × SI	β_{5k}	0.044	
	Reduced 4	Female	β_1	< 0.001	119.4
		Age	β_{2k}	< 0.001	
		SI	β_3	0.016	
		Female × SI	β_4	0.779	
		Age × SI	β_{5k}	0.134	
	Reduced 5	Female	β_1	< 0.001	-12.1
		Age	β_{2k}	< 0.001	
		SI	β_3	0.006	
		Female × Age	β_{4k}	< 0.001	
	Reduced 6	Female	β_1	< 0.001	66.1
		Age	β_{2k}	< 0.001	
		SI	β_3	0.018	
		Female × SI	β_4	0.75	
	Reduced 7	Female	β_1	< 0.001	106.3
		Age	β_{2k}	< 0.001	
		SI	β_3	0.016	
		Age × SI	β_{4k}	0.132	
	Reduced 8	Female	β_1	< 0.001	53.1
		Age	β_{2k}	< 0.001	
		SI	β_3	0.018	
Other causes	Full	Female	β_1	< 0.001	5.7
		Age	β_{2k}	< 0.001	
		SI	β_3	< 0.001	
		Female × Age	β_{4k}	0.086	
		Female × SI	β_5	0.895	
		Age × SI	β_{6k}	0.75	
		Female × SI	β_{7k}	0.884	
	Reduced 1	Female	β_1	< 0.001	-51
		Age	β_{2k}	< 0.001	
		SI	β_3	< 0.001	
		Female × Age	β_{4k}	< 0.001	
		Female × SI	β_5	0.895	
		Age × SI	β_{6k}	0.745	
	Reduced 2	Female	β_1	< 0.001	-113.6
		Age	β_{2k}	< 0.001	
		SI	β_3	< 0.001	
		Female × Age	β_{4k}	< 0.001	
		Female × SI	β_5	0.894	
	Reduced 3	Female	β_1	< 0.001	-64.8
		Age	β_{2k}	< 0.001	
		SI	β_3	< 0.001	
		Female × Age	β_{4k}	< 0.001	
		Age × SI	β_{5k}	0.744	
	Reduced 4	Female	β_1	< 0.001	-3.7
		Age	β_{2k}	< 0.001	

Cause of death	Model	Effect	Parameter	p-value	AIC
		SI	β_3	< 0.001	
		Female × SI	β_4	0.908	
		Age × SI	β_{5k}	0.837	
	Reduced 5	Female	β_1	< 0.001	-127.4
		Age	β_{2k}	< 0.001	
		SI	β_3	< 0.001	
		Female × Age	β_{4k}	< 0.001	
	Reduced 6	Female	β_1	< 0.001	-65.7
		Age	β_{2k}	< 0.001	
		SI	β_3	< 0.001	
		Female × SI	β_4	0.907	
	Reduced 7	Female	β_1	< 0.001	-17.3
		Age	β_{2k}	< 0.001	
		SI	β_3	< 0.001	
		Age × SI	β_{4k}	0.836	
	Reduced 8	Female	β_1	< 0.001	-79.3
		Age	β_{2k}	< 0.001	
		SI	β_3	< 0.001	

B.12 Parameter estimates for model in Equation 3.7.

Cause of death	Effect	Parameter	Estimate	SE	p-value	σ_j^2	σ_{res}^2
COVID-19	Intercept	β_0	-3.682	0.468	<0.001	0	1.159
	Female	β_1	-0.428	0.513	0.039		
	Age 25–44	β_{22}	0.931	0.448	<0.001		
	Age 45–64	β_{23}	3.377	0.445	<0.001		
	Age 65–74	β_{24}	4.932	0.445	<0.001		
	Age 75–84	β_{25}	6.047	0.445	<0.001		
	Age 85+	β_{26}	7.199	0.445	<0.001		
	SI	β_3	0.027	0.005	<0.001		
	Female × Age 25–44	β_{42}	0.277	0.618	0.655		
	Female × Age 45–64	β_{43}	-0.217	0.607	0.721		
	Female × Age 65–74	β_{44}	-0.349	0.607	0.565		
	Female × Age 75–84	β_{45}	-0.099	0.607	0.871		
	Female × Age 85+	β_{46}	0.019	0.607	0.975		
External causes	Intercept	β_0	0.113	0.063	0.082	0.001	0.05
	Female	β_1	-0.859	0.065	0.039		
	Age 25–44	β_{22}	1.375	0.065	<0.001		
	Age 45–64	β_{23}	1.832	0.065	<0.001		
	Age 65–74	β_{24}	2.197	0.065	<0.001		
	Age 75–84	β_{25}	2.954	0.065	<0.001		
	Age 85+	β_{26}	4.175	0.065	<0.001		
	SI	β_3	-0.003	0.001	<0.001		
	Female × Age 25–44	β_{42}	-0.338	0.092	<0.001		
	Female × Age 45–64	β_{43}	-0.035	0.092	0.702		
	Female × Age 65–74	β_{44}	0.268	0.092	0.004		
	Female × Age 75–84	β_{45}	0.375	0.092	<0.001		
	Female × Age 85+	β_{46}	0.639	0.092	<0.001		
Heart and vascular diseases	Intercept	β_0	-2.353	0.076	<0.001	0.0001	0.075
	Female	β_1	0.056	0.095	0.561		
	Age 25–44	β_{22}	2.125	0.828	<0.001		
	Age 45–64	β_{23}	4.425	0.828	<0.001		
	Age 65–74	β_{24}	5.881	0.828	<0.001		
	Age 75–84	β_{25}	7.006	0.828	<0.001		
	Age 85+	β_{26}	8.325	0.828	<0.001		
	SI	β_3	-0.001	0.001	0.195		
	Female × Age 25–44	β_{42}	-0.863	0.124	<0.001		
	Female × Age 45–64	β_{43}	-0.953	0.124	<0.001		
	Female × Age 65–74	β_{44}	-0.788	0.124	<0.001		

	Female × Age 75–84	β_{45}	-0.463	0.124	<0.001		
	Female × Age 85+	β_{46}	-0.123	0.124	0.32		
Infectious diseases	Intercept	β_0	-2.315	0.144	<0.001	0	0.1291
	Female	β_1	-0.38	0.206	0.066		
	Age 25–44	β_{22}	0.038	0.162	0.814		
	Age 45–64	β_{23}	1.674	0.147	<0.001		
	Age 65–74	β_{24}	3.315	0.147	<0.001		
	Age 75–84	β_{25}	4.269	0.147	<0.001		
	Age 85+	β_{26}	5.422	0.147	<0.001		
	SI	β_3	-0.002	0.001	0.1847		
	Female × Age 25–44	β_{42}	0.598	0.251	0.018		
	Female × Age 45–64	β_{43}	-0.069	0.23	0.763		
	Female × Age 65–74	β_{44}	-0.323	0.23	0.162		
	Female × Age 75–84	β_{45}	0.043	0.23	0.854		
	Female × Age 85+	β_{46}	0.505	0.23	0.029		
Mental & behavioral disorders	Intercept	β_0	-3.35	0.145	<0.001	0	0.098
	Female	β_1	-0.056	0.169	0.74		
	Age 25–44	β_{22}	1.151	0.145	<0.001		
	Age 45–64	β_{23}	2.845	0.143	<0.001		
	Age 65–74	β_{24}	3.761	0.143	<0.001		
	Age 75–84	β_{25}	5.256	0.143	<0.001		
	Age 85+	β_{26}	6.808	0.143	<0.001		
	SI	β_3	-0.005	0.001	<0.001		
	Female × Age 25–44	β_{42}	-0.652	0.194	<0.001		
	Female × Age 45–64	β_{43}	-0.947	0.192	<0.001		
	Female × Age 65–74	β_{44}	-0.38	0.192	0.049		
	Female × Age 75–84	β_{45}	0.089	0.192	0.644		
	Female × Age 85+	β_{46}	0.311	0.192	0.106		
Neoplasms	Intercept	β_0	-1.478	0.053	<0.001	0	0.042
	Female	β_1	-0.228	0.06	0.039		
	Age 25–44	β_{22}	1.697	0.059	<0.001		
	Age 45–64	β_{23}	4.257	0.059	<0.001		
	Age 65–74	β_{24}	5.67	0.059	<0.001		
	Age 75–84	β_{25}	6.403	0.059	<0.001		
	Age 85+	β_{26}	7.0.58	0.059	<0.001		
	SI	β_3	-0.002	0.001	0.006		
	Female × Age 25–44	β_{42}	0.362	0.084	<0.001		
	Female × Age 45–64	β_{43}	0.01	0.084	0.904		
	Female × Age 65–74	β_{44}	-0.207	0.084	0.014		
	Female × Age 75–84	β_{45}	-0.31	0.084	<0.001		

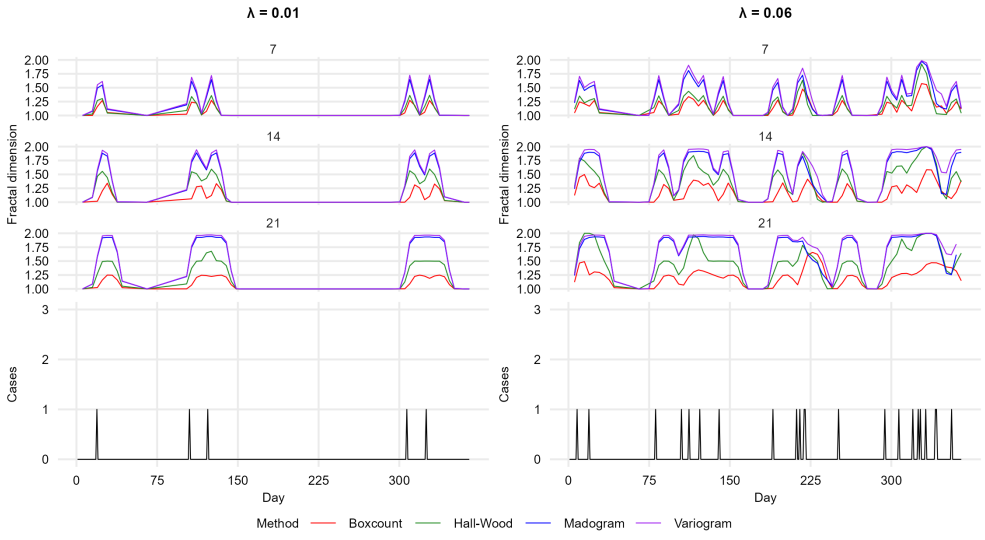
	Female × Age 85+	β_{46}	-0.394	0.084	<0.001		
Other causes	Intercept	β_0	0.436	0.043	<0.001	0	0.028
	Female	β_1	-0.156	0.048	0.001		
	Age 25–44	β_{22}	0.328	0.048	<0.001		
	Age 45–64	β_{23}	2.262	0.048	<0.001		
	Age 65–74	β_{24}	3.648	0.048	<0.001		
	Age 75–84	β_{25}	4.657	0.048	<0.001		
	Age 85+	β_{26}	5.869	0.048	<0.001		
	SI	β_3	-0.004	0.001	<0.001		
	Female × Age 25–44	β_{42}	-0.405	0.068	<0.001		
	Female × Age 45–64	β_{43}	-0.382	0.068	<0.001		
	Female × Age 65–74	β_{44}	-0.394	0.068	<0.001		
	Female × Age 75–84	β_{45}	-0.244	0.068	<0.001		
	Female × Age 85+	β_{46}	0.031	0.068	0.649		

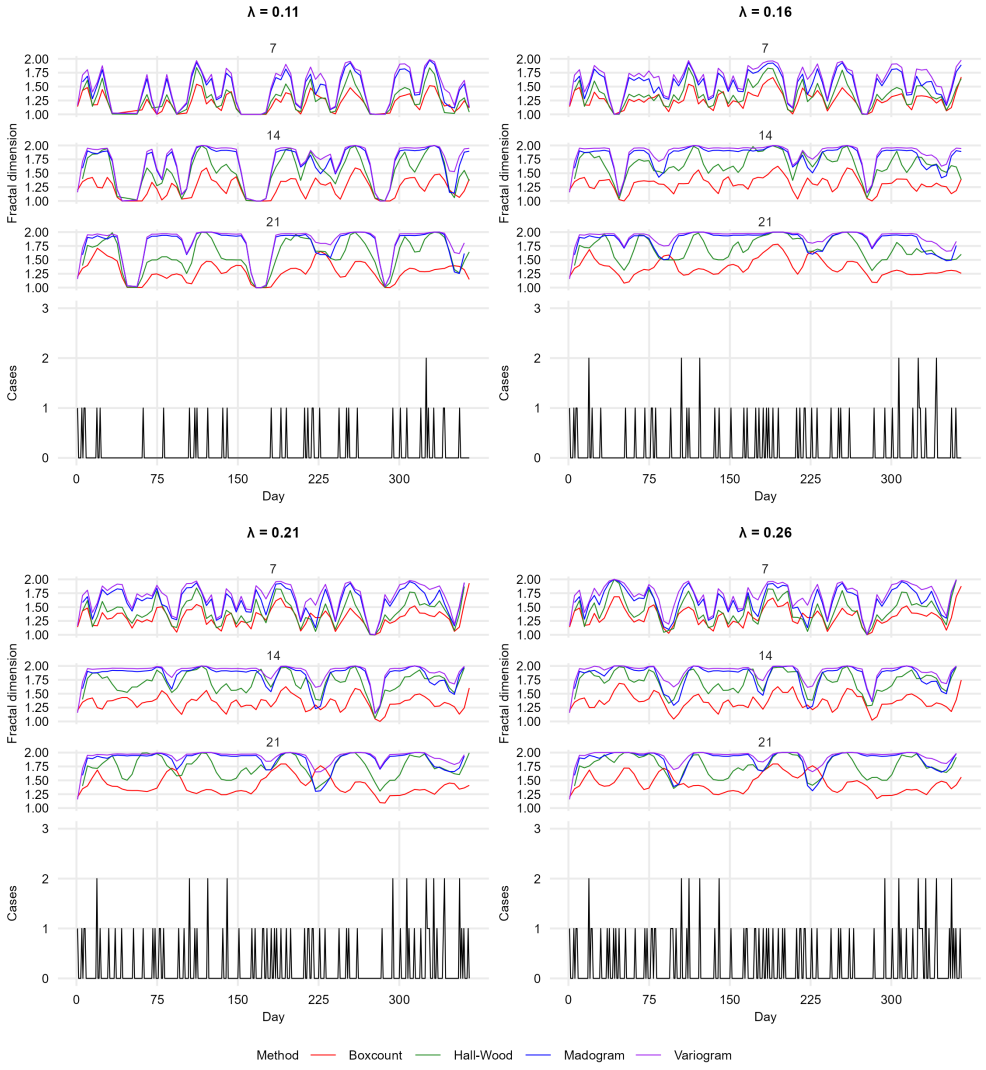
APPENDIX C

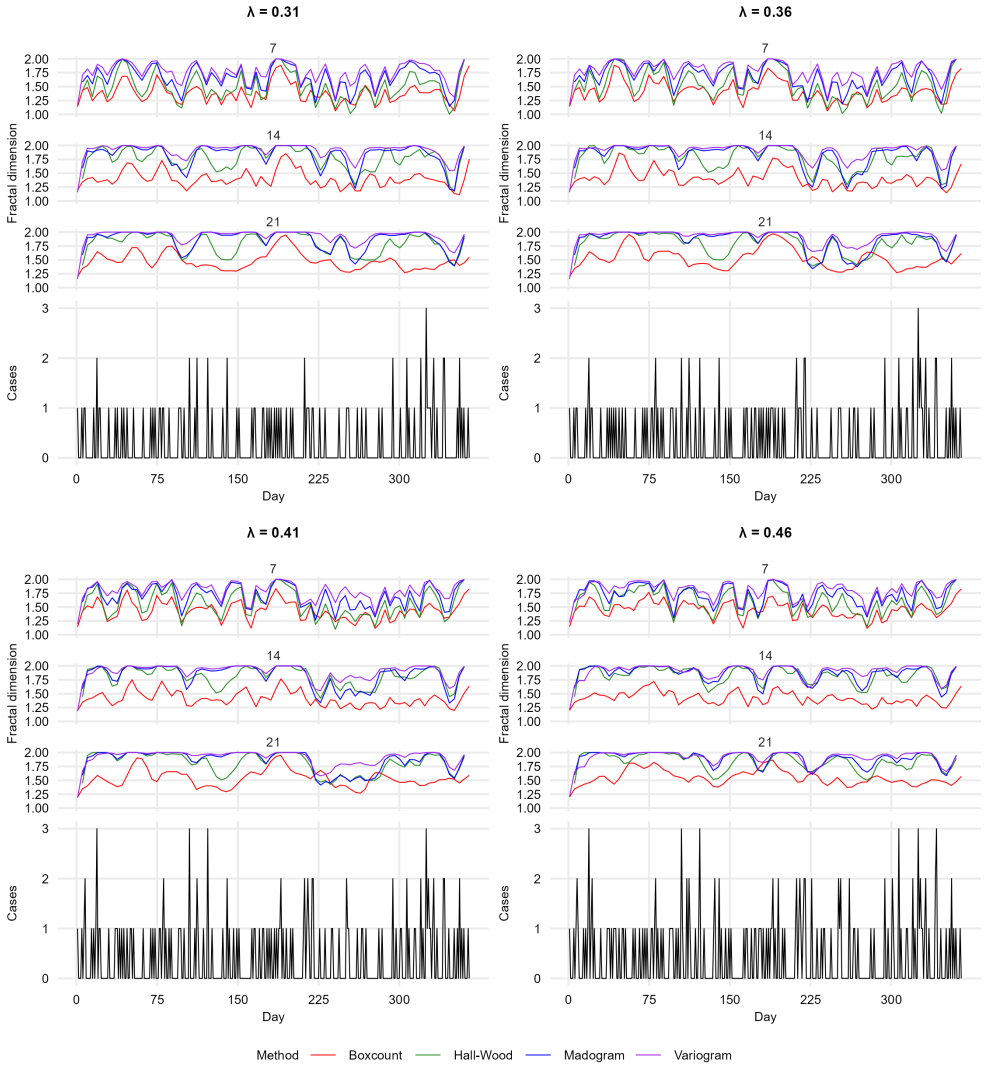
APPENDIX FOR CHAPTER 4

C.1 Simulated daily COVID-19 incidence.

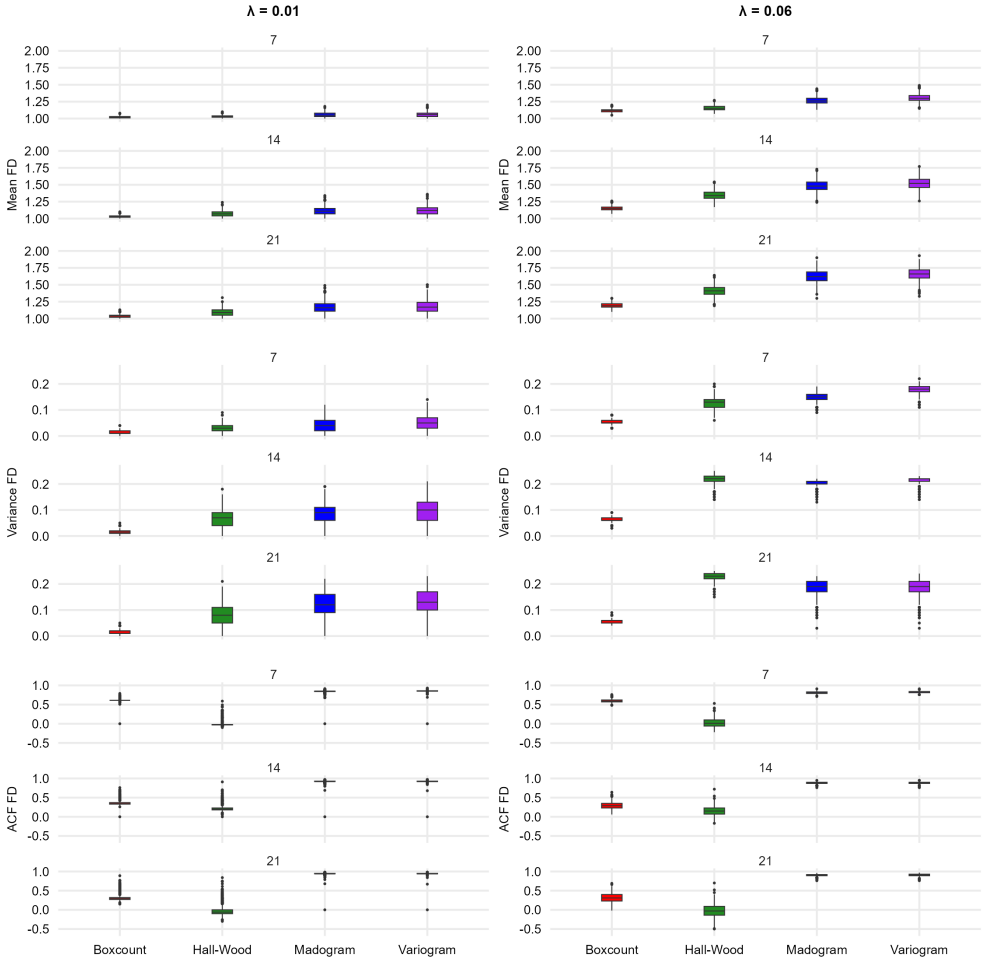
Smoothed local fractal dimension curves are depicted on top of the incidence curve.



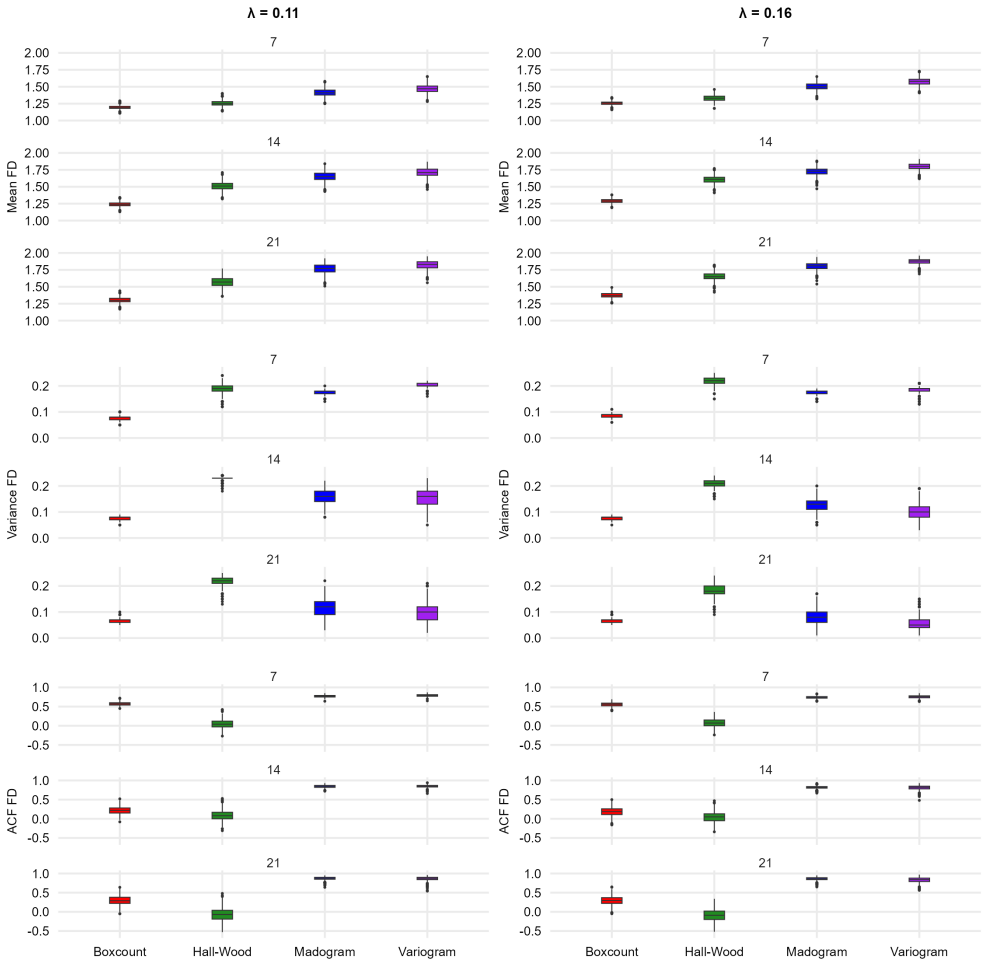

λ = 0.11
λ = 0.16
λ = 0.21
λ = 0.26
7
14
21
Fractal dimension
Cases
Day
0
75
150
225
300
Method
Boxcount
Hall-Wood
Madogram
Variogram

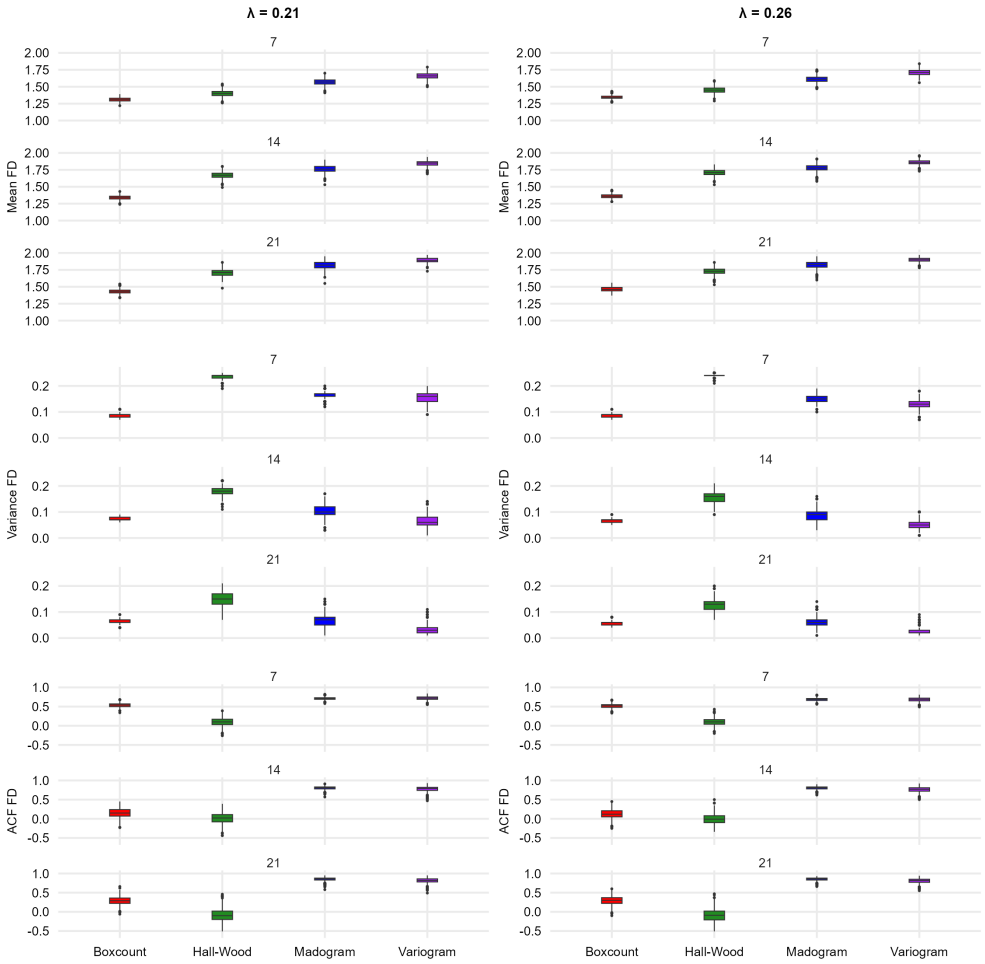

λ = 0.31
λ = 0.36
λ = 0.41
λ = 0.46
7
14
21
Fractal dimension
Cases
Day
Method
Boxcount
Hall-Wood
Madogram
Variogram

C.2 Summary statistics of the local fractal dimension curve after 1,000 replications.

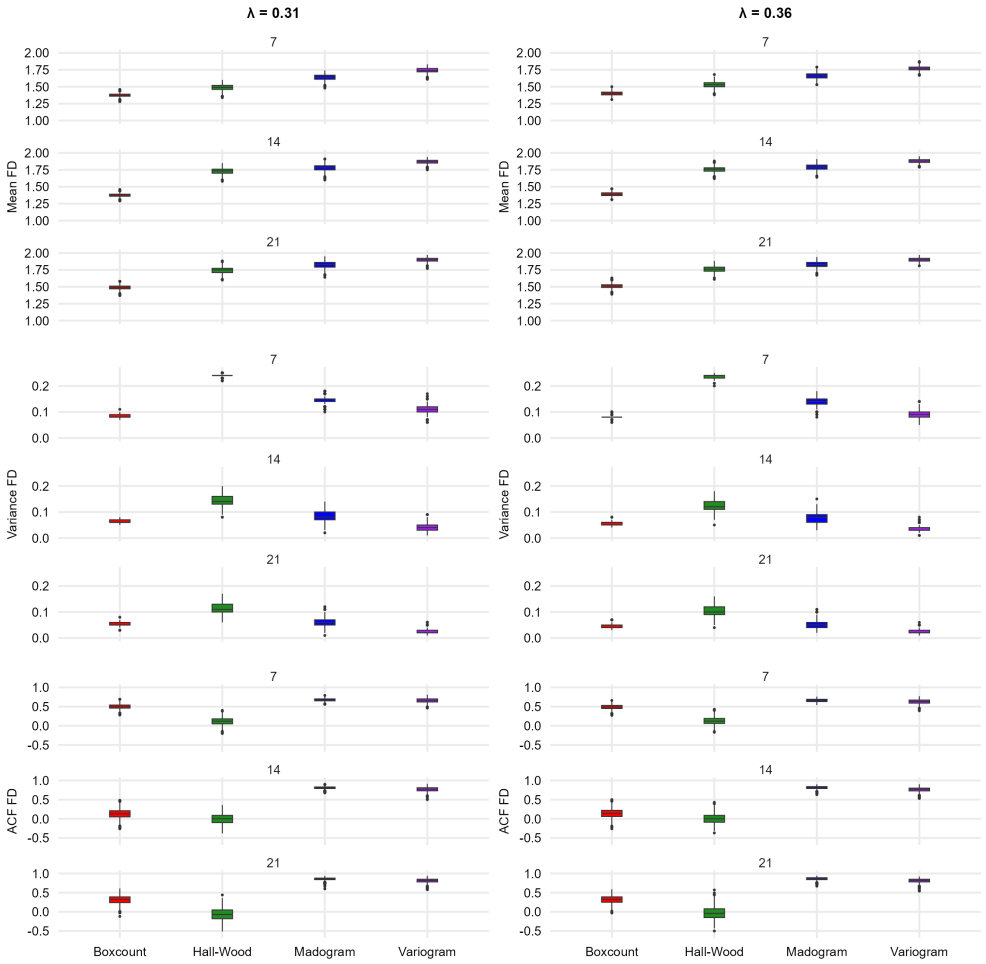


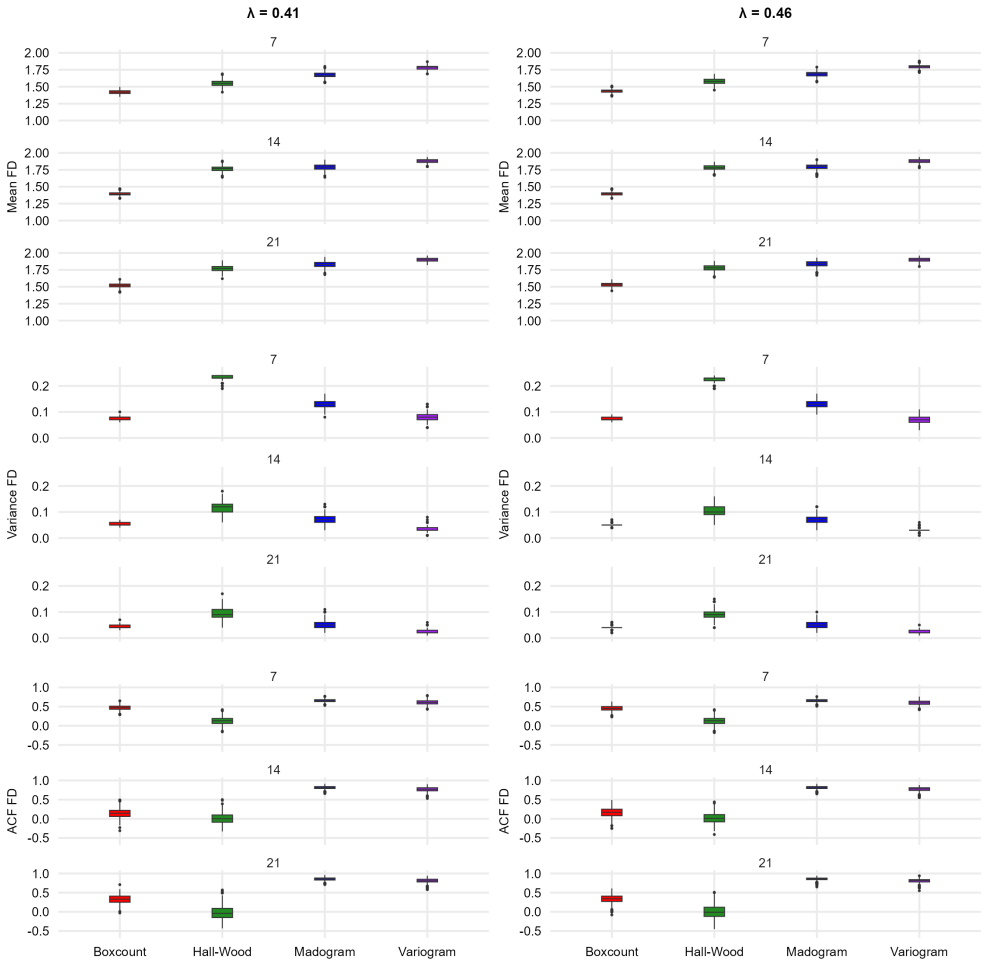
Appendix C.2. Summary statistics of the local fractal dimension curve after 1,000 replications. 159



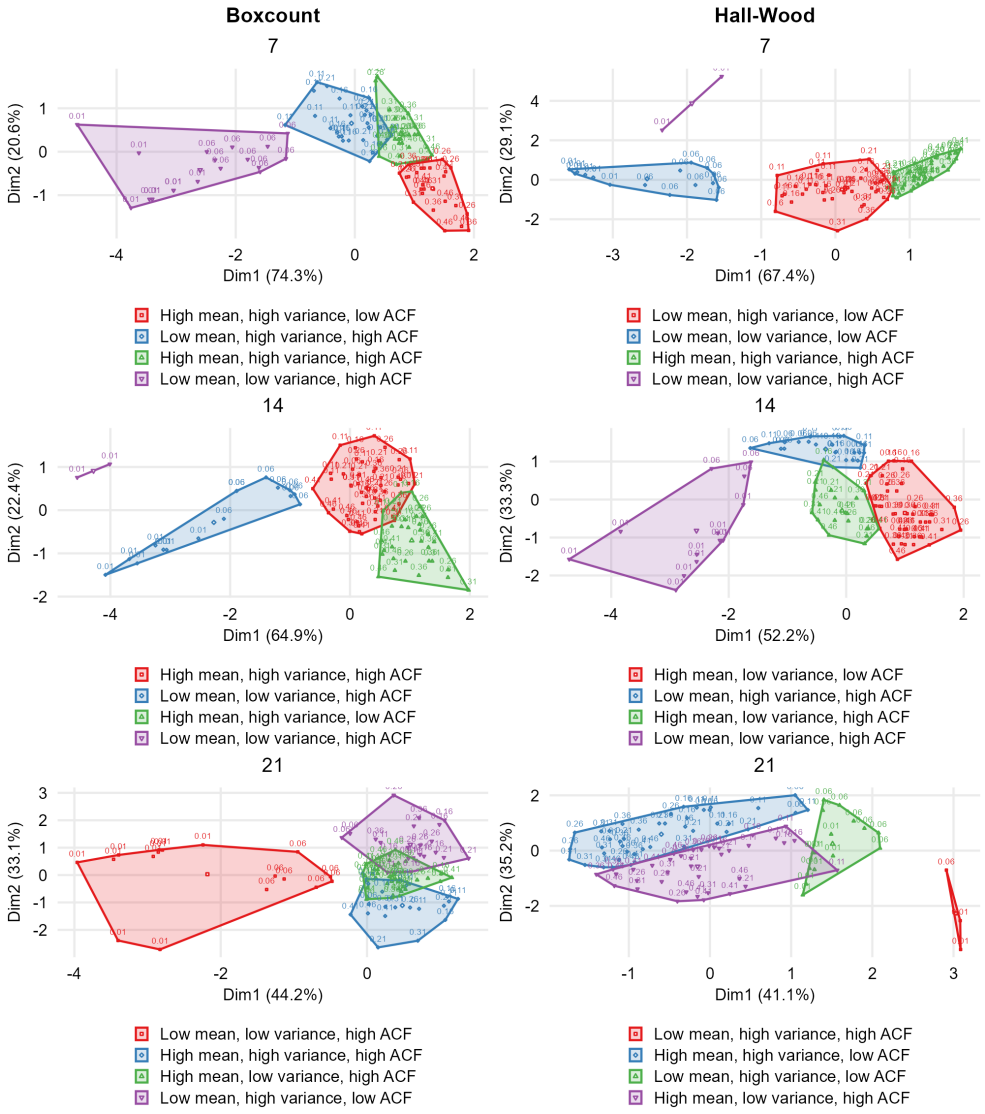

λ = 0.21
λ = 0.26
7
14
21
Mean FD
Variance FD
ACF FD
Boxcount
Hall-Wood
Madogram
Variogram

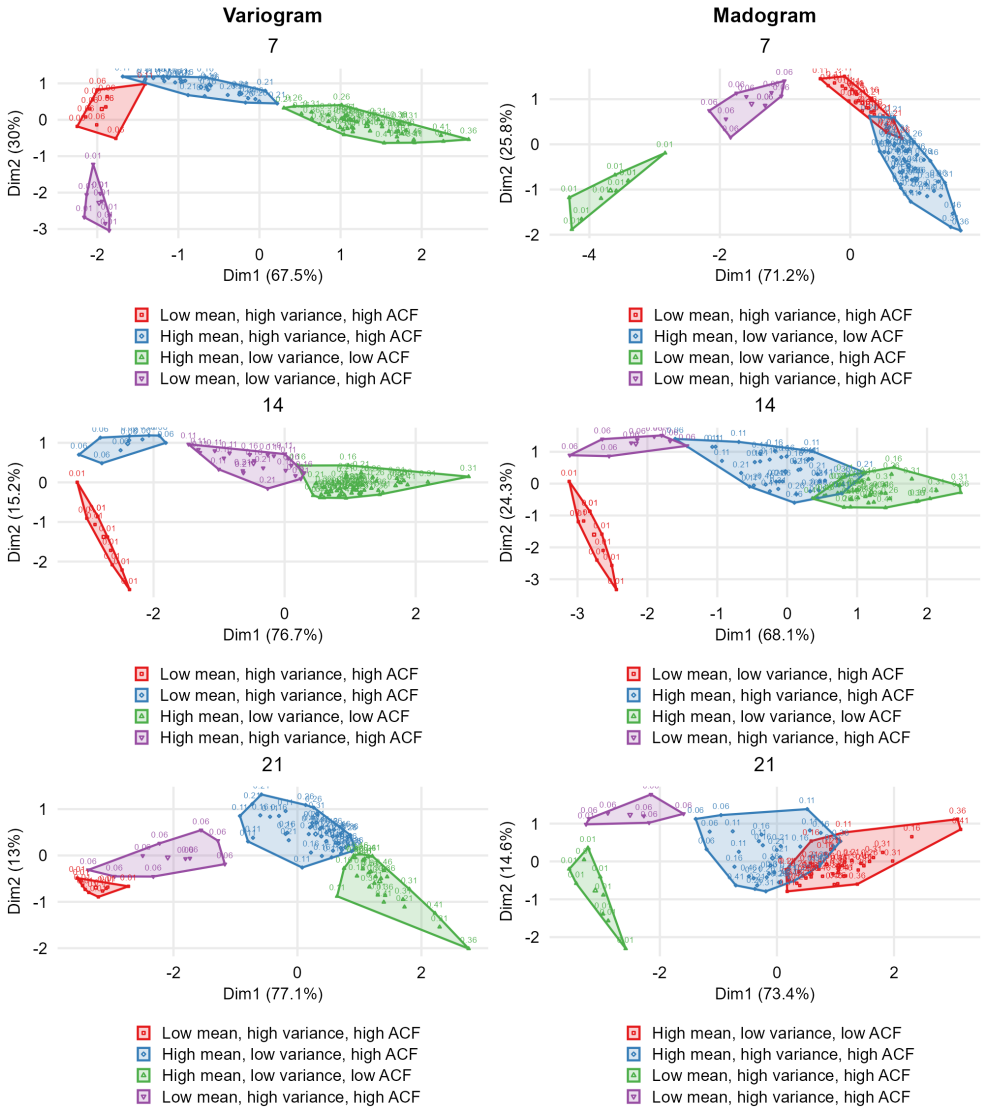
Appendix C.2. Summary statistics of the local fractal dimension curve after 1,000 replications. 161



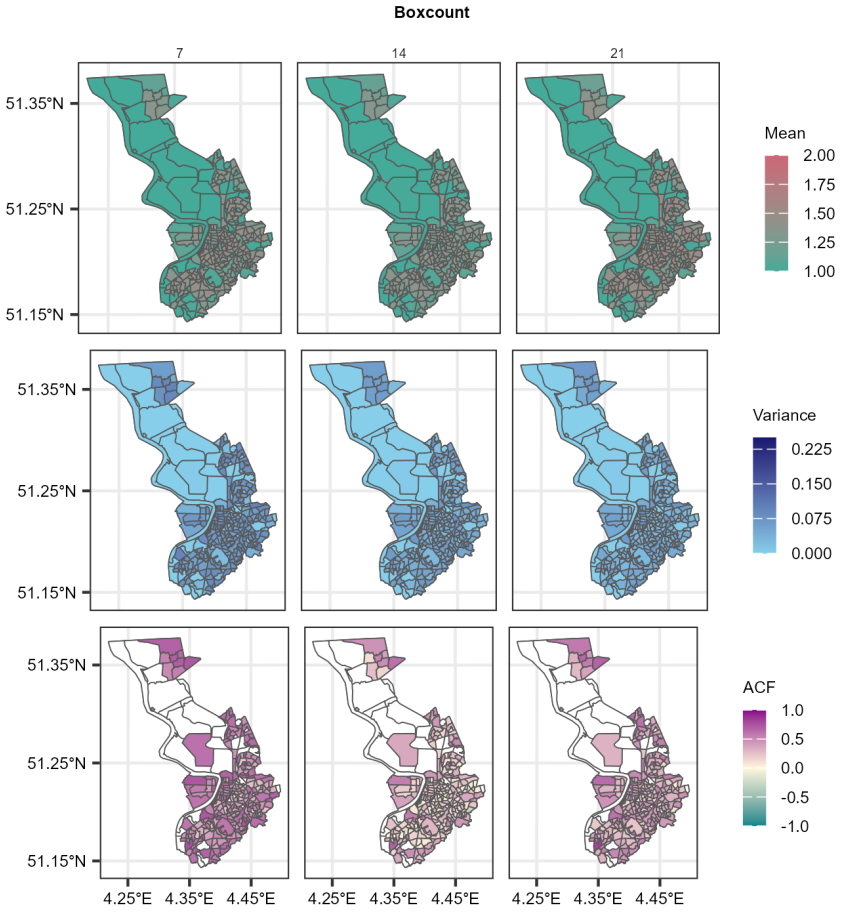

λ = 0.41
λ = 0.46
Mean FD
Variance FD
ACF FD
7
14
21
Boxcount
Hall-Wood
Madogram
Variogram

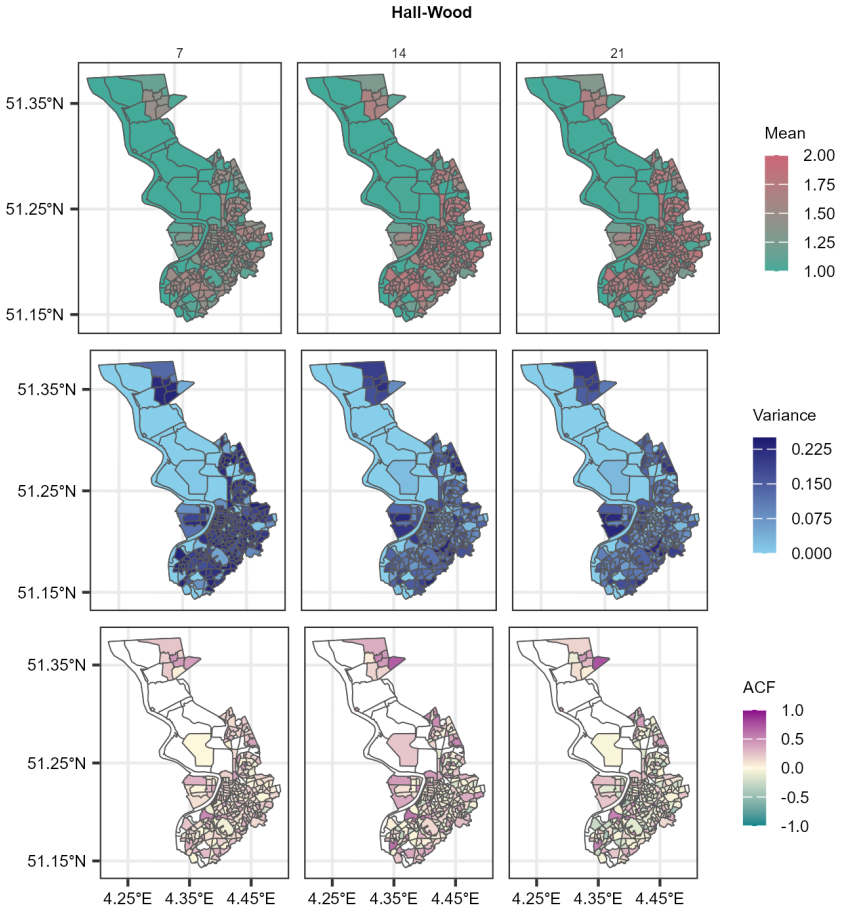
C.3 Clusters detected using k-means clustering based on results from 10 replications.

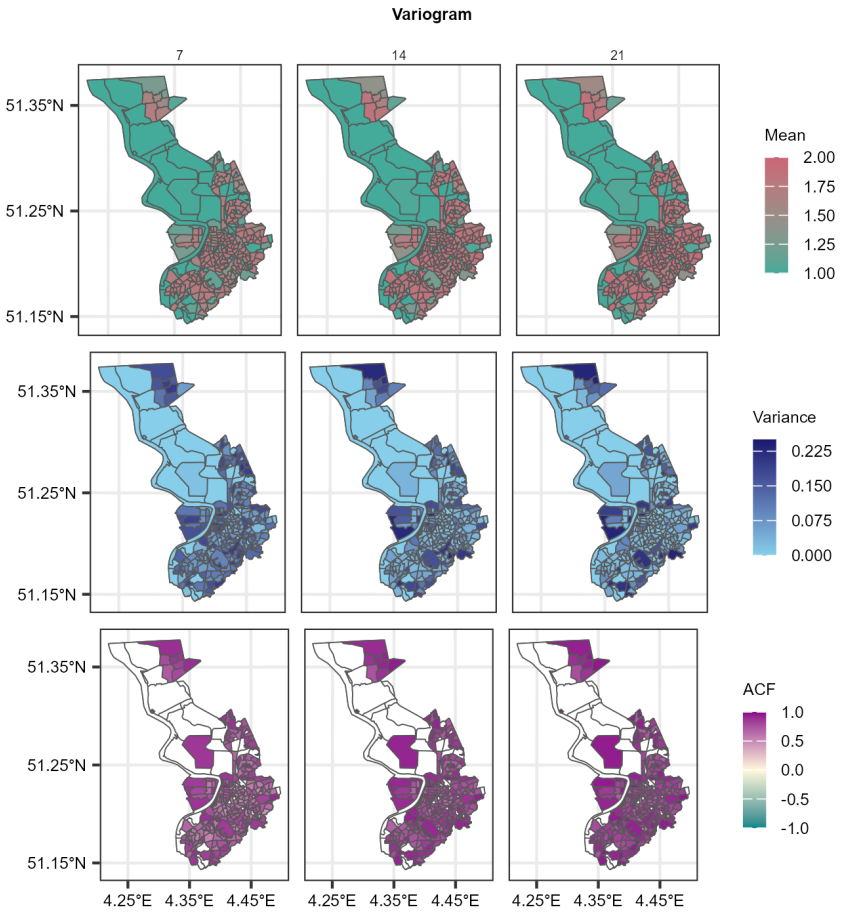


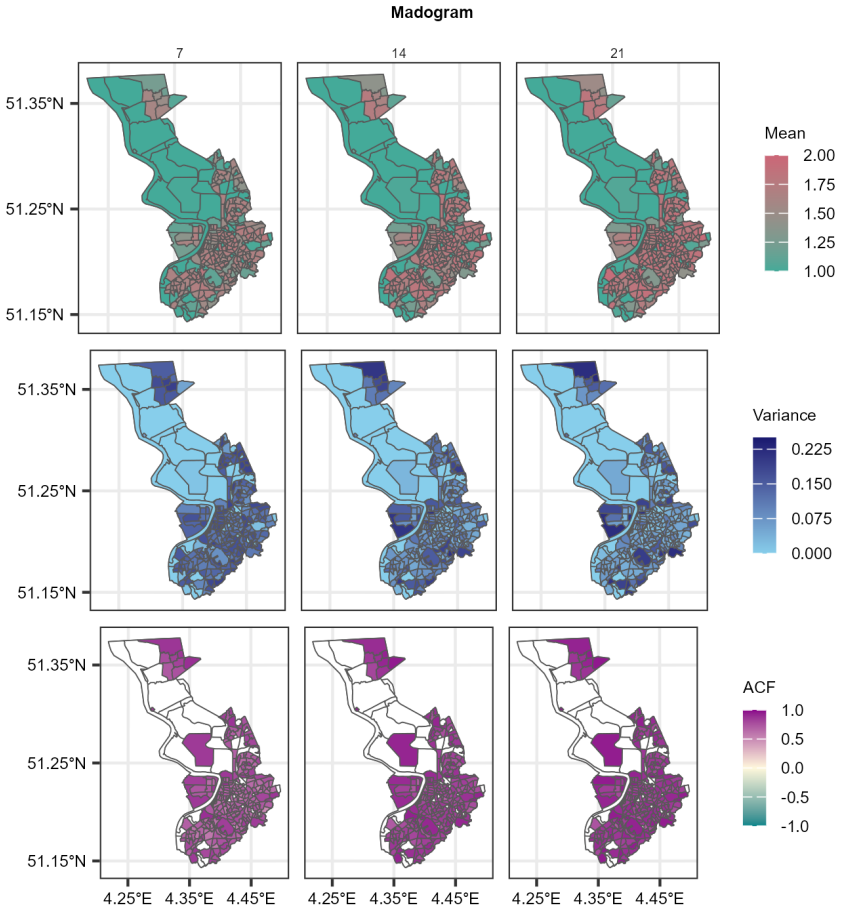

Variogram
7
Dim2 (30%)
Dim1 (67.5%)
Low mean, high variance, high ACF
High mean, high variance, high ACF
High mean, low variance, low ACF
Low mean, low variance, high ACF
Madogram
7
Dim2 (25.8%)
Dim1 (71.2%)
Low mean, high variance, high ACF
High mean, low variance, low ACF
Low mean, low variance, high ACF
Low mean, high variance, high ACF
14
Dim2 (15.2%)
Dim1 (76.7%)
Low mean, high variance, high ACF
Low mean, high variance, high ACF
High mean, low variance, low ACF
High mean, high variance, high ACF
14
Dim2 (24.3%)
Dim1 (68.1%)
Low mean, low variance, high ACF
High mean, high variance, high ACF
High mean, low variance, low ACF
Low mean, high variance, high ACF
21
Dim2 (13%)
Dim1 (77.1%)
Low mean, high variance, high ACF
High mean, low variance, high ACF
High mean, low variance, low ACF
Low mean, high variance, high ACF
21
Dim2 (14.6%)
Dim1 (73.4%)
High mean, low variance, low ACF
High mean, high variance, high ACF
Low mean, high variance, high ACF
Low mean, high variance, high ACF

C.4 Fractal dimension indicators for each statistical sector in the municipality of Antwerp.

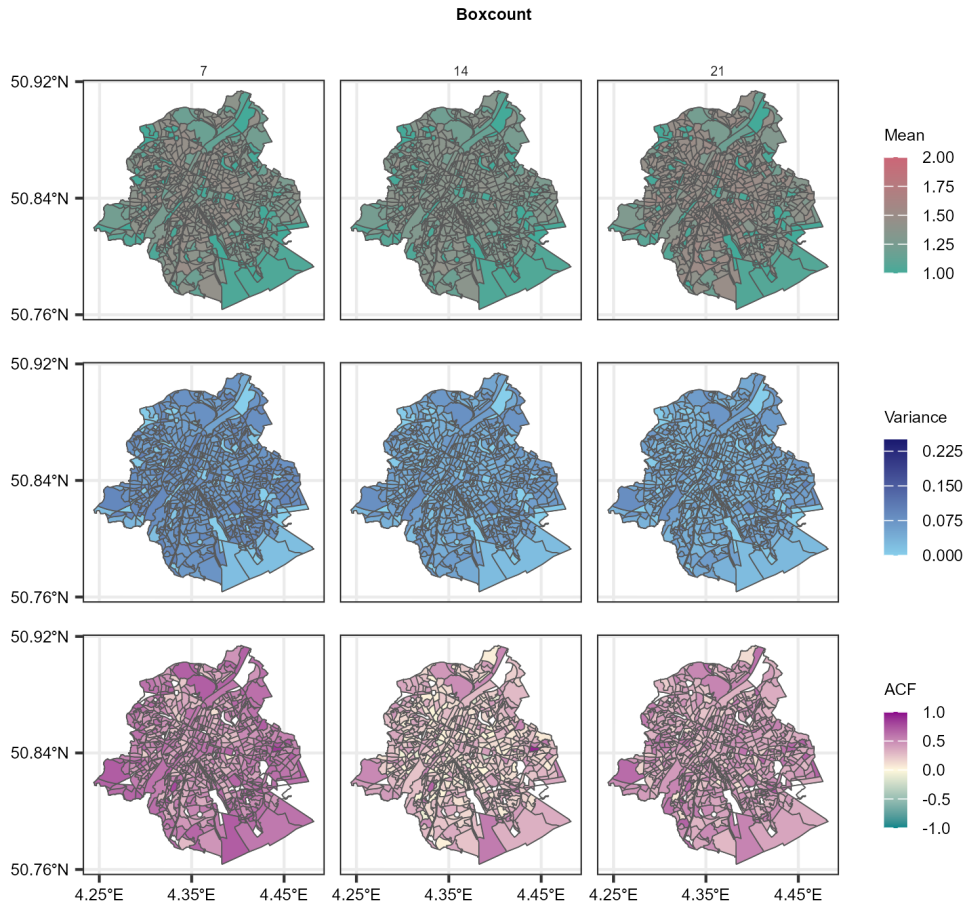


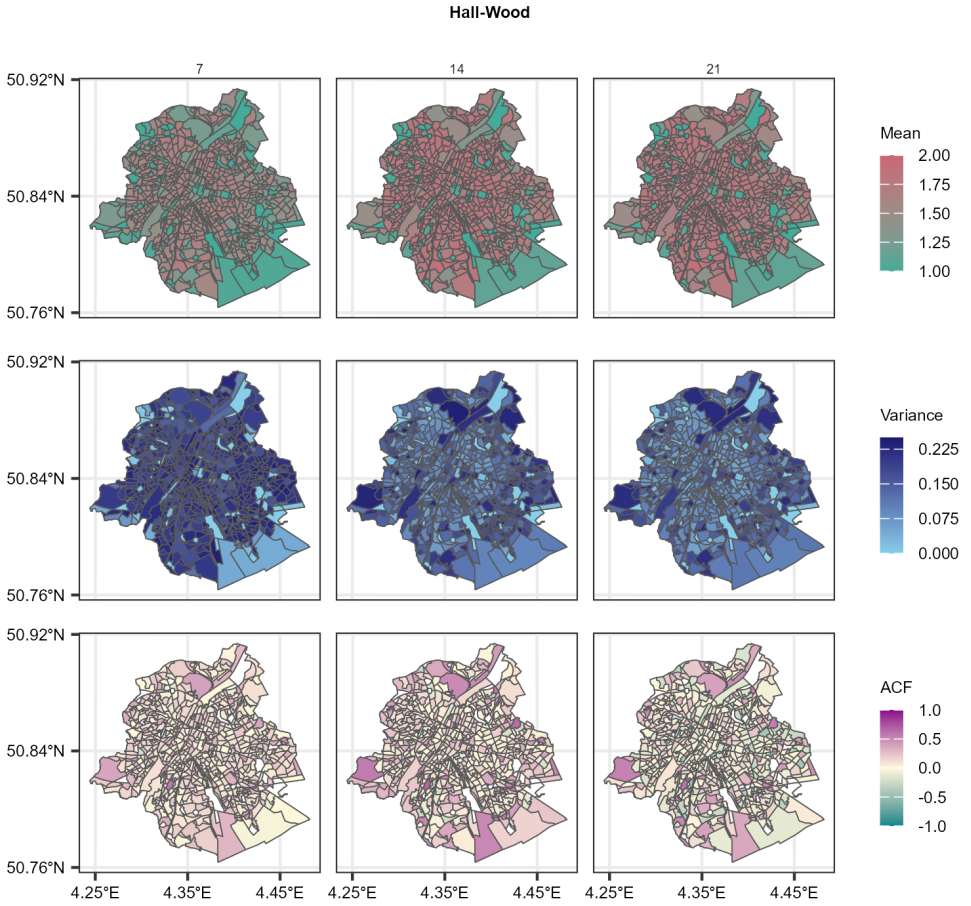

Hall-Wood
7
14
21
51.35°N
51.25°N
51.15°N
Mean
2.00
1.75
1.50
1.25
1.00
Variance
0.225
0.150
0.075
0.000
ACF
1.0
0.5
0.0
-0.5
-1.0
4.25°E
4.35°E
4.45°E

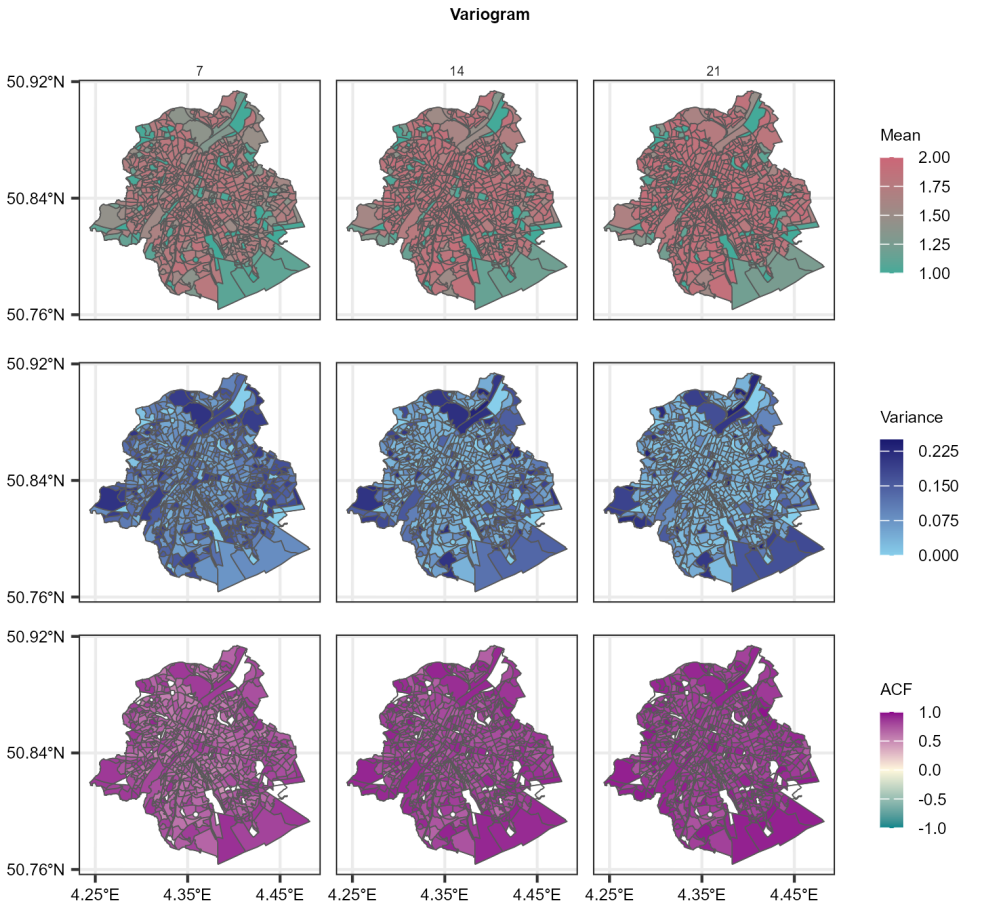

Variogram
7
14
21
51.35°N
51.25°N
51.15°N
4.25°E
4.35°E
4.45°E
Mean
2.00
1.75
1.50
1.25
1.00
Variance
0.225
0.150
0.075
0.000
ACF
1.0
0.5
0.0
-0.5
-1.0

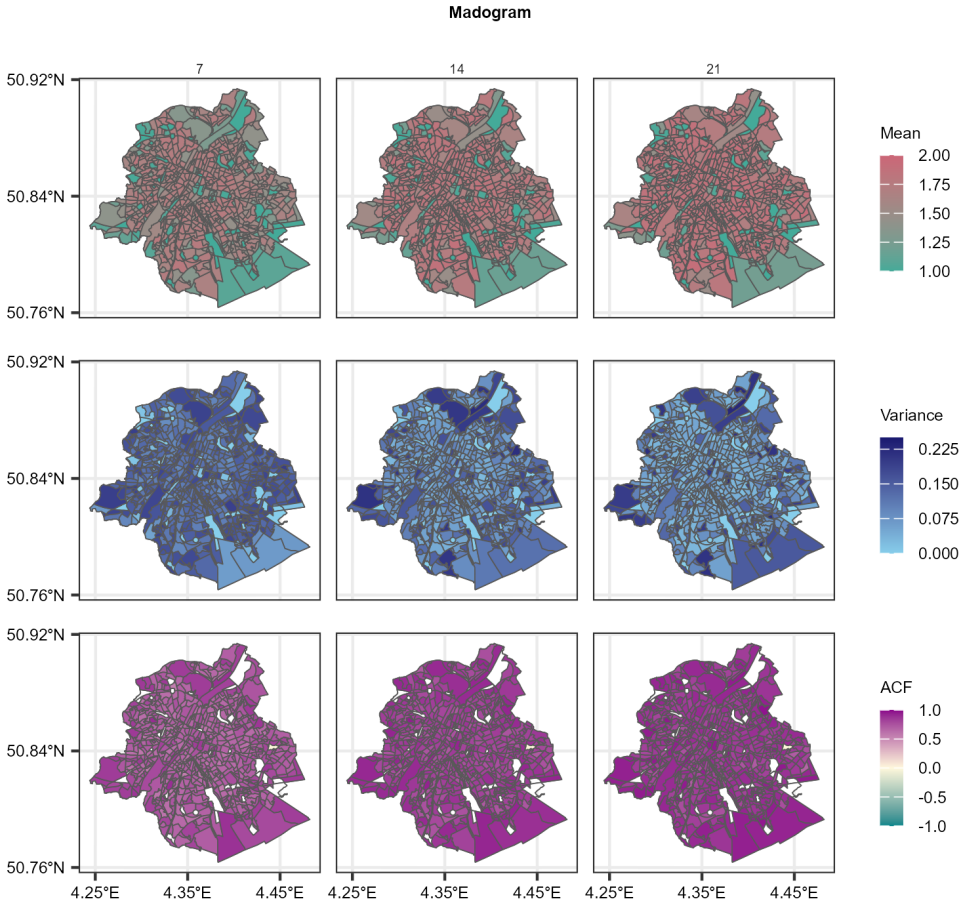

Madogram
7
14
21
51.35°N
51.25°N
51.15°N
Mean
2.00
1.75
1.50
1.25
1.00
Variance
0.225
0.150
0.075
0.000
ACF
1.0
0.5
0.0
-0.5
-1.0
4.25°E
4.35°E
4.45°E

C.5 Fractal dimension indicators for each statistical sector in the Brussels-Capital region.

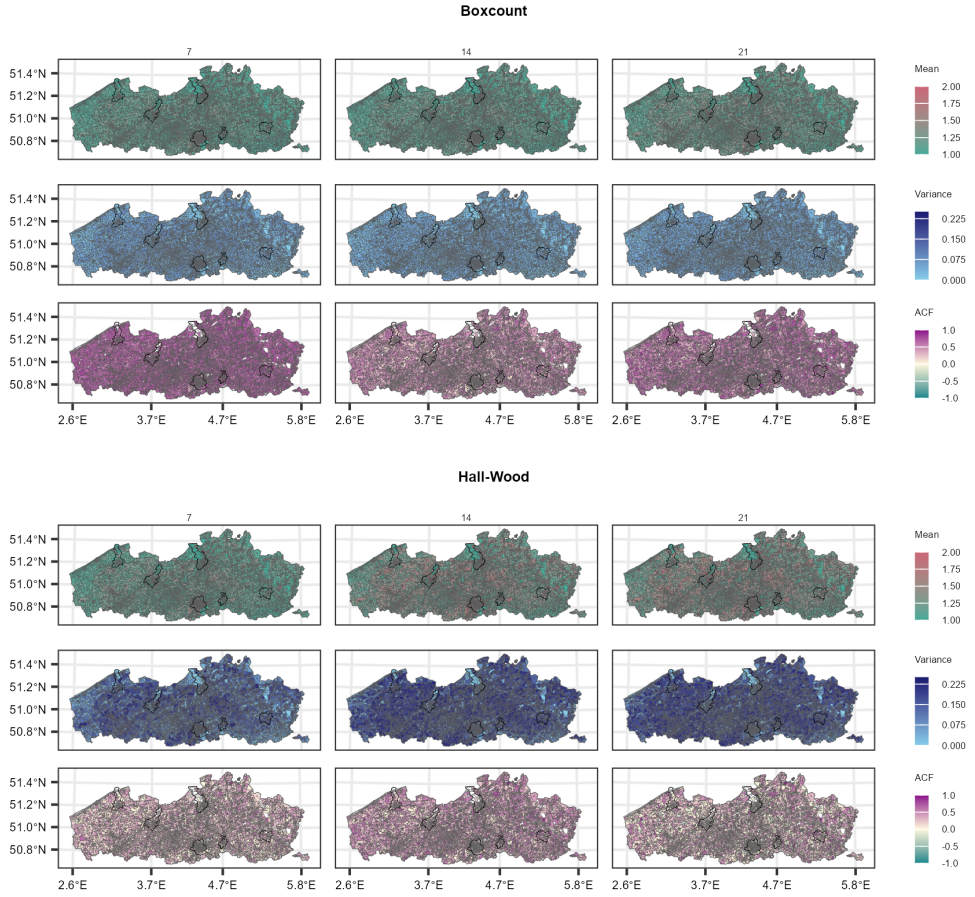


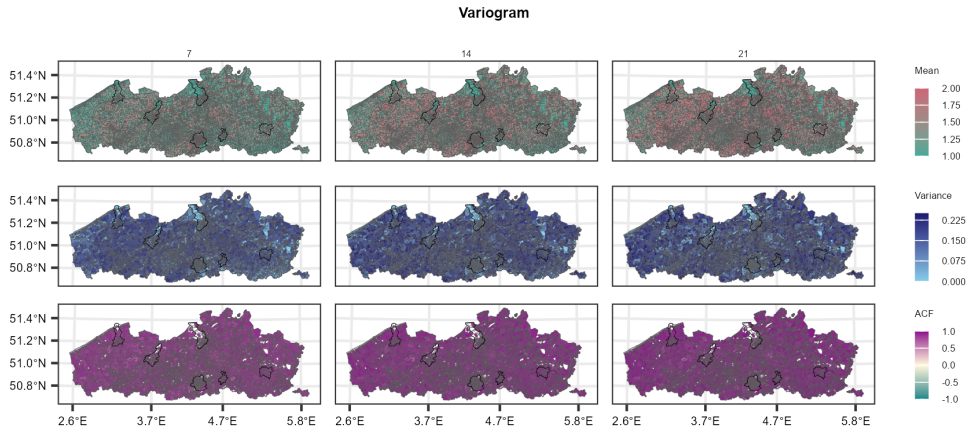

Hall-Wood
7
14
21
50.92°N
50.84°N
50.76°N
4.25°E
4.35°E
4.45°E
Mean
2.00
1.75
1.50
1.25
1.00
Variance
0.225
0.150
0.075
0.000
ACF
1.0
0.5
0.0
-0.5
-1.0

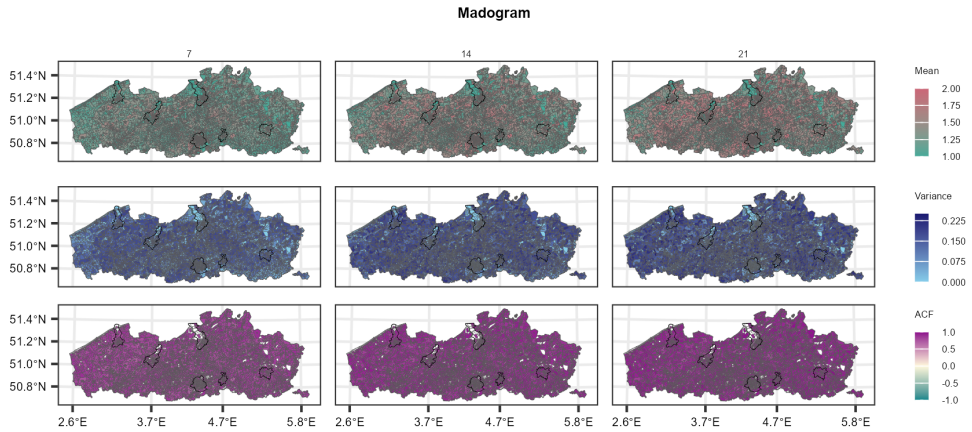

Variogram
7
14
21
50.92°N
50.84°N
50.76°N
4.25°E
4.35°E
4.45°E
Mean
2.00
1.75
1.50
1.25
1.00
Variance
0.225
0.150
0.075
0.000
ACF
1.0
0.5
0.0
-0.5
-1.0


Madogram
7
14
21
50.92°N
50.84°N
50.76°N
4.25°E
4.35°E
4.45°E
Mean
2.00
1.75
1.50
1.25
1.00
Variance
0.225
0.150
0.075
0.000
ACF
1.0
0.5
0.0
-0.5
-1.0

C.6 Fractal dimension indicators for each statistical sector in the Flemish and Brussels-Capital Regions.




Variogram
7
14
21
51.4°N
51.2°N
51.0°N
50.8°N
2.6°E
3.7°E
4.7°E
5.8°E
Mean
2.00
1.75
1.50
1.25
1.00
Variance
0.225
0.150
0.075
0.000
ACF
1.0
0.5
0.0
-0.5
-1.0


Madogram
7
14
21
51.4°N
51.2°N
51.0°N
50.8°N
2.6°E
3.7°E
4.7°E
5.8°E
Mean
2.00
1.75
1.50
1.25
1.00
Variance
0.225
0.150
0.075
0.000
ACF
1.0
0.5
0.0
-0.5
-1.0

C.7 Canonical loading based on longer sliding windows.

Region	Sliding window	Variable	Boxcount	Hall-Wood	Variogram	Madogram
Flanders	14	Population size	-0.992	-0.991	-0.99	-0.99
		Population density	-0.686	-0.687	-0.686	-0.686
		Shannon index	-0.362	-0.373	-0.375	-0.376
		Older age population	0.277	0.276	0.266	0.268
		Median income	0.297	0.299	0.303	0.302
		Vaccination coverage	0.155	0.153	0.145	0.146
		Satisfaction	-0.075	-0.074	-0.088	-0.084
		Trust in the federal government	-0.327	-0.337	-0.338	-0.339
		Trust in the regional government	-0.257	-0.271	-0.281	-0.281
		Mean FD	-0.905	-0.904	-0.881	-0.884
		Variance FD	-0.43	-0.135	0.537	0.322
		ACF FD	0.618	0.563	0.832	0.596
	21	Population size	-0.991	-0.991	-0.991	-0.991
		Population density	-0.686	-0.686	-0.681	-0.683
		Shannon index	-0.37	-0.373	-0.371	-0.373
		Older age population	0.274	0.271	0.252	0.255
		Median income	0.302	0.299	0.3	0.299
		Vaccination coverage	0.152	0.149	0.137	0.139
		Satisfaction	-0.075	-0.079	-0.1	-0.095
		Trust in the federal government	-0.33	-0.336	-0.328	-0.332
		Trust in the regional government	-0.263	-0.273	-0.279	-0.281
		Mean FD	-0.916	-0.889	-0.857	-0.857
		Variance FD	-0.227	0.074	0.798	0.693
		ACF FD	0.403	0.426	0.839	0.571
Brussels-Capital	14	Population size	-0.99	-0.989	-0.972	-0.983
		Population density	-0.613	-0.619	-0.658	-0.643
		Shannon index	-0.251	-0.256	-0.29	-0.252
		Older age population	0.304	0.317	0.278	0.286
		Median income	0.306	0.333	0.349	0.351
		Vaccination coverage	0.249	0.264	0.213	0.251
		Mean FD	-0.573	-0.416	-0.101	-0.137
		Variance FD	0.699	0.991	0.871	0.888
		ACF FD	0.468	0.363	0.425	0.122
	21	Population size	-0.985	-0.986	-0.956	-0.973
		Population density	-0.629	-0.632	-0.686	-0.671
		Shannon index	-0.268	-0.256	-0.296	-0.263
		Older age population	0.331	0.321	0.245	0.262
		Median income	0.346	0.337	0.345	0.347
		Vaccination coverage	0.256	0.255	0.195	0.242
		Mean FD	-0.549	-0.326	-0.067	-0.121
		Variance FD	0.858	0.996	0.814	0.847
		ACF FD	0.299	0.143	0.453	0.192

SUMMARY

The COVID-19 pandemic profoundly reshaped multiple aspects of daily life, from societal norms to research priorities and public health perspectives. In its aftermath, the urgency of preparing for and responding to future large-scale health crises has become a central concern for many countries, including Belgium. This underscores the importance of critically evaluating and refining methodological approaches to ensure accurate estimation, nuanced interpretation, and actionable insights that can guide policy makers and the broader public health community.

Chapters 2 and 3 explored hierarchical models, specifically Bayesian spatial models implemented through INLA and linear mixed models as a framework for analyzing spatiotemporal COVID-19 data. These models revealed important demographic and contextual drivers of the pandemic, including age, sex, income, population density, student population size, and human mobility. Additional influencing factors, such as vaccination coverage and the stringency of government interventions, were also highlighted. Together, these findings illustrate the multifaceted nature of epidemic spread and the need for methods capable of integrating variables operating across social, spatial, and temporal dimensions.

Hierarchical models remain one of the most flexible and powerful tools for capturing these complexities. Their ability to incorporate spatial dependence, temporal trends, and individual-level heterogeneity makes them invaluable in epidemiological research. However, this flexibility comes at a price: building complex hierarchical models requires careful consideration of distributional assumptions, rigorous assessment of model fit, or appropriate specification of priors in the Bayesian context. Moreover, computational demands grow substantially with model complexity, raising practical challenges in real-time surveillance. Overly complex models can also suffer from identifiability issues, convergence failures, or singularities, emphasizing the importance of balancing sophistication with interpretability and computational feasibility.

Chapter 4 introduced fractal dimension analysis as a complementary tool to characterize epidemic dynamics. Unlike hierarchical models, fractal-based approaches focus on quantifying the complexity of time-series data without imposing strong assumptions about the underlying spatial mechanisms. By summarizing epidemic dynamics through indicators such as mean, variance, and autocorrelation of local fractal dimension curves, this method enables rapid characterization of epidemic patterns at very fine administrative scales. This is particularly valuable when data are noisy, fragmented, or rapidly evolving, conditions frequently encountered during emerging outbreaks. The proposed fractal-based clustering framework further demonstrates the potential of this approach to identify vulnerable areas and provide actionable insights for targeted interventions.

Together, the hierarchical modeling framework and the fractal dimension approach illustrate the value of combining established and innovative methodologies. Hierarchical models excel at capturing structured relationships and quantifying the effects of covariates, while fractal analysis offers an efficient way to summarize and compare epidemic complexity across space and time. Integrating these approaches can enrich epidemiological analysis, improve outbreak surveillance, and ultimately strengthen preparedness for future health crises.

SAMENVATTING

De COVID-19-pandemie heeft verschillende aspecten van het dagelijks leven ingrijpend veranderd, variërend van maatschappelijke normen tot onderzoeksprioriteiten en publieke gezondheidszorg. De urgentie om zich voor te bereiden op en te reageren bij toekomstige grootschalige gezondheidscrisis is een centrale zorg geworden voor veel landen, waaronder België. Dit benadrukt het belang van een kritische evaluatie en verdere verfijning van methodologische benaderingen om robuuste schattingen, genuanceerde interpretaties en beleidsrelevante inzichten te faciliteren die beleidsmakers en de bredere publieke gezondheidssector kunnen ondersteunen.

In hoofdstukken 2 en 3 worden hiërarchische modellen, met name Bayesiaanse ruimtelijke model geïmplementeerd via INLA en lineair mixed model, onderzocht in het kader van spatiotemporele COVID-19-gegevens. Deze modellen brachten belangrijke demografische en contextuele determinanten van de pandemie aan het licht, waaronder leeftijd, geslacht, inkomen, bevolkingsdichtheid, de omvang van de studentenpopulatie en menselijke mobiliteit. Daarnaast worden ook factoren zoals vaccinatiegraad en overheidsmaatregelen benadrukt. Gezamenlijk illustreren deze bevindingen het veelzijdige karakter van epidemische verspreiding en de noodzaak van methoden die in staat zijn om variabelen te integreren die op sociaal, ruimtelijk en temporeel niveau opereren.

Hiërarchische modellen behoren tot de meest flexibele en krachtige methoden voor complexe gegevens te analyseren. Hun vermogen om ruimtelijke afhankelijkheid, temporele trends en individuele heterogeniteit te integreren speelt een belangrijke rol in vele epidemiologische onderzoeken. Deze flexibiliteit kent echter ook een keerzijde: het opbouwen van complexe hiërarchische modellen vereist zorgvuldige aandacht voor verschillende voorwaarden, onder andere verdelingsaannames, rigoureuze toetsing van model fit of een adequate specificatie van *priors* in een Bayesiaanse context. Bovendien nemen de computa-

tionele vereisten aanzienlijk toe met de complexiteit van het model, wat praktische uitdagingen met zich meebrengt voor *real-time surveillance*. Overmatig complexe modellen kunnen daarnaast lijden onder problemen zoals identificeerbaarheid, convergentiefalen of singulariteiten, wat het belang benadrukt van een evenwichtige afweging tussen modelsophisticatie, interpreteerbaarheid en computationele haalbaarheid.

In Hoofdstuk 4 wordt fractale dimensie analyse geïntroduceerd als een complimentaire methode om de dynamiek van epidemieën te beschrijven. In tegenstelling tot hiërarchische modellen richten fractaal-gebaseerde benaderingen zich op het kwantificeren van de complexiteit van tijdreeksgegevens, zonder sterke aannames te beschouwen over onderliggende ruimtelijke mechanismen. Door epidemische dynamiek samen te vatten via indicatoren zoals het gemiddelde, de variantie en de autocorrelatie van lokale fractale dimensiecurves, kunnen epidemische patronen relatief snel worden gekarakteriseerd op zeer fijne administratieve niveaus. Dit is bijzonder waardevol in situaties waarin gegevens gefragmenteerd of sterk veranderlijk zijn, kenmerken die vaak voorkomen bij opkomende uitbraken. De voorgestelde methode toont bovendien het potentieel om kwetsbare gebieden te identificeren en beleidsrelevante inzichten te genereren voor gerichte interventies.

Het hiërarchische model en de fractale dimensie analyse illustreren samen de meerwaarde van het combineren van gevestigde en innovatieve methodologische benaderingen. Waar hiërarchische modellen uitblinken in het vastleggen van gestructureerde verbanden en het kwantificeren van covariaateffecten, biedt fractale analyse een efficiënte manier om de complexiteit van epidemieën in ruimte en tijd samen te vatten en te vergelijken. De integratie van deze methoden kan epidemiologische analyses verrijken, surveillancesystemen versterken en uiteindelijk bijdragen aan een betere paraatheid voor toekomstige gezondheidscrisis.