ID N. AIDR01/16886

UNIVERSITÀ CAMPUS BIO-MEDICO DI ROMA

DEPARTMENT OF ENGINEERING

UNIVERSITÀ DEGLI STUDI DI TORINO

DEPARTMENT OF COMPUTER SCIENCE

Italian National Ph.D. in Artificial Intelligence
Health and Life Sciences
XXXVII Cycle

Modeling framework for managing epidemiological scenarios

Supervisors

Paolo Castagno

Matteo Sereno

Candidate

Daniele Baccega

March, 2025

To myself.

Acknowledgements

I would like to express my gratitude to my supervisors, Prof. Matteo Sereno and Dr. Paolo Castagno, for allowing me to pursue my Ph.D. and for their invaluable guidance, consistent support, and patience throughout this journey, which began with my Master's thesis in 2020. I also extend my thanks to Prof. Eugenio Guglielmelli and Prof. Paolo Soda for their excellent coordination of the Ph.D. program. On the same line, I am grateful to Dr. Antonio Fernández Anta for providing me with the opportunity to spend time abroad at IMDEA Networks Institute in Madrid. This enriching experience enabled me to collaborate with esteemed researchers like Dr. Jose Aguilar, Dr. Juan Marcos Ramirez, and, even if only online, Prof. Carlos Baquero. Lastly, I would like to thank Prof. Marco Beccuti, Prof. Attilio Fiandrotti, and Dr. Simone Pernice for their valuable guidance and support throughout this journey.

A special thanks to the Q-Bio group for all the enjoyable lunches, the laughter, and the persistent support (in random order: Elena, Nicola, Irene, Daniela, Simone, Marco Jr., Marco Sr., Francesca, Sandro, Giulia, Giulia, Riccardo, Francesco, Roberta, and Aaron). I would also like to extend my sincere gratitude to the Global Computing Group at IMDEA Networks Institute for the wonderful time spent in Madrid (in random order: Sergio, Diego, Antonio, Jose, Junlang, Ari, Nico, Juan Marcos, and Javier). A special thank you to Sergio for teaching me how to play ping pong during breaks, and to Junlang and Diego for joining me in the games. I hope to see you all again soon! Last but not least, I would like to extend a special thanks to my colleagues from the XXXVII cycle of the Ph.D. in AI. We had a wonderful week together in Rome during the Ph.D. school.

I would like to thank my family for their resolute support throughout the years and for being a source of strength during the challenging times of my Ph.D. Finally, I am deeply grateful to my partner, Alice, for their patience, support, and constant encouragement every day. Without you, this journey would not have been possible.

I would also like to acknowledge the use of AI-based tools, such as ChatGPT, to assist in improving the clarity and flow of certain sections of this thesis.

Published and submitted content

The ideas and investigations of this thesis resulted in the following publications and submissions:

Articles published in journal:

1. **Baccega, D.**, Castagno, P., Fernández Anta, A., and Sereno, M. *Enhancing COVID-19 forecasting precision through the integration of compartmental models, machine learning and variants.* Sci Rep 14, 19220 (2024). doi: `https://doi.org/10.1038/s41598-024-69660-5` [39].

    - This work is fully included and its content is reported in Chapter 3.
    - The author discussed, developed, and tested the framework, and conceived and conducted the experiments. The author fully participated in the article's writing.

2. Saglietto, A.[†], **Baccega, D.**[†], Esposito, R., Anselmino, M., Dusi, V., Fiandrotti, A. and De Ferrari, G.M. (2024). *Convolutional neural network (CNN)-enabled electrocardiogram (ECG) analysis: a comparison between standard twelve-lead and single-lead setups.* Front. Cardiovasc. Med. 11:1327179. doi: `https://doi.org/10.3389/fcvm.2024.1327179` [232].

    - This work is not included in the thesis. The abstract is included in the Appendix C.
    - The author discussed, developed, and tested the network, and conceived and conducted the experiments. The author fully participated in the article's writing.

3. **Daniele Baccega**, Simone Pernice, Pietro Terna, Paolo Castagno, Giovenale Moirano, Lorenzo Richiardi, Matteo Sereno, Sergio Rabellino, Milena Maria Maule, and Marco Beccuti. *An Agent-Based Model to Support Infection Control Strategies at School.* In: Journal of Artificial Societies and Social Simulation 25.3 (2022), p. 2. URL:

[†]Co-first authors

http://jasss.soc.surrey.ac.uk/25/3/2.html. doi: https://doi.org/10.18564/jasss.4830 [41].

- This work is fully included and its content is reported in Chapter 2.
- The author discussed, developed, and tested the agent-based model, and conceived and conducted the experiments. The author partially participated in the article's writing.

Articles in preparation:

4. **Daniele Baccega**, and Terrone, Irene and Heywood, Peter and Chisholm, Robert and Richmond, P. and Contaldo, Sandro Gepiro and Bosio, Lorenzo and Pernice, Simone and Beccuti, Marco, *Forge4flame: An Intuitive Dashboard for Designing GPU Agent-Based Models to Simulate Infectious Disease Spread.* doi: http://dx.doi.org/10.2139/ssrn.5194584. [42]

   - This work is fully included and its content is reported in Chapter 5.
   - The author discussed, developed, and tested the models, and conceived and conducted the experiments. The author is fully participating in the article's writing.

5. **Daniele Baccega**, Jose Aguilar, Carlos Baquero, Antonio Fernández Anta, Juan Marcos Ramirez. *Social Compliance with NPIs, Mobility Patterns, and Reproduction Number: Lessons from COVID-19 in Europe.* medRxiv (2025): 2025-01. doi: https://doi.org/10.1101/2025.01.10.25320334. [38]

   - This work is fully included and its content is reported in Chapter 4.
   - The author discussed, developed, and tested the tool, and conceived and conducted the experiments. The author is fully participating in the article's writing.

Articles published in conference proceedings:

6. Marzio Pennisi, Giuliana Franceschinis, **Daniele Baccega**, Simone Pernice, Irene Terrone. *Exploiting GPU computing for effective Agent-Based simulation: initial experiments.* 2025 33rd Euromicro International Conference on Parallel, Distributed, and Network-Based Processing (PDP). IEEE, 2025. [201]

   - This work is not included in the thesis.

- The author did not contribute to the development. The author partially participated in the writing of the article.

7. **Baccega, Daniele**, Simone Pernice, Paolo Castagno, Matteo Sereno, and Marco Beccuti. *Living Along COVID-19: Assessing Contention Policies Through Agent-Based.* Computational Intelligence Methods for Bioinformatics and Biostatistics: 18th International Meeting, CIBB 2023, Padova, Italy, September 6–8, 2023, Revised Selected Papers. Vol. 14513. Springer Nature, 2025. [40].

    - This work is fully included and its content is reported in Chapter 2.
    - The author discussed, developed, and tested the agent-based model, and conceived and conducted the experiments. The author fully participated in the article's writing.

8. Sandro Gepiro Contaldo, Emanuele Pietropaolo, Lorenzo Bosio, Simone Pernice, Irene Terrone, **Daniele Baccega**, Yuting Wang, Rahul Kumar Sahoo, Giuseppe Rizzo, Alessia Visconti, Paola Berchialla, and Marco Beccuti. *PHeP: TrustAlert Open-Source Platform for Enhancing Predictive Healthcare with Deep Learning.* medRxiv (2024): 2024-08. doi: `https://doi.org/10.1101/2024.08.22.24311205`. In the 19th Conference on Computational Intelligence Methods for Bioinformatics & Biostatistics (CIBB 2024). The extended version of the article is accepted in the proceedings of the conference Springer Lecture Notes in Computer Science (LNCS) [86].

    - This work is not included in the thesis. The abstract is included in the Appendix C.
    - The author did not contribute to the development. The author partially participated in the article's writing.

Articles presented at conferences:

9. **Daniele Baccega**, Paolo Castagno, Antonio Fernández Anta, and Matteo Sereno. *Enhancing COVID-19 Forecasting Precision through the Integration of Compartmental Models, Machine Learning and Variants.* 7th epiDAMIK ACM SIGKDD International Workshop on Epidemiology meets Data Mining and Knowledge Discover, Barcelona (2024). URL: `https://openreview.net/forum?id=v89eTCg3EO`.

    - This work is fully included and its content is reported in Chapter 3.
    - The author discussed, developed, and tested the framework, and conceived and conducted the experiments. The author fully participated in the article's writing.

Abstract accepted at conferences:

10. Emanuele Koumantakis, Giuseppe Rizzo, Mauro Dragoni, Marco Beccuti, Paola Berchialla, for the **TrustAlert Group**[1]. *An Integrated Platform for Early Warning and Forecasting of Health Emergencies: the TrustAlert Project.* In the American Medical Informatics Association (AMIA) annual symposium, San Francisco, CA (2024). URL: `https://amia.secure-platform.com/symposium/solicitations/102001/sessiongallery/94150`.

    - This work is not included in the thesis.
    - The author did not contribute to the development. The author partially participated in the writing of the abstract.

11. Andrea Saglietto[†], **Daniele Baccega**[†], Matteo Anselmino, Veronica Dusi, Roberto Esposito, Attilio Fiandrotti and Gaetano Maria De Ferrari. *Artificial intelligence-enabled single-lead ECG versus standard 12-lead ECG.* Accepted as abstract in 19° Congresso Nazionale AIAC, Bologna (2023). URL: `https://aiac2023.webaimgroup.eu/wordpress/wp-content/uploads/aiac23-programmadefinitivo.pdf`.

    - This work is not included in the thesis. The abstract is included in the Appendix C.
    - The author discussed, developed, and tested the network, and conceived and conducted the experiments. The author fully participated in the writing of the abstract.

Tutorial accepted at conferences:

12. **Daniele Baccega**, Marco Beccuti, Simone Pernice, Irene Terrone. *Epidemic Modeling Frameworks for Studying Infectious Disease Dynamics.* Accepted in 13th IEEE International Conference on Healthcare Informatics (IEEE ICHI 2025).

    - This work is not included in the thesis.
    - The author fully participated in the writing of the tutorial proposal.

[1] I'm a member of the TrustAlert Group.
[†] Co-first authors

Abstract

Computational epidemiology is an interdisciplinary research area that aims to develop and use computational models to understand and control the spread of a disease in a population. Computational models and algorithms have been successfully used to evaluate different intervention strategies, including decisions in the field of public health such as the definition of vaccination and antiviral timings, as well as interventions aimed at limiting social contacts like the closure of schools, public spaces, and the suspension of socio-economic activities. In the last twenty years, different pandemics such as Dengue, SARS, MERS, swine flu, Ebola, Zika, and COVID-19 have highlighted the difficulties in managing such a global emergency. In particular, the outbreak of COVID-19 represents the most striking example of a global pandemic that has seen the proliferation of epidemic studies that analyzed various aspects related to the spread of infectious diseases. Although these studies entail several limitations and assumptions, they played a fundamental role in supporting policy-makers in the choice of management strategies for such emergencies.

In particular, a continuous monitoring system helps anticipate, manage, and simulate the dynamics of health crisis medical needs. The goal is to advance the preparedness and response capabilities of territorial public health agencies in emergencies. To achieve this goal, it is necessary to establish a comprehensive, integrated platform, to provide early warnings, monitoring, and forecasting tools to public health response agencies and local healthcare services for anticipating medical needs. We can distinguish three main core objectives: *i)* to early detect the onset of a health crisis, such as the spread of cases due to specific pathologies, *ii)* to assess the impact of the health crisis at the territorial level, and *iii)* to forecast scenarios integrating macro- and micro-level models of different types. Moreover, a continuous monitoring system is crucial for socio-economic stability by enabling early detection of issues, optimizing resource allocation, and facilitating informed decision-making. By providing real-time data and fostering data-driven innovation, it contributes to increased productivity, competitiveness, and overall economic resilience. The sheer amount of data to handle brings several challenges, including data heterogeneity, time dependence, temporal/spatial multi-resolution, incompleteness, and inaccuracy. These aspects need to

be taken into account while developing incremental and scalable data analysis techniques to learn the information needed to design and validate epidemiological models.

The aim of this Ph.D. thesis is the development of modeling tools, techniques, and guidelines useful to provide timely information on the future effects of political interventions on the progress of the contagion and to assist the actions of policy-makers by allowing them to evaluate and compare different actions to the possible outcome. We begin with a general introduction to computational epidemiology and continuous monitoring systems, followed by a discussion of the contributions and the related state-of-the-art. The focus then shifts to each specific contribution. First, we present an agent-based model designed to evaluate the efficacy of various non-pharmaceutical interventions and vaccination strategies within a well-defined closed environment (we focused on a generic middle school as a case study). This model assesses the infection risks associated with school activities and examines the effects of potential control strategies following the introduction of infectious cases. Next, we introduce Sybil, an integrated framework that combines two different types of approaches: a data-centric approach (a machine learning-based predictive model) and an analytical approach (a variant-aware compartmental model) to enhance prediction accuracy and explainability taking advantage of the relative stability of disease characteristic indices to make future projections. Following this, we analyze the relationship between government-imposed restrictions, mobility patterns, and the effective reproduction number R_t of COVID-19. Using data from major platforms, we examine how movement trends respond to policy changes, revealing cultural differences in social compliance between Northern and Southern European countries, which influence disease transmission dynamics. Then, we present *Forge4Flame (F4F)*, a user-friendly dashboard that addresses a key limitation of agent-based models: the challenge of constructing realistic environments, particularly when modeling complex spaces such as offices, schools, supermarkets, and more. This tool automatically generates the necessary code, simplifying the modeling process without the need for any computational expertise. Finally, the concluding chapter summarizes the key findings and outlines potential directions for future research.

Contents

1 Introduction **22**
1.1 Epidemiology . . . 23
1.1.1 Computational epidemiology . . . 24
1.1.2 Brief pandemic history . . . 25
1.1.3 COVID-19 . . . 26
1.2 Continuous monitoring system . . . 27
1.2.1 Early warning system . . . 29
1.2.2 Computational models . . . 30
1.2.3 Knowledge extraction . . . 34
1.2.4 Combination of multiple approaches . . . 38
1.2.5 Adapting pandemic solutions to broader challenges . . . 38
1.3 Contribution . . . 40
1.3.1 Contribution 1: *Living along COVID-19: assessing contention policies through Agent-Based Models* . . . 41
1.3.2 Contribution 2: *Sybil: enhancing COVID-19 forecasting precision with variant-aware integrated models* . . . 41
1.3.3 Contribution 3: *Social compliance with NPIs, mobility patterns, and reproduction number: lessons from COVID-19 in Europe* . . . 42
1.3.4 Contribution 4: *Forge4Flame: an intuitive dashboard for designing GPU Agent-Based Models to simulate the spread of infectious diseases* 43
1.4 State-of-the-art . . . 44
1.4.1 *Living along COVID-19: assessing contention policies through Agent-Based Models* . . . 45
1.4.2 *Sybil: enhancing COVID-19 forecasting precision with variant-aware integrated models* . . . 47
1.4.3 *Social compliance with NPIs, mobility patterns, and reproduction number: lessons from COVID-19 in Europe* . . . 52

1.4.4 *Forge4Flame: an intuitive dashboard for designing GPU Agent-Based Models to simulate the spread of infectious diseases* 56
1.5 Chapters description 60

2 Living along COVID-19: assessing contention policies through Agent-Based Models 62
2.1 Model 64
2.1.1 School environment 64
2.1.2 Contagion processes 65
2.1.3 Non-pharmaceutical interventions 68
2.2 Verification and validation 70
2.3 Results 73
2.3.1 First set: *Basic Countermeasures for Risk Mitigation* 74
2.3.2 Second set: *Combined Impact of Various NPIs on School Safety* . . . 79
2.4 Discussion 90
2.5 Conclusion 92
2.6 Availability of data and software code 93
2.6.1 How to reproduce the results 93

3 Sybil: enhancing COVID-19 forecasting precision with variant-aware integrated models 94
3.1 Sybil framework 96
3.1.1 Compartmental analytical model 96
3.1.2 Prophet predictive model 99
3.1.3 Data 101
3.2 Results 103
3.2.1 Italy 103
3.2.2 Errors 110
3.2.3 Austria 112
3.2.4 Errors 116
3.2.5 Italy with eleven variants (V = 11) 117
3.2.6 Other European countries 120
3.2.7 Fixed recovery rate 123
3.2.8 USA data 131
3.2.9 Continuous monitoring 136
3.3 Discussion 141
3.4 Conclusion 143

3.5 Availability of data and software code 144
3.5.1 How to reproduce the results 144

4 Social compliance with NPIs, mobility patterns, and reproduction number: lessons from COVID-19 in Europe 146
4.1 Data preparation 149
4.1.1 Data 149
4.1.2 Pre-processing 151
4.1.3 Preliminary data analysis 154
4.2 Our approach 157
4.3 Results 161
4.3.1 Patterns 161
4.3.2 Clustering 163
4.3.3 Regression Models 173
4.4 Discussion 175
4.5 Conclusions 177
4.6 Ethical declaration 178
4.7 Availability of data and software code 179
4.7.1 How to reproduce the results 179

5 Forge4Flame: an intuitive dashboard for designing GPU Agent-Based Models to simulate the spread of infectious diseases 181
5.1 Background 185
5.1.1 FLAME GPU 2 185
5.1.2 Docker 186
5.1.3 Slurm 187
5.2 Dashboard 188
5.2.1 Forge4Flame (F4F) 188
5.2.2 Architecture 190
5.3 Results 192
5.3.1 Case study: a general middle school environment 192
5.3.2 Comparison and validation: NetLogo vs. FLAME GPU 2 194
5.3.3 From macro- to micro-simulation 196
5.4 Discussion 201
5.5 Conclusion 202
5.6 Availability of data and software code 203

6 Conclusion and Future Works **204**
6.1 Future works 207

Appendices **209**

A Extended data analysis on NPIs, mobility, infection rates and regression models **209**
A.1 Extended data analysis 209
A.2 Extended results on regression models 218

B Forge4Flame User Guide **233**
B.1 How to download 233
B.1.1 Repository 233
B.1.2 Forge4Flame (F4F) 234
B.1.3 FLAME GPU 2 235
B.1.4 Dockers 235
B.1.5 Slurm 236
B.2 How to run 236
B.2.1 F4F 236
B.2.2 FLAME GPU 2 237
B.2.3 Docker Compose 239
B.2.4 Slurm 239
B.2.5 How to reproduce the results 241
B.3 Forge4Flame application 241
B.3.1 Homepage 241
B.3.2 Canvas 242
B.3.3 Rooms 245
B.3.4 Agents 247
B.3.5 Resources 251
B.3.6 Infection 253
B.3.7 What-If 256
B.3.8 Configuration 264
B.3.9 Run 264
B.3.10 Settings 266
B.3.11 Post Processing 267
B.4 School 273

B.5 Link between F4F and FLAME GPU 2 281

C Other Works **285**

C.1 Convolutional neural network (CNN)-enabled electrocardiogram (ECG) analysis: a comparison between standard twelve-lead and single-lead setups 285

C.2 PHeP: TrustAlert Open-Source Platform for Enhancing Predictive Healthcare with Deep Learning 289

List of Figures

1.1 John Snow's 1854 cholera map. 23
1.2 Continuous monitoring system. 27
1.3 Early warning system. 29
1.4 Computational models. 31
1.5 Knowledge extraction. 35
1.6 Road map of the thesis. 60

2.1 Graphical representation of a generic school. 64
2.2 The COVID-19 progression model. 66
2.3 R_0 comparison with state-of-the-art. 70
2.4 ABM validation. 71
2.5 NPIs vs no NPIs. 75
2.6 Screening policies comparison. 77
2.7 Number of days before detection of an infected student. 78
2.8 Cost function. 78
2.9 Cumulative distribution function (CDF) in low circulation. 82
2.10 Cumulative distribution function (CDF) in medium circulation. 83
2.11 Cumulative distribution function (CDF) in high circulation. 84
2.12 Area under the CDFs. 85
2.13 Distributions of infections in low circulation. 86
2.14 Distributions of infections in medium circulation. 87
2.15 Distributions of infections in high circulation. 88

3.1 Schematic representation of Sybil's steps. 95
3.2 Evolution of some variables in Italy. 104
3.3 First scenario in Italy. 105
3.4 Sybil applied on global infection, recovery and fatality rates of the first scenario in Italy. 106

3.5 Second scenario in Italy. 107
3.6 Third scenario in Italy. 108
3.7 Fourth scenario in Italy. 109
3.8 Fifth scenario in Italy. 109
3.9 Sixth scenario in Italy. 110
3.10 Evolution of some variables in Austria. 113
3.11 First scenario in Austria. 114
3.12 Second scenario in Austria. 115
3.13 Third scenario in Austria. 116
3.14 Fourth scenario in Austria. 116
3.15 Evolution of some variables in Italy (V = 11). 118
3.16 Evolution of the $I_v(\tilde{t})$ compartments in Italy moving the training window by one day (V = 11) 118
3.17 Evolution of the $\beta_{BA.1}(\tilde{t})$ and $\beta_{BA.2}(\tilde{t})$ infection rates in Italy moving the training window by one day (V = 11) 119
3.18 Results on Belgium. 121
3.19 France. 122
3.20 First two scenarios in Italy (fixed recovery rate). 124
3.21 Third scenario in Italy (fixed recovery rate). 125
3.22 Austria (fixed recovery rate). 126
3.23 Belgium (fixed recovery rate). 127
3.24 France (fixed recovery rate). 128
3.25 Germany (fixed recovery rate). 129
3.26 Sweden (fixed recovery rate). 130
3.27 Colorado, USA. 132
3.28 New York, USA. 133
3.29 West Virginia, USA. 134
3.30 Evolution of $I_v(\tilde{t})$ compartment in Italy (continuous monitoring system). . . 136
3.31 Evolution of $I_v(\tilde{t})$ compartment in the first ascending scenario in Italy (continuous monitoring system). 137
3.32 Evolution of $\beta_{Omicron}(\tilde{t})$ compartment in the first ascending scenario in Italy (continuous monitoring system). 138
3.33 Evolution of $I_v(\tilde{t})$ compartment in the second ascending scenario in Italy (continuous monitoring system). 139
3.34 Evolution of $\beta_{Alpha}(\tilde{t})$ compartment in the second ascending scenario in Italy (continuous monitoring system). 140

4.1 Temporal evolution of each variable. 155
4.2 Schematic representation of our study. 158
4.3 Heatmaps display country ratios. 162
4.4 Silhouette score computed using K-means. 164
4.5 Clusters derived using the average values with K-means. 165
4.6 Clusters derived using both average values and standard deviations with K-means. 166
4.7 Dimension of the spline basis vector. 168
4.8 Cluster separation measure (fDB) and total tightness. 169
4.9 Two, five, and six clusters for the $StringencyOnMobility$ ratio. 171
4.10 Two, five, and six clusters for the $MobilityOnInfRates$ ratio. 172
4.11 Two, five, and six clusters for the $StringencyOnInfRates$ ratio. 173

5.1 Diagram representation of the novel modeling paradigm. 182
5.2 Overview of F4F. 188
5.3 F4F architecture. 191
5.4 School and F4F. 192
5.5 ITP applied the scenario without NPIs. 195
5.6 ITP applied the scenario with NPIs. 196
5.7 Graphical schema of the *from macro to micro* approach. 197
5.8 Early warning results. 199
5.9 Percentage reduction in the infected population compared to the baseline. . . 200

A.1 Evolution of Facebook mobility variables, complement of the stringency index and infection rates for some countries. 210
A.2 Evolution of Facebook mobility variables, complement of the stringency index and infection rates for the remaining countries. 211
A.3 Evolution of average Facebook mobility, complement of the stringency index and infection rates for some countries. 212
A.4 Evolution of average Facebook mobility, complement of the stringency index and infection rates for the remaining countries. 213
A.5 Evolution of Google mobility variables, complement of the stringency index and infection rates for some countries. 214
A.6 Evolution of Google mobility variables, complement of the stringency index and infection rates for the remaining countries. 215
A.7 Evolution of average Google mobility, complement of the stringency index and infection rates for some countries. 216

A.8 Evolution of average Google mobility, complement of the stringency index and infection rates for the remaining countries. 217
A.9 Comparison between the ground truth and the predicted values from NPIs to mobility in Finland. 218
A.10 Comparison between the ground truth and the predicted values from NPIs to mobility in Norway. 219
A.11 Comparison between the ground truth and the predicted values from NPIs to mobility in Sweden. 220
A.12 Comparison between the ground truth and the predicted values from NPIs to mobility in Belgium. 221
A.13 Comparison between the ground truth and the predicted values from NPIs to mobility in Ireland. 222
A.14 Comparison between the ground truth and the predicted values from NPIs to mobility in the United Kingdom. 223
A.15 Comparison between the ground truth and the predicted values from NPIs to mobility in France. 224
A.16 Comparison between the ground truth and the predicted values from NPIs to mobility in Germany. 225
A.17 Comparison between the ground truth and the predicted values from NPIs to mobility in Austria. 226
A.18 Comparison between the ground truth and the predicted values from NPIs to mobility in Italy. 227
A.19 Comparison between the ground truth and the predicted values from NPIs to mobility in Greece. 228
A.20 Comparison between the ground truth and the predicted values from NPIs to mobility in Spain. 229
A.21 Comparison between the ground truth and the predicted values from NPIs to mobility in Portugal. 230
A.22 Comparison between the ground truth and the predicted values from mobility to infection rates in Finland, Norway, Sweden, Belgium, Ireland, and the United Kingdom. 231
A.23 Comparison between the ground truth and the predicted values from mobility to infection rates in France, Germany, Austria, Italy, Greece, Spain, and Portugal. 232

B.1 The Homepage page. 242
B.2 The Canvas page. 243

B.3 The Canvas page with empty canvas. . . . 244
B.4 The Canvas page with *floor0*. . . . 244
B.5 The Canvas page with *floor1*. . . . 245
B.6 The Rooms page. . . . 246
B.7 Definition of a room. . . . 246
B.8 Selection of rooms color. . . . 247
B.9 The Agents page. . . . 247
B.10 An example of the definition of an agent. . . . 248
B.11 An example of a determined flow. . . . 249
B.12 An example of a random flow. . . . 250
B.13 An example of a daily rate entry flow (slot 1). . . . 251
B.14 An example of a daily rate entry flow (slot 2). . . . 251
B.15 An example of a time window entry flow. . . . 251
B.16 The Resource page 252
B.17 The Set resources panel. . . . 252
B.18 Example of resources. . . . 253
B.19 The Infection page. . . . 254
B.20 The SEIRDS compartmental model. . . . 254
B.21 The What-If page. . . . 256
B.22 Setting ventilation globally. . . . 257
B.23 Setting ventilation specifically. . . . 257
B.24 Setting mask policy globally. . . . 257
B.25 Setting mask policy specifically. . . . 257
B.26 Setting vaccination policy globally. . . . 258
B.27 Setting vaccination policy specifically. . . . 258
B.28 Setting swab policy globally. . . . 259
B.29 Setting swab policy specifically. . . . 259
B.30 Setting quarantine policy globally. . . . 259
B.31 Setting quarantine policy specifically. . . . 260
B.32 Setting external screening policy globally. . . . 260
B.33 Setting external screening policy specifically. . . . 260
B.34 Saved countermeasures in Countermeasures. . . . 261
B.35 Example of filter on saved countermeasures. . . . 261
B.36 Defining the virus parameters. . . . 261
B.37 Setting random initial infected. . . . 262
B.38 Setting initially infected globally. . . . 262

B.39 Setting initially infected specifically. 262
B.40 Outside contagion panel. 263
B.41 Example of a plot generated after uploading an outside contagion file. 263
B.42 Saved countermeasures in Virus parameters. 263
B.43 The Configuration page. 264
B.44 The Run page. 264
B.45 Choose how to run FLAME GPU 2 simulations. 265
B.46 Pop-up displayed when the *Run* button is clicked. 266
B.47 The Settings page. 266
B.48 The Post Processing page. 267
B.49 Example of query. 268
B.50 Plots generated after selecting a listed directory. 269
B.51 The 2D visualization section. 271
B.52 The 2D visualization with room's names and one agent type. 271
B.53 The 2D visualization with room's area and one agent. 272
B.54 The 2D visualization with cumulative virus concentration. 272
B.55 The Contact Matrix panel. 273
B.56 *floor0* of the school environment visualized in the canvas of *F4F*. 275
B.57 *floor1* of the school environment visualized in the canvas of *F4F*. 275
B.58 *floor2* of the school environment visualized in the canvas of *F4F*. 276
B.59 Determined flow of agent *student_1A*. 276
B.60 Random flow of agent *student_1A*. 276
B.61 Entry flow of agent *student_1A*. 277
B.62 Determined flow of agent *teacher_0*. 277
B.63 Random flow of agent *teacher_0*. 277
B.64 Entry flow of agent *teacher_0*. 277
B.65 Infection parameters used to validate the school model. 278
B.66 Example of query on school. 278
B.67 Plots on school. 279
B.68 2D visualization of school with virus concentration. 280
B.69 Contact matrix on school. 280
B.70 All possible edges of the graph for *floor0*. 281
B.71 3D visualization in FLAME GPU 2 of the example used in the guide. 282
B.72 3D visualization in FLAME GPU 2 of the school. 282

C.1 Detailed CNN architecture. 286

C.2 Distribution of PTB-XL ECG diagnoses. 286
C.3 AUC percentage variation among single-lead (D1) and 12-lead scenarios. . . 287
C.4 Average AUC with 95% CI. 288
C.5 Overview of the infrastructure. 289

List of Tables

2.1 Model parameters. 73
2.2 Scenarios first set. 74
2.3 Model parameters first set. 75
2.4 Scenarios second set. 80
2.5 Model parameters second set. 80

3.1 RMSE for the first three scenarios considered in Italy. 111
3.2 RMSE for the last three scenarios considered in Italy. 112
3.3 RMSE for the scenarios considered in Austria. 117
3.4 RMSE for the scenarios considered in USA. 135

4.1 Spearman correlation among selected variables. 156
4.2 Average RMSE and NMAE across five folds for XGBoost models 174
4.3 Summary of key findings. 175

5.1 Shared model parameters in FLAME GPU 2 and NetLogo for comparison. . 193
5.2 Comparison between NetLogo and FLAME GPU 2. 194
5.3 Early warning scenarios. 198

C.1 Sample characteristics. 290
C.2 Prediction's accuracy. 290

Chapter 1

Introduction

Mankind has faced various sudden (earthquakes, tornadoes, hurricanes, tsunamis, floods) and slow (climate change, deforestation, environmental degradation, financial crises, wars, infectious diseases) **emergencies**. In recent decades, the intensity and geographic scope of these crises have dramatically increased, requiring a detailed reassessment of hazard, vulnerability, risk, and disaster management. In today's interconnected world, even localized events, like an epidemic or an environmental disaster, can quickly become global concerns, disrupting societies on multiple levels, leading to loss of life, economic collapse, displacement, and long-term instability, affecting supply chains, public health, and international relations. This is especially true if disaster mitigation, preparedness, response, and recovery are handled in an ad hoc and inappropriate manner. In these cases, they can cause more harm than benefit [144, 157].

Fighting emergencies demand a timely and coordinated response, guided by well-defined policies and procedures under the leadership of **policy-makers**—the key decision-makers during crises. A foundational element of effective disaster management is the integration of a **continuous monitoring system**, which delivers real-time data and analytics to track evolving conditions, anticipate emerging risks, and support proactive, informed decision-making. By providing actionable insights, such systems enable policy-makers to dynamically adapt strategies, allocate resources efficiently, and respond promptly to urgent needs.

Additionally, these systems must provide robust ***what-if*** **analysis** capabilities, allowing policy-makers to simulate various scenarios and assess the potential impact of different interventions before implementation. This capability helps in evaluating the effectiveness of proposed actions, optimizing decision-making, and preparing for potential scenarios. These capabilities are essential for achieving critical objectives during emergencies, such as safeguarding lives, restoring essential services and infrastructure, protecting vulnerable populations, and enabling swift recovery efforts. Ultimately, the success of disaster management—whether

Figure 1.1: John Snow's 1854 cholera map vividly illustrated the link between a cluster of cholera cases (red circles) and a contaminated water pump (in blue) on Broad Street, London, revolutionizing the understanding of disease transmission. Image from [247].

for routine incidents or large-scale emergencies—relies on the seamless organization, communication, and decision-making of leadership [144].

In this chapter, we first introduce epidemiology and computational epidemiology, the central focus of this thesis (Section 1.1). Next, we shift our attention to continuous monitoring systems and their key components (Section 1.2). Following this, we outline the main contributions of this thesis (Section 1.3) and review the state of the art (Section 1.4). Finally, a brief overview of the thesis structure is provided (Section 1.5).

1.1 Epidemiology

This thesis focuses on slow emergencies and, specifically, on infectious diseases like epidemics and pandemics. **Epidemiology** is the study of determinants, occurrence, distribution, and control of health and disease in a defined population. It is a descriptive science that involves determining rates, or quantifying how often a disease occurs within a particular population.

The most commonly studied rate is the *attack rate*, also known as the *incidence proportion*, which is calculated by dividing the number of disease cases by the population in which these cases have occurred. Epidemiology can provide a detailed description of a disease and various factors related to its occurrence, even before the cause of the disease is identified [59].

The roots of epidemiology can be traced back to ancient times, but it emerged as a formal scientific discipline in the 19th century, largely due to the work of John Snow in London [247]. Snow's investigation of the 1854 cholera outbreak provided empirical evidence for the link between contaminated water and disease transmission, laying the foundation for modern epidemiological methods—Figure 1.1. Today, epidemiology has evolved far beyond its origins, developing sophisticated analytical techniques and embracing new technologies to explore a wide range of health issues, from infectious diseases and chronic conditions to behavioral and environmental health risks—in fact, there has been a transition from traditional epidemiology to **computational epidemiology**. Epidemiology is multidisciplinary and integrates principles from bio-statistics, medicine, social and environmental sciences, sociology, genetics, and biology to address complex public health challenges and design effective disease prevention and control strategies [108].

1.1.1 Computational epidemiology

Computational epidemiology is an interdisciplinary research area that aims to develop and use computational models and algorithms to understand and control the spread of a disease in a population [174, 175]. The essence of computational epidemiology lies in using these tools to assist policy-makers in making informed decisions to anticipate, manage, and simulate an incoming emergency. This research area comprises a wide range of approaches to address various health problems [59, 192, 163]:

i) *Descriptive epidemiology* aims to study the distributions of diseases and determinants offering a way for organizing and analyzing data to describe variations in disease recurrence across different populations, geographical areas, and periods (i.e., person, place, and time).

ii) *Analytical epidemiology* aims to examine the associations between exposures (such as risk and protective factors for diseases) and outcomes (such as incidence, prevalence, and mortality) and is used to test research hypotheses regarding potential causal effects and/or associations between exposures and outcomes.

iii) *Experimental epidemiology* aims to develop an experimental model in which one or more selected factors are manipulated and the impact of this manipulation is then observed to

either confirm or refute the hypothesis.

Moreover, the area of computational epidemiology is increasingly recognizing the effects of social-structural factors on states of health. The primary focus of *social epidemiology* is to examine how society and various forms of social organization affect the health and well-being of individuals and populations by investigating the frequency, distribution, and social determinants of health. Consequently, social epidemiology extends beyond the analysis of individual risk factors, incorporating the study of the broader social context in which health and disease occur. Research has demonstrated that factors such as socioeconomic status, education, and access to healthcare play a crucial role in shaping health outcomes [54, 134, 178, 30].

Computational epidemiology extends beyond infectious diseases to encompass chronic conditions, behavioral health issues, and environmental factors. For instance, the field has contributed to understanding the risk factors for chronic diseases such as cardiovascular disease, cancer, and diabetes. Epidemiological research has identified key determinants of these conditions, including lifestyle factors such as diet, physical activity, and smoking, as well as genetic predispositions and environmental exposures. This knowledge has informed public health campaigns, clinical guidelines, and policy decisions aimed at reducing the burden of chronic diseases and improving population health [155, 55].

1.1.2 Brief pandemic history

In recent decades, the field of computational epidemiology has expanded to address emerging and re-emerging health threats, including antibiotic resistance [119], the impacts of climate change [287], and pandemics [266]. As we are more interested in the latter, the 20^{th} and 21^{st} centuries have witnessed several significant pandemics [210]. Below is a summary:

- *Spanish flu (1918–1919)*: caused by the *H1N1* influenza virus, it infected a third of the world and killed 20 to 50 million people. Unlike typical flu, it hit young adults the hardest. It spread rapidly during World War I, with no vaccines or treatments available, prompting lasting advancements in public health.

- *Asian Flu (1957)*: attributed to the *H2N2* influenza virus, this pandemic emerged in East Asia before rapidly spreading across the globe. It caused extensive health crisis and substantial disruption, resulting in an estimated 1 to 2 million deaths worldwide.

- *Hong Kong Flu (1968)*: triggered by the *H3N2* influenza virus, this pandemic infected millions worldwide and led to around 1 million deaths, highlighting the persistent threat of influenza viruses.

- *Ebola (1976)*: it was first identified in the Democratic Republic of Congo and has since caused several outbreaks, primarily in Central and West Africa. It spreads through direct contact with infected bodily fluids, leading to severe hemorrhagic fever, with high mortality rates in many outbreaks. Researchers have worked on treatments and vaccines to combat the virus, notably the rVSV-ZEBOV vaccine, which has proven effective in controlling recent outbreaks.

- *HIV/AIDS (1980s)*: this pandemic marked another profound global health crisis, as the virus spread rapidly across continents, leading to over 36 million deaths to date. The pandemic transformed public health practices, promoting global awareness of sexually transmitted infections, blood safety, and the critical need for antiretroviral therapies.

- *Swine Flu (2009)*: caused by the *H1N1* influenza virus, the Swine Flu pandemic rapidly spread to nearly every country across the globe. Within its first year, it led to prevalent illness and an estimated 200.000 deaths. The virus particularly affected younger populations, and its rapid transmission highlighted the need for global public health responses, including vaccination campaigns and travel restrictions.

- *SARS (2002-2003) and MERS (2012)*: the Severe Acute Respiratory Syndrome (SARS) and Middle East Respiratory Syndrome (MERS) outbreaks demonstrated the global threat of coronaviruses. SARS, originating in China, spread to 26 countries, causing around 800 deaths, while MERS, first detected in Saudi Arabia, had a higher fatality rate of 35%, primarily affecting the Middle East but reaching 27 countries.

1.1.3 COVID-19

The **COVID-19** pandemic, which began in late 2019 and ended in May 2023, has led to widespread illness, economic disruption, and over 7 million deaths globally to date, demonstrating the far-reaching impact of emerging infectious diseases in a highly interconnected world. This pandemic has highlighted the critical role of computational epidemiology in monitoring disease outbreaks, assessing the effectiveness of public health measures, and guiding vaccination strategies. The integration of advanced technologies, such as digital health tools and big data analytics, has enhanced the ability of epidemiologists to track and analyze health trends in real-time, providing valuable insights for timely and effective responses to public health crises [97].

1.2 Continuous monitoring system

The impact of emergencies can be reduced by improving the international exchange of information and developing **continuous monitoring and early warning** systems. A continuous monitoring system (CMS) helps in forecasting potential surges, managing containment efforts, and assessing the effectiveness of public interventions. This approach not only aids in containing the emergency but also ensures efficient resource allocation and informs ongoing policy decisions, ultimately contributing to global security. This solution plays a crucial role by integrating data from diverse sources, such as public reports, hospitals, and digital tools like mobile apps and social media platforms. Leveraging recent technological advancements—including big data analytics, artificial intelligence, and Geographic Information Systems (GIS)—has greatly enhanced the CMS's capabilities, enabling real-time tracking of emergency trends, identifying critical hot spots, and predicting potential future outbreaks [63].

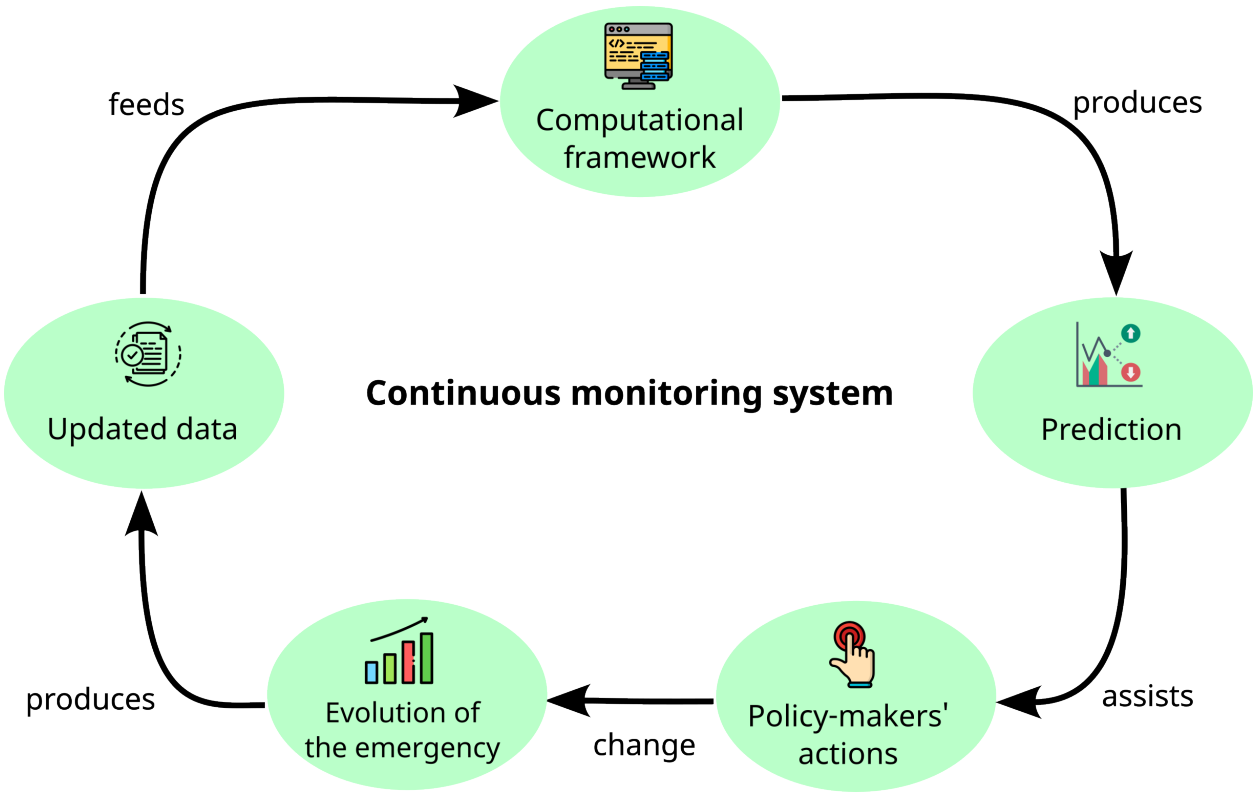


Figure 1.2: A continuous monitoring system utilizes a computational framework to analyze emergency dynamics, generating predictions to guide public decisions and policies. As interventions are implemented, the system monitors the emergency progress through continuous data collection on selected metrics and updates its predictions based on this ongoing data.

Specifically, a CMS integrates computational modeling, data analytics, and adaptive decision-making—Figure 1.2 shows an example of a general schema. Here is an overview of a CMS that could be implemented during an infectious disease emergency, like a pandemic:

this system begins with *i)* collecting the necessary data to *ii)* feed into a computational framework, which may use artificial intelligence to analyze the dynamics of the disease, such as transmission rates, affected populations and regions, and potential future scenarios; based on this analysis, *iii)* the model generates predictions that *iv)* inform public health decisions and policy-makers actions, such as implementing or lifting restrictions, modifying travel regulations, or optimizing vaccine distribution. As these interventions are put into effect, *v)* the monitoring system tracks the evolution of the pandemic through ongoing data collection on key metrics like new cases, hospitalizations, recoveries, and fatalities. This updated data is continuously *vi)* fed back into the computational framework, which re-calibrates its predictions to reflect the current situation. This feedback loop allows for real-time adjustments to public strategies, ensuring a flexible and responsive approach to emergency management. By maintaining a dynamic connection between predictive modeling, policy actions, and up-to-date data, a continuous monitoring system enables timely and effective responses to the ever-changing landscape of the emergency.

Government interventions can be categorized into two distinct groups: pharmaceutical and non-pharmaceutical. Pharmaceutical interventions (PIs) include medical treatments aimed at preventing, managing, or healing diseases, such as vaccines, antiviral medications, monoclonal antibodies, immunomodulators, and antibiotics. On the other hand, non-pharmaceutical interventions (NPIs) do not involve medical treatments but focus on strategies such as social distancing, quarantine and isolation, screening campaigns, travel restrictions, mask-wearing, hygiene measures, remote work, closure of educational institutions and non-essential businesses, limitations on public gatherings, and contact tracing. These interventions are implemented to prevent or control the spread of diseases, especially during outbreaks, by reducing transmission through changes in behaviors, environments, and community interactions. For instance, during the COVID-19 pandemic, the implementation of NPIs has varied significantly across different regions and periods, reflecting diverse epidemiological, political, and socio-economical contexts [204].

A continuous monitoring system must facilitate **early warnings** to enhance the ability to prevent worst-case scenarios (Section 1.2.1). These systems utilize a variety of tools and techniques to enhance their effectiveness including **computational models** (Section 1.2.2) that simulate scenarios like disease outbreaks or natural disasters and **knowledge extraction** methods (Section 1.2.3) to analyze large datasets and identify key patterns using **(big) data analysis** and **artificial intelligence (AI)** to detect anomalies and predict emerging risks in real-time. These techniques enable the generation of timely alerts, allowing for a more effective and proactive response to new emergencies. In the following sections, we will provide a detailed description of the tools utilized in this thesis, highlighting their functionalities,

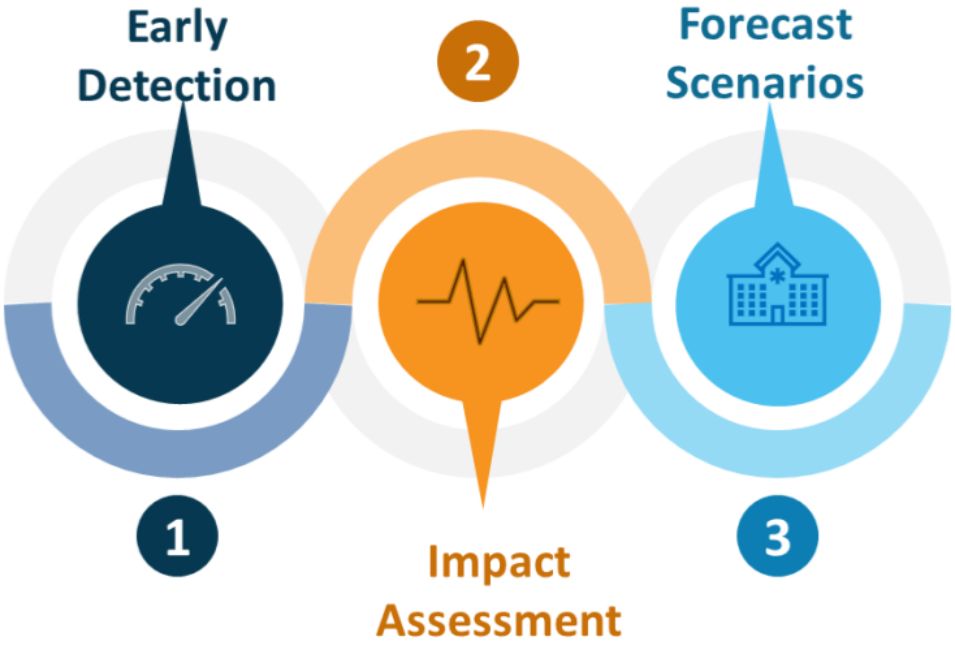


Figure 1.3: Primary objectives of an early warning system: early detection, impact assessment, and forecast scenarios. Logo of the TrustAlert [212] project.

along with their advantages and limitations.

1.2.1 Early warning system

A continuous monitoring system enables the development of an early warning system, a critical tool for coordinating international efforts to anticipate and effectively manage such emergencies. They rapidly share data, standardize reporting, and facilitate global collaboration, ensuring an effective and timely response [214]. In particular, the primary objectives of an early warning system are as follows—Figure 1.3:

- **Early detection**: minimize the delay between detecting a new emergency and implementing public interventions, enhance surveillance systems to swiftly identify and respond to unusual patterns, and improve diagnostic capabilities to prevent the worst scenarios. This step could leverage a near-real-time stream of information acquired by external services, like the GDELT Project [159], and curated information extracted from emergency department records.

- **Impact assessment**: measure the potential burden on healthcare systems to ensure preparedness and capacity, evaluate economic impacts (such as disruptions to businesses and the labor market), inform economic recovery plans, and assess social consequences—

including effects on education and mental health—, to develop targeted support services. This step could involve mapping population vulnerabilities to evaluate the local impact of the emergency, alongside processing real-time data from regional activity flows.

- **Forecast scenarios**: anticipate the trajectory of an outbreak and provide timely insights into its potential spread, duration, and peak, enable proactive planning and allocation of resources to areas at greatest risk, assess the potential economic and social impact of the emergency, allowing for early preparation and mitigation strategies, and improve models for predicting the effects of various interventions and responses, leading to more informed public decisions. This step could involve forecasting scenarios at multiple levels of granularity to support decision-making, including conducting cost-effectiveness analyses of different intervention strategies.

The cost of implementing an early warning system can exceed its benefits when it is primarily utilized to mitigate rare events, such as those with a low probability of occurrence, or to address false alarms. For example, if the financial damage associated with such an event significantly exceeds the cost of the system on an annual basis, the investment may be deemed inefficient. An early warning system would be economically viable only if the frequency of the event was higher, the potential damages were substantially greater, or the cost of the system was considerably reduced [230]. However, the increasing prevalence of extreme events and emergencies, driven by climate change, highlights the growing importance of such systems. These extreme events are often accompanied by significant economic, social, and environmental costs, further underscoring the necessity of investing in early warning mechanisms to mitigate their impacts [265].

1.2.2 Computational models

Computational models are sophisticated tools that use mathematical equations and algorithms to simulate complex systems and processes. These models replicate real-world phenomena by integrating data and theoretical principles, allowing researchers and decision-makers to simulate, analyze, predict, and understand various scenarios and outcomes. Computational models are essential for unveiling patterns and guiding informed decisions in complex and evolving environments [252, 67, 225]. They offer valuable predictive insights and facilitate scenario analysis, but their accuracy depends on data quality and assumptions, and they may require significant computational resources.

In modeling mathematical and statistical epidemiology, the primary distinction is drawn between **deterministic and stochastic models**. Deterministic models, such as systems of ordinary differential equations (ODEs), represent how average quantities evolve within a

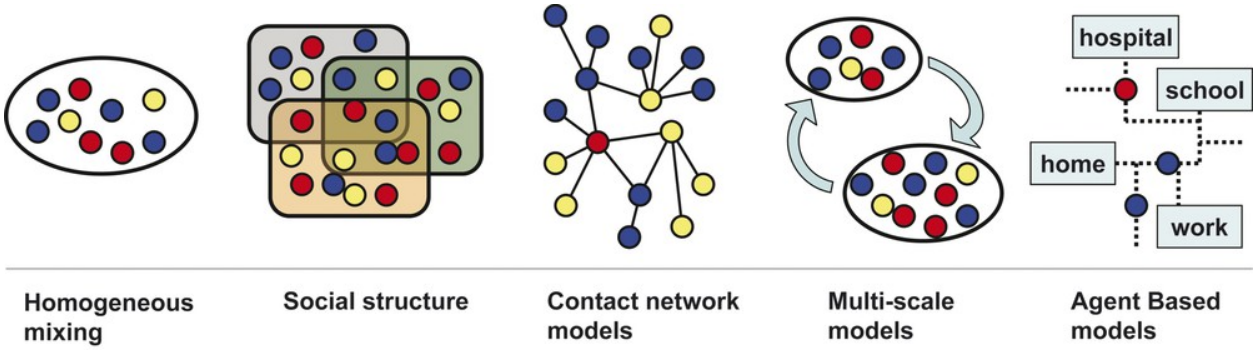


Figure 1.4: Different types of computational models are employed to study social interactions and their effects. Image from [83].

population. These models assume smooth and predictable progressions and do not explicitly account for randomness. Deterministic models are computationally effective and are ideal for looking at large populations or when overall trends are the focus. Nonetheless, they might not capture random variations, particularly in small populations or at the early stages of an outbreak. On the other hand, stochastic models introduce randomness, making them crucial for systems where uncertainty and variability are significant. The stochastic temporal development of such a system is described by Continuous-Time Markov Chains (CTMCs), which depict the probabilities associated with different system states over time. Due to the state-space explosion issue, solving CTMCs analytically is not feasible for complex systems. Instead, numerical methods like the Stochastic Simulation Algorithm (SSA) [112] are used to simulate trajectories that reflect the CTMCs' solution, accurately incorporating system randomness. To mitigate SSA's high computational demands, approximation strategies like the τ-leaping algorithm [113] offer a balance between accuracy and efficiency by combining stochastic and deterministic characteristics.

Both deterministic and stochastic models can be utilized for **macro- and micro-simulations**, based on research goals, system specifics, and data accessibility. Macro-simulations emphasize the aggregate behavior of the system. Here, deterministic ODEs-based compartmental models such as Susceptible-Infected-Recovered (SIR) models often illustrate broad trends. For scenarios where randomness is crucial, like rare events or small populations, stochastic macro-simulations using CTMCs or SSA provide a more precise depiction. Micro-simulations focus on the detailed interactions between individual components, for example, people or pathogens. Typically, these are implemented as Agent-Based Models (ABMs), where each agent follows defined rules and interacts with others, leading to emergent system-wide behaviors. Deterministic simplifications in ABMs can streamline interactions in extensive systems, whereas stochastic versions preserve the ability to model variability and uncommon events, which are vital for understanding the diversity in disease transmission.

Figure 1.4 illustrates five different categories of computational models, ranging from macro

to micro, each with progressively increasing computational costs [83]:

- **Homogeneous mixing**: assumes that every individual in a population is equally likely to interact with every other individual and treats the population as a well-mixed group without any structure, where interactions happen randomly and uniformly.
- **Social structure**: accounts for the fact that individuals belong to different social groups or communities, based on factors such as age, economic status, region, and other demographic characteristics. It represents multiple overlapping groups, where each group has its internal interactions—represented by homogeneous mixing—, and considers the differences in contact patterns within and between these groups, reflecting more realistic social structures.
- **Contact network models**: represent the population as a network, where nodes represent individuals, and edges represent the interactions or connections between them. They capture the complex patterns of interactions, allowing for a more detailed understanding of how diseases or information spread through social networks.
- **Multi-scale models**: integrate multiple levels of interaction, such as individual, group, and population levels, and consider the effects across these different scales. For example, the model may combine homogeneous mixing for some groups with network-based or structured mixing for others, providing a more nuanced approach to modeling real-world interactions.
- **Agent-based models (ABMs)**: simulate the actions and interactions of individual agents (such as people or organizations) within specific environments. They consider different locations or settings (e.g., city, neighborhood, hospital, school, work, home) and model how agents move between these locations and interact based on defined rules. This approach captures individual behavior and decision-making, allowing for a very detailed simulation of complex social systems.

Compartmental models are widely recognized and employed to study the spreading of an infection in a population [60, 261]. They are easy to explain: the model contains several compartments, each representing a specific sub-population, and uses rates to move individuals from one compartment to another. The simplest compartmental model to mimic the evolution of an infection in a population requires only two compartments and one rate; the two compartments represent the only two possible states of the individuals in the population—either susceptible (S) or infected (I)—and one transition rate that represents the pace at which one infected individual infects a susceptible one. Because of the two states

characterizing this model, it is commonly known as the Susceptible-Infected (SI) model. To capture critical features of the disease under study, an SI model may be insufficient, requiring a more complex model to represent the dynamics of infection better. More advanced models can incorporate additional factors such as incubation periods, vaccination effects, virus variants, and demographic considerations like age groups and symptom severity. They can also account for human mobility, behavioral changes, and public health interventions, such as movement restrictions, offering a more comprehensive understanding of the epidemic.

Compartmental models are easy to understand and interpret, are computationally efficient, can be used to stratify the population at different levels (e.g., by age, place, health status), and provide a comprehensive mathematical representation of epidemic progression. This level of transparency simplifies explaining and interpreting the results to various stakeholders, such as policy-makers, healthcare professionals, and the general public. Consequently, it fosters a better understanding of the situation and facilitates more informed decision-making. However, despite these advantages, these models are not without their challenges. They rely heavily on a set of assumptions about the system being modeled, which may not always hold in real-world scenarios. Additionally, the parameter estimation process presents significant challenges, as accurate data collection and validation are crucial for the reliability of the models.

Agent-Based Models (ABMs) [94] refer to a category of computational models invoking the dynamic actions, reactions, and intercommunication protocols among the agents in a shared environment, to evaluate their design and performance and derive insights on their emerging behavior and properties. ABMs are defined as data structures, where agents are modeled by a set of variables and interaction rules stated consistently with the programming language used for the simulations, in which they act and adapt their internal and external behavior. The modeler defines these rules to obtain realistic simulations of behavior within the system under consideration—these rules can be simple (e.g., backward induction on an extensive game form) or more complex (e.g., based on heuristics derived from cognitive psychology or neuroscience). These models can also include a description of the spatial domain, either discrete or continuous. In such cases, agents might be allowed to move throughout the environment interacting with it and other agents depending on their physical proximity by receiving inputs and responding with actions based on those inputs [172, 94].

ABMs can capture complex behaviors allowing a realistic representation of interactions and discovering global patterns (emerging behaviors) that arise from local interactions. They are flexible and adaptable tools that allow the definition of agents with different characteristics, behaviors, and decision-making processes. Additionally, they can easily incorporate changing environments and evolving rules. On the other hand, ABMs are computationally expensive

especially when simulating large populations or complex interaction rules, and hard to calibrate and validate—real-world benchmarks are not always available or easily measurable and these models have many parameters and rules. Moreover, they require lots of assumptions, often produce results that are only available through simulation, and lack analytical solutions—it is possible to formalize an agent-based model through an analytical model only under specific assumptions [50].

Another approach is to **integrate** micro- and macro-simulations, which offers the potential to enhance both robustness and reliability by leveraging the strengths of each scale. This integrated method allows for a more comprehensive analysis, capturing fine-grained details through micro-simulations while maintaining the broader, systemic insights provided by macro-simulations. As a result, it surpasses the limitations of individual simulation paradigms, providing a more accurate and resilient modeling framework for studying complex infectious disease dynamics.

1.2.3 Knowledge extraction

Data collected while monitoring the evolution of an emergency is heterogeneous, uncertain, and incomplete. For instance, in the context of infectious diseases, this data includes but is not limited to, information on reported cases, fatalities, recoveries, vaccination rates, hospitalizations, intensive care unit admissions, people's mobility, and the implementation of public health measures [170, 116].

From this data, various metrics can be calculated to monitor the progression of the emergence, such as incidence—the fraction of new cases in a population over a specific period—, prevalence—the fraction of the total number of cases (new and existing) in a population over a specific period—, case fatality rate (CFR)—the proportion of people who have been diagnosed with the disease and end up dying of it—, infection fatality rate (IFR)—the proportion of deaths among all infected individuals, including all asymptomatic and undiagnosed subjects—, basic reproduction number (R_0)—expected number of cases directly generated by one case in a population where all individuals are susceptible to infection—, effective reproduction number (R_t)—average number of new infections caused by a single infected individual at time t in the partially susceptible population—, hospitalization rate, and outbreak size.

However, some metrics can be challenging to compute, while others may be impossible to determine without sufficient or accurate data. For these reasons, these metrics are typically estimated using data collected from a sample of the affected population. This approach allows for estimates when complete data from the entire population is unavailable or difficult

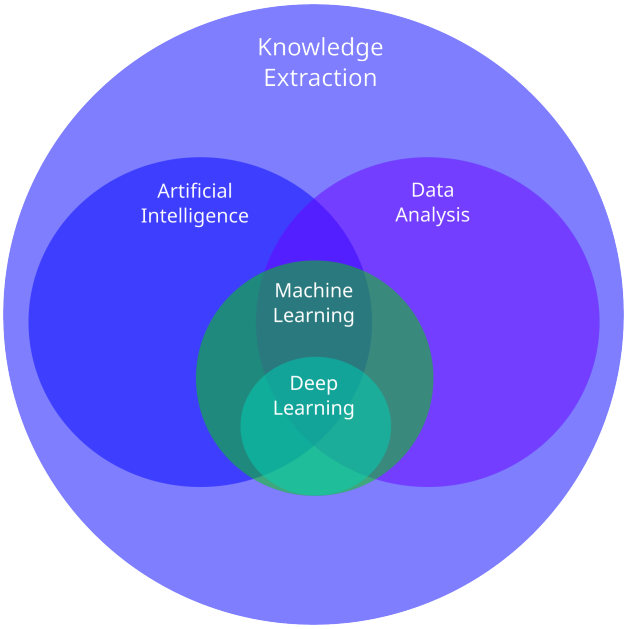


Figure 1.5: Knowledge extraction encompasses data analysis and artificial intelligence (AI), with AI including machine learning, and machine learning further containing deep learning as its most specialized subset.

to obtain. Specifically, **knowledge extraction** techniques can be employed to achieve this goal. Using (big) data analysis and artificial intelligence, knowledge extraction techniques can extract knowledge from structured (relational databases, CSV, XML) and unstructured (text, documents, images) sources and uncover hidden relationships, enabling the refinement and adaptation of the used models [264]—Figure 1.5.

Data analysis

Data analysis is essential in managing emergencies, as it provides valuable insights that inform public health decisions and help establish a continuous monitoring system. According to [242], various analytical procedures *"provide a way of drawing inductive inferences from data and distinguishing the signal (the phenomenon of interest) from the noise (statistical fluctuations) present in the data"*. In the epidemiological context, it involves collecting, processing, and interpreting data from various sources such as hospitals, laboratories, and public health agencies. For example, during the COVID-19 pandemic, governments relied on health data from testing centers and hospitals to track infection rates, hospitalizations, and fatalities.

A common analytical approach is descriptive analysis [158], which summarizes data using statistical metrics like averages and trends to understand the scale and spread of an outbreak. This approach helps identify patterns, such as rising infection rates or declining recovery times. Geo-spatial analysis [95] also plays a key role in mapping the geographic spread of

disease. Visual representations such as heatmaps were widely used during COVID-19 to highlight outbreak hotspots and allocate medical resources effectively. Another important technique is contact tracing [100], which tracks the spread of a disease through social networks and surveys. Health authorities use this method to identify and notify individuals exposed to infected persons, helping to contain the outbreak. Predictive models take these analyses further by using historical data to forecast future trends through methods like time-series forecasting and machine learning—see the next section for more details.

Data analysis is vital in managing epidemics by enabling evidence-based decision-making and identifying disease trends. It helps detect outbreaks early, allocate resources efficiently, and target interventions to reduce disease spread. Continuous monitoring allows for evaluating and refining response strategies in real-time. Despite challenges like ensuring data accuracy, maintaining privacy, and providing timely insights, data analysis remains essential for effective epidemic management and minimizing public health impacts.

Artificial intelligence

Artificial intelligence (AI) plays a crucial role in enhancing the effectiveness and efficiency of a continuous monitoring system, enabling real-time tracking, analysis, and response to dynamic data streams. By incorporating AI, these systems can autonomously make decisions, detect anomalies, and optimize processes without human intervention.

AI techniques, such as machine learning (ML), natural language processing (NLP), computer vision, and reinforcement learning offer CMSs the ability to analyze vast amounts of data continuously, learning from historical patterns and making predictions about future behavior [87, 142]. Specifically, in the context of an epidemic:

- Machine learning (ML) can predict outbreak patterns, identify infection hotspots, and detect emerging health threats by analyzing historical health data. For example, ML models were employed during the COVID-19 pandemic to forecast infection surges and allocate healthcare resources effectively using techniques like auto-regressive and time-series predictive models, random forests, and clustering. These models are trained on past data, such as infection rates and population density, to predict future trends and inform decision-making. However, ML models rely heavily on high-quality data, and incomplete or biased datasets can compromise their predictive accuracy. Moreover, many models operate as *black boxes*, making their decision-making processes difficult to interpret [56, 284].

- Natural language processing (NLP) can analyze public health communications, monitor social media for outbreak-related discussions, and detect misinformation. During the

Zika virus outbreak, NLP tools like sentiment analysis and topic modeling helped extract critical data from online reports, enabling faster response planning. By processing vast amounts of social media data and news articles, NLP techniques helped track public perception and concerns, ensuring timely and targeted public health messaging. Nonetheless, NLP models can struggle with language nuances, context interpretation, and the detection of evolving slang, which can limit their effectiveness in real-time monitoring [270].

- Computer vision can analyze medical images such as chest X-rays to detect conditions such as pneumonia linked to respiratory infections. Techniques like image classification using convolutional neural networks (CNNs) were used during COVID-19 to facilitate rapid and accurate patient diagnoses. Transfer learning methods allowed pre-trained models to be adapted for COVID-19-related tasks, speeding up diagnosis in high-demand settings. However, computer vision models require large labeled datasets for training, which may not always be available. Errors in image interpretation could lead to misdiagnoses, raising concerns about reliability in critical healthcare contexts [160].

- Reinforcement learning can optimize epidemic response strategies by learning from past actions. For example, reinforcement learning methods like Q-learning and Deep Q-Networks (DQN) can help adjust vaccination campaigns by reallocating doses based on real-time infection rates and vaccination coverage. By dynamically adjusting strategies based on observed outcomes, reinforcement learning ensures that resources are distributed efficiently to achieve maximum impact in controlling the epidemic. Nevertheless, these models require extensive computational resources and can be challenging to deploy in regions with limited technical infrastructure. Additionally, their reliance on historical data can limit adaptability when faced with unprecedented scenarios [191].

By leveraging these AI technologies, CMSs can significantly improve their ability to monitor complex environments, respond to challenges in real-time, and optimize performance autonomously. The integration of AI not only enhances the decision-making process but also enables predictive and prescriptive capabilities, making systems more adaptive and intelligent. However, the success of AI-based epidemic management depends on addressing potential limitations, including data quality issues, ethical concerns about data privacy, and the need for transparent, interpretable models.

1.2.4 Combination of multiple approaches

We began by exploring the concept of continuous monitoring and early warning systems, highlighting their importance in various applications. From there, we delved into the range of tools and technologies essential for setting up such systems. In particular, we focused on the role of computational models and methods for knowledge extraction, which include advanced (big) data analysis techniques and the integration of artificial intelligence to enhance system accuracy and effectiveness. Each technique has its own set of advantages and disadvantages when evaluated in isolation. For instance, computational models can provide accurate simulations and predictions but may require extensive data and computational power. On the other hand, knowledge extraction techniques can uncover hidden patterns in large datasets, yet they may struggle with incomplete or noisy data. Similarly, artificial intelligence can offer highly adaptive and efficient solutions but may introduce challenges in terms of interpretability and the need for continuous training. However, integrating different models allows us to leverage their strengths and mitigate their weaknesses. In particular, an integrated model consists of two or more independent components, each capable of functioning autonomously and often developed independently by different research teams. The benefit of integration lies in its ability to re-purpose existing code and models for analyzing more complex systems. This approach simplifies the analysis by allowing the overall system to be examined and modeled in manageable segments [268].

1.2.5 Adapting pandemic solutions to broader challenges

This thesis focuses on infectious diseases and computational epidemiology due to the outbreak of the COVID-19 pandemic that started in 2019. In the last few years, we faced an unprecedented global crisis that required rapid and innovative solutions. The tools and methodologies we developed during that time were initially designed to address the urgent need for monitoring and responding to the pandemic. However, the systems we put in place may be versatile and capable of addressing challenges beyond the health sector. The continuous monitoring tools, in particular, are applicable to a variety of other situations where real-time data collection and analysis are critical. Moreover, the combination of methodologies we employed, such as data integration, real-time analytics, and predictive modeling, can be effectively applied in different domains. For example, in the field of cybersecurity, these same approaches can help identify threats and vulnerabilities, enabling organizations to detect and respond to incidents in real-time. Similarly, in city planning, these approaches can assist in optimizing urban infrastructure, traffic flow, and resource allocation by analyzing real-time data on population movement and environmental factors. This adaptability demonstrates

the broader value of these tools, which extend beyond their original use case in the pandemic to other critical areas requiring vigilance and proactive intervention.

1.3 Contribution

This thesis focuses on **computational epidemiology** with the main aim to **forecast scenarios** using a combination of **computational models** and **knowledge extraction** techniques, concentrating on COVID-19 as a case study. The main motivation is to develop evidence-based recommendations for policy-makers creating guidelines—based on *data* and *model predictions*—to assist them in implementing the most effective and sustainable public health measures during a pandemic. These guidelines will focus on optimizing intervention strategies, allocating resources efficiently, and ensuring that policies are adaptable to evolving pandemic conditions. By incorporating continuous monitoring systems, the process will integrate real-time data collection and ongoing model predictions to inform decisions, ensuring that policies are responsive to changing circumstances and based on the latest available evidence. In particular, this thesis encompasses the research and contributions made throughout the three-year duration of the Ph.D. program:

i) We study the evolution of COVID-19 spread in a well-defined closed environment using an agent-based model to assess the efficacy of different Non-Pharmaceutical Interventions (NPIs)—such as social distancing, room's air ventilation, masks, quarantine, and screening campaigns—and vaccination strategies—including vaccine distribution methods and coverage levels—, showing how these affect transmission rates influencing the spread of the disease in such environments and investigating how the interplay between NPIs and vaccination affects overall outbreak control [41, 40].

ii) We predict the future evolution of a pandemic using a novel framework that integrates two distinct approaches: a compartmental model and a machine learning-based predictive model. This hybrid framework also incorporates data on different variants of the virus under study, enhancing its accuracy [39].

iii) We explore the relationship among country-imposed restrictions, people's mobility, and the effective reproduction number R_t of COVID-19. By leveraging restrictions and mobility data from major platforms, we analyze how movement trends vary in response to policy changes and restrictions. This approach enables us to understand the dynamics of disease spread across various countries and demographics, identifying key behavioral patterns and responses to interventions.

iv) We address a significant limitation of Agent-Based Models (ABMs) used for infectious disease modeling: the challenge of creating realistic well-defined closed environments, particularly for complex spaces such as offices, schools, hospitals, or urban areas. These

models often require extensive manual setup, which can be time-consuming and error-prone. To overcome this, we have developed a tool that simplifies the definition of ABMs without requiring coding.

The following sections provide a brief overview of each contribution, highlighting the key findings.

1.3.1 Contribution 1: *Living along COVID-19: assessing contention policies through Agent-Based Models*

We use **Agent-Based Models (ABMs)** to evaluate the efficacy of various NPIs and vaccination strategies in controlling COVID-19 transmission in a well-defined closed environment. As a case study, we consider a generic middle school environment and assess the infection risk associated with school activities, along with the effects of potential control measures following the introduction of infectious cases [41, 40]. Key findings include:

- The combination of mask-wearing and room air ventilation can significantly reduce virus transmission, with simulations demonstrating a 79% decrease in cumulative infections compared to a scenario without NPIs.
- Screening campaigns play a crucial role in reducing transmission by identifying and isolating cases early, but their effectiveness depends on rapid implementation that overcomes the spread of the virus.
- Vaccination campaigns covering less than 40% of the susceptible population are insufficient to prevent major outbreaks.
- Proper mask usage is crucial, even during periods of rapidly increasing infection rates.
- High virus circulation outside schools poses a greater challenge, but combining masks with screening and quarantine policies can still mitigate the spread.

1.3.2 Contribution 2: *Sybil: enhancing COVID-19 forecasting precision with variant-aware integrated models*

Sybil [39] is an integrated framework that combines a machine learning-based predictive model with a variant-aware compartmental model to improve the accuracy and explainability of disease spread forecasts. By integrating a data-centric machine learning approach with an analytic compartmental model, Sybil is designed to enhance public health decision-making,

enabling proactive measures such as vaccination campaigns, travel advisories, and community engagement programs. Key findings from the development and application of Sybil include:

- Reliable predictions can be achieved even in the face of substantial changes in disease dynamics, offering robust forecasting capabilities for real-time adjustments to control strategies.
- The framework requires less training data compared to traditional models, making it accessible for implementation in scenarios where data is limited or incomplete.
- The ability to incorporate and track multiple virus variants allows for a more nuanced understanding of transmission dynamics, crucial for adjusting strategies in response to new variants.
- Sybil consistently outperforms a variety of state-of-the-art approaches, demonstrating superior performance in prediction accuracy.

1.3.3 Contribution 3: *Social compliance with NPIs, mobility patterns, and reproduction number: lessons from COVID-19 in Europe*

We examine the relationship between government-imposed restrictions, mobility patterns, and the effective reproduction number R_t of COVID-19. We analyze mobility and restrictions data from major platforms to understand how movement trends respond to policy changes, providing insights into disease transmission dynamics across diverse countries and demographics. Our focus is on behavioral patterns and intervention responses. Specifically, we analyze data from several European countries, including Finland, Norway, Sweden, Belgium, Ireland, the United Kingdom, France, Germany, Austria, Italy, Greece, Spain, and Portugal. Key findings include:

- Identify a strong correlation between the considered restrictions and mobility variables, contrasted by a weaker correlation between mobility variables and infection rates.
- Identify cultural patterns by clustering European countries using K-means [125] and CONNECTOR [203], specifically:
  - Observe a clear separation between Northern and Southern Europe in terms of social compliance with NPIs.
  - Found no such distinction when social interactions were the dominant factor.

- Applied XGBoost [77] to estimate missing mobility data and infer missing R_t values, representing an innovative methodological contribution.
- Highlighted the importance of tailoring public health strategies to cultural and regional differences, particularly regarding compliance with NPIs.

1.3.4 Contribution 4: *Forge4Flame: an intuitive dashboard for designing GPU Agent-Based Models to simulate the spread of infectious diseases*

We implement ***Forge4Flame (F4F)***, a user-friendly dashboard—developed in R Shiny—designed to simplify the definition of an agent-based model for FLAME GPU 2 [223, 222] ABMs, automatically generating the necessary code. This tool enables users to define the simulation environment, agent interactions, disease models, and *what-if* scenarios for interventions such as NPIs and vaccination strategies, all through a graphical interface without coding. Key findings include:

- First graphical interface to design and run agent-based model simulations on GPUs—this tool does not have an equivalent counterpart in the existing literature.
- A significant speed-up is demonstrated using FLAME GPU 2 over NetLogo, with performance compared across identical scenarios.
- We leverage Slurm [288] to run simulations on HPC infrastructure, highlighting the scalability and performance benefits of using FLAME GPU 2.
- We integrate *F4F* with a previously published compartmental model [202] developed to simulate COVID-19 spread in the Piedmont region of Italy to enhance our case study.

1.4 State-of-the-art

In this section, we explore the state-of-the-art relevant to our research. We begin by examining continuous monitoring and early warning systems. Then, we analyze the state-of-the-art approaches related to our various contributions—see Section 1.3—, emphasizing the novelties and advancements it brings compared to existing literature.

Continuous monitoring systems have been developed and applied across a diverse range of contexts, showing their versatility and potential for real-time data collection and analysis. For instance, in the field of environmental science, they have been employed to monitor honeybee colonies [180], soil conditions [105], and plant health [82], providing critical insights into ecosystem dynamics. In healthcare, continuous monitoring has been instrumental in tracking glucose levels for diabetes management [227], while in urban environments, CMSs have been used to monitor various aspects of city life, such as air quality, traffic, and infrastructure [51].

Moreover, in the field of epidemiology, continuous monitoring systems have been essential in effectively managing infectious diseases. They have been used to track disease transmission in real-time, evaluate the effectiveness of public health interventions, and forecast future outbreaks [179, 167, 17, 99]. During the COVID-19 pandemic, these systems facilitated the collection of epidemiological data, enabling health authorities to make data-driven decisions on testing, vaccination campaigns, and movement restrictions. Similarly, wearable health devices have allowed continuous monitoring of individual health metrics such as body temperature, heart rate, and oxygen saturation, supporting early detection of infections [33, 152]. These applications highlight the critical role of CMSs in enhancing pandemic preparedness, improving response strategies, and ultimately mitigating the impact of global health crises.

Early warning systems have been developed across a variety of contexts to predict and mitigate the impacts of various hazards. For example, systems for flash floods [162, 153] and natural disasters [294, 48] help prepare communities for imminent threats. In finance, they have been used to detect early signs of financial crises [64, 150], while earthquake prediction systems [19, 135, 23] provide valuable lead time for preparedness. Similarly, early warnings for tornadoes [253] and asteroid impacts [262] aim to reduce harm during extreme events.

In the public health domain, early warning systems for infectious diseases [186, 259, 18] are crucial for detecting outbreaks and enabling timely interventions. These systems use real-time data from diverse sources, such as healthcare facilities, social media, and environmental monitoring, to track the spread of diseases and predict potential outbreaks. For example, during the COVID-19 pandemic, early warning systems helped identify emerging clusters of infections, guiding public health measures like social distancing and quarantine policies [190, 44]. Additionally, they support surveillance of disease vectors, such as mosquitoes in the case

of malaria and Zika virus, by analyzing environmental and climatic factors that influence their spread [137, 69]. By enabling rapid response, these systems not only help to contain the immediate effects of infectious disease outbreaks but also improve long-term strategies for prevention and control. They exemplify how timely information can significantly reduce the risk and impact of health crises.

Nevertheless, continuous monitoring and early warning systems are often not easily explainable, particularly to policy-makers, due to their complexity and the large volumes of data they generate. The intricate nature of models and the lack of transparency in how they process data can create barriers to trust and hinder effective decision-making. Policy-makers, who rely on clear and actionable insights to guide public health strategies or resource allocation, may struggle to interpret the outputs without a sufficient understanding of the underlying models and assumptions. This challenge highlights the growing need for tools that not only provide accurate predictions but also present results in a comprehensible and transparent manner.

Addressing this issue is one of the core aims of our work. We aim to bridge the gap between complex monitoring systems and policy-makers by developing user-friendly dashboards, clear visualizations, and simplified models. These tools are designed to make simulation results more accessible, enabling informed decision-making and fostering greater trust in the models used to shape public health interventions and other critical policies.

1.4.1 *Living along COVID-19: assessing contention policies through Agent-Based Models*

Several studies have explored the spread of COVID-19 in school environments and evaluated the effectiveness of various NPIs. Liao et al. [161] develop a Contact Network Agent-Based Model (CN-ABM) to simulate COVID-19 transmission in school environments and assess how vaccination rates, community risk levels, and teaching modalities influence the epidemic curve. The model uses daily activity patterns to create contact networks and track the health status of agents. Without vaccination, keeping student occupancy under 50% can significantly flatten the epidemic curve, with 25% on-campus students yielding the best result (peak infection rate $< 1\%$). With vaccination, a combination of 75% on-campus students and at least a 45% vaccination rate can suppress the curve, with the optimal result achieved at a 65% vaccination rate.

Abdollahi et al. [13] assess the impact of school closures and self-isolation of symptomatic COVID-19 cases in Ontario. The agent-based model shows that school closures alone reduced attack rates by 7.2–12.7% and ICU admissions by 3.3–6.7%. However, combining school

closures with 20% self-isolation of symptomatic cases resulted in greater reductions, with attack rates and ICU admissions dropping by over 6.3% and 9.1%, respectively.

Phillips et al. [207] use an agent-based model to assess the effects of different reopening strategies for childcare centers and primary schools during COVID-19. It finds that grouping siblings together reduces outbreak size and student days lost. Larger class sizes significantly increase outbreak sizes, with each doubling of class size more than doubling the outbreak and lost student days.

Panovska et al. [197] model the impact of school re-openings in the United Kingdom during the COVID-19 pandemic, considering various reopening strategies and testing levels. It finds that with effective contact tracing and testing (up to 87% of symptomatic individuals), a second COVID-19 wave can be prevented. Full-time school reopening would require 75% of symptomatic cases to be tested and isolated, while a part-time system requires 65%. Results show that without these measures, reopening schools would likely lead to a second wave, peaking in late 2020 or early 2021.

Tatapudi et al. [256] use an agent-based model to examine the impact of school re-openings on COVID-19 spread in an urban U.S. region. A full (100%) reopening leads to an 8.15% increase in reported cases, while partial re-openings (75% and 50%) result in smaller, statistically insignificant increases. The model also suggests that relaxing safety protocols (mask use, sanitizing, etc.) could lead to a slight 2.14% rise in infections.

Hernandez et al. [128] model the spread of COVID-19 in Yucatán, Mexico's schools, where classes have remained virtual. Using agent-based simulations with social networks (family, friends, study groups), the study evaluates the risks of returning to in-person learning. The results show that isolating students and reducing class sizes are effective strategies to minimize transmission.

While these studies align with certain observations, none have fully assessed the combined impact of various NPIs, external factors, and different COVID-19 variants in a general school environment. While some studies examine specific NPIs like school closures, social distancing, and vaccination, they often do so in isolation. Additionally, the influence of external factors such as state/regional transmission rates and healthcare differences, as well as the impact of emerging variants, remains under-explored. Key novelties of the study include:

- *Detailed modeling of school interactions*: the model simulates complex interactions between students and teachers, incorporating social network dynamics and the timing of school activities like classes, breaks, and transitions. This provides a realistic framework for understanding how daily school routines contribute to disease spread.

- *Testing of multiple interventions*: the study evaluates various infection control strategies,

including physical distancing, mask usage, vaccination, screening, and quarantine policies, assessing their effectiveness in reducing COVID-19 transmission within the school environment.

- *School-specific parameters*: the model includes school-specific parameters such as class sizes, spatial layouts, and community structures, allowing for a tailored analysis of how different school settings might respond to similar interventions.

- *Consideration of external factors*: the study accounts for external factors like outside contagion and external screening practices. By incorporating broader state/regional transmission rates and the influence of testing protocols, the model reflects real-world dynamics, where external COVID-19 prevalence affects the likelihood of outbreaks within schools.

- *Modeling of different variants*: the study also simulates the impact of different COVID-19 variants on school transmission dynamics, providing a nuanced understanding of how emerging variants, with varying levels of transmissibility and virulence, might alter the effectiveness of interventions.

1.4.2 *Sybil: enhancing COVID-19 forecasting precision with variant-aware integrated models*

In the context of COVID-19 predictions, numerous methodologies are available for predicting the future trajectory of an epidemic, leveraging diverse modeling approaches [115]. Machine Learning (ML) and Deep Learning (DL) models, including Convolutional Neural Network (CNN), Recurrent Neural Network (RNN) featuring Long Short-Term Memory (LSTM) or Gated Recurrent Unit (GRU) cells, and multivariate CNN, have emerged as highly prominent approaches for forecasting. In the literature, several reviews are closely related to our work. Rodríguez et al. [229] provided an exhaustive review of data-driven computational methods for epidemic forecasting powered by machine learning advancements, leveraging diverse data sources from biological to behavioural, while Sayeed et al. [238] provided an in-depth critique and analysis of the machine-learning approaches that have been implemented by researchers to predict Covid-19. Rahimi et al. [216] reviewed and analyzed key machine learning forecasting models for COVID-19, featuring a scientometric analysis of COVID-19 data and a comparison of various solution approaches. Finally, Kamalov et al. [143] reviewed and analyzed deep learning methods for COVID-19 forecasting, presenting a model-based taxonomy, evaluating the performance of 53 selected studies, and identifying deficiencies and areas for future improvement.

In particular, regarding ML models, Alabdulrazzaq et al. [16] investigated and validated the accuracy of an Auto-Regressive Integrated Moving Average (ARIMA) model over a relatively long period of time using Kuwait COVID-19 data as a case study. Zhao et al. [292] compared the predictions obtained using an ARIMA model, a Multiple Linear Regression (MLR) model, and Prophet [257] with the aim to predict the daily worldwide confirmed cases of COVID-19 during the Omicron outbreak, finding out that ARIMA has superior prediction performance compared to MLR and Prophet models. Rendana et al. [219] used ARIMA to forecast the number of cases, using Spearman correlation to analyze the relationship between B.1.1.7 cases and meteorological variables such as temperature, humidity, rainfall, sunshine, and wind speed. Battineni et al. [49] used Prophet to predict the epidemic trend and derive an epidemic curve in the USA, Brazil, India, and Russia, while Wang et al. [272] combined a logistic model with Prophet with the same objective using COVID-19 data for Brazil, Russia, India, Peru, Indonesia, and the whole world. Sardar et al. [237] compared an ARIMA model, Prophet, a Generalized Linear Model Elastic Net (GLMNet), and a RF for predicting the confirmed cases in the South Asian Association for Regional Cooperation (SAARC) countries also finding out that ARIMA is superior to the other models. Finally, a comparison among a SVM, Prophet, and a Linear Regression (LR) model was used by Gupta et al. [122] to predict the active rate, the death rate, and the cured rate in India showing that Prophet provided the best predictions with vast uncertain and small data sets.

Concerning DL models, Sah et al. [233] compared the results obtained using Prophet, an ARIMA model, and a stacked LSTM-GRU model with the aim to forecast the number of confirmed and active cases in India, revealing that the proposed stacked LSTM-GRU model outperforms all the other models. Shahid et al. [240] used an ARIMA model, a Support Vector Regression (SVR) model, an LSTM model, and a Bi-LSTM model to make predictions on confirmed cases, deaths, and recoveries in the top-10 countries affected by COVID-19, showing that Bi-LSTM outperforms in terms of endorsed indices. ArunKumar et al. [29] used GRU and LSTM cells, ARIMA, and Seasonal ARIMA (SARIMA) to forecast the trends in the same countries, finding out that for most of the countries, deep learning models outperformed statistical models, but for some countries statistical models outperformed deep learning models. An RNN with LSTM cells and variants such as Deep LSTM, Convolutional LSTM, and Bi-LSTM was used by Arora et al. [28] with the aim to predict the number of positive cases in India, discovering that Bi-LSTM gives the best results. Tang et al. [254] introduced a novel multivariate time series forecasting model, named Densely Connected Flexible Dendritic Neuron Model (DFDNM), designed to predict daily and weekly positive COVID-19 cases with enhanced accuracy and adaptability, by considering factors like social behavior, vaccination, viral variety, and infectivity, incorporating high flexibility and dense shortcut connections to

enhance nonlinear capacity, learning efficiency, and long-term prediction, and demonstrating strong performance on multiple datasets. Jin et al. [141] developed a new neural forecasting model, called Attention Crossing Time Series (ACTS), that makes forecasts by comparing patterns across time series obtained from multiple regions, enabling data-driven forecasting without modeling disease dynamics. ACTS outperforms leading COVID-19 forecasting models in 13 out of 18 tests, including predictions of new cases, hospitalizations, and deaths. Wang et al. [271] proposed Deep Learning Based Epidemic Forecasting with Synthetic Information (DEFSI), an epidemic forecasting framework that integrates the strengths of artificial neural networks and causal methods. This framework enables accurate high-resolution forecasting from low-resolution observations by integrating within-season and between-season data through a two-branch neural network, constructing region-specific training datasets via multi-agent simulations, and demonstrating superior performance in forecasting Influenza-Like Illness (ILI) at both state and county levels. Ayyoubzadeh et al. [37] used Google Trends data to predict COVID-19 incidence in Iran, highlighting public concerns reflected in search trends and the potential for improved model performance with more data. They compared this approach to past epidemic prediction work and suggested combining Google Trends with other data sources for better accuracy, marking the first use of Google search data for COVID-19 forecasting in Iran. Tariq et al. [255] compared various deep learning models, including LSTM, Bi-LSTM, CNN, hybrid CNN-LSTM, Multi Layer Perceptron (MLP), and RNN, to forecast COVID-19 cases in the United Arab Emirates (UAE) and Malaysia, using Bayesian optimization to enhance model performance and identify the most effective models for these regions to guide public health strategies. Finally, a comparison among a CNN, a Multivariate CNN (MCNN), LSTM, and GRU cells was conducted by Nabi et al. [187] with the aim to predict future outbreak scenarios in Brazil, Russia, and the United Kingdom, showing that uni-variate and MCNNs outperform RNN.

Despite their increasing popularity and their capacity to outperform conventional techniques such as auto-regressive models, including ARIMA and SARIMA [61], these data-centric approaches exhibit several notable limitations that warrant attention. First and foremost, the predictions generated by these advanced methods are not easily explainable. The shortage, or even lack, of such property makes it challenging for users to comprehend the underlying factors and mechanisms driving the forecasts, which can undermine trust and hinder the interpretability of the results. Additionally, these sophisticated approaches often struggle to accurately predict significant changes in trends, such as abrupt peaks in data or the sudden emergence of new variants or outbreaks. Such shortcomings are particularly problematic in scenarios where timely and precise forecasting is critical for effective decision-making and intervention strategies. Consequently, while these data-centric methods hold great promise

and have demonstrated superior performance in many instances, their limitations must be carefully considered and addressed to leverage their potential in practical applications fully.

On the other hand, compartmental models are analytic approaches specifically tailored to reproduce the evolution of an infection in a population which take into account different aspects related to an epidemic, like vaccinations and variants [68, 183, 118]. In particular, Caldwell et al. [68] provided important insights on the COVID-19 pandemic like calculating the herd immunity threshold, verifying that nascent vaccines can prevent severe disease, determining optimal vaccine allocation strategies and that Variants of Concern (VoC) are more transmissible and lethal than previously circulating strains. Miller et al. [183] demonstrated that systematic properties of population dynamics in infectious diseases can inform the likelihood of the random process of mutation. Gostic et al. [118] provided suggestions for how to mitigate some technical challenges and highlight open problems in effective reproduction number (R_t) estimation. Moreover, extended versions of compartmental models were used to improve the accuracy of the predictions in different countries taking into account different age groups, symptoms, hospitalizations, intensive care unit admissions, and vaccinations [109, 57, 188, 217], undetected infections and human mobility between regions [291], quarantine and governmental intervention measures [173, 188, 217].

Stochastic transmission models and Bayesian models were also used to study the transmission of the COVID-19 disease [126, 154]. These models account for the inherent uncertainty in disease transmission by integrating random variables like individual interactions, chance encounters, and variations in infectiousness, offering a more nuanced understanding of how diseases spread and allowing for more robust predictions in complex and dynamic real-world scenarios.

These analytic models incorporate a wide range of factors, including population demographics, rates of infection, recovery, and mortality, thereby providing a comprehensive mathematical representation of epidemic progression [60, 261]. This level of transparency simplifies the process of explaining and interpreting the results to various stakeholders, such as policy-makers, healthcare professionals, and the general public. Consequently, it fosters a better understanding of the situation and facilitates more informed decision-making. However, despite these advantages, these analytic approaches are not without their challenges. They rely heavily on a set of assumptions about the system being modeled, which may not always hold true in real-world scenarios. Additionally, the process of parameter estimation presents significant challenges, as accurate data collection and validation are crucial for the reliability of the models. Indeed, several COVID-19 forecasting studies incorporate data-centric and analytical approaches. Qian et al. [213] combined a Bayesian model that integrates Gaussian processes with a Susceptible-Exposed-Infected-Recovered (SEIR) compartmental model to

assess the impact of COVID-19 lockdown policies globally, enabling scenario analyses and policy evaluations. It combines the solid mechanistic foundations of SEIR models with the flexible data-driven modeling of machine learning, trained end-to-end via stochastic variational inference. The model provides projections of COVID-19 fatalities under different lockdown and reopening strategies, offering valuable insights for policy decisions. Arik et al. [26] proposed a novel approach that integrates machine learning into compartmental disease modeling, specifically the SEIR model, to predict COVID-19 progression enhancing forecasting accuracy by combining traditional epidemiological models with advanced machine learning techniques, while in [27] they utilized the framework to forecast COVID-19 cases, deaths, and hospitalizations, demonstrating robust performance across diverse U.S. and Japanese regions. Rodríguez et al. [228] introduced Epidemiologically-Informed Neural Networks (EINNs), a framework that integrates mechanistic epidemic models with data-driven neural networks to enhance both short-term and long-term epidemic forecasting accuracy. EINNs effectively combine theoretical epidemiological dynamics with AI's capacity to process diverse data sources, overcoming the limitations of traditional models in handling complex, multi-modal information. This approach provides robust predictions and insights into epidemic trends, supporting informed public health decision-making. These last four works [213, 26, 27, 228] share similarities with Sybil as they integrate compartmental models (such as SEIR) with machine learning to improve forecasting accuracy, while also prioritizing model transparency. However, none of these studies incorporate variants in their models, and they require extensive computational resources and a lengthy training phase, unlike Sybil, which avoids these limitations.

Other studies have integrated data-centric and analytical approaches to enhance epidemic forecasting. For instance, a combination of compartmental models and ARIMA with the aim to develop informed mitigation strategies and inform the preparedness of policy-makers and healthcare workers [149], a logistic model and Prophet [272], a compartmental model and a DL model (DNN/RNN-LSTM) to estimate the stochastic transmission parameters of the compartmental model with a discrete time series construct [98], a compartmental model, a machine learning model and a Bayesian model for empirically grounded COVID-19 predictions [274], and several other combinations [215, 229].

Several limitations exist in the literature on disease prediction models. Few studies incorporate the evolving nature of variants, which hinders the ability to predict the spread of new strains. Additionally, while many studies focus on either data-centric or analytical approaches, there is limited research that combines both, which could provide more reliable predictions. By addressing key limitations of both machine learning/deep learning and compartmental models, Sybil provides a promising approach to COVID-19 forecasting that

strikes a balance between predictive accuracy, transparency, and accessibility. This makes it a valuable tool for guiding public health interventions. Specifically, Sybil distinguishes itself from existing approaches by:

- *Combining strengths*: the model integrates the advantages of data-centric and analytical approaches for more accurate predictions.
- *Reliable predictions*: Sybil consistently provides reliable forecasts, even when faced with major changes in the disease diffusion process, including the emergence of new variants or shifts in transmission dynamics.
- *Data efficiency*: unlike machine learning and deep learning models, Sybil requires less extensive training data, making it practical in data-limited situations.
- *Variant evolution analysis*: the framework allows for the examination of the evolution of different COVID-19 variants, enhancing its predictive depth.
- *Reproducibility*: the model ensures that its results can be reproduced, boosting its reliability and value for future research.
- *Open-source accessibility*: its open-source code is publicly available, fostering transparency and enabling other researchers to use, refine and build upon the model.

1.4.3 *Social compliance with NPIs, mobility patterns, and reproduction number: lessons from COVID-19 in Europe*

The effect of government-imposed restrictions on people's mobility and the consequent repercussions on the spread of the disease has been explored in several studies [52, 290].

Snoeijer et al. [246] investigate the relative impact of NPIs on mobility changes across multiple countries using Apple and Google mobility data. They found that lock-downs, states of emergency, business and public service closures, and school closures had the largest effect on reducing mobility. Chi-square and cluster analyses revealed that NPIs like school closures and business shutdowns were highly correlated in timing and implementation, suggesting that the observed effects of these interventions may be amplified by their simultaneous enforcement.

Askitas et al. [31] estimate the average dynamic impact of various interventions on COVID-19 incidence and individuals' movement patterns by developing a statistical model that accounts for the simultaneous implementation of multiple interventions using daily data from 175 countries. Their findings indicate that restrictions on private gatherings and the closure of schools and workplaces significantly reduced COVID-19 infections. In contrast,

restrictions on internal movement and public transport had no effects, while international travel restrictions yielded only short-lived effects.

Banholzer et al. [47] evaluate the effectiveness of seven NPIs aimed at controlling the spread of SARS-CoV-2, including school closures and bans on gatherings. Using a semi-mechanistic Bayesian hierarchical model and data from the first wave of COVID-19 across 20 countries, the findings reveal that bans on large gatherings were the most effective in reducing new infections, followed by venue and school closures. In contrast, stay-at-home and work-from-home orders were the least effective.

Cartenì et al. [70] quantify the impact of mobility habits on the spread of COVID-19 in Italy using a multiple linear regression model. The results showed that mobility patterns significantly contributed to explaining the number of infections, alongside other factors such as daily testing rates and environmental variables like air pollution and temperature. These findings highlight the role of mobility, in combination with other key variables, in influencing the transmission of the virus.

Shao et al. [243] examine the relationship between ambient temperature and COVID-19 transmission by examining human mobility data from 47 countries. The findings show that higher temperatures are negatively correlated with transmission rates but positively correlated with human mobility, which increases transmission. This suppression effect suggests that, while warmer weather may reduce transmission rates, it can also lead to increased outdoor activities, potentially worsening COVID-19 spread. Effective control measures are crucial, particularly as temperatures rise, to address the ongoing transmission despite warmer weather in some regions.

Bryant et al. [5] introduce a Bayesian model that estimates daily COVID-19 deaths based on mobility patterns in response to NPIs across Europe. Using Google mobility data from five categories, the model finds strong correlations between mobility shifts and death rates, with reductions in grocery and pharmacy mobility significantly lowering the basic reproduction number R_0.

Ilin et al. [9] use simple statistical models—using the same data as Bryant et al. [5] and Facebook data—to show that NPIs significantly reduce mobility, which can effectively predict infections based on data from countries like China, France, Italy, South Korea, and the USA.

Garcìa-Cremades et al. [8] explore the wide-ranging impacts of the COVID-19 pandemic, highlighting that, until herd immunity is achieved, measures like quarantines and social distancing are necessary to reduce mobility, despite their economic repercussions. The study evaluates various prediction models to create a decision support system for policymakers. It proposes consensus strategies to enhance model accuracy and introduces a multivariate model incorporating Google mobility data for better forecasting of 14-day case incidence.

Guan et al. [7] analyze anonymized health and mobility data from Israel to develop predictive models for daily new COVID-19 cases and test positivity rates. Models that incorporate mobility data significantly outperformed those that do not, reducing RMSE by 17.3% for new cases and 10.2% for positivity rates. The findings indicate high accuracy in predicting outbreak severity, offering a valuable tool for policymakers to implement targeted interventions in the post-vaccination era.

Barros et al. [11] investigate the effectiveness of NPIs on COVID-19 transmission by analyzing their impact on the effective reproduction number R_t and human mobility, used as a measure of policy adherence. Using a causal inference approach, the study examines five NPIs—confinement, school closures, mask mandates, events, and work restrictions—across 113 countries up to June 2020. The findings reveal that all NPIs, except mask-wearing, significantly influenced mobility trends, with school and events having the greatest effect on social distancing. Moreover, school closures, mask mandates, and work-from-home policies resulted in a sustained reduction in R_t.

Badr et al. [4] examine the impact of social distancing on COVID-19 transmission across the 25 most affected U.S. counties from January to April 2020. Using anonymized cell phone mobility data, the researchers found a strong correlation between reduced mobility and slower COVID-19 case growth rates, with Pearson correlation coefficients above 0.7 in 20 counties. Mobility dropped by 35–63%, with the effect on transmission typically observed 9–12 days later, consistent with the virus's incubation period. The study also found that behavioral changes in mobility began before official stay-at-home orders were issued, indicating a proactive public response.

Tokey [10] analyzes the relationship between human mobility and COVID-19 transmission across U.S. counties from March to August 2020. It explores spatiotemporal mobility patterns and their correlation with infection rates, controlling for socio-demographic factors and policy changes. The findings indicate that mobility decreased until the first COVID-19 wave peaked, then began to rise. Spatial models reveal a negative correlation between infection rates and mobility indicators, such as miles traveled and out-of-county trips. Higher infection areas had more online workers, fewer older and educated individuals, and higher poverty rates.

Adeniyi et al. [15] analyze a COVID-19 compartmental model that incorporates public compliance with health rules and sanitation, aiming to determine the effective reproduction number R_t and system stability. It shows that the system is stable under certain conditions and that combining adherence to rules and sanitation is the most cost-effective strategy for controlling the virus.

Chung et al. [81] investigate how trust in social NPIs influences travel intentions during a pandemic using data from South Korea, one of the few countries that did not impose mobility

restrictions after the COVID-19 outbreak. The findings indicate that trust in social NPIs mediates the relationship between efforts to cope with travel restrictions and the intention to travel during the pandemic.

Painter et al. [196] find that residents in U.S. Republican counties are less likely to fully comply with stay-at-home orders compared to those in Democratic counties—using geolocation and debit card transaction data. On the other hand, Democratic counties are more likely to shift to remote spending after such orders are implemented. While factors like COVID-19 risk and geography are considered, political beliefs appear to be a key factor influencing compliance with government mandates, with political alignment potentially explaining these partisan differences.

Fianlly, Van Rooij et al. [267] examine factors influencing Americans' compliance with stay-at-home measures using a survey of 570 participants from 35 states. It finds that fear of authorities reduces compliance, while personal ability, self-control, and intrinsic motivations like moral support and social norms promote compliance. The findings highlight key factors for improving compliance to mitigate the virus.

In general, these works highlight the crucial role of NPIs and mobility patterns in controlling the spread of COVID-19 infections. Key interventions like school closures, lockdowns, and restrictions on gatherings were effective in reducing mobility and, consequently, transmission rates. However, the effectiveness of specific measures varied, with some studies finding that restrictions on internal movement or public transport had limited or short-lived effects. The correlation between reduced mobility and lower infection rates was a recurring theme, underscoring mobility as a key factor in the pandemic's dynamics. Some studies also noted that mobility began declining before formal interventions, indicating proactive public behavior. Environmental factors, such as temperature, interacted with mobility and transmission patterns, with higher temperatures potentially increasing mobility, which could counteract reductions in transmission due to weather. Moreover, some studies highlight that compliance with COVID-19 measures is influenced by various factors. Trust in social measures and political beliefs can significantly impact adherence to restrictions, while personal ability, self-control, and intrinsic motivations like moral support play crucial roles. Combining rule adherence with sanitation is found to be particularly effective in controlling the virus.

While numerous studies have examined the impact of NPIs on mobility and the spread of COVID-19, most focus on specific regions, interventions, or mobility data sources, independently. There is a notable lack of comprehensive analyses exploring how mobility patterns across different countries respond to government-imposed and self-imposed restrictions, and how these patterns influence the effective reproduction number R_t over time, shaped by cultural patterns, variations in social compliance and interpersonal interactions. Moreover,

no study has estimated missing mobility data from restrictions and missing R_t values from mobility. The key novelties of this work are as follows:

- *Cultural patterns in mobility:* identification of cultural patterns by analyzing how mobility and the R_t are influenced by both official restrictions and self-imposed limitations (in the absence of formal measures), and variations in social compliance and interaction across different countries.
- *Clustering countries:* apply clustering methods such as K-means [125] and CONNECTOR [203], examining the distinctions between Northern and Southern European countries in terms of their social compliance with restrictions, while also highlighting no distinctions in social interactions.
- *Missing data estimation*: utilize XGBoost regression models [77] to:
  - Estimate missing mobility data using restriction indicators.
  - Infer missing R_t values from mobility patterns.

1.4.4 *Forge4Flame: an intuitive dashboard for designing GPU Agent-Based Models to simulate the spread of infectious diseases*

Agent-Based Models (ABMs) are a powerful tool used to model complex systems in various fields like economics [124, 107, 236], biology [25, 248, 58], social sciences [111, 103, 221], engineering [14, 88, 251], natural disasters [244, 78, 184], and epidemiology [41, 40, 165, 146, 245, 91, 147, 130, 117, 110, 293].

Specifically, in the field of epidemiology, Lorig et al. [166] conducted a systematic review of ABMs for COVID-19 transmission, focusing on how they can be used to investigate various interventions. The review analyzes these models by comparing the interventions explored, model inputs, individual characteristics, and disease states. Key research questions include the purpose of ABMs in COVID-19 transmission studies, how individual behaviors and decision-making are modeled, the simulated population size and area, the modeling of transmission processes, the representation of disease progression, the interventions investigated, and the trustworthiness of the models for decision-makers.

Kerr et al. [146] introduced COVID-19 Agent-based Simulator (Covasim), an open-source, stochastic agent-based simulator designed to model COVID-19 transmission dynamics and assess the impact of various interventions. By incorporating individual-level attributes such as age, sex, and comorbidities, Covasim enables detailed simulations of epidemic trends and

healthcare resource needs. The model's flexibility allows users to tailor simulations to specific populations and intervention strategies, facilitating informed decision-making in public health responses.

Silva et al. [245] introduced a SEIR agent-based model to simulate COVID-19 spread and assess the economic impact of seven social distancing scenarios. It finds that lockdowns are most effective in reducing infections but come with high economic costs, while combining face masks with partial isolation offers a balanced approach. The study also provides evidence that vertical isolation is ineffective.

Cuevas et al. [91] introduced an agent-based model to assess COVID-19 transmission risks in facilities, considering individual mobility, interactions, and contagion susceptibility. Unlike traditional models, it captures heterogeneous behaviors within shared spaces, allowing for scenario testing to identify effective coexistence strategies. The approach provides a more detailed, dynamic representation of disease spread in confined environments, aiding in informed decision-making for public health measures.

Khalil et al. [147] proposed a stochastic multi-agent model to simulate person-to-person contact during a pandemic influenza in Egypt. The model aims to understand transmission dynamics, assess potential impacts, and evaluate the effectiveness of control strategies. Using WHO-recommended parameters and 2006 Egypt census data, it creates a population structure and agent attributes. The study reviews various epidemiological modeling approaches, discusses control strategies, and analyzes the results of the modeling experiments.

Hinch et al. [130] presented OpenABM-Covid19, an agent-based model to simulate COVID-19 transmission and assess interventions like contact tracing, lockdowns, and physical distancing. Using UK census data, it models daily interactions, asymptomatic infections, and age-dependent disease progression. The model evaluates the effectiveness of various interventions and offers computational efficiency, transparency, and flexibility for policymakers.

Gomez et al. [117] introduced INFEKTA, an agent-based model designed to simulate the spread of infectious diseases like COVID-19. INFEKTA differs from existing models by considering individual characteristics, such as age, medical conditions, and daily routines, including transportation. It generates a complex network of interactions based on population density and includes the effects of NPIs like social distancing. The model aims to understand disease transmission dynamics at the individual level and assess the impact of interventions on controlling outbreaks.

Galvão Filho et al. [110] presented a novel approach to enhance the efficiency of individual-based epidemiological simulations by utilizing CUDA-compatible GPUs. This method addresses the significant computational demands of simulating large populations in models like SIR. The study demonstrates that even with simple SIR models, GPU-based parallelism

significantly reduces processing time, making large-scale simulations more feasible.

Zhou et al. [293] developed a behavior-driven agent-based model to simulate influenza transmission in public indoor venues, like shopping malls. It aims to fill a gap in existing epidemiological models, which focus more on homes and workplaces, by investigating disease dynamics in fine-scale environments where social interactions are less predictable. The study evaluates how factors like timing, number, and role of initial cases impact transmission and assesses the effectiveness of interventions such as vaccination and social distancing. The model provides insights for policymakers to tailor local strategies for controlling disease spread in specific settings.

Finally, Shamil et al. [241] presented an agent-based model to simulate the transmission dynamics of COVID-19 in a city. The model is validated using data from Ford County, Kansas, and then applied to a scaled-down version of New York City. It explores the effects of protective measures, such as lockdowns, and the role of digital contact tracing via smartphones. The results show that if over 60% of the population is traceable through smartphones, the R_t falls below 1 within three weeks. With 75% smartphone coverage, R_t drops below 1 even sooner, and new infections are eliminated within three months.

A major limitation of ABMs is the challenge of creating realistic well-defined closed environments such as offices, schools, or hospitals. These tools require manual setup, which can be time-consuming and technically demanding, limiting the flexibility and accessibility of ABMs, especially for users without specialized knowledge, making it harder to apply to large-scale, real-world scenarios. We focus the attention on ABMs in the context of infectious diseases to study virus transmission within well-defined closed environments from small to medium size.

In the field of agent-based modeling, numerous tools, such as NetLogo [285], are widely utilized. In [12], the authors provided a very exhaustive list of ABMs, focusing the attention, for each of them, on two different aspects: *i)* the effort to develop the model (simple/easy, moderate, complex/hard), and *ii)* the scalability (lightweight/small scale, medium scale, high/large scale, and extreme scale). NetLogo [285]—simple/easy and medium-scale—is one of the most widely used software platforms for developing and operating ABMs. Its popularity stems from its simplicity, extensive libraries, and user-friendly interface, making it suitable for educational purposes and advanced research.

Our experience with NetLogo [41, 40] has shown that high- or large-scale models are particularly well-suited for representing well-defined, closed environments. While there are many other valid agent-based modeling tools available (such as Repast Simphony [189] and Repast HPC [84], the widely used, free, and open-source ABM environment, Mason [169], a fast and easily extensible discrete-event multi-agent simulation toolkit in Java, or Mesa [145],

a Python-based framework for agent-based modeling, facilitating the creation and analysis of complex agent-based simulations), these tools primarily run on CPUs. This can limit their computational efficiency, especially for complex simulations, such as those modeling disease spread in enclosed spaces. To handle such intricate ABMs, the use of GPUs becomes essential. To address this challenge, P. Richmond et al. presented FLAME GPU 2 [222, 223], a state-of-the-art tool that leverages GPU acceleration to enhance the computational speed and efficiency of agent-based modeling significantly. We chose this tool for the following reasons: *i)* it is open source and is constantly updated by the developer team, *ii)* it exploits GPUs (CUDA) and provides a high-level C++ for CUDA as a programming language (without the need to write low-level CUDA code), *iii)* it is scalable and portable (high-/large-scale modeling environment), we can easily deploy and run analyses on HPC and/or GPU environments, and *iv)* it offers a useful 3D visualization interface.

Despite its advantages, FLAME GPU 2 can still be complex to use, which may discourage many potential users, particularly those in the epidemiological community who may not possess extensive programming expertise. To overcome this barrier, in synergistic collaboration with the FLAME GPU 2 developers, a graphical interface that simplifies the modeling process has been developed, called *Forge4Flame (F4F)*, making it more accessible to a broader audience of researchers and public health professionals. *Forge4Flame (F4F)* is a user-friendly dashboard, developed in Shiny R, designed to simplify the definition of an ABM environment for FLAME GPU 2 [222, 223] agent-based models, automatically generating the necessary code to design and execute ABMs.

To the best of our knowledge, Heiland et al. [127] developed an open-source, physics-based multicellular simulation framework for biological models, while Peng et al. [199] introduced FLARES, a novel GPU-accelerated framework designed to enhance the performance of agent-based infectious disease transmission models. However, these tools differ from the approach presented because they do not allow the definition of the ABM environment. Notably, **this newly developed tool does not have an equivalent counterpart in the existing literature**. *Forge4Flame (F4F)*'s key strengths can be summarized as follows:

- *User-friendly modeling: F4F* simplifies the modeling process, making it accessible to users without prior coding experience.
- *Interdisciplinary collaboration:* facilitates seamless cooperation among diverse research fields, improving the efficiency of disease model development and analysis.
- *Accelerated workflow:* enhances the speed of model development and result evaluation, supported by Slurm [288] for efficient distributed computing on HPC systems.

- *Integration of modeling scales:* bridges macro-level population models with micro-level individual simulations, enabling comprehensive analyses of disease dynamics.

1.5 Chapters description

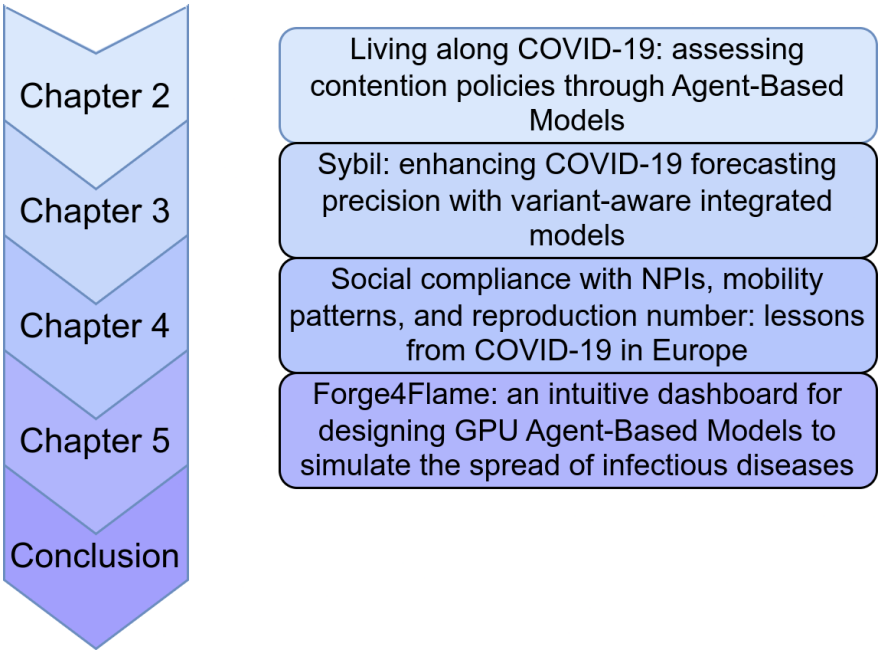


Figure 1.6: Road map of the thesis.

This thesis is organized into five chapters, besides the introduction—see Figure 1.6:

- In Chapter 2 we present the agent-based model used to study the efficacy of different NPIs and vaccination strategies, considering a generic middle school environment as a case study and evaluating the risk of infection associated with school activities and the effect of possible control strategies after the introduction of infective cases [41, 40].

- In Chapter 3 we present Sybil [39], an integrated framework that combines a machine learning-based predictive model—a data-centric approach—and a variant-aware compartmental model—an analytic approach—capable of providing improved prediction accuracy and explainability.

- In Chapter 4 we examine the relationship among NPIs, people's mobility, and the effective reproduction number R_t of COVID-19, understanding how movement trends vary in response to policy changes and restrictions, and how these trends impact the

spread of the disease across various countries and demographics, identifying behavioral patterns and estimating missing mobility data based on NPIs, as well as infer missing R_t values from mobility.

- In Chapter 5 we introduce *Forge4Flame (F4F)*, a user-friendly dashboard—developed in R Shiny—designed to simplify the definition of an agent-based model—for FLAME GPU 2 [222, 223] ABMs—automatically generating the necessary code.
- In Chapter 6, we provide a comprehensive summary of the entire thesis, revisiting the key objectives, findings, and contributions discussed throughout. This final chapter also reflects on the implications of the results, the limitations encountered, and potential areas for future work.

In the appendix, we provide additional results related to a couple of published works not presented in the thesis, including those by Saglietto et al. [232] and Contaldo et al. [86].

Chapter 2

Living along COVID-19: assessing contention policies through Agent-Based Models

This chapter examines how ABMs can assist policy-makers in implementing effective countermeasures and monitoring emergencies, like pandemics. In particular, we develop an agent-based model of a generic middle school environment to study the efficacy of different non-pharmaceutical interventions (NPIs) and vaccination strategies in a well-defined closed environment, evaluating the risk of infection associated with school activities and the effect of possible control strategies after the introduction of infective cases [41, 40].

School closures were often considered unavoidable control measures to reduce the transmission of COVID-19 because educational settings were considered potentially high-risk activities. Schools are characterized by close-contact settings, with students spending many hours close to each other in confined and enclosed spaces. In addition, compared with adults, children and adolescents infected with COVID-19 tend to be more commonly asymptomatic or have mild, non-specific symptoms, escaping early identification and isolation and potentially favoring the spread of the virus both inside and outside the school [85].

By late spring 2020, 80% of EU countries had fully or partially closed preschools, 90% had closed primary schools, and 100% secondary schools or higher education establishments [101]. However, following the first two pandemic waves, it was generally observed that COVID-19 transmission in school settings was not the primary determinant of community transmission, leading to the hypothesis that cases observed in school settings mainly reflect COVID-19 transmission in the general population [102, 168]. Furthermore, school closures are not cost-free and can, in turn, have dramatic social, economic, and health consequences. School

closures can have detrimental effects on children's physical and mental health and well-being and have the potential to increase inequality. School systems also provide significant social support, particularly for single-parent and economically disadvantaged households [171]. In light of emerging epidemiological evidence and the potential social impact of school closures, governments implemented various NPIs in school settings, such as face masks, restricted entry to school, daily symptom checking, and reduced class sizes. These policies aimed to reduce the risk of transmission, leaving closure as an extreme and last resource.

Large school outbreaks were frequently observed, particularly when prevention strategies were not fully enforced or adhered to [138, 250]. Moreover, the emergence of more transmissible SARS-CoV-2 variants, such as Delta and Omicron, coupled with a higher proportion of susceptible individuals among children, exhibited a high degree of immune evasion and likely contributed to increased infections among this age group [224]. The advent of vaccines has significantly contributed to reducing the number of deaths [275]. However, in the school setting, children were not initially covered by the vaccine, and even after that, their vaccine coverage grew slowly. As a result, maintaining a combination of NPIs remains crucial to minimizing the risk of COVID-19 transmission in educational settings. To strengthen mitigation efforts, both passive (syndromic) and active (molecular) surveillance strategies have been recommended to support case isolation, contact tracing, and quarantine measures in schools.

In this chapter, we first outline the model employed in our study, including a description of the agent-based model environment, the employed contagion processes, and the considered NPIs (Section 2.1). Next, we verify and validate the model (Section 2.2) and present two sets of results (Section 2.3). Finally, we discuss (Section 2.4), draw conclusions about the key findings (Section 2.5), and show how to replicate the results (Section 2.6).

2.1 Model

We propose an agent-based model, developed in the multi-agent programmable modeling environment NetLogo version 6.1.1 [285], that includes *i)* a detailed description of the environment and the population dynamics within such environment—including both how agents move within the physical spaces and social contact structure of the population—, *ii)* the COVID-19 progression model—modeled as a Susceptible-Exposed-Infected-Recovered (SEIR) compartmental model—, and *iii)* the non-pharmaceutical interventions (NPIs) available to contain the infection spreading.

2.1.1 School environment

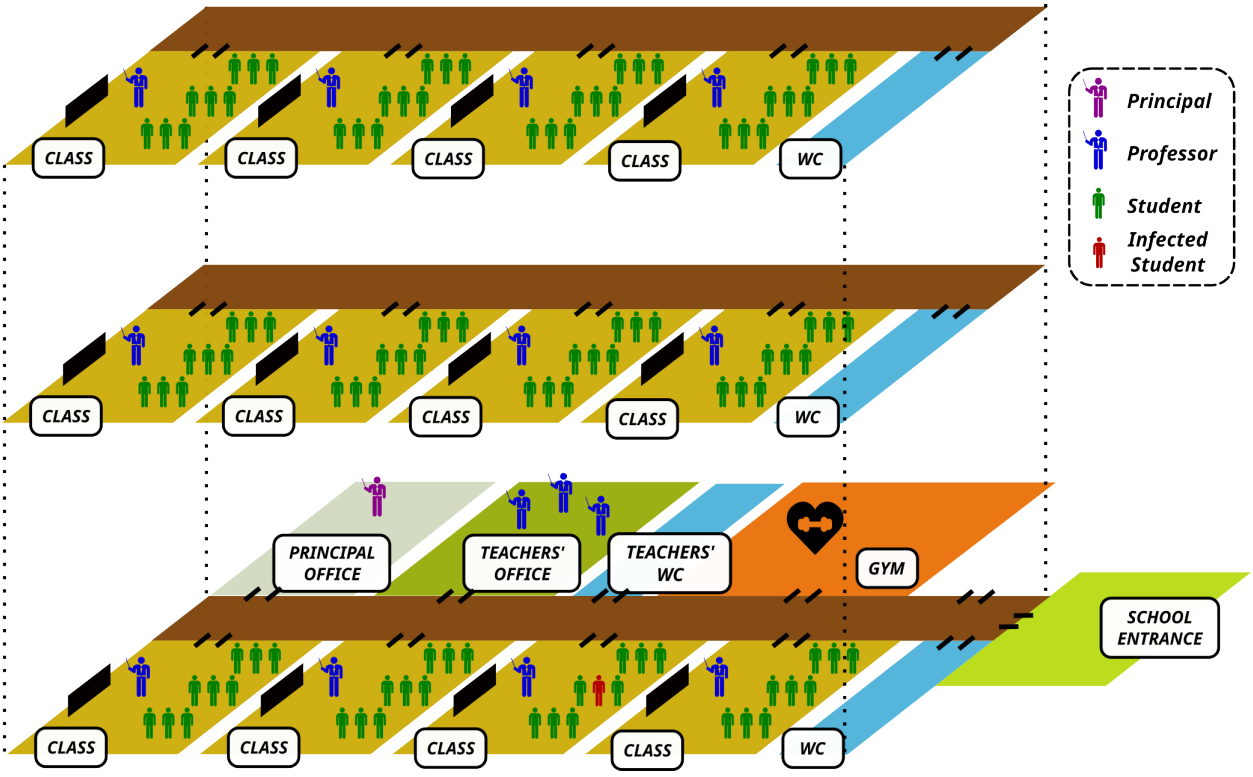


Figure 2.1: Graphical representation of a generic school featuring 12 classrooms distributed across three floors, along with common areas such as the entrance, corridors, bathrooms, stairs, gym, teachers' office, and principal's office. Each classroom includes twenty-four desks, evenly spaced apart, along with a teacher's desk and a blackboard at the front. The school population consists of a total of 290 individuals, including 240 students, 49 teachers, and a principal.

As an example scenario, graphically depicted in Figure 2.1, we considered a generic middle school with 12 classrooms (arranged on three floors) with shared spaces (i.e., entrance, corridors, bathrooms, stairs, gym, teachers' office, and principal's office). Each floor consisted of one corridor connecting all the rooms and one bathroom only for students. The school

entrance, the gym, the teacher's office, the principal's office, and the teachers' bathroom are all on the first floor. Each classroom is characterized by twenty-four desks (equally distanced from each other), one teacher's desk, and a blackboard. Students and personnel—overall 290 people divided into 240 students, 49 teachers, and a principal—move across the school according to their role and schedule. According to the organization of middle schools in Italy, students stay in the same classroom and with the same classmates all day long. To limit social interaction we defined the behavior of the agent as follows. While in classes, students stay at their desks unless the teacher asks them to go to the blackboard. Further movements are not allowed. Outside the classrooms, the only activities allowed are the movements toward the gym for a class and towards the principal's office. Teachers move from classroom to classroom according to the class schedule. They can also go to the teachers' office when they have no scheduled classes in the upcoming hour. Teachers and students can go to the bathroom on the same floor. Moreover, during breaks, students can stay in their classrooms or hallways near their classrooms and go to the bathroom. Table 2.1 reports all the rates associated with agents' behaviors.

Students' schedules foresee 30 classes of 50 minutes each, from Monday to Friday. Students arrive at school from 7:35 a.m., with the first class beginning at 8:10 a.m. A school day consists of six classes of 50 minutes each and two breaks. The first break is between the second and the third classes, while the second one is after the fourth class. At the beginning of each simulation students are allocated to classrooms (i.e., the first twenty agents are allocated to one classroom, and so on), and this association lasts for the whole simulation. Equivalently, timetables for teachers are chosen according to global scheduling, which is an adapted version of real teachers' scheduling. Such scheduling defines the weekly scheduling of classes for all the teachers.

2.1.2 Contagion processes

The next component of the agent-based model is the COVID-19 progression model, which is based on the Susceptible-Exposed-Infected-Recovered (SEIR) [202] epidemiological model, showed in Figure 2.2 where the population is divided into four compartments: *i)* the Susceptible (S) population consists of healthy individuals who have not yet been exposed to the disease, *ii)* the Exposed (E) population includes those who have been exposed and are infected, but are not yet infectious, *iii)* the Infected (I) population comprises individuals who can transmit the disease to others, and *iv)* the Recovered (R) population includes those who have recovered from the disease. Each agent is labeled with a specific disease progression status, whose parameters are reported in Table 2.1. In particular, we labeled each agent

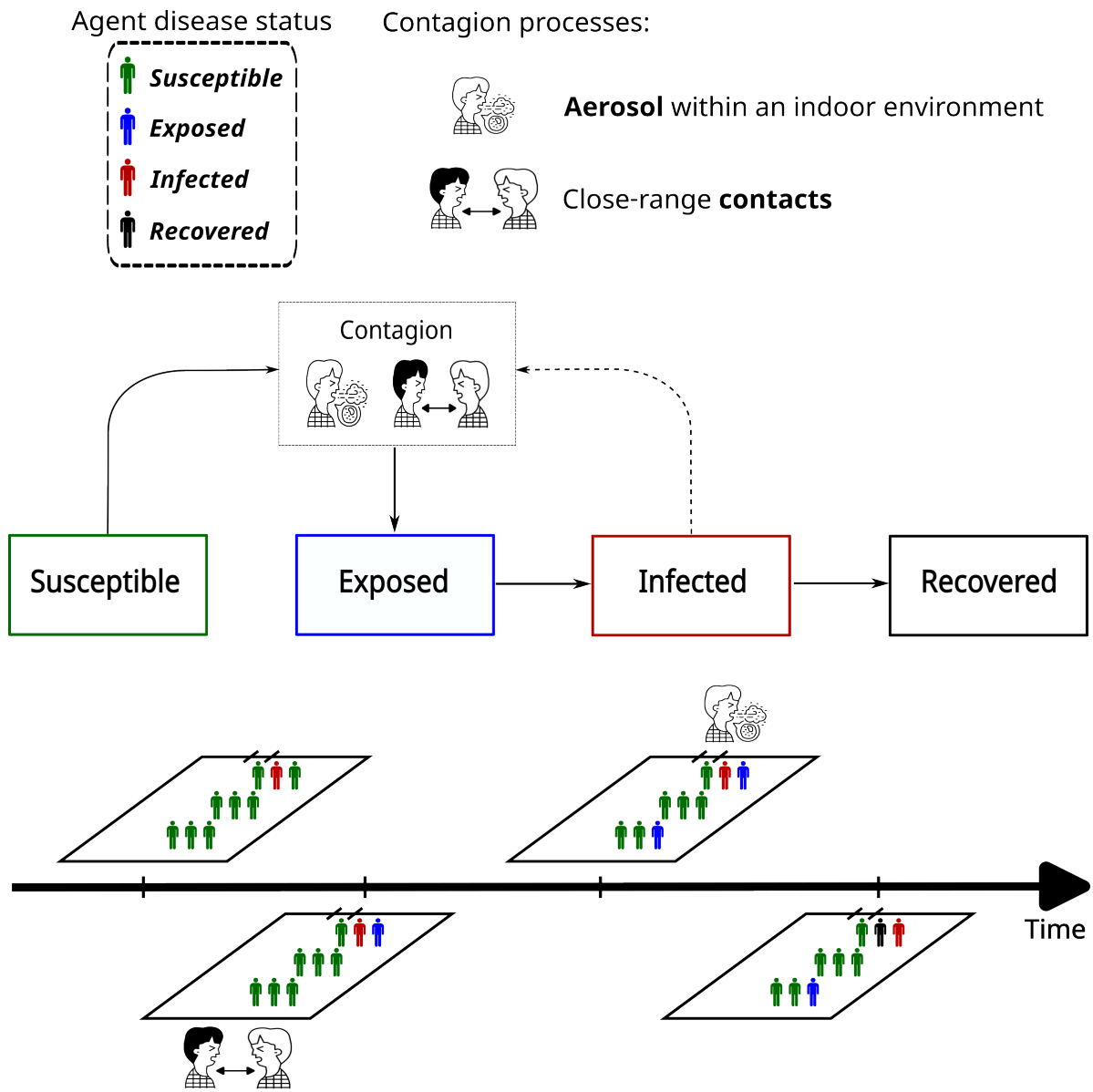


Figure 2.2: The COVID-19 progression model (SEIR)—the dashed line indicates that contagion occurs via close-range contacts with an infected individual [133] or through aerosol transmission [114]—with an example of its the evolution. We considered four time points: t_1, t_2, t_3, and t_4 ($t_1 < t_2 < t_3 < t_4$). At time t_1, all agents are susceptible except one infected. By time t_2, a nearby agent becomes exposed due to close-range transmission. At time t_3, a distant agent becomes exposed via aerosol transmission. Finally, at time t_4, the initially infected agent recovers, the close-range contact exposed agent becomes infected, and the aerosol-exposed agent remains exposed.

with one of the four statuses of the disease progression, updating it as follows: *i)* from the susceptible status to the exposed one—to model the virus exposure—, *ii)* from the exposed status to the infected one—when the incubation period ends, and agents become contagious—, *iii)* from the infected status to the recovered one—when the infectious period ends. The virus exposure is modeled through two different drivers: either close-range contacts [133] between susceptible and infected agents or through aerosol [114] within an indoor environment.

Considering the former case, a probability of successful contagious contact is associated with each susceptible agent (S_i) when it stays in an area ($\mathbf{A}$) of $4.41m^2$ around an infected agent. This probability is defined by the following function [133]:

$$P_{S_i} = c_r \frac{C_{S_i}}{\mathbf{A}} \tag{2.1}$$

where c_r is the *contamination risk*, and C_{S_i} is the total time (in minutes) during which the S_i agent stays in the area of an infected one.

Differently, the pathogen's spread by aerosol is computed from the *quanta* concentration in each room. A quanta is defined as the dose of airborne droplet nuclei required to cause infection in 63% of susceptible people [62]. Therefore, based on [114], we count the number of quanta inhaled in a specific room (e.g., classrooms, bathrooms, etc.) by each agent to obtain the probability of being infected without entering directly in contact with an infected agent, which is defined by the following function:

$$P_{S_i} = 1 - e^{-\frac{N_{\text{virus}}}{k_p}} \tag{2.2}$$

where k_p represents the reciprocal of the probability that a single pathogen will initiate the response [114, 273], and N_{virus} is the quantity of inhaled quanta by each susceptible agent in a time interval Δt. Its value is described by solving in Δt the differential Equation 2.3:

$$dN_{virus} = (1 - \eta_{mask} f_{\text{mask}})\ C(t) Q_{inh}\ dt \tag{2.3}$$

where η_{mask} is the efficacy of the mask, f_{mask} is the proportion of agents that wear correctly the mask, $C(t)$ is the quanta concentration at time t and Q_{inh} is the inhalation rate. Observe that quanta concentration at time t mainly depends on the number of infected individuals in the room, mask usage, the intensity of occupants' activities, and the room air ventilation.

The arising of new variants required to take into account in the model the presence of several pathogens, with different infection characterizations—namely Alpha, Beta, Delta, Omicron BA.1, and Omicron BA.2 variants were introduced. In particular, COVID-19 variants are versions of the SARS-CoV-2 virus with specific genetic mutations. Some variants may exhibit changes in transmissibility, severity, and resistance to immunity. We used the results of [200] to adjust the transmissibility of the selected variant.

In Figure 2.2, we illustrate an example of the transitions between different statuses over four time points: t_1, t_2, t_3, and t_4 ($t_1 < t_2 < t_3 < t_4$). At time t_1, all agents in the room are susceptible, except for one infected. At time t_2, an agent close to the infected agent becomes exposed due to close-range transmission [133]. At time t_3, another agent, located

further away from the initially infected individual, becomes exposed through the aerosol transmission [114]. Finally, at time t_4, the initially infected agent has recovered, the agent previously exposed at close-range transmission has become infected, and the agent exposed through aerosol transmission remains exposed.

2.1.3 Non-pharmaceutical interventions

Eventually, several NPIs policies can be implemented to prevent the infection from spreading within the school. A contention policy might require the application of one or more NPIs, compounding both actions aimed at reducing contact between students and their exposure to the infection.

In particular, the considered NPIs are distancing requirements, room air ventilation, testing campaigns, masks, and quarantine. As a baseline for all the experiments, we assumed that desks are equally distanced from each other, and if masks are required, students use them correctly and for all the time they are at school. Likewise, if room air ventilation is enabled, it operates continuously at 3 Air Changes per Hour (ACH)—ACH refers to the number of times the total air volume in a room or space is completely replaced within one hour. This airflow is ascribable for keeping the door and windows of a classroom open (without ventilation systems), to reduce the quanta concentration in the classrooms.

Various screening test strategies were incorporated into the model to help control COVID-19 spread in the school by enabling early detection of infected students and placing them in quarantine. Specifically, we considered the following four screening test strategies [43]: *i) none (NN)* where no screening policy is performed, *ii) test all (TA)* where all the participants are tested every week, *iii) round robin (RR)* where the participants in each classroom are divided into groups—we considered four groups—, ensuring that over a month, every participant is tested, with one group being tested each week, and *iv) sparse round robin (SRR)* which is similar to RR, but the tests are divided into two weekly days.

Once the infection has been discovered, three actions were considered: *no quarantine (NQ)*, *class quarantine (CQ)*, or *student quarantine (SQ)*. Specifically, NQ does not trigger any intervention when new infections are discovered, while CQ sparks a class quarantine only when the virus circulation within the classroom is over a threshold n_q, and *SQ* sparks the quarantine of single infected students when we discover them. CQ and SQ are the quarantine policies used in the Piedmont Region in November/December 2021 [208] and January/February 2022, respectively. Specifically, an infected student in quarantine does not contribute to the infection spreading. Thus, using the CQ policy, if one or more infections are detected in a given classroom, the entire class is tested. The entire classroom goes into

quarantine if the total number of infections reaches or exceeds n_q. If the threshold is not met, only the infected students are quarantined, while the remaining students are retested after d_{test} days, following the same quarantine procedure described earlier.

Eventually, we considered four cases for the use of masks: *i) no masks (NM)* where masks are not required, *ii) surgical all the time (SATT)* requiring to wear of the surgical mask *iii) FFP2 when required (FWR)* and *iv) always on (AO)* requiring to wear no masks and surgical masks, respectively. If n_m infections are detected in a classroom, these two last policies mandate the use of FFP2 masks by all students and teachers in that classroom for d_{FFP2} days.

Moreover, to account for interactions with the environment outside the school, we considered external screening campaigns and infections due to personal interactions in other environments, such as at home. In particular, two different probabilities of external screening are applied to students every day: the probability of testing a student due to sporting activities that involve screening campaigns (e.g., playing soccer) and the probability of testing a student outside the school due to the presence of some symptoms.

The efficacy of a given contention policy strongly depends on the virus circulation. The infection prevalence in the general population should be reflected in the spread dynamics within the school. Specifically, the model must account for the fact that any student will likely catch the infection outside the school. Therefore, we introduced the infection outside the school as a stochastic arrival process, so that the intensity of people getting the infection outside the school reflects the infection's prevalence in the population. In particular, a probability is applied to each susceptible agent every day, representing the possibility of infection outside the school.

Additionally, we analyzed a specific pharmaceutical intervention, examining how varying proportions of vaccinated individuals and different levels of vaccine efficacy could impact COVID-19 spread within the school. We define vaccine efficacy as the fraction of the vaccinated population that is effectively immune to infection, while the remaining portion remains susceptible despite having been vaccinated. Indeed the teachers teaching in various classrooms might be an important vector of diffusion among classrooms playing the role of shortcuts in the sense of Watts-Strogatz model of random graphs [276].

2.2 Verification and validation

Model validation and verification are two essential steps to make a model suitable for supporting high-consequence decision-making. The former aims to evaluate how the model accurately represents the typical behavior of the system under study. To validate the logic describing the complex phenomena represented in the model, we analyzed our model under extreme parameter configuration. For instance, we verified that if no infected agents are initially introduced into the system, then no outbreak will ever occur. In the second step, the model verification, we ensured that the parameters fed to the model would produce the expected outcomes. For instance, we verified that the average exposition time and infection time obtained by our model are in agreement with the literature, see Table 2.1.

Different scenarios considering only one of the two virus infection drivers (via contact or aerosol) were performed to identify the primary infection driver in the model. In accordance with the information released by the World Health Organization (WHO), we observed that aerosol transmission is the primary infection driver. The first two boxplots in the first row of Figure 2.3 report the R_0 variability over 200 simulation traces for the two drivers, clearly showing the greater strength of aerosol transmission compared to the transmission due to direct contacts.

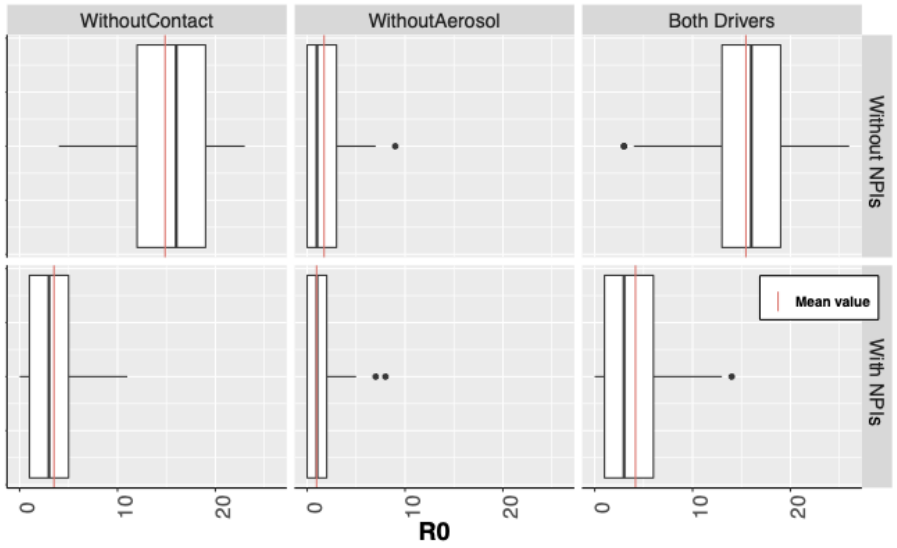


Figure 2.3: Box-plots representing the variability of the number of new infected students generated from one initial infected student (i.e., R_0), considering 200 simulations for each scenario.

Instead, during the model validation phase, we measured the predictive accuracy of our model comparing its outcomes with empirical data and results already present in the literature. First, we validated the aerosol transmission in our model with respect to that

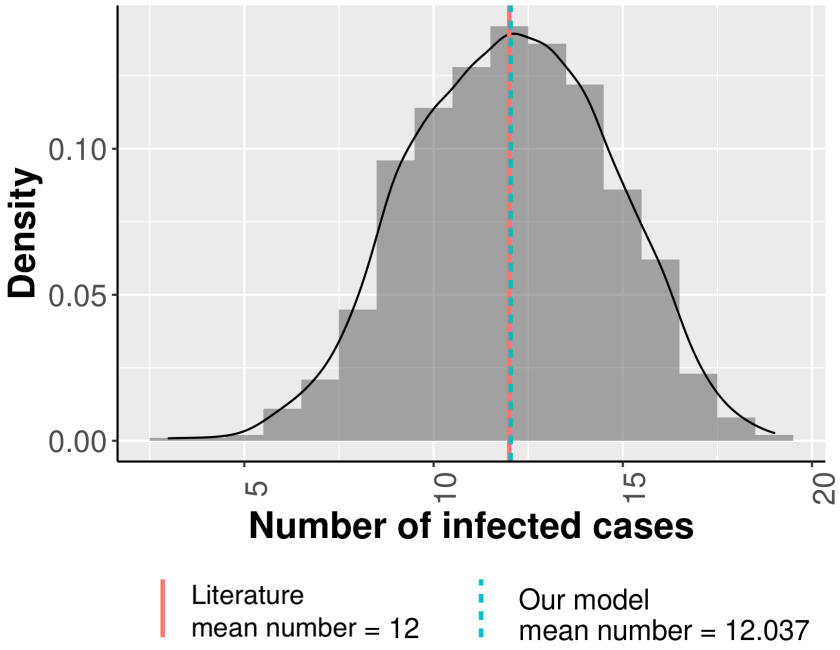


Figure 2.4: Distribution of the number of infected students computed from 1000 simulation experiments representing an event lasting 2 hours with 20 attendees and one infected speaker.

proposed in [114] reproducing the same settings in our agent-based model. Indeed, Figure 2.4 compares the number of infected individuals after two hours of lessons in a room with 20 participants and with only the teacher being infected obtained by our model with respect to that reported by [114]. In detail, the red line and the histogram represent the mean and the empirical distribution of such a measure derived by 1000 simulation traces of our model, while the dashed blue line represents the mean computed in [114]. From this, it is clear that the two measures are strongly in accordance.

Secondly, to evaluate the overall infection process, we decided to verify the accuracy of the basic reproduction number R_0 estimated by our model. Unfortunately, to the best of our knowledge, we did not find any other work studying R_0 in a similar context, therefore we decided to compare our results with those proposed in [226, 156], in which the basic R_0 is estimated considering the data generated by the COVID-19 outbreak in Diamond Princess Cruise Ship. Of course, the considered environment and the age of the population in these two works are quite different in our context, however, we believe that similar remarks on R_0 values could be applied also in our model. Specifically, both these works strongly suggest the R_0 of this disease in the enclosed areas should be higher compared with the one estimated by WHO ranging from 2 to 3. Indeed, both works report that an R_0 between 2 and 3 would not be able to account for the soaring number of infection cases reported in a short period in

the Diamond Princess Cruise Ship. Thus, these works using different modeling approaches estimated greater R_0 values (i.e. ranging from 4 to 7 in [156] and 15 in [226]) justifying it by the fact that people in tightly enclosed environments can easier gather and enter in contact.

The R_0 values obtained by our model for different scenarios are shown in Figure 2.3. The box plots in Figure 2.3 illustrate the variability of R_0 across 200 simulation traces, for three different scenarios—considering both with and without NPIs. The scenarios are as follows: *(i)* **Without Contact**, where close-range contacts are excluded, *(ii)* **Without Aerosol**, where aerosol-based transmission is not considered, and *(iii)* **Both Drivers**, where both contact and aerosol transmission are included. These results align with the considerations outlined in [226, 156].

2.3 Results

The results presented here aim to study the spread of COVID-19 infections in a school and to evaluate the effectiveness of NPIs and vaccination policies in such an environment. To study the efficacy of NPIs and vaccination policies in schools, we considered both active and passive interventions: the former requires actions from the healthcare system, while the latter only requires the students to behave according to some specific guidelines. The active interventions consist of vaccinations and testing campaigns. As passive interventions we considered social distancing as a common baseline behavior, rooms' air ventilation, and using masks—surgical or FFP2—and quarantine as possible contention policies.

We divided this study into two different sets of experiments, with Table 2.1 showing the shared parameters: *i)* in the first set of experiments—named *Basic Countermeasures for Risk Mitigation*—we evaluated the efficacy of multiple interventions aimed at mitigating risk, including the use of face masks, room air ventilation, various screening policies with differing participant numbers, and the effectiveness of the vaccination for teachers and the principal—initially excluding the consideration of student vaccination—, while *ii)* in the second set of experiments—named *Combined Impact of Various NPIs on School Safety*—, we focused on evaluating the combined effects of different screening, masks, and quarantine policies including varying the effectiveness of vaccination and assessing different fractions of vaccinated students, assuming that teachers and the principal are consistently vaccinated.

All the results presented here have been computed on high-performance computing infrastructures described in [20, 21], and all the parameters characterizing the model are summarized in the Tables 2.1, 2.3 and 2.5. Furthermore, the time in the simulation experiments is discrete and each tick corresponds to one minute.

Parameter	Value	Reference
Number of classrooms	12	Assumed
Students per classroom	20	Assumed
Average number of days of incubation	3	[289], [194]
Average number of days of infection	7	[66]
Ventilation	3 ACH	Assumed
Fraction of population wearing a mask (when required)	100%	Assumed
Distance learning	0%	Assumed
Probability of going to the bathroom (per minute)	$8e^{-4}$	Assumed
Probability of going to the blackboard (per minute)	$6e^{-4}$	Assumed
Probability of going to the principal office (per minute)	$1.4e^{-4}$	Assumed
Probability of going into the hall during interval	0.532	Assumed
Contamination risk (c_r)	0.024	[133]
Aerosol contagion process k_p parameter	410	[114]
Duration of the lessons in minutes	50	Assumed

Table 2.1: Model parameters shared by the two sets of experiments.

2.3.1 First set: *Basic Countermeasures for Risk Mitigation*

In the first set of experiments, we analyzed 21 different scenarios obtained as reported in Table 2.2 considering whether or not to use face masks and air ventilation in the room, and varying the screening policy, the number of participants in the screening, and the effectiveness of the vaccination for teachers and the principal—in this set of experiments, we did not take into account the vaccination coverage among school students because until the end of 2021 a large part of the pediatric population (the vaccination campaign had not targeted subjects $<$ 12 years of age), and even when it will be launched, vaccine hesitancy of children's parents may result in a high proportion of unvaccinated children.

When an infected student is identified through a screening strategy, they are placed in quarantine for 14 days [218] to prevent further transmission—see Table 2.3 for the used parameters. This quarantine duration is designed to align with established guidelines, ensuring sufficient time for recovery and minimizing the risk of spreading the infection within the community.

	Mask		Ventilation		Screening Policy				Screening participants				Vaccination Effectiveness	
	NM	*SATT*	No	Yes	*NN*	*TA*	*RR*	*SRR*	0%	25%	50%	100%	75%	100%
1	×		×		×				×					×
2		×		×	×				×					×
3		×		×		×				×				×
4		×		×		×					×			×
5		×		×		×						×		×
6		×		×			×			×				×
7		×		×			×				×			×
8		×		×			×					×		×
9		×		×				×		×				×
10		×		×				×			×			×
11		×		×				×				×		×
12		×		×	×				×				×	
13		×		×		×				×			×	
14		×		×		×					×		×	
15		×		×		×						×	×	
16		×		×			×			×			×	
17		×		×			×				×		×	
18		×		×			×					×	×	
19		×		×				×		×			×	
20		×		×				×			×		×	
21		×		×				×				×	×	

Table 2.2: Summary of the first set of scenarios simulated.

Let us observe that, to make screening strategies independent of the starting day, we considered the average values obtained by repeating the simulations varying the starting day of the strategy (Monday, Tuesday, etc). For each of these strategies, we evaluated their impact by ranging the number of participants in the screening test from 100% to 25%. Moreover, we considered two different levels of efficacy in teacher/principal vaccination: i) the vaccine is completely effective, and ii) the vaccine is effective only in 75% of the cases. We also

Parameter	Value	Reference
Number of initial infected (students)	1	Assumed
Vaccinated people	Teachers and the principal	Assumed
Vaccine efficacy	100 / 75%	Assumed
Students vaccination coverage	0%	Assumed
Duration of quarantine in days	14	[218]
Virus variant	First COVID-19 variant	Assumed
Mask efficacy	59%	[200]
Mask type	Surgical	Assumed
Screening adhesion	0 / 25 / 50 / 100%	Assumed
Screening policy	NN / TA / RR / SRR	[43]
Mask policy	NM / SATT	Assumed
Quarantine policy	SQ	Assumed
First probability of external screening (per day)	0	Assumed
Second probability of external screening (per day)	0	Assumed
Probability of outside contagion	0	Assumed
Simulated days for each run	35	Assumed
Number of runs for each experiment	1000	Assumed

Table 2.3: Specific model parameters used in the first set of experiments.

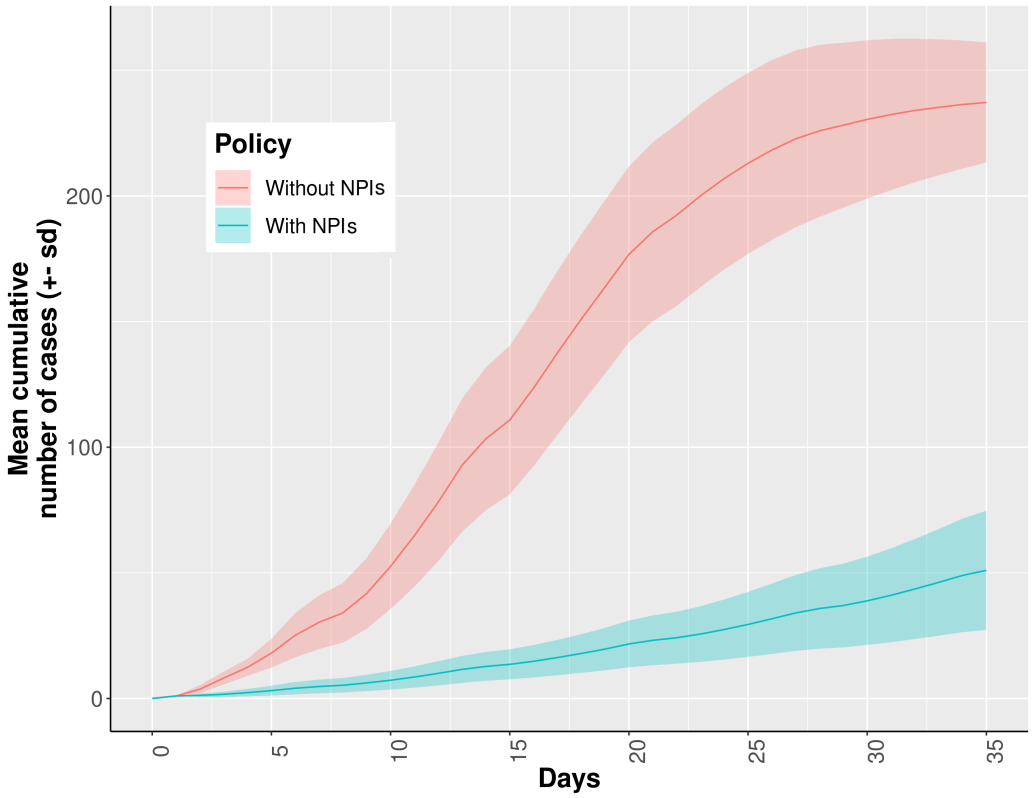


Figure 2.5: Cumulative number of infected students considering 35 days. The solid lines represent the mean trajectory considering 1000 simulations, while the band around the mean depends on the standard deviation. Finally, the color defines the used policy: blue in case of using NPIs (masks and room air ventilation), red otherwise.

assume in all scenarios, except the first one, that all people who enter the school wear face masks ($SATT$ policy). This design choice is twofold: it allows us to model a more realistic scenario and mitigates the effect of the screening campaign. Moreover, we are considering

only asymptomatic infections that can be detected by screening tests.

In the first experiment, we studied how the use of two NPIs, i.e. face masks and ventilation of rooms, can mitigate the spread of the infection in a school. Figure 2.5 shows day by day the cumulative number of infected students both when face masks and ventilation of the rooms are not considered (i.e., red line) and when they are taken into account (i.e. blue line). It is straightforward to notice the higher number of infected students when these two interventions are not implemented. Indeed, after 35 days most of the students (i.e. 237 out of a total of 240) were exposed. Differently, the use of these two NPIs allows for reducing the average cumulative number of infected students from 237 to 50 cases. Moreover, in the third column of Figure 2.3 the R_0 variability obtained by simulations with and without these two interventions is plotted. The use of face masks and ventilation of the rooms reduce the average R_0 from 15 to 4. Thus, these results highlight that these two NPIs are a mandatory requirement to slow down the infection making screening campaigns effective.

In the rest of this section, in line with the above results, we consider the use of face masks ($SATT$ policy) and ventilation of the rooms as a baseline scenario, without screening (NN policy). For the next batch of experiments, a third non-pharmaceutical intervention representing a screening campaign is introduced.

Figure 2.6 compares our baseline scenario with different scenarios in which the three screening policies described in Section 2.1.3 are implemented (i.e., TA, RR and SRR) varying on the rows the fraction of participants to the screening campaign and on the columns the vaccination efficacy. Focusing on the scenarios with all the students involved in the screening campaign (i.e., 100%) and 100% vaccination effectiveness, i.e. the first plot on the left, it is straightforward to note that all the screening policies reduce the incidence of infection in the population with respect to NN. Furthermore, although the RR policy is the one that does the worst, it is capable of halving the cumulative number of infections at the end of the observation period. A slight improvement can be obtained by splitting the screening test performed in the RR policy on two different days of the week as in the SRR policy: this design choice allows us to bring forward the unveiling of some infections. Indeed, anticipating the discovery of one infected reduces the period in which their classmates are exposed to such an infected individual, thus reducing their probability of getting infected. Similar observations can be derived when the number of students taking part in the screening test is reduced (i.e. only 75% and 25% respectively) as shown in Figure 2.6, first column, and second and third rows. In particular, these results highlight that even with a very limited number of tests it is possible to control the diffusion of the infection. For instance, in the case of student participation of 25% the policy RR, testing only five students for a week, can halve the total number of infected students w.r.t. NN. Moreover, Figure 2.7 shows the distribution of

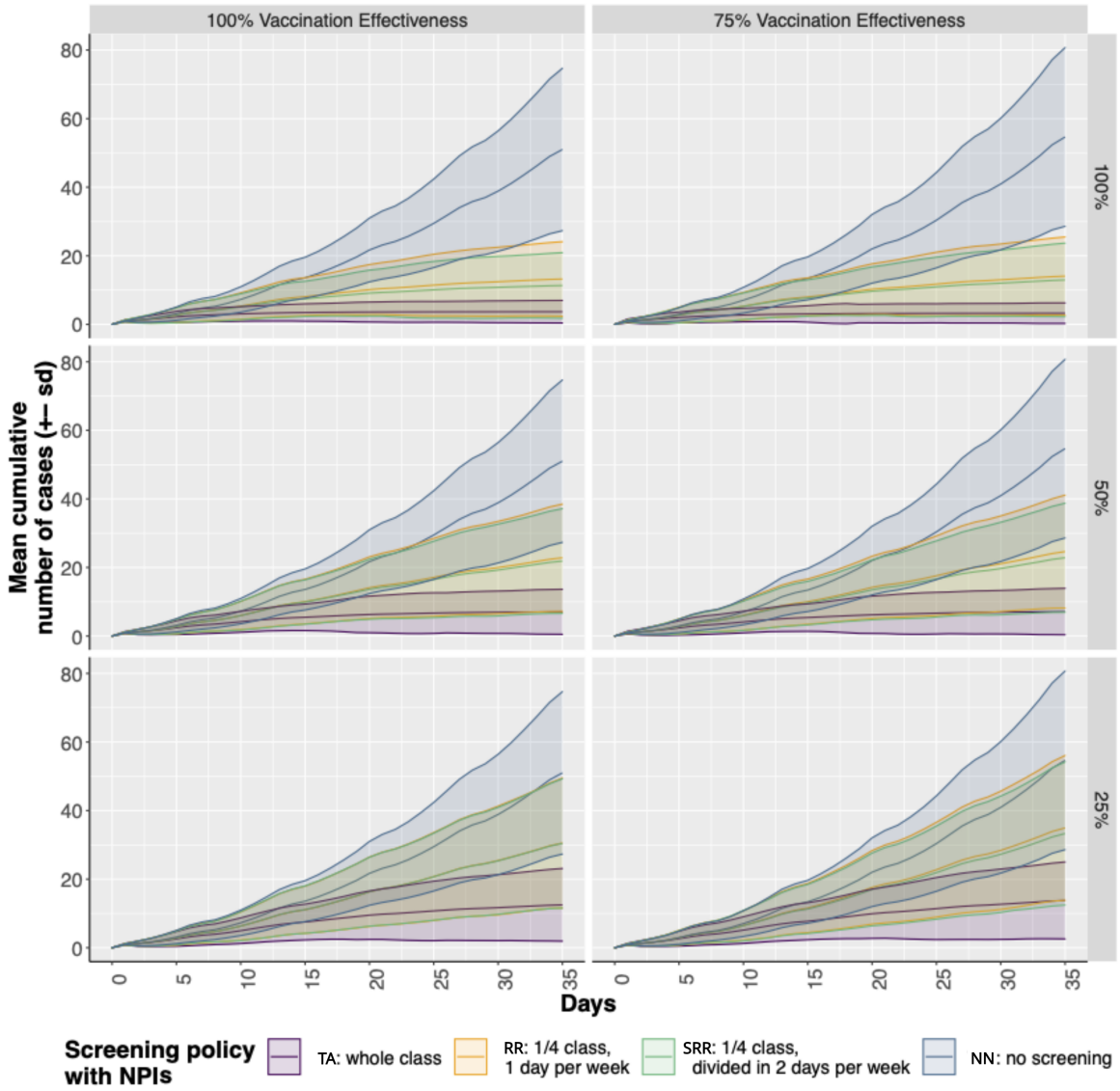


Figure 2.6: A comparison in terms of infected individuals between the scenarios considering only the use of face masks and ventilation of the rooms with respect to the scenarios in which the three screening policies TA, RR and SRR are implemented varying on the rows the number of participants to the screening campaign and on the columns the vaccination efficacy.

the average number of days needed to detect an infected student for the different screening policies and varying the number of participants to the screening policies. From the results shown, we can state clearly that the policy TA turns out the one with the lowest total number of infected individuals.

However, for policy-makers, the number of infected individuals could be only one of the many aspects to be considered when making a decision. According to this, we carried out a cost-benefit analysis defining a cost function that takes into account the costs of new

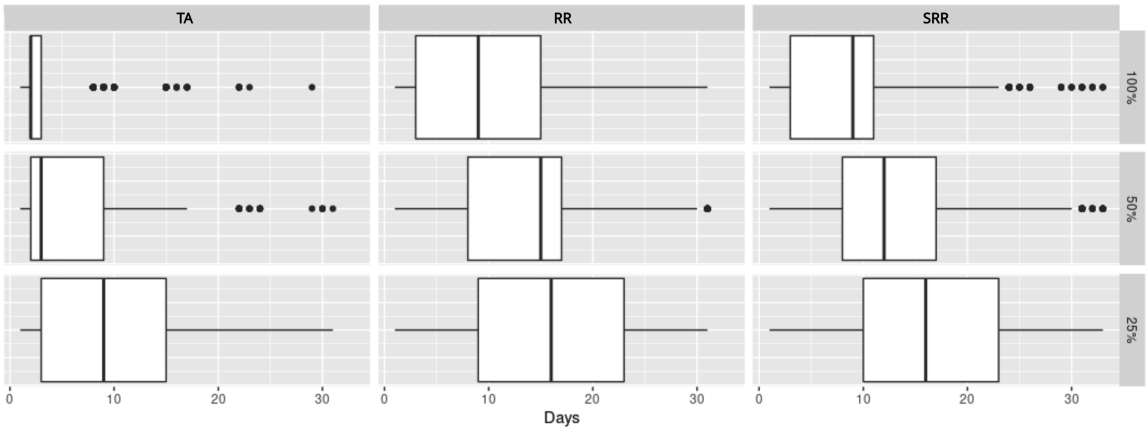


Figure 2.7: Boxplot reporting the observed number of days before detection of an infected student, with both different vaccination policies (from left to right) and adhesion percentages (from top to bottom).

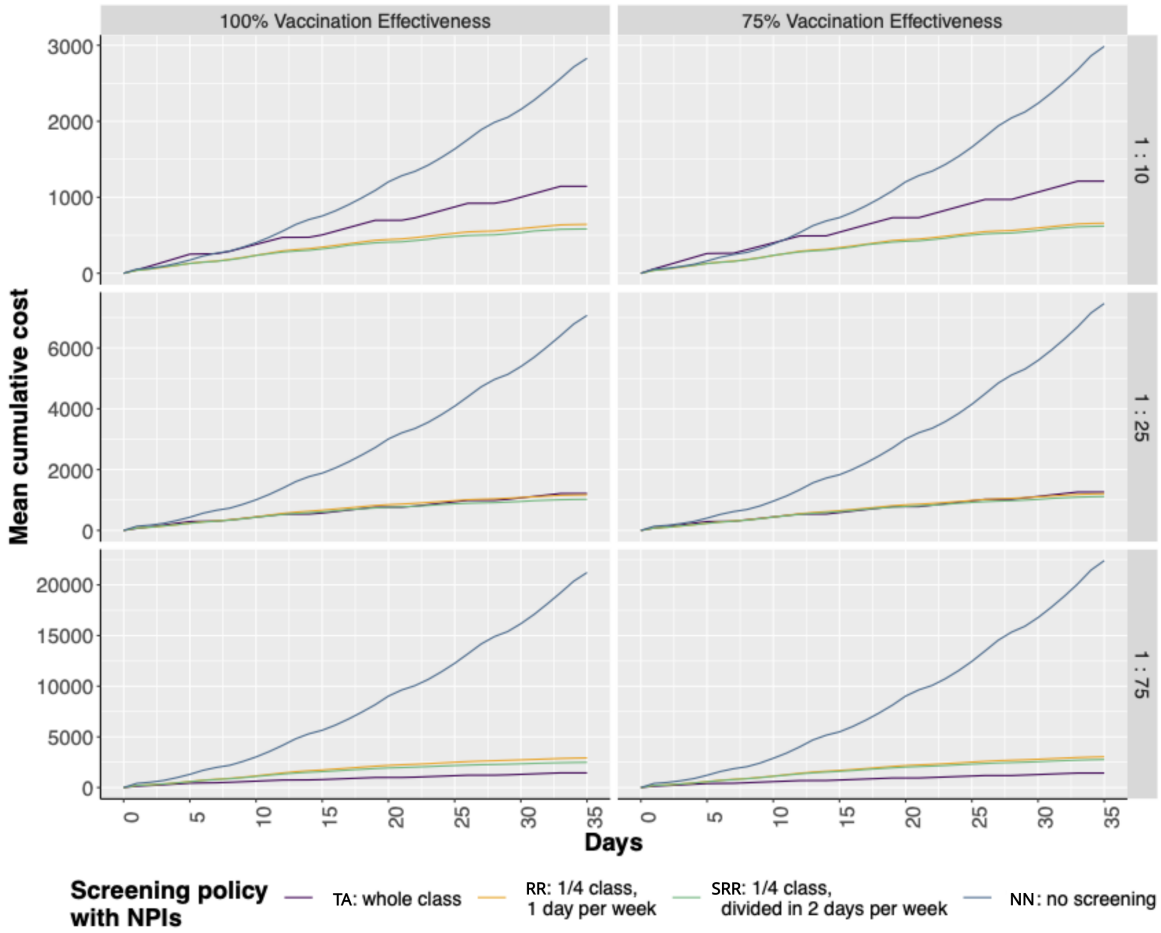


Figure 2.8: Cost function evaluation for different screening policies, with both different vaccination effectiveness (from left to right) and $\frac{k_1}{k_2}$ (from top to bottom).

infections and the overall cost (monetary and psychological) of performing swabs as follows:

$$\chi(t) = k_1 \tau n_i(t) + k_2 n_s(t), \tag{2.4}$$

where k_1 and k_2 are constants, τ is the meantime necessary to unveil a newly infected individual, and $n_i(t)$, $n_s(t)$ are the number of infected cases at time t and the number of students under screening at time t respectively. Thus, this cost function measures the performance of a particular screening policy in terms of the meantime of identification τ, the number of infected cases, and the number of swabs. Figure 2.8 reports the assessment of this cost function varying the $k_2 : k_1$ ratio for each considered screening policy and the levels of efficacy in the teachers' vaccination. In detail, from upper to bottom the $k_2 : k_1$ ratio decreases increasing the costs of new infections, while from left to right efficacy in the teachers' vaccination is reduced by 25%.

It is interesting to note that independent of the level of efficacy in the teacher's vaccination when the ratio between the two costs k_2 and k_1 is greater than $\frac{1}{25}$ (i.e., the leftmost plot) the policies performing a smaller number of swabs are more profitable with respect to the TA policy. Differently, when this ratio is smaller than $\frac{1}{25}$, meaning more importance on the impact of new infections increases, TA becomes the most effective.

2.3.2 Second set: *Combined Impact of Various NPIs on School Safety*

In the second set of experiments, we assumed that, in all scenarios, the principal and all teachers were vaccinated. We considered two screening policies (namely, NN and RR), three quarantine policies (namely, NQ, CQ, and SQ), three mask policies (namely, NM, FWR, and AO), four levels of vaccine efficacy (100%, 70%, 50%, and 30%), four levels of student vaccination coverage (70%, 40%, 10%, and 0%), and three levels of virus circulation (low, medium, and high). Table 2.4 summarizes the analyzed scenarios, with 32 scenarios for each combination of students' vaccination coverage and virus circulation levels, resulting in a total of 384 scenarios. Table 2.5 summarizes the parameters used in this second set of experiments.

Specifically, we considered eight clusters of results for each combination of vaccine efficacy, students' vaccination coverage, and virus circulation levels (see Table 2.4): *i)* scenarios 1 to 4 represent our baseline, modeling a policy in which NPIs countermeasures are not implemented, *ii)* scenarios 5 to 8 account for screening policy RR and quarantine policy CQ, *iii)* scenarios 9 to 12 for quarantine policy CQ alone, *iv)* scenarios 13 to 16 for quarantine policy SQ alone, *v)* scenarios 17 to 20 for mask policy FWR and quarantine policy CQ, *vi)* scenarios 21 to 24 for mask policy FWR and quarantine policy SQ, *vii)* scenarios 25 to 28 for mask policy AO

	Quarantine Policy			Screening Policy		Mask policy			Vaccination Effectiveness			
	NQ	*CQ*	*SQ*	*NN*	*RR*	*NM*	*FWR*	*AO*	30%	50%	70%	100%
1	×			×		×			×			
2	×			×		×				×		
3	×			×		×					×	
4	×			×		×						×
5		×			×	×			×			
6		×			×	×				×		
7		×			×	×					×	
8		×			×	×						×
9		×		×		×			×			
10		×		×		×				×		
11		×		×		×					×	
12		×		×		×						×
13			×	×		×			×			
14			×	×		×				×		
15			×	×		×					×	
16			×	×		×						×
17		×		×			×		×			
18		×		×			×			×		
19		×		×			×				×	
20		×		×			×					×
21			×	×			×		×			
22			×	×			×			×		
23			×	×			×				×	
24			×	×			×					×
25		×		×				×	×			
26		×		×				×		×		
27		×		×				×			×	
28		×		×				×				×
29			×	×				×	×			
30			×	×				×		×		
31			×	×				×			×	
32			×	×				×				×

Table 2.4: Summary of the scenarios simulated in the second set for each student's vaccination coverage (70%, 40%, 10% and 0%) and circulation level (low, medium and high).

Parameter	Value	Reference
Initial infected	0	Assumed
Vaccinated people	Teachers, the principal, and students	Assumed
Vaccine efficacy	100 / 70 / 50 / 30%	Assumed
Students vaccination coverage	70 / 40 / 10 / 0%	Assumed
Duration of quarantine in days	10	[208]
Virus variant	Omicron BA.1	Assumed
Mask efficacy	59% (surgical) / 90% (FFP2)	[200]
Mask policy	NM / FWR / AO	Assumed
Quarantine policy	NQ / CQ / SQ	[208]
Screening adhesion	100%	Assumed
Screening policy	NN / RR	[43]
First probability of external screening (per day)	0.0075	Assumed
Second probability of external screening (per day)	0.03	Assumed
Probability of outside contagion	0.0004 (low circulation) / 0.008 (medium circulation) / 0.02 (high circulation)	Assumed
d_{test}	5	[208]
d_{FFP2}	10	Assumed
n_q	5	[208]
n_m	1 (FWR) / 5 (AO)	Assumed
Simulated days for each run	60	Assumed
Number of runs for each scenario	500	Assumed

Table 2.5: Specific model parameters used in the second set of experiments.

and quarantine policy CQ, and *viii)* scenarios 29 to 32 for mask policy AO and quarantine policy SQ.

Figures 2.9-2.11 show the cumulative distribution function (CDF) of the total number of infected agents at the final time for different vaccine efficacy—decreasing moving from left to right—, several levels of students vaccination coverage—decreasing moving from top to bottom—and different NPIs combinations.

Figure 2.9 shows the results obtained under a low virus circulation level. The first setting shows the baseline (quarantine NQ, screening NN, mask policy NM) and allows us to understand the effect of the vaccination alone. In the case of 100% vaccine efficacy, increasing the fraction of vaccinated students reduces the probability of encountering a high number of infected agents after 60 days. As expected, decreasing the vaccine efficacy increases the probability of a higher number of infected agents. Comparing this initial setting with the second one, we observe that the introduction of screening and quarantine policies (quarantine CQ, screening RR, mask policy NM) leads to a reduction in the probability of having a higher number of infections after 60 days. Also, it can be noted that vaccination campaigns covering less than 40% of the population provide poor results in terms of number of infections. It is important to notice that the most relevant results are obtained with higher vaccine efficacy: the testing process gets less effective as infections become more likely to happen, meaning that the virus spreads faster than the pace at which testing is performed.

The last six settings illustrate the effects of the NM, FWR, and AO mask policies in combination with CQ and SQ quarantine policies, along with NN screening. Comparing the second setting to the third one (quarantine CQ, screening NN, mask policy NM) is clear that the screening RR has a negligible impact for the used quarantine policy. Focusing the attention on quarantine policies, we can state that CQ works much better than SQ, especially in the settings with FWR and AO mask policies. Regarding mask policies, we can conclude that the AO mask policy gives the best results (especially combined with the CQ quarantine policy). This is because the AO mask policy provides a boost in the protection of about 50% by forcing the use of FFP2 masks—see Table 2.5. In addition, opposite to what happens in the case of testing, the impact of the AO mask policy increases as the vaccine efficacy decreases. Indeed, despite how fast the infection spreads, masks always provide the same level of protection.

Figures 2.10 and 2.11 show the same eight settings under a medium and high virus circulation level. Once again, we find that increasing the fraction of vaccinated students reduces the probability of observing a high number of infected agents after 60 days while decreasing vaccine efficacy increases this probability. These results indicate that containing the spread is more challenging under medium and high virus circulation levels compared to

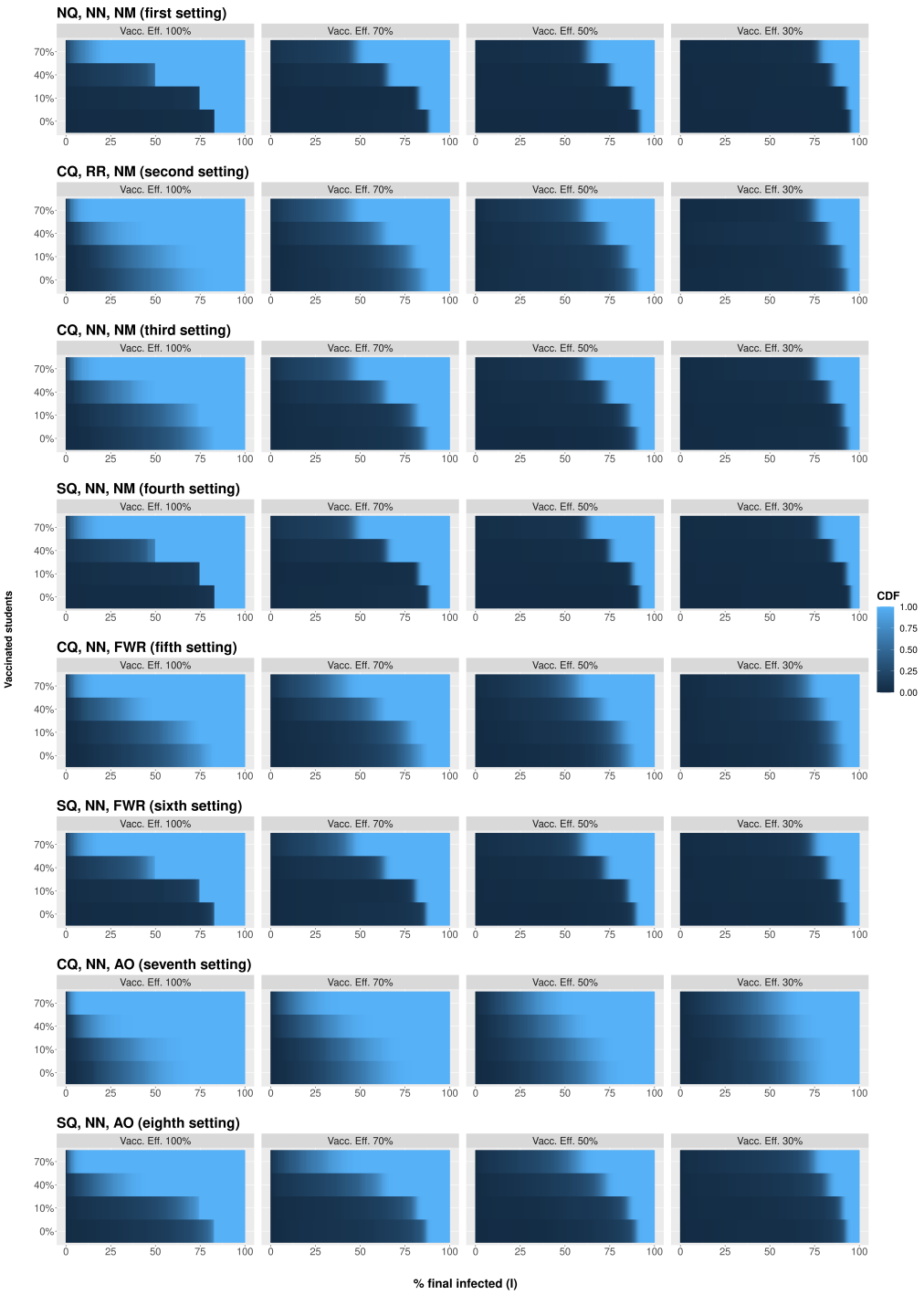


Figure 2.9: Cumulative distribution function (CDF) of the total number of infected agents at the final time for each considered scenario under a low virus circulation level. For each scenario—identified by the setting (on the rows), the vaccine efficacy (on the columns), and the fraction of vaccinated students (on the rows of each subplot)—we run 500 simulations—see Table 2.5. Lighter CDFs indicate more effective countermeasures. The first, second, and seventh settings are the most noteworthy. The first setting (NQ, NN, NM) shows that higher infections are associated with lower vaccine efficacies and lower fractions of vaccinated students. The second setting (CQ, RR, NM) shows that a fraction of vaccinated students $< 40\%$ provides poor results in terms of levels of protection—the testing process gets less effective as infections become more likely to happen. Finally, the seventh setting (CQ, NN, AO) shows that masks are more effective with lower vaccine efficacies—masks always provide the same level of protection.

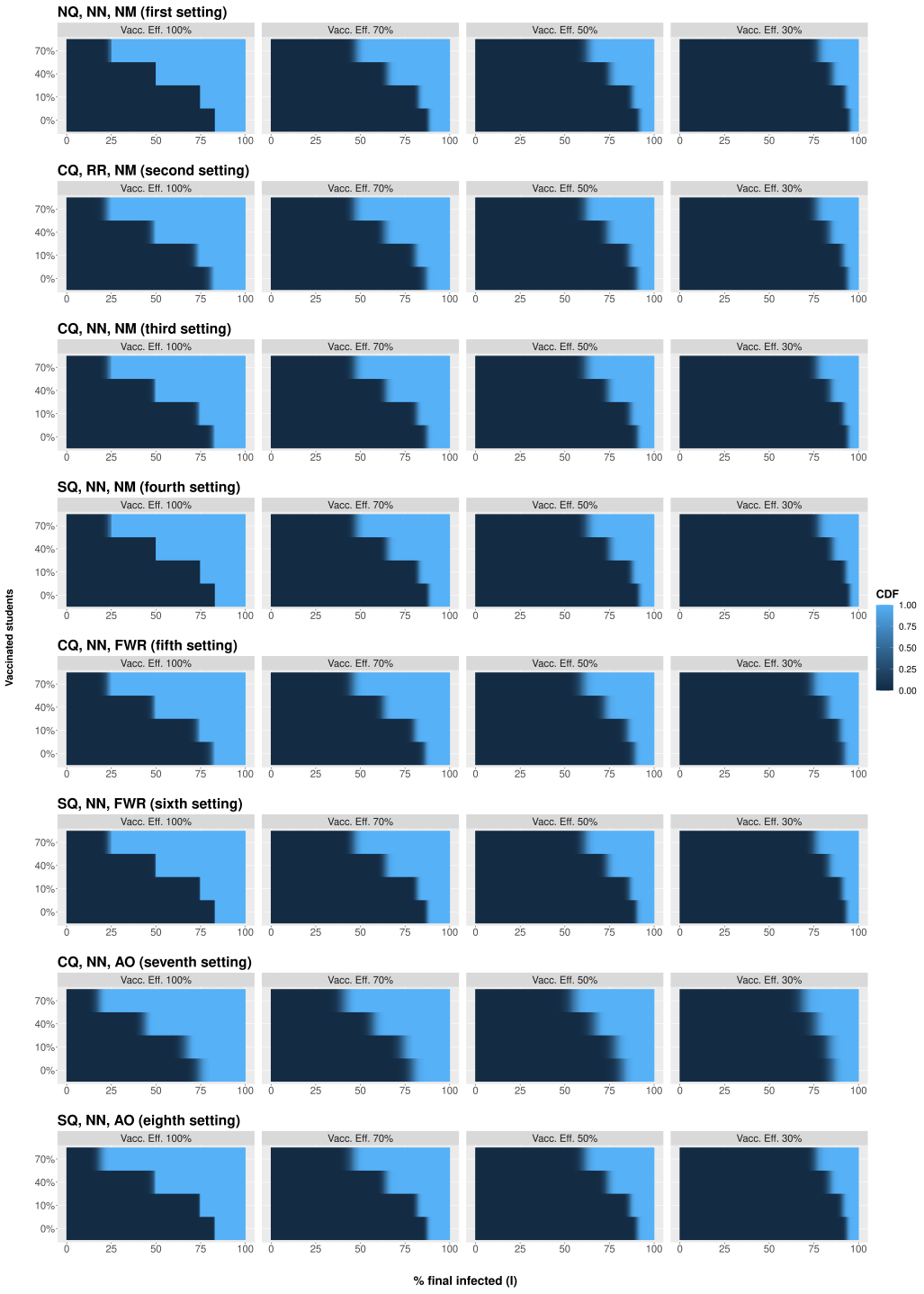


Figure 2.10: Cumulative distribution function (CDF) of the total number of infected agents at the final time for each considered scenario under a medium virus circulation level. For each scenario—identified by the setting (on the rows), the vaccine efficacy (on the columns), and the fraction of vaccinated students (on the rows of each subplot)—we run 500 simulations—see Table 2.5. Lighter CDFs indicate more effective countermeasures. We can state that higher infections are associated with lower vaccine efficacies and lower fractions of vaccinated students and a combination of masks and quarantine policies allows to reduce the final number of infected agents.

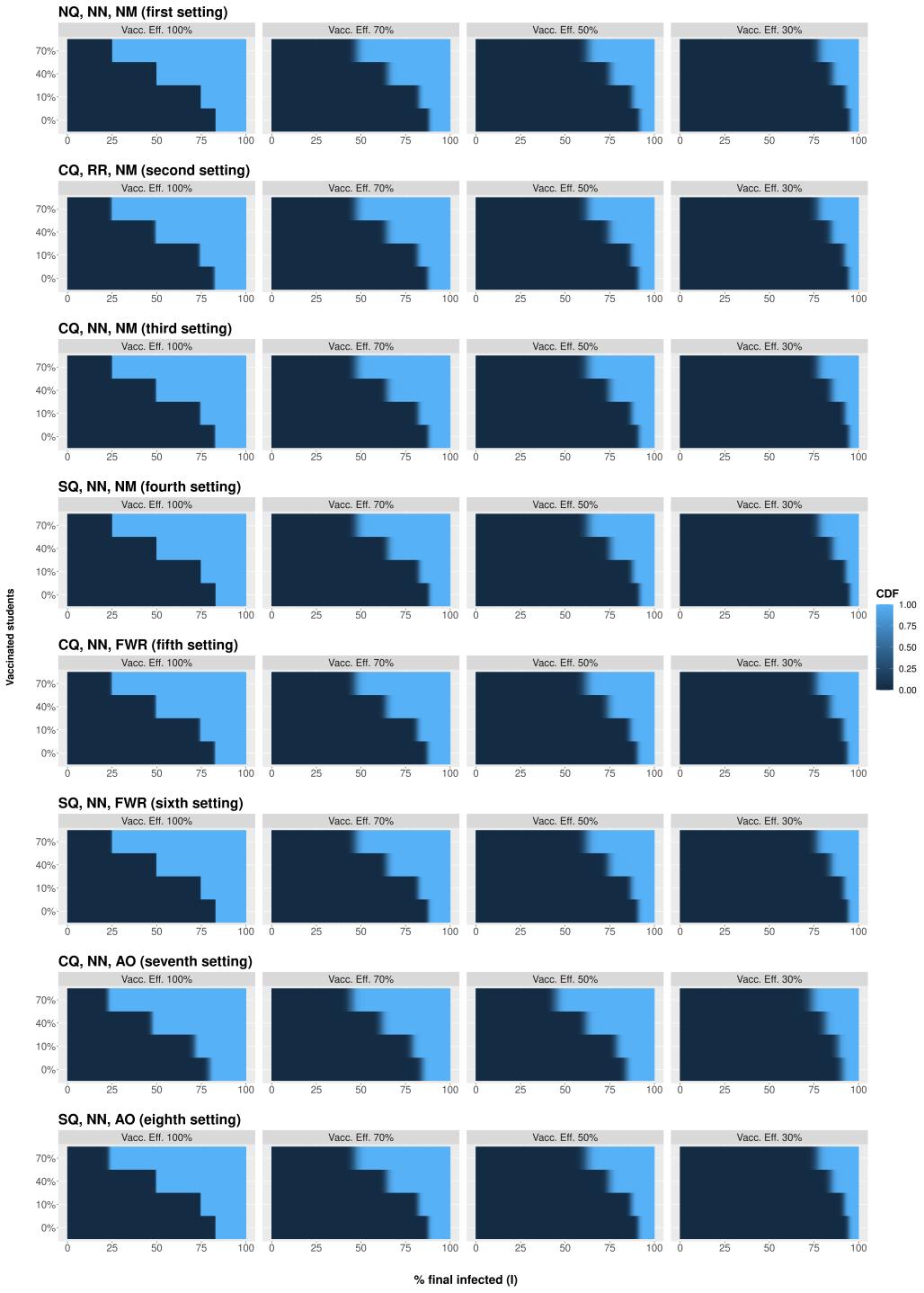


Figure 2.11: Cumulative distribution function (CDF) of the total number of infected agents at the final time for each considered scenario under a high virus circulation level. For each scenario—identified by the setting (on the rows), the vaccine efficacy (on the columns), and the fraction of vaccinated students (on the rows of each subplot)—we run 500 simulations—see Table 2.5. Lighter CDFs indicate more effective countermeasures. We can state that higher infections are associated with lower vaccine efficacies and lower fractions of vaccinated students and a combination of masks and quarantine policies allows to reduce the final number of infected agents.

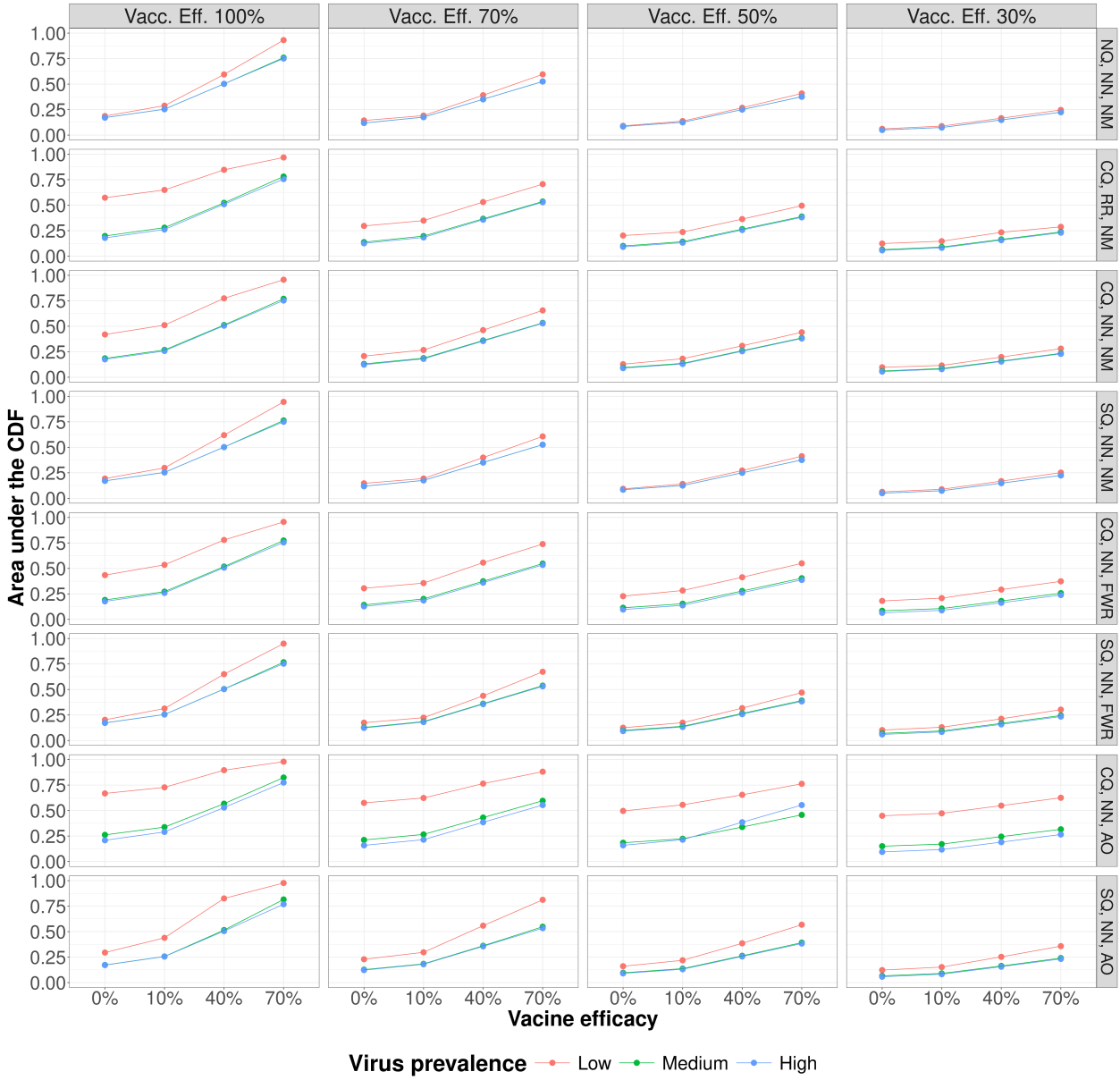
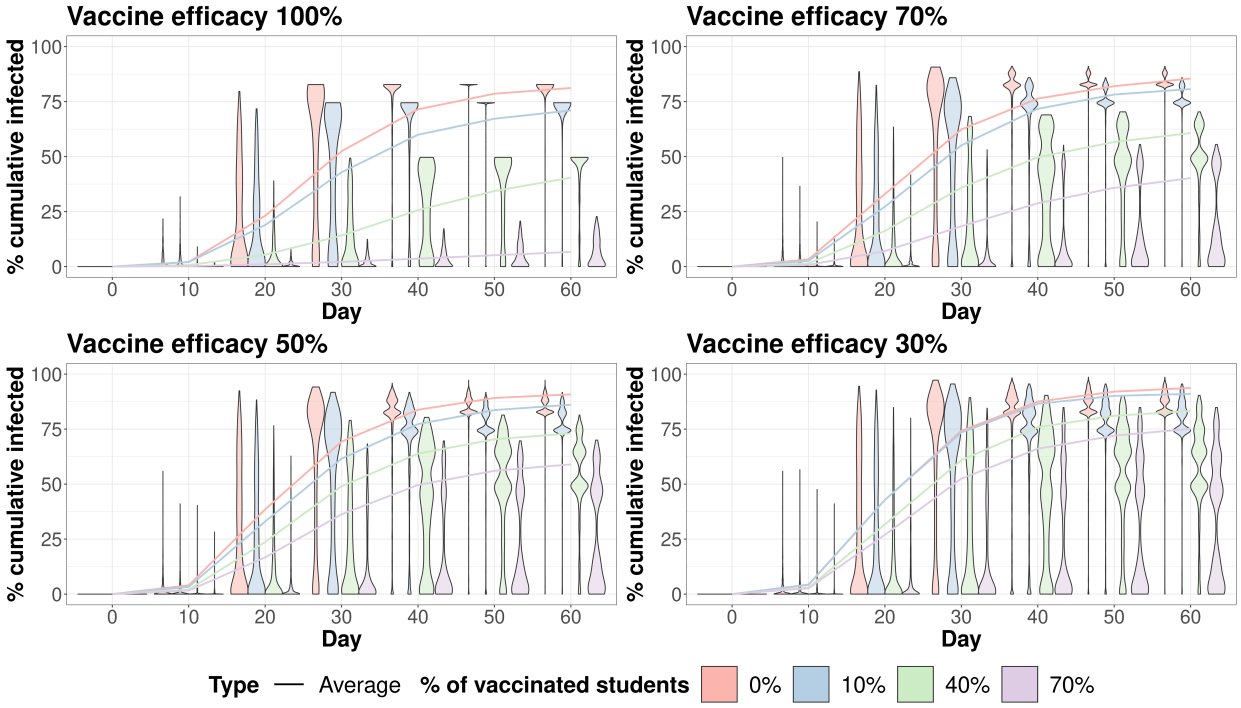


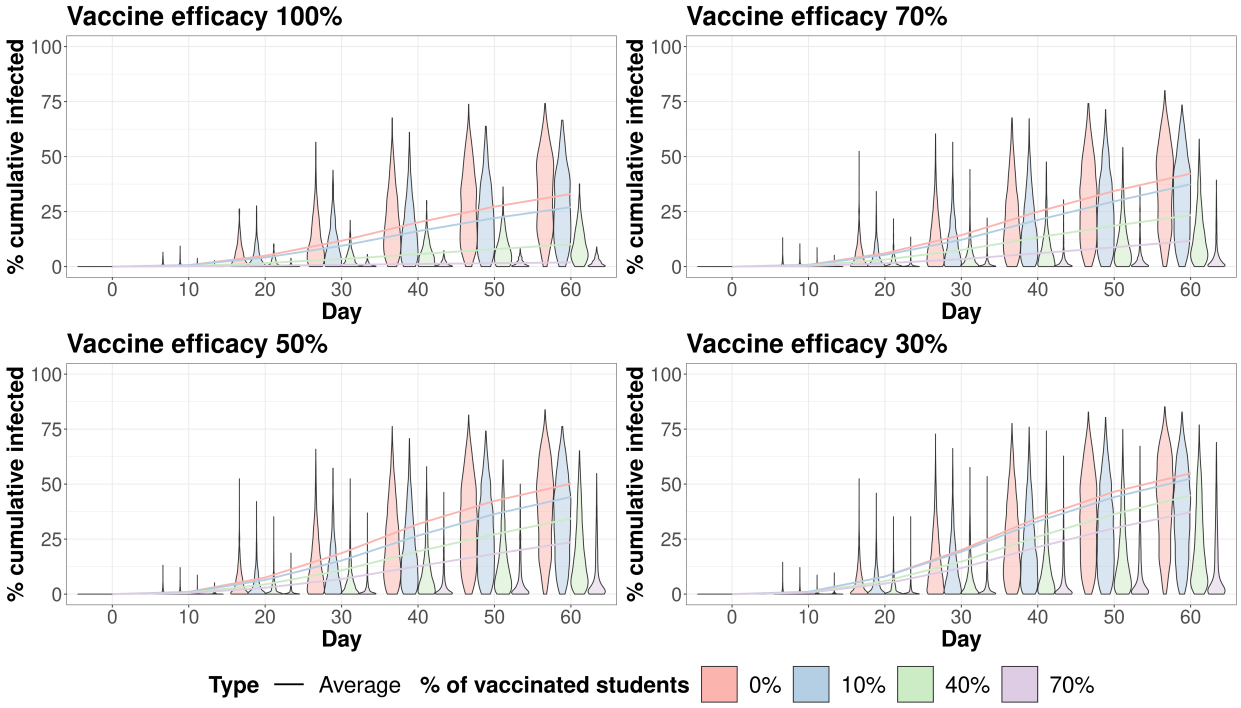
Figure 2.12: Area under the CDFs—normalized between 0 and 1—showed in Figure 2.9—low virus circulation, in red—, Figure 2.10—medium virus circulation, in green—, and Figure 2.11—high virus circulation, in blue. For each scenario—identified by the setting (on the rows), the vaccine efficacy (on the columns), the fraction of vaccinated students (on the columns of each subplot), and the virus circulation level (distinguished by the colors)—, the values represent the effectiveness of the implemented countermeasures in the considered scenario. Higher values indicate greater effectiveness of the countermeasures. Virus prevalence appears to influence the area under the CDF differently across the scenarios, with low prevalence often having higher values compared to medium and high prevalence, particularly in scenarios with higher vaccine efficacy.

a low virus circulation level. However, a combination of masks and quarantine policies can effectively reduce the final number of infected agents. Moreover, Figure 2.12 displays the area under the CDF (normalized between 0 and 1) for each scenario, making it easier to compare the effects across different levels of virus circulation.

Figures 2.13-2.15 illustrate the distribution of infections over time and the average number

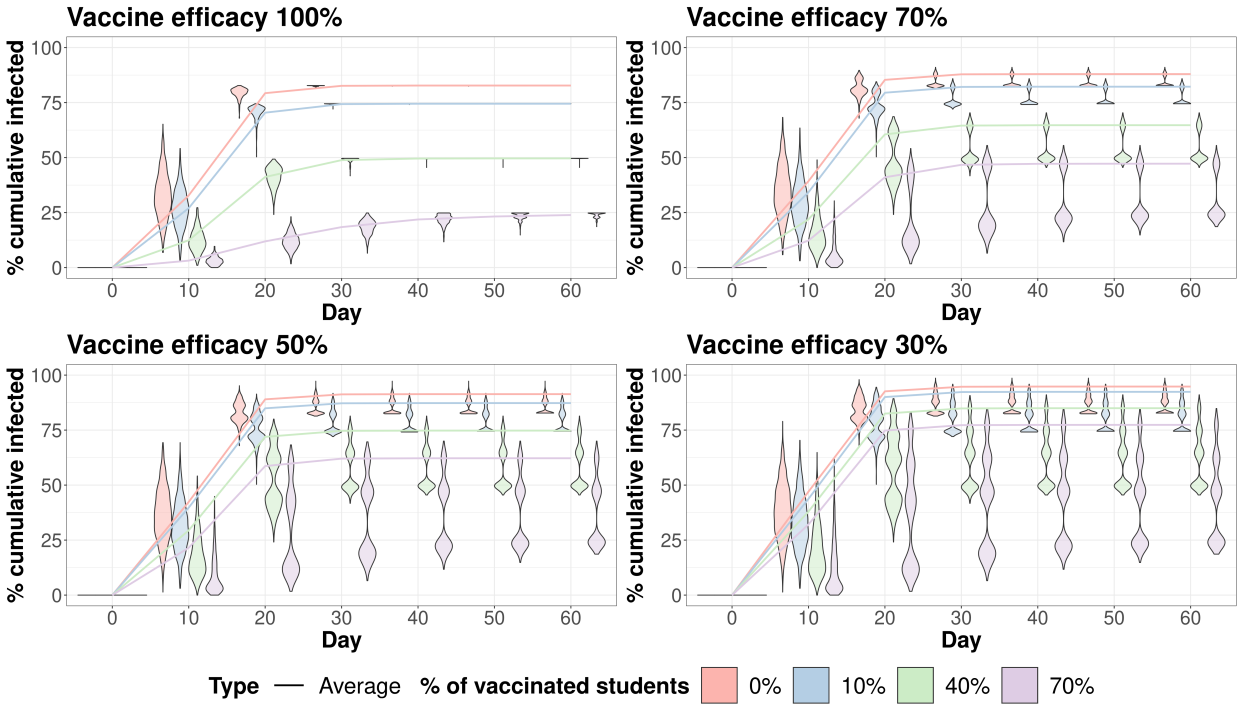


(a) First setting (NQ, NN, NM). Various probability masses emerge with two main outcomes: the infection fades or explodes in a major outbreak.

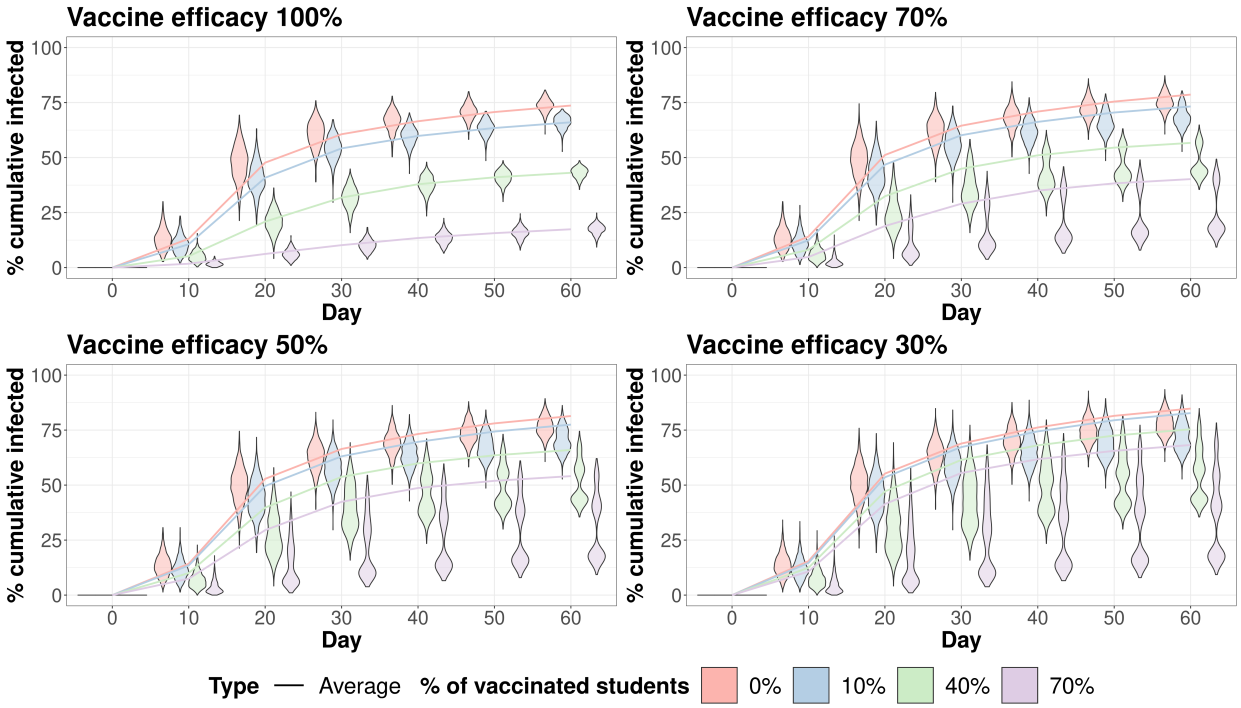


(b) Seventh setting (CQ, NN, AO). The outcome leading to a major outbreak is no longer present: masks and quarantine slow down the infection and contain the risk of a major outbreak.

Figure 2.13: Distributions of infections over time and the average number of infections for each considered scenario in the first and seventh settings under a low virus circulation level. For each scenario—the sub-figures illustrate two distinct settings, with each subplot representing a different vaccine efficacy level and the varying colors in each subplot correspond to different fractions of vaccinated students—we run 500 simulations—see Table 2.5.

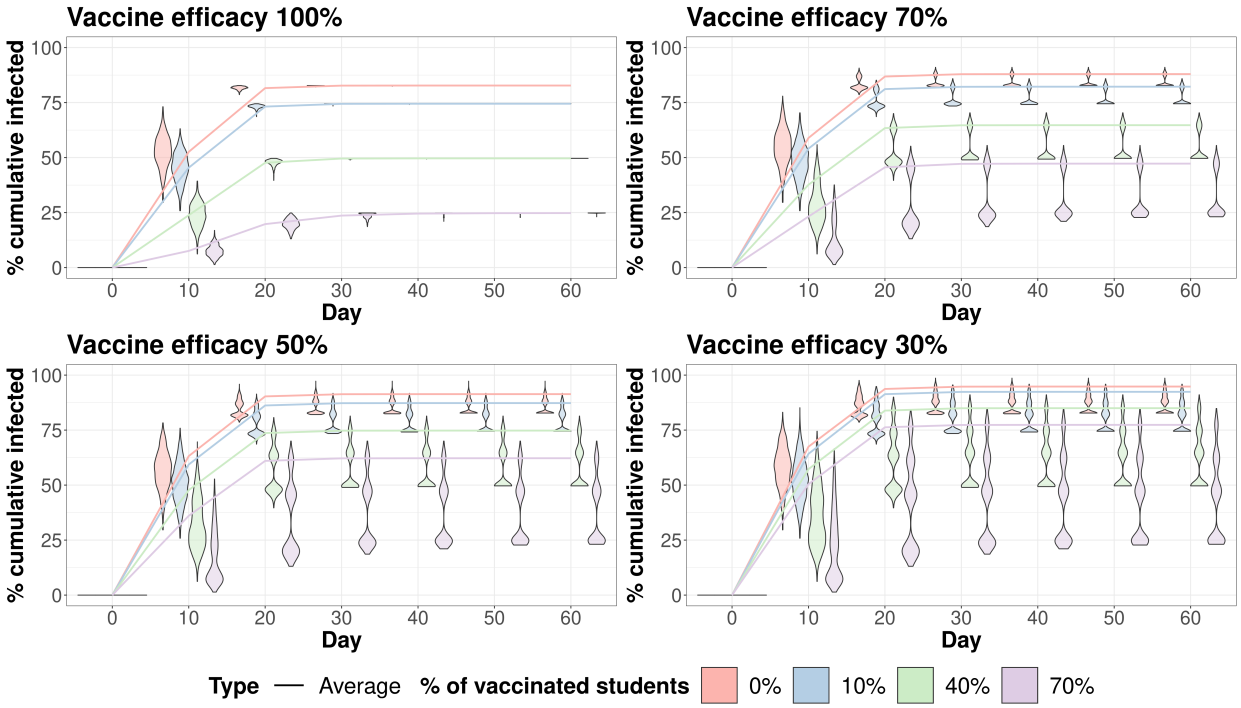


(a) First setting (NQ, NN, NM).

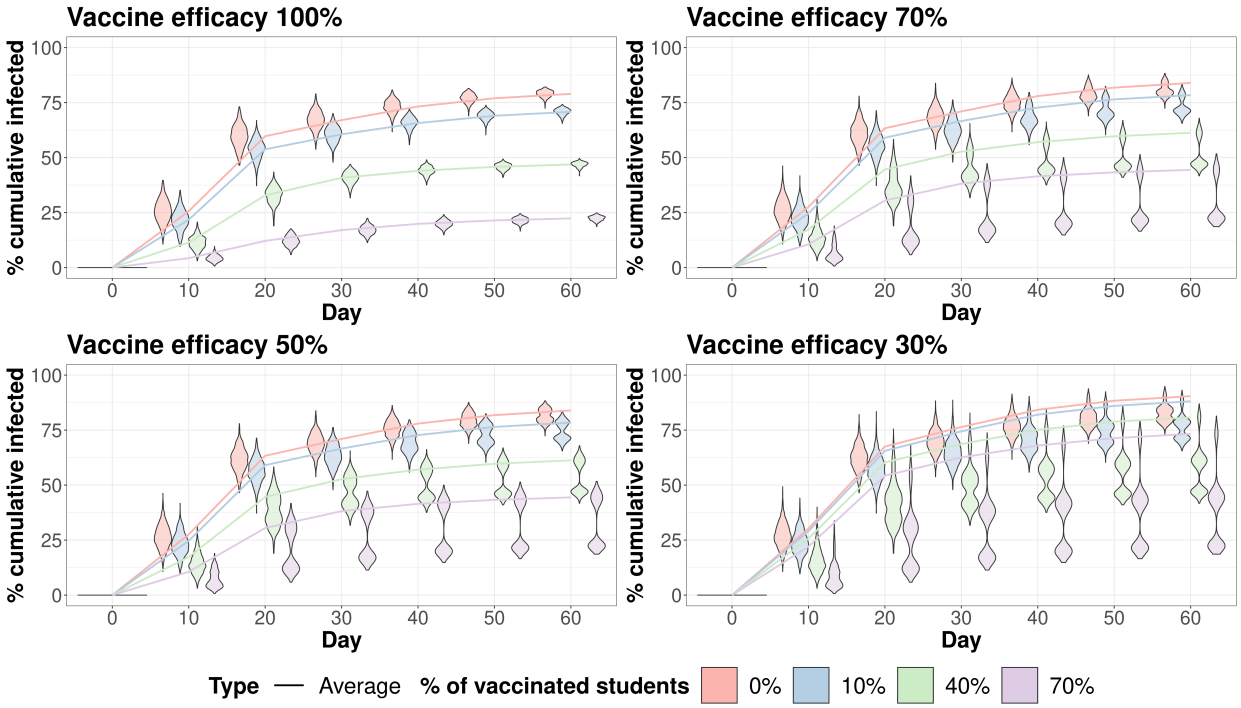


(b) Seventh setting (CQ, NN, AO).

Figure 2.14: Distributions of infections over time and the average number of infections for each considered scenario in the first and seventh settings under a medium virus circulation level. For each scenario—the sub-figures illustrate two distinct settings, with each subplot representing a different vaccine efficacy level and the varying colors in each subplot correspond to different fractions of vaccinated students—we run 500 simulations—see Table 2.5. The vanishing of the infection is no longer present and different probability masses emerged. The most favorable outcome is represented by the probability mass with the lowest mean. A combination of masks and quarantine policies allows to reduce the final number of infections.



(a) First setting (NQ, NN, NM).



(b) Seventh setting (CQ, NN, AO).

Figure 2.15: Distributions of infections over time and the average number of infections for each considered scenario in the first and seventh settings under a high virus circulation level. For each scenario—the sub-figures illustrate two distinct settings, with each subplot representing a different vaccine efficacy level and the varying colors in each subplot correspond to different fractions of vaccinated students—we run 500 simulations—see Table 2.5. The vanishing of the infection is no longer present and different probability masses emerged. The most favorable outcome is represented by the probability mass with the lowest mean. A combination of masks and quarantine policies allows to reduce the final number of infections.

of infections. Representing the evolution of the infection distribution allows us to understand how interventions contribute to containing the spread of the infection and, more significantly, how they shift the probability of a negative outcome towards a more convenient one in the presence of different probability masses. In particular, Figure 2.13a clearly shows two main outcomes under a low virus circulation level: the infection fades or explodes in a major outbreak. Such behavior is easily observed in the violin plot, where the probability mass gets either at zero or between the susceptible population size and half of the susceptible population size—intended as the number of agents not covered by the vaccine immunity. The impact of introducing the AO mask policy and the CQ quarantine policy is clear when we consider Figure 2.13b. Indeed, here the probability mass is shifted towards a lower number of infections at each time point. Also, the outcome leading to a major outbreak is not present anymore. This result clearly shows the efficacy of the NPIs countermeasure: although masks and quarantine are not enough to stop the spread of the infection they allow to slow it down and contain, if not avoid, the risk of a major outbreak under a low virus circulation level.

Figures 2.14a, 2.14b, 2.15a and 2.15b show the possible outcomes under a medium and high virus circulation level. In particular, it is straightforward that the best outcome, i.e., the vanishing of the infection, is no longer present. With a vaccine efficacy less or equal to 70% different probability masses emerge from the results. In particular, the best outcome is represented by the probability mass with the lowest mean. From these figures, it is clear that a combination of masks and quarantine policies allows to reduce the final number of infections. Specifically, the CQ quarantine policy and the AO mask policy represent the best outcome also under a medium and high virus circulation level. More generally, even if we no longer have the extinction of the infection, the use of NPIs makes it possible to slow down its spread. We can conclude by observing that NPIs are effective when they manage to outrun the infection, which is why increasing virus circulation decreases effectiveness.

2.4 Discussion

In this study, we evaluated the impact of different non-pharmaceutical intervention policies on the reduction of COVID-19 infections in the school environment using an agent-based model approach. The computational modeling allowed us to pinpoint several useful takeaways for policy-makers. Overall, we observed that the combination of different NPIs can drastically reduce virus transmission in the school environment after the introduction of an infected case. Our model shows that, in the absence of containment measures, the introduction of a single infective case can lead to the infection of the majority of school students. However, the combination of mask-wearing and room air ventilation can drastically reduce, on average, viral transmission. In our simulation, after the introduction of one infective student, the use of these two NPIs decreased the cumulative number of infected students by 79% (from 237 to 50). In this study, we also tested the effect of three screening policies under different participation levels (25%, 50%, 100%). Our model suggests that screening can further reduce viral transmission through early identification and isolation of infective cases. In addition, we found that even with low participation, screening can still have a not negligible impact in terms of COVID-19 cases avoided.

One of the interesting features of the proposed model is that it takes into account students', teachers', and the principal's behaviors both in the classroom and in the shared spaces, allowing the simulation of the whole school environment rather than the infection dynamics of a single class. In addition, the model incorporates two different ways of viral transmission, namely the direct viral transmission occurring through close-range contacts and the viral quanta diffusion through aerosol. Our findings are consistent with the literature results. For instance, in the validation analysis, our model suggests that the viral transmission in the school environment is mainly due to the aerosolization of viral quanta as already suggested by other studies [114]. In addition, the estimated effects of wearing masks and indoor ventilation are consistent with the results found in other modeling studies [80].

Different screening plans in the school environment have been compared through a stochastic compartmental model by Baccini and Cereda [43] and found that testing half of the class on rotation every week was better than testing all students every two weeks, allowing earlier detection of cases, especially when the rate of contagion in the class is high. Less expensive strategies, such as a weekly test of one-fourth of the students gave comparable results in terms of sensitivity. The latter is one of the strategies tested using our model. Mass testing and school screening have been criticized [177] by experts, mainly based on the insufficient precision of the tests and the risk that false reassurance of a negative test may lead to more risky behavior that would have occurred without the test.

However, screening campaigns similar to those modeled by this study have been adopted and implemented in the real world [269, 53]. In North West Italy, a regional screening program was implemented aimed at second- and third-grade students in first-level secondary schools based on voluntary participation in the first quarter of 2021 [106]. In the study period, the cumulative incidence in schools participating in the program was nearly 4 times lower than that in non-participating schools. Models such as the one proposed here may provide insight into under which circumstances school screenings may result in fewer quarantined classes and a lower loss of school days.

Moreover, we compared the effects of wearing masks all the time and the implementation of a testing protocol, using two quarantine policies used in the Piedmont region of Italy in 2021 and 2022. First of all, vaccination campaigns covering less than 40% of the susceptible population provide poor results in protecting from a major outbreak. Furthermore, we showed that testing campaigns are effective as far as the testing process is faster than the virus diffusion. On the other hand, wearing masks properly is an effective countermeasure also when the infection speeds up.

Eventually, the introduction of an infection external to the school allows us to consider different settings of infection prevalence—where in this work we consider three different levels of infection's prevalence in the general population. In particular, in a setting with a medium or high circulation of the virus, the results show that it is more challenging to contain the spread for a low virus circulation level, but a combination of masks and quarantine policies allows to reduce the final number of infections.

2.5 Conclusion

We developed an agent-based model to study the evolution of SARS-CoV-2 spread in a generic middle school. We used this model to showcase different NPIs approaches and vaccination strategies to contain the spreading of the virus. The significance of risk reduction through the policies assessed in this work is potentially relevant to public health authorities and school administrators at both the local and national levels.

Implementation of standard preventive measures such as mask-wearing and environmental ventilation should be considered necessary in phases of low viral circulation. In contrast, active preventive measures such as screening campaigns can be seen as additional measures to be implemented to reduce viral transmission in the school environment during high viral circulation periods to avoid massive school closures.

NPIs are useful for infection, especially in the early stages, and can help to keep activities open even in scenarios with higher circulation, but other public health instruments must support them.

While this study offers valuable insights, there are several limitations that should be acknowledged. First, certain assumptions in the model—such as external screening campaigns and infection prevalence—may oversimplify real-world dynamics. Second, although the findings offer meaningful implications for specific environments, their applicability to other school settings or broader public spaces may be limited due to differences in cultural, social, and infrastructural factors.

Future work could refine the model by incorporating more detailed data on screening campaigns and infection prevalence to better reflect real-world dynamics. Additionally, expanding the model to include diverse school settings and public spaces would improve its generalization. Some of these limitations are addressed in Chapter 5.

2.6 Availability of data and software code

The ABM developed for the analyses presented in this work and the configurations presented above are available at `https://github.com/qBioTurin/epischool-abm/`.

2.6.1 How to reproduce the results

Requirements

You need to have docker installed on your computer, for more info see this document: `https://docs.docker.com/engine/installation/`. Ensure your user has the right to run docker (without the use of sudo). To create the docker group and add your user in a Unix system:

- Create the docker group:

```
$ sudo groupadd docker
```

- Add your user to the docker group:

```
$ sudo usermod -aG docker $USER
```

- Log out and log back in so that your group membership is re-evaluated.

Reproduce

To reproduce the results in the first set—Section 2.3.1—clone the repository (`https://github.com/qBioTurin/epischool-abm.git`) and run:

```
$ cd epischool-abm
$ git checkout JASSS
$ ./reproduce.sh
```

To reproduce the results in the second set—Section 2.3.2—run (inside the $epischool-abm$ directory):

```
$ git checkout CIBB
$ ./reproduce_extended.sh
```

Chapter 3

Sybil: enhancing COVID-19 forecasting precision with variant-aware integrated models

This chapter outlines computational models to generate predictions within the context of continuous monitoring and early warning systems—Figure 1.2. Accurate forecasting is essential for effective public health planning, resource allocation, and decision-making. It helps develop proactive strategies like vaccination campaigns, travel advisories, and community engagement, enhancing preparedness to control the spread of infectious diseases and minimize their global impact. Specifically, we present **Sybil** [39], an integrated framework that merges a machine learning-based predictive model with a variant-aware compartmental model aimed at enhancing prediction accuracy and explainability. Sybil takes advantage of the relative stability of disease characteristic indices to make future projections and employs a straightforward, widely recognized analytical model to analyze infection dynamics. Sybil stands out from existing literature through its *i)* ability to deliver accurate forecasts even amidst significant changes in the diffusion process, and *ii)* reduced reliance on extensive training data. Additionally, the framework offers *iii)* the capability to examine the evolution of various infection variants and *iv)* ensures the replicability of its results. Moreover, *v)* the open-source code is publicly accessible online.

In this chapter, we begin by outlining the framework developed in our study, including the compartmental model, the machine learning-based predictive model, and the data used (Section 3.1). Next, we present various predictions as results (Section 3.2). Finally, we discuss (Section 3.3), draw conclusions about the key findings (Section 3.4), and show how to replicate the results (Section 3.5).

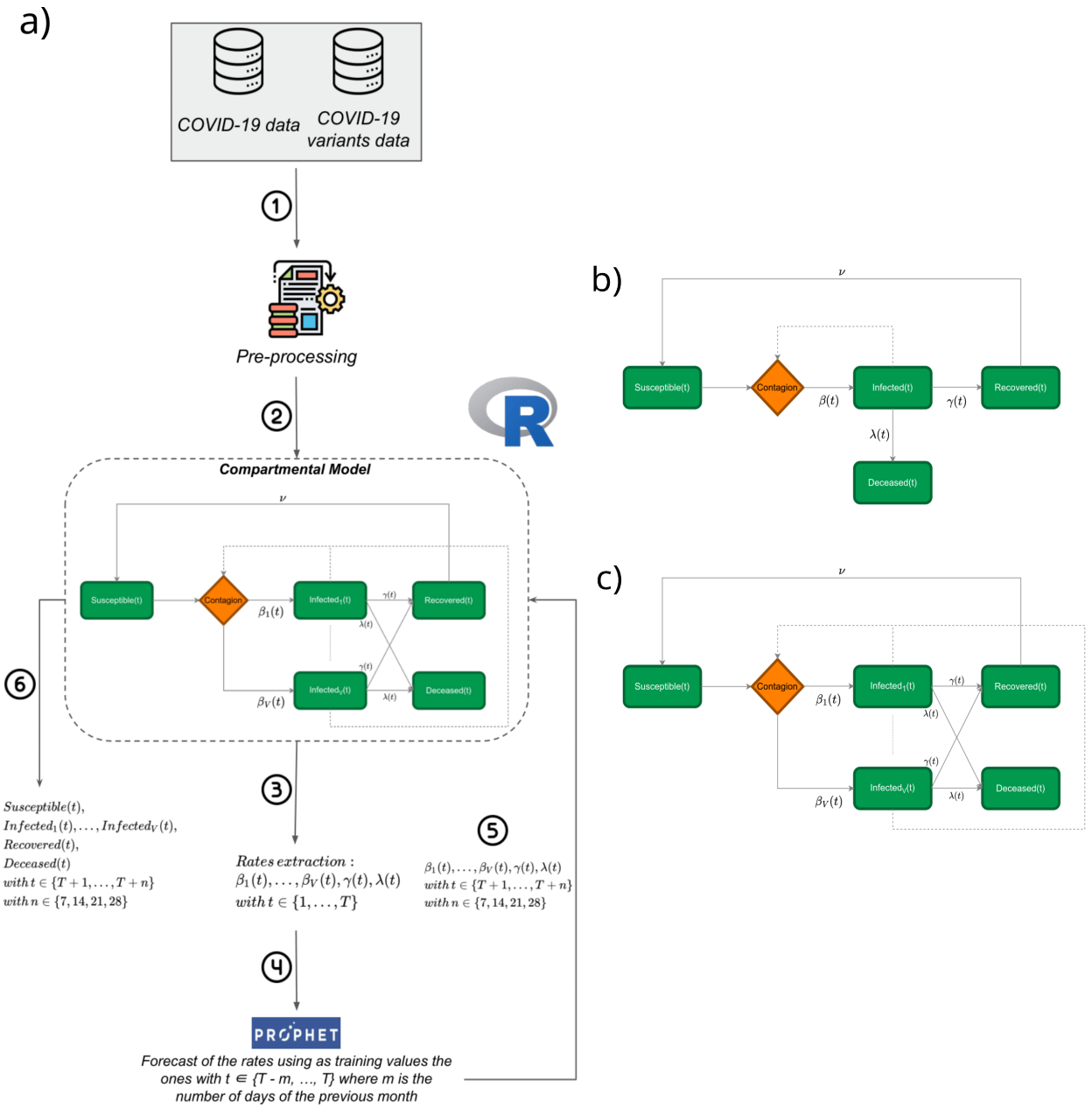


Figure 3.1: *(a)* Schematic representation of Sybil's steps. Starting from COVID-19 data of a specific country *1)* we pre-processed and *2)* injected them into the SI^{V}RDS model reconstructing the evolution of the epidemic in the considered country. Then, *3)* we extracted the infection rates $\beta_v(t)$ for each variant v, the recovery rates $\gamma(t)$ and the fatality rates $\lambda(t)$ with t ∈ {1, ..., T} (where t = 1 represents the first day with available data and t = T represents the forecasting point) and *4)* we applied Prophet [257] using the rates of the previous month (with t ∈ {T - m, ..., T}, where m represents the number of days of the previous month) as training data to predict the rates one, two, three and four weeks in the future (with t ∈ {T + 1, ..., T + n} and n ∈ {7, 14, 21, 28}). Finally, *5)* we injected these predicted rates inside the SI^{V}RDS model to obtain *6)* the evolution of Susceptible(t), Infected$_1$(t), ..., Infected$_V$(t), Recovered(t) and Deceased(t) with t ∈ {T + 1, ..., T + n}) (with n ∈ {7, 14, 21, 28}). *(b)* SIRDS compartmental model. $\beta(t)$, $\gamma(t)$, $\lambda(t)$ and ν refer to infection, recovery, fatality, and end-of-immunization rates, respectively (which are all time-dependent, but the end-of-immunization rate). *(c)* SI^{V}RDS compartmental model. $\beta_1(t)$, ..., $\beta_V(t)$ refer to the infection rates of the V considered variants. Dashed lines in *(b)* and *(c)* refer to the fact that an infection is due to contact among susceptible and infected individuals.

3.1 Sybil framework

We propose an integrated framework aimed at providing accurate and explainable predictions for epidemic spreads. Sybil combines a simple compartmental model with a machine learning-based predictive model to forecast the future progression of infection, even in the presence of multiple virus strains. At the core of Sybil, there is a simple analytical model that has dual functionality. In the first stage of Sybil's operation, the analytical model is used to derive the values of critical model parameters from the surveillance data. Then, these parameters' values—ascribable to the reproduction number over time, R_t—are used as training data for the ML component of Sybil. Based on that knowledge, it predicts the future values for the key parameters, which are then plugged back into the analytical model. Then, Sybil computes the future evolution of daily infections using the analytical model with these key parameters' future values (see Figure 3.1-*(a)* for a schematic representation of Sybil's steps).

The effectiveness and accuracy of Sybil's forecasts are demonstrated through two methods: *i)* comparing them against real data gathered from active surveillance of the pandemic in several European and U.S. states, and *ii)* comparing them against forecasts generated by various state-of-the-art approaches, including Prophet [257], ARIMA / SARIMA [61], Neural Prophet [263], LSTM [131] and GRU [79] neural networks, EpiNow2 [235], and the framework by Watson et al. [274]. In the latter case, we predict the evolution of different active variants over the specified period.

3.1.1 Compartmental analytical model

The first component of Sybil is the analytical model. Specifically, the model used is a Susceptible-Infected-Recovered-Deceased-Susceptible (SIRDS) compartmental model with reinfections—represented in Figure 3.1-*(b)*—that is described by the system of Ordinary Differential Equations (ODEs) in Equation 3.1.

$$\begin{aligned}
\frac{dS}{dt} &= -\beta(t)\frac{S(t)I(t)}{N} + \nu R(t) \\
\frac{dI}{dt} &= \beta(t)\frac{S(t)I(t)}{N} - \gamma(t)I(t) - \lambda(t)I(t) \\
\frac{dR}{dt} &= \gamma(t)I(t) - \nu R(t) \\
\frac{dD}{dt} &= \lambda(t)I(t)
\end{aligned} \tag{3.1}$$

In this system, β represents the infection rate, γ the recovery rate, λ the fatality rate, ν the end-of-immunization rate, and N the total population. We discretized this continuous

system of ODEs obtaining a system of difference equations that express how the sizes of the compartments vary over time on a discrete scale—Equation 3.2; t is a continuous variable, while $\tilde{t}$ is a discrete variable that takes as values multiples of $\Delta\tilde{t}$.

$$\begin{aligned}
\frac{S(\tilde{t}+\Delta\tilde{t})}{\Delta\tilde{t}} &= S(\tilde{t}) - \beta(\tilde{t})\frac{S(\tilde{t})I(\tilde{t})}{N} + \nu R(\tilde{t}) \\
\frac{I(\tilde{t}+\Delta\tilde{t})}{\Delta\tilde{t}} &= I(\tilde{t}) + \beta(\tilde{t})\frac{S(\tilde{t})I(\tilde{t})}{N} - \gamma(\tilde{t})I(\tilde{t}) - \lambda(\tilde{t})I(\tilde{t}) \\
\frac{R(\tilde{t}+\Delta\tilde{t})}{\Delta\tilde{t}} &= R(\tilde{t}) + \gamma(\tilde{t})I(\tilde{t}) - \nu R(\tilde{t}) \\
\frac{D(\tilde{t}+\Delta\tilde{t})}{\Delta\tilde{t}} &= D(\tilde{t}) + \lambda(\tilde{t})I(\tilde{t})
\end{aligned} \tag{3.2}$$

The rates are time-dependent—meaning that they may vary at each time step, with the time step corresponding to one day—i.e. $\Delta\tilde{t} = 1$, which is the frequency of the observations in the datasets—Equation 3.3. The only exception is end-of-immunization rate ν, which is assumed to be $\nu = \frac{1}{180}$ since on average the immunization due to infection is estimated to be lost after 180 days [277].

$$\begin{aligned}
S(\tilde{t}+1) &= S(\tilde{t}) - \boldsymbol{\beta(\tilde{t})}\frac{S(\tilde{t})I(\tilde{t})}{N} + \nu R(\tilde{t}) \\
I(\tilde{t}+1) &= I(\tilde{t}) + \boldsymbol{\beta(\tilde{t})}\frac{S(\tilde{t})I(\tilde{t})}{N} - \boldsymbol{\gamma(\tilde{t})}I(\tilde{t}) - \boldsymbol{\lambda(\tilde{t})}I(\tilde{t}) \\
R(\tilde{t}+1) &= R(\tilde{t}) + \boldsymbol{\gamma(\tilde{t})}I(\tilde{t}) - \nu R(\tilde{t}) \\
D(\tilde{t}+1) &= D(\tilde{t}) + \boldsymbol{\lambda(\tilde{t})}I(\tilde{t})
\end{aligned} \tag{3.3}$$

Using the surveillance data—possibly after a pre-processing phase (step *1)* of Figure 3.1-*(a)*)—, Sybil computes the evolution of the infection process by solving the SIRDS model in Equation 3.4, where $\Delta I(\tilde{t})$ represents the new infections at time $\tilde{t}$, $\Delta R(\tilde{t})$ the new recoveries at time $\tilde{t}$ and $\Delta D(\tilde{t})$ the new deceases at time $\tilde{t}$ (with S(0) = N and I(0) = R(0) = D(0) = 0)—step *2)* of Figure 3.1-*(a)*.

However, obtaining all the required parameters to solve equations in Equation 3.3 is not straightforward. Indeed, surveillance data does not provide the transition rates—namely, the bold elements in the system of difference equations—and hence a further step is required: using Equation 3.5 (obtained from Equation 3.3), we can estimate the daily infection, recovery, and fatality rates—step *3)* of Figure 3.1-*(a)*.

$$\begin{aligned} S(\tilde{t}+1) &= S(\tilde{t}) - \Delta I(\tilde{t}) + \nu R(\tilde{t}) \\ I(\tilde{t}+1) &= I(\tilde{t}) + \Delta I(\tilde{t}) - \Delta R(\tilde{t}) - \Delta D(\tilde{t}) \\ R(\tilde{t}+1) &= R(\tilde{t}) + \Delta R(\tilde{t}) - \nu R(\tilde{t}) \\ D(\tilde{t}+1) &= D(\tilde{t}) + \Delta D(\tilde{t}) \end{aligned} \tag{3.4}$$

$$\begin{aligned} \lambda(\tilde{t}) &= \frac{D(\tilde{t}+1) - D(\tilde{t})}{I(\tilde{t})} && \text{if } I(\tilde{t}) > 0, \text{ else } \lambda(\tilde{t}) = 0 \\ \gamma(\tilde{t}) &= \frac{R(\tilde{t}+1) - R(\tilde{t}) + \nu R(\tilde{t})}{I(\tilde{t})} && \text{if } I(\tilde{t}) > 0, \text{ else } \gamma(\tilde{t}) = 0 \\ \beta(\tilde{t}) &= \frac{I(\tilde{t}+1) - I(\tilde{t}) + \gamma(\tilde{t})I(\tilde{t}) + \lambda(\tilde{t})I(\tilde{t})}{S(\tilde{t})I(\tilde{t})}N && \text{if } S(\tilde{t})I(\tilde{t}) > 0, \text{ else } \beta(\tilde{t}) = 0 \end{aligned} \tag{3.5}$$

Bringing in variants into the model of Equation 3.3— again, possibly after a pre-processing phase (step *1)* of Figure 3.1-*(a)*)—requires introducing one additional compartment to account for infections caused by each virus strain, hence obtaining an $\mathrm{SI^{V}RDS}$ model, where the superscript at the I stands for the maximum number of virus strains included in the model (see Figure 3.1-*(c)*). After introducing further compartments, also additional rates are necessary. Specifically, instead of a global infection rate $\beta(\tilde{t})$, there are V different infection rates—one for each variant—for each time step $\tilde{t}$. Making a simplifying assumption, Sybil assumes that the evolution of the $I_v(\tilde{t})$ compartment—for each variant v and for each time step $\tilde{t}$—is computed starting from the evolution of the $I(\tilde{t})$ compartment of the model without variants and the daily proportion of the considered variant; that is,

$$I_v(\tilde{t}) = I(\tilde{t})\pi_v(\tilde{t}) \tag{3.6}$$

where $\pi_v(\tilde{t})$ is the proportion of infections due to variant v at time $\tilde{t}$. Moreover, we assumed no correlations among variants and that being recovered from a variant v makes a person immune from all variants (and after $\frac{1}{\nu}$ days, on average, a person will be susceptible again).

$$I_v(\tilde{t}+1) = I_v(\tilde{t}) + \boldsymbol{\beta_v(\tilde{t})}\frac{S(\tilde{t})I(\tilde{t})}{N} - \gamma(\tilde{t})I_v(\tilde{t}) - \lambda(\tilde{t})I_v(\tilde{t}) \qquad v \in \{1, ..., V\} \tag{3.7}$$

$$\beta_v(\tilde{t}) = \frac{I_v(\tilde{t}+1) - I_v(\tilde{t}) + \gamma(\tilde{t})I_v(\tilde{t}) + \lambda(\tilde{t})I_v(\tilde{t})}{S(\tilde{t})I(\tilde{t})}N \qquad v \in \{1, ..., V\} \tag{3.8}$$

Equation 3.7 describes the evolution of the I_v compartments—step *2)* of Figure 3.1-*(a)*. Here, we used the previously computed global recovery and fatality rates, and $\beta_v(\tilde{t})$ is a column vector with the time-dependent infection rates of V variants. In particular, we took infections from the $I(\tilde{t})$ compartment instead of the $I_v(\tilde{t})$ compartments to avoid approximation errors. This simplification is necessary because, especially during the early stage of a variant, we have small values in the $I_v(\tilde{t})$ compartments that could cause issues with Sybil. Consequently, in Equation (5) we have a vector of V rates $\beta_v(\tilde{t})$ rather than a matrix of V×V rates. Moreover, even if we were to use a matrix, we could only compute the main diagonal values, as we lack the necessary ground-truth data to distinguish recoveries and deaths for each variant. Such detailed information, including specific recovery and fatality rates for each variant, is not available in the data sets we used. Finally, Equation 3.8 (obtained from Equation 3.7) describes how to compute the infection rates $\beta_v(\tilde{t})$ for the v variant—step *3)* of Figure 3.1-*(a)*.

It is worth noticing that in devising infection and fatality rates from surveillance data, we incorporate all the effects due to the different virus strains, the impact of contention policies, and the impact of people's mobility, even if such details are not explicitly detailed in the model.

3.1.2 Prophet predictive model

The second component of the Sybil framework is Prophet [257], an open-source framework developed by Facebook for time series forecasting. It is based on an additive model design and it has been conceived to have intuitive parameters that can be adjusted without knowing the details of the underlying model. The modeling approach of Prophet combines the strengths of both statistical modeling and machine learning techniques; it utilizes a generalized additive model that incorporates piece-wise linear trends, nonlinear growth, and seasonality adjustments using a Fourier series. This flexible modeling approach enables Prophet to capture simple and complex data patterns. In particular, it consists of three main model components: trend, seasonality, and holidays combined in the following equation:

$$y(\tilde{t}) = g(\tilde{t}) + s(\tilde{t}) + h(\tilde{t}) + \epsilon_{\tilde{t}} \tag{3.9}$$

where $g(\tilde{t})$ is the trend function that models non-periodic changes, $s(\tilde{t})$ represents periodic changes (e.g., daily, weekly, and yearly seasonality), and $h(\tilde{t})$ represents the effects of holidays

which occur on potentially irregular schedules over one or more days. The error term $\epsilon_{\tilde{t}}$ represents any characteristic changes that are not accommodated by the model. Furthermore, Prophet makes it possible to estimate uncertainty in trend forecasts. It employs Markov Chain monte Carlo (MCMC) to generate many plausible future trajectories. The MCMC procedure randomly samples from the posterior distribution of the model parameters, allowing for a range of possible outcomes. These sampled parameter sets are then used to generate multiple forecast trajectories.

As explained in Section 3.1.1, from Equations 3.5 and 3.8 we extracted the infection rates $\beta_v(\tilde{t})$ for each variant v, the recovery rates $\gamma(t)$ and the fatality rates $\lambda(t)$ (with t $\in$ {1, ..., T}, where t = 1 represents the first day with available data and t = T represents the forecasting point). Starting from these rates, we applied Prophet to predict the values of the rates one, two, three and four weeks in the future (with t $\in$ {T + 1, ..., T + n} and n $\in$ {7, 14, 21, 28}) using as training data the values of the rates of the previous month (with t $\in$ {T - m, ..., T}, where m represents the number of days of the previous month)—step *4)* in Figure 3.1-*(a)*—(we also tried to use more months, one year and the whole history as training data, but the performance of Sybil have worsened considerably due to the high variability present over long periods). Finally, we injected these new values into the SI$^{\text{V}}$RDS model—step *5)* in Figure 3.1-*(a)*—obtaining the evolution of each compartment one, two, three and four weeks in the future—step *6)* in Figure 3.1-*(a)*.

Why Prophet?

We chose Prophet as the machine learning component of Sybil for its flexibility, non-linearity, robustness, and user-friendliness. It is designed for time series forecasting and is particularly effective for data with strong seasonal patterns and historical trends. It uses an additive model to fit non-linear trends with daily, weekly, and yearly seasonality, and it can also incorporate holiday effects. This tool is robust to missing data and outliers, and it provides automated, fast, and tunable forecasts.

We also conducted a comparative analysis between Prophet and ARIMA [61]. Specifically, we performed preliminary work using ARIMA, which resulted in comparable outcomes to those achieved with Prophet. However, it is important to note that ARIMA requires a substantial amount of domain knowledge to effectively select and interpret parameters, such as the maximum orders of differencing, auto-regressive components, moving average components, and others. Moreover, ARIMA is susceptible to significant trend errors when there are changes in trend near the cutoff period and may struggle to capture seasonality [257].

3.1.3 Data

All datasets used in this work are open-source and publicly available. The surveillance data used for the considered countries—namely, Italy, Austria, Belgium, France, Germany, Sweden, and the U.S.—in Section 3.2 section are available in the COVID-19 R library [121, 120]—for the results presented for Italy and Austria we used a snapshot dated November 22$^{\text{nd}}$, 2023, for Belgium, France, Germany, and Sweden a snapshot dated May 2$^{\text{nd}}$, 2024, and for the U.S. a snapshot dated June 13$^{\text{th}}$, 2024. This data comprises all the data reported in Figures 3.2 to 3.34 plus other data not relevant to the present study—such as the number of vaccinations, tests, hospitalizations and people in intensive care, information about applied countermeasures, data on mobility, etc. In particular, from the COVID-19 Data Hub [121, 120], we used the following data (from February 2020 to May 2023): the cumulative number of cases, the cumulative number of recoveries, the cumulative number of deaths, and the total population. Starting from these, we computed the daily new cases, the daily new recoveries, and the daily new deaths we used in Equation 3.4.

Concerning European countries, information about active virus strains comes from the European Center for Disease Control (ECDC) [1, 148]—for the results presented for Italy, Austria, Belgium, France, Germany, and Sweden we used a snapshot dated July 25$^{\text{th}}$, 2023—, while for U.S., this information comes from the Centers for Disease Control and Prevention (CDC) [71]—for the results presented, we used a snapshot dated June 13$^{\text{th}}$, 2024. Here, ECDC and CDC collected the results of serological tests, and there is information about a significant number of variant lineages, such as B.1.1.7, BA.1, BA.2, P.1, XBB, and many others. ECDC data comprises, for each week, the number of sequenced cases (which are a small portion of the new cases) and, for each lineage, the number and the proportion of sequenced cases that had that particular lineage. CDC data comprises, for each week, lineage and Health and Human Services (HHS) U.S. region (and for the whole USA), the proportion of the lineage in the population. To use such data, we aggregated these lineages in four main variant families with the aim of using the WHO variants labels: Alpha, Delta, Omicron, and *Other* [132]—see Section 3.2.5 for some results obtained using a different aggregation of variants in Italy. In particular, the *Other* variant comprises the initial SARS-CoV-2 lineage, all the other lineages (e.g., Beta, Gamma, Kappa), and some noisy values in the surveillance data.

Since variants' diffusion data is provided weekly, in a pre-processing phase, we expanded such data to devise approximated daily values. Specifically, we employed splines—piece-wise-defined mathematical functions that use multiple polynomial segments to create a smooth and continuous curve—to obtain daily values. We used these daily values in Equation 3.6.

We want to highlight that Sybil strongly depends on the availability of data. In [121, 120] there are only a few nations with a complete time series (with a daily step) for the used data. For this reason, we extended Sybil to include a pre-processing step (using splines) to fill in missing data and to move from a weekly to a daily step in case of availability of data with a weekly step—without altering too much of the information present in the time series.

In the following sections, we mainly used time-dependent recovery rates thanks to the availability of well-reported official data—on cases, recoveries, and fatalities—for Italy and Austria. We also reported results on Belgium and France using time-dependent recovery rates—although the official data were not well reported. Moreover, surveillance data available for many countries worldwide reports incorrect data on recoveries or does not report data to devise daily recovery rates. Hence, to dispense with data on recoveries, we have also reported results obtained with fixed recovery rate with mean $\frac{1}{\gamma}$ equals to 14 days [65]—same for each variant—considering more European countries—i.e., Italy, Austria, Belgium, France, Germany, and Sweden—and some U.S. states. The results obtained are very good and show the robustness of the Sybil approach.

3.2 Results

We test the efficacy and the robustness of Sybil taking into account different European countries—i.e., Italy, Austria, Belgium, France, Germany, and Sweden—and some U.S. states. Accurate forecasts strongly depend on the regularity of the data to predict: it is pretty easy to foresee the future evolution of some dimensions increasing—or equally decreasing—linearly. Conversely, predicting the behavior in the presence of sudden changes in those quantities is much harder. Unfortunately, outbreaks, as well as peak infection fading, exhibit such a behavior. We compared the predictions obtained using Sybil in Italy and Austria with the predictions obtained with the plain use of different state-of-the-art approaches—i.e., Prophet [257], ARIMA / SARIMA [61], Neural Prophet [263], LSTM [131] and GRU [79] neural networks, and EpiNow2 [235]—to forecast the evolution of the number of daily infections. In the figures, we reported the comparison between Sybil and Prophet since the latter is a component of the former, but we computed the Root Mean Squared Error (RMSE) with the standard deviation between the predictions obtained with Sybil and all the considered state-of-the-art approaches and the ground truth—we reported these values in the figures legends and tables. Moreover, we conducted a comparative analysis with the study of Watson et al. [274] considering some U.S. states.

3.2.1 Italy

Surveillance data from Italy spanning a period from February 2020 to May 2023 is used. Figure 3.2-*a)* shows the evolution of the $I(\tilde{t})$ compartment obtained using Equation 3.4. Figures 3.2-*b)*, 3.2-*c)* and 3.2-*d)* show the evolution of infection, recovery and fatality rates respectively, extracted using Equation 3.5. During this period, there are a significant number of lineages, some coexisting and others with an evolutionary advantage taking over. Figures 3.2-*e)* and 3.2-*f)* show the daily proportion of each variant and the evolution of the $I_v(\tilde{t})$ compartments obtained using Equation 3.7, respectively. Figure 3.2-*g)* shows the evolution of infection rates for each variant.

Vertical dashed lines in Figure 3.2-*f)* mark the time point picked to evaluate the accuracy of Sybil in Italy. The first two and the fourth selected points are placed in the ascending part of an outbreak but close enough to the peak: for accurate predictions, Sybil must successfully reproduce a change in the concavity of the function. The third selected point is placed just before the start of an outbreak to show that Sybil is also able to predict a new emerging variant/outbreak. Lastly, the fifth and sixth selected points are placed during a period of relative stability.

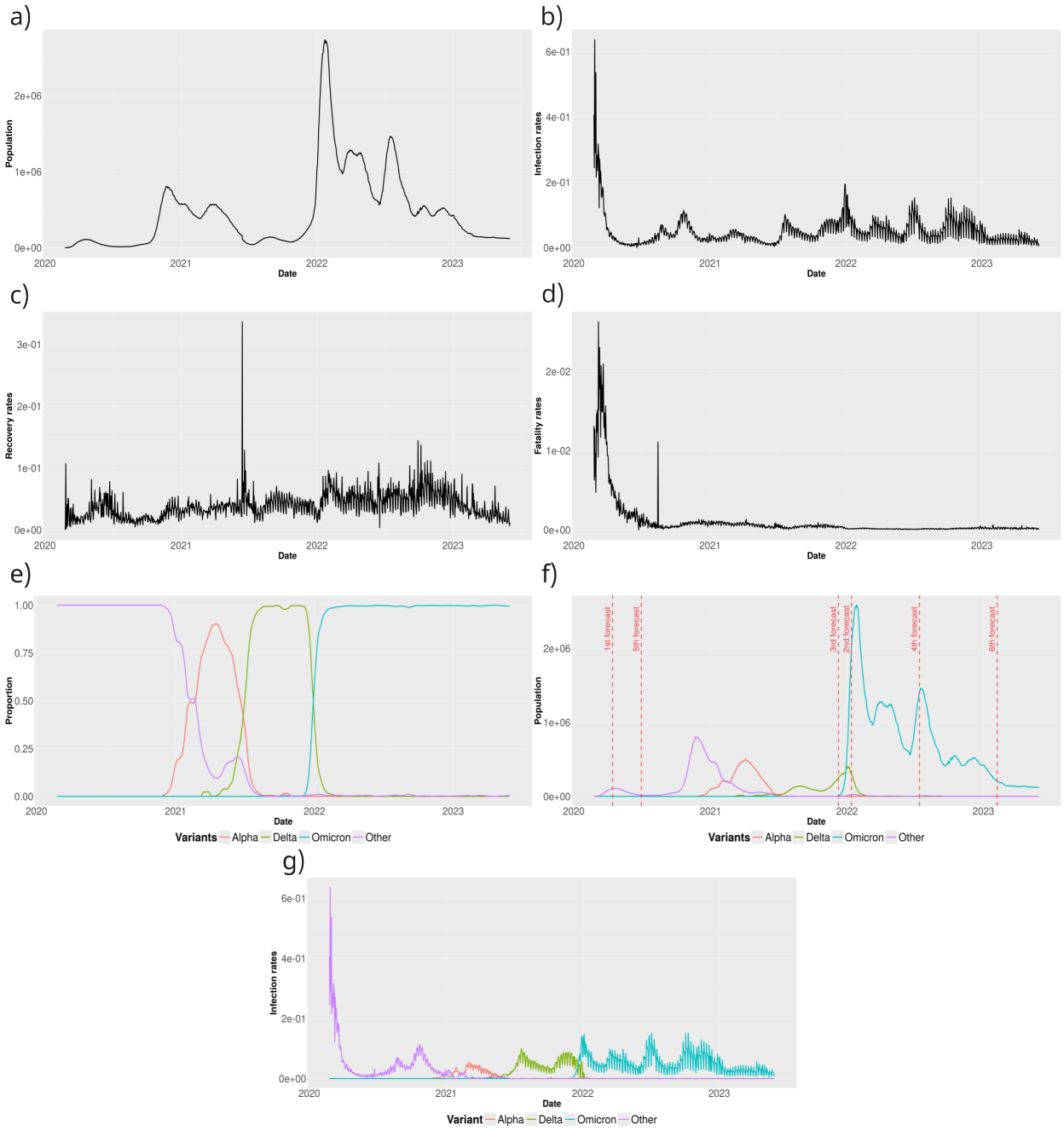


Figure 3.2: The figure shows the evolution of *a)* $I(\tilde{t})$ compartment, *b)* infection rates $\beta(\tilde{t})$, *c)* recovery rates $\gamma(\tilde{t})$, and *d)* fatality rates $\lambda(\tilde{t})$, *e)* the proportion of the aggregated variants, *f)* the evolution of $I_v(\tilde{t})$ compartments with the six considered forecast scenarios—vertical dashed lines mark the selected dates to test Sybil's forecasting—, and *g)* the evolution of infection rates $\beta_v(\tilde{t})$ from February 2020 to May 2023 in Italy.

The first forecast scenario corresponds to the first infection wave, which in Italy started in February 2020. Only one SARS-CoV-2 strain was detected during this initial wave, namely the *Other* variant of Figure 3.2-*f)*. Figure 3.3 reports four different forecasts, all starting on April 15$^{\text{th}}$, 2020, and spanning several forecast windows—from one to four weeks. For each forecasting window, we compare predictions obtained using Sybil (green line) and

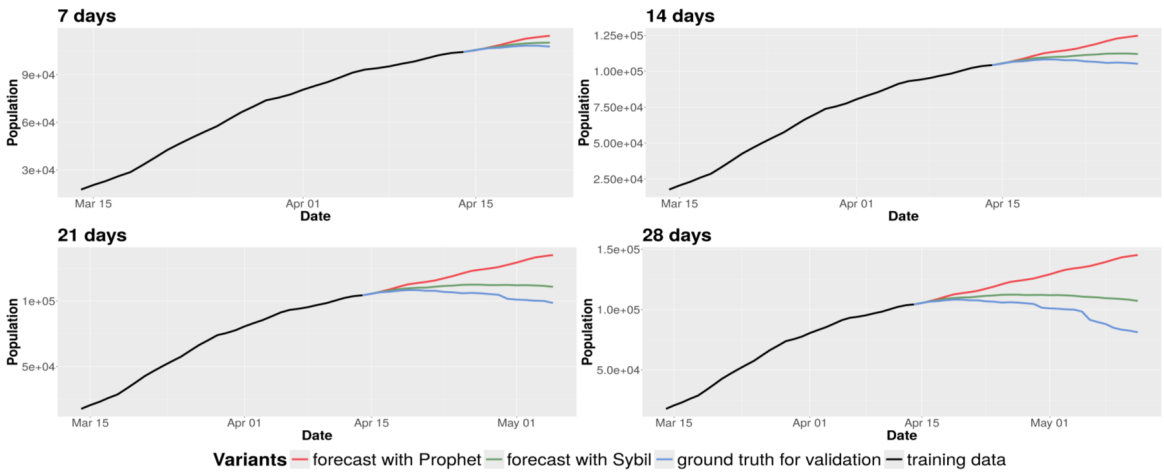


Figure 3.3: Evolution of infections in the first scenario in Italy using Sybil (green line) and Prophet (red line) adopting the same period as training data (black line) comparing and contrasting the predictions against the surveillance data for the period spanning the forecasting window (blue line). We forecast starting from April 14$^{\text{th}}$, 2020. The RMSEs between the ground-truth and the Sybil forecast are 1381 (1 week), 4141 (2 weeks), 7149 (3 weeks), and 13197 (4 weeks). The RMSEs between the ground-truth and the Prophet forecast are 3893 (1 week), 10851 (2 weeks), 19615 (3 weeks), and 32593 (4 weeks).

the standard approach used in the literature (red line), which requires using the selected forecasting approach—Prophet in our case—to project the number of infections in the coming weeks. The two approaches use the same period as training data (black line) for a fair comparison. They are compared and contrasted against the surveillance data for the period spanning the forecasting window (blue line). Comparing predictions obtained with the plain use of Prophet to forecast the evolution of the number of daily infections, we notice that it fails to predict the plateau characterizing the highest part of the peak. Considering Sybil's predictions, we see that starting from short-term predictions—i.e., seven days—it catches the decreasing trend. Increasing the forecasting window, Sybil always provides an excellent approximation of the future evolution of the daily infections, but for four weeks. In this latter case, Sybil's predictions diverge in magnitude from the ground truth, but not in trend. It provides a qualitative indication of the future evolution of the pandemic, which eventually fades. Figures 3.4-*a)*, 3.4-*b)*, and 3.4-*c)* show the infection rate $\beta(\tilde{t})$, the recovery rate $\gamma(\tilde{t})$, and the fatality rate $\lambda(\tilde{t})$ for the observation period. Here, we clearly see that the values for the rate are noisy; what we see is not a smooth line, but a saw-tooth function with many peaks and valleys. Nonetheless, the range of variation of that function is limited and with a clear trend, which Sybil easily learns and replicates.

The second scenario considers a period spanning the highest peak in the data set. Here, we consider the period from December 13$^{\text{th}}$, 2021 to January 13$^{\text{th}}$, 2022 as training data, and we forecast the daily infections for the period January-February 2022 (starting from January 14$^{\text{th}}$). In Italy, there were three active variants within this time window: Omicron, Delta, and the *Other* variant. Figure 3.5-(a) shows the number of daily infections for the three active

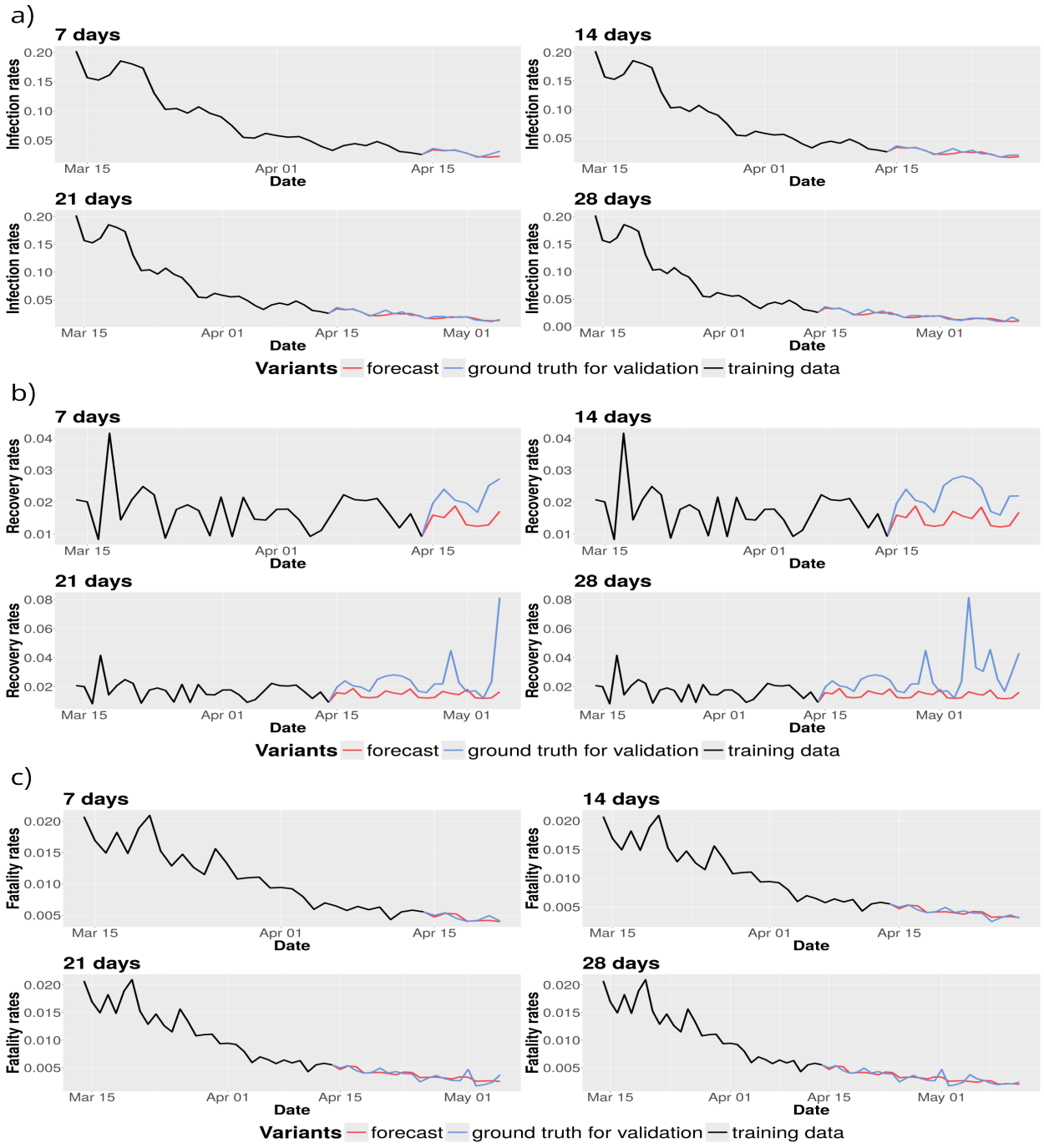


Figure 3.4: The figure shows Sybil applied on *a)* global infection rates $\beta(\tilde{t})$, *b)* global recovery rates $\gamma(\tilde{t})$, and *c)* global fatality rates of the first scenario in Italy in which we forecast starting from April 14$^{\text{th}}$, 2020 (the red line shows the prediction, while black and blue lines represent the training data and the ground-truth values extracted from the surveillance data, respectively).

variants and compares the ground truth (dashed line) with Sybil's forecasts for one to four weeks. Again, forecasts are highly accurate for seven to twenty-one days long predictions and slightly anticipate the peak's descendent phase at four weeks. In Figure 3.5-*(b)*, the two approaches are contrasted against the ground truth. Like the first scenario, the Prophet's

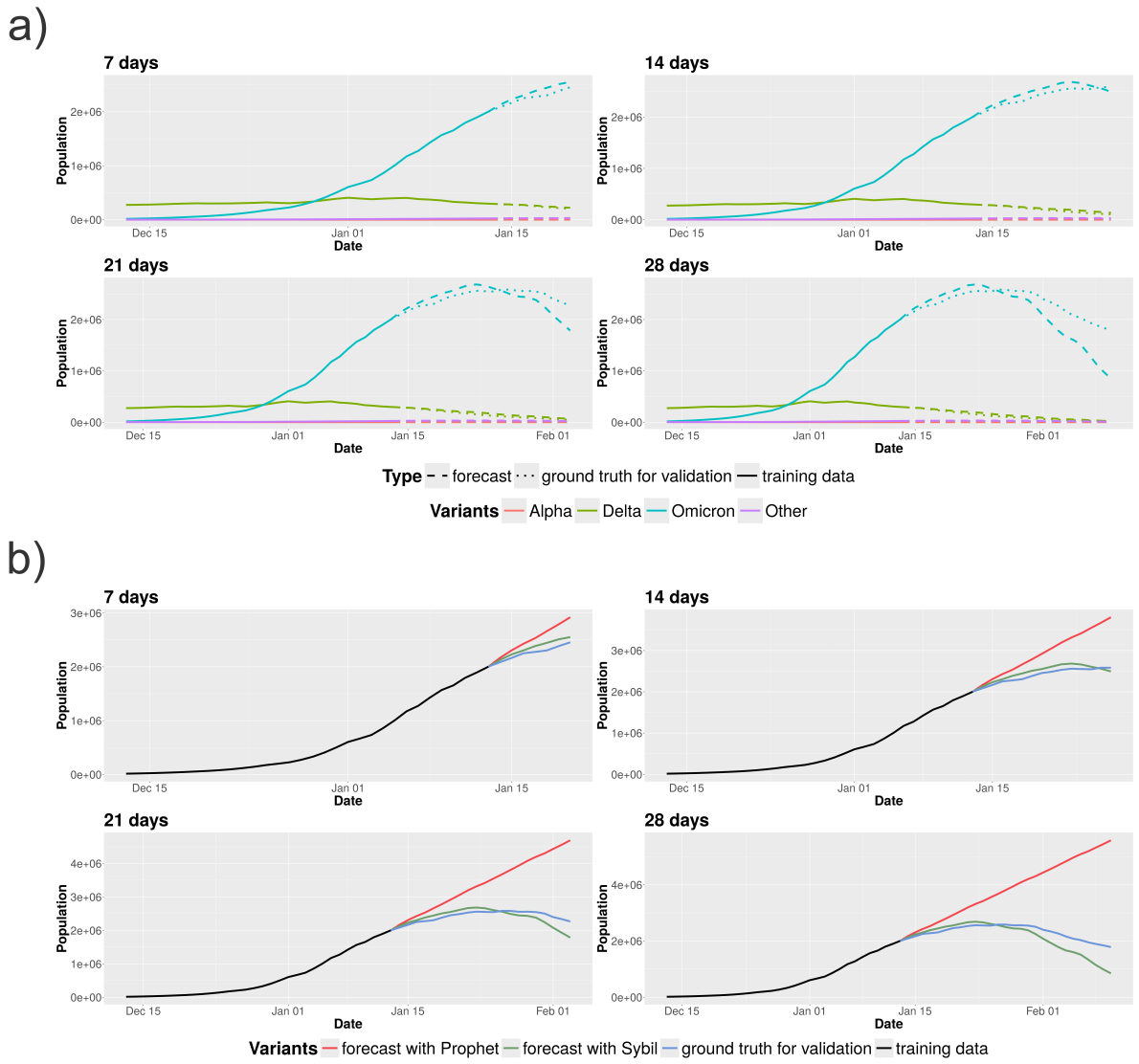


Figure 3.5: The figures show *(a)* the evolution of infections using Sybil in the second scenario in Italy in which we forecast starting from January 13$^{\text{th}}$, 2022 (the dashed line shows the prediction, while solid and dotted lines represent the training data and the ground-truth values extracted from the surveillance data, respectively) and *(b)* the comparison between Sybil (green line) and Prophet (red line) on the number of infections for the Omicron variant using the same period as training data (black line) comparing and contrasting the predictions against the surveillance data for the period spanning the forecasting window (blue line). The RMSEs between the ground-truth and the Sybil forecast are 96883 (1 week), 99589 (2 weeks), 190196 (3 weeks), and 380341 (4 weeks). The RMSEs between the ground truth and the Prophet forecast are 300266 (1 week), 676381 (2 weeks), 1226726 (3 weeks), and 1946285 (4 weeks).

predictions do not capture the peak. Not only do Prophet's predictions get far from the real data—they grow while the infection fades—but, in this scenario, they do not even provide a valid qualitative prediction.

The third point picked to test Sybil's predictions falls just before the explosion of the largest outbreak, caused by the Omicron variant. In particular, Figure 3.6 shows how the one-week forecast changes moving the training window from December 18$^{\text{th}}$, 2021 to

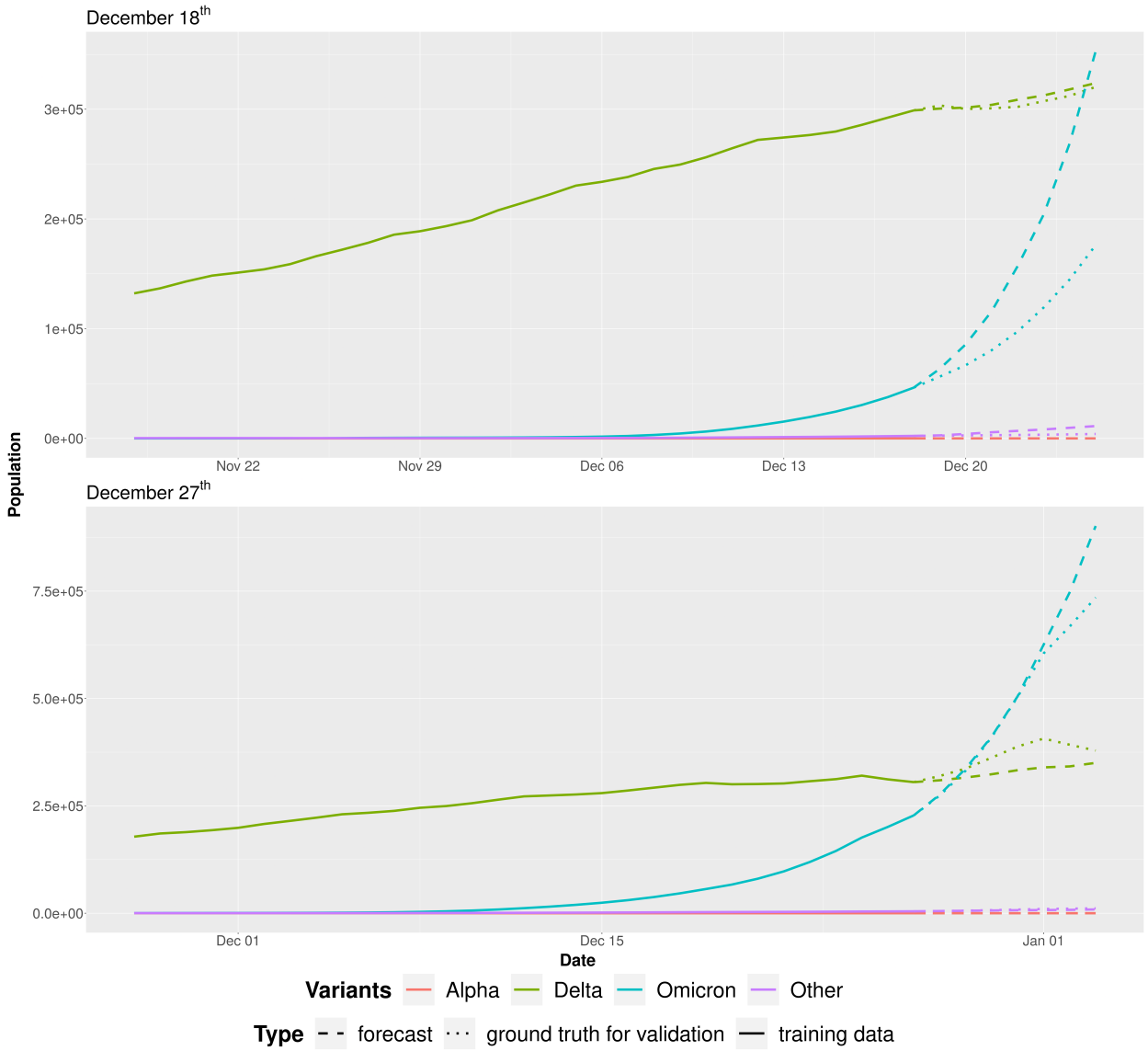


Figure 3.6: Evolution of infections using Sybil on the third scenario in Italy. In the first plot, we forecast starting from December 18$^{\text{th}}$, 2021, while in the second one, we forecast starting from December 27$^{\text{th}}$, 2021. Both the plots refer to a forecast one week into the future. The dashed line shows the prediction, while solid and dotted lines represent the training data and the ground-truth values extracted from the surveillance data, respectively. In the December 18$^{\text{th}}$, 2021 forecast, the RMSEs between the ground-truth and the Sybil forecast after one week is 89604, while between the ground-truth and the Prophet forecast is 79959. In the December 27$^{\text{th}}$, 2021 forecast, the RMSEs between the ground-truth and the Sybil forecast after one week is 68624, while between the ground-truth and the Prophet forecast is 239562.

December 27$^{\text{th}}$, 2021. In this scenario, the ascending trajectory is very steep, and Sybil finds it harder to calibrate with respect to the previous cases. Here, the Omicron variant is a new emerging variant and Sybil initially foresees a more aggressive exponential growth. Looking at Figure 3.6 we can state that Sybil nailed the qualitative prediction, but needed more data to calibrate it. We can also see that Sybil captures well the prediction on the other active variant—the Delta variant. Here we are showing the best results obtained in this scenario.

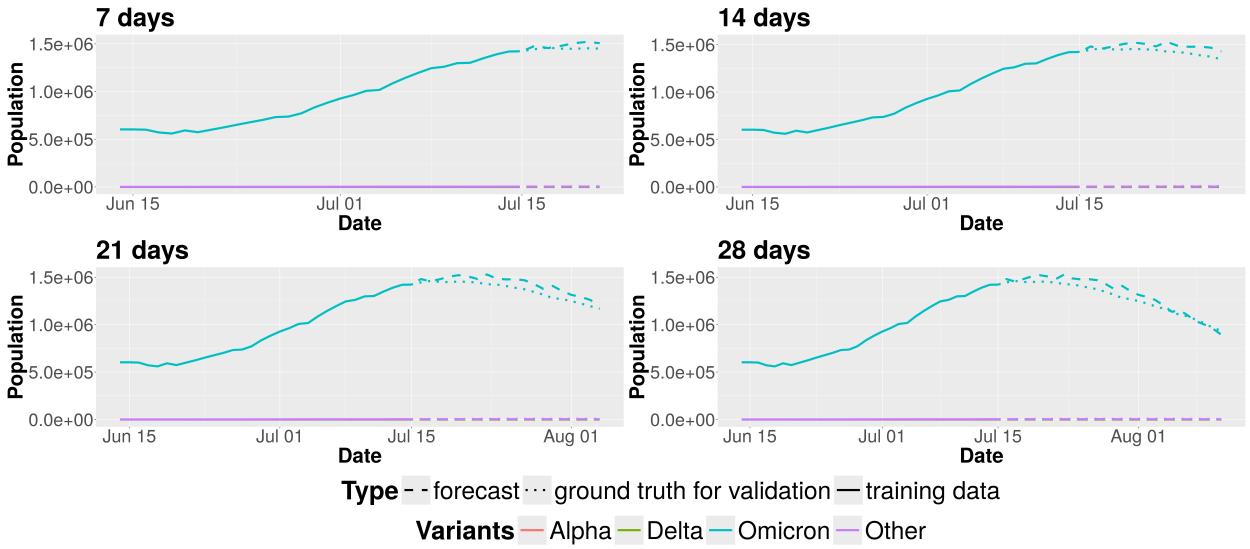


Figure 3.7: Evolution of the $I_{Omicron}(\tilde{t})$ compartment using Sybil on the fourth scenario in Italy in which we forecast starting from July 14$^{\text{th}}$, 2022 (the dashed line shows the prediction, while solid and dotted lines represent the training data and the ground-truth values extracted from the surveillance data, respectively).

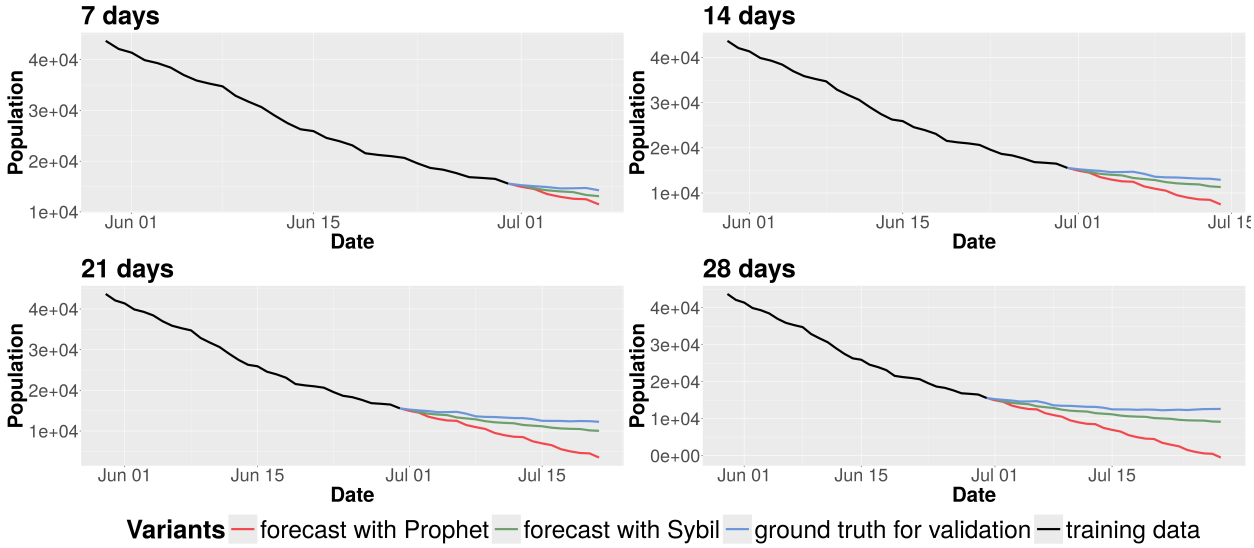


Figure 3.8: The figure refers to the fifth scenario in Italy in which we forecast starting from June 30$^{\text{th}}$, 2020 and shows the evolution of the $I_{Other}(\tilde{t})$ compartment using Sybil (green line) and Prophet (red line) using the same period as training data (black line) comparing and contrasting the predictions against the surveillance data for the period spanning the forecasting window (blue line).

In section 3.2.9, we provide more details on this scenario and show another case where an explosive outbreak occurs, discussing the convenience of establishing a continuous monitoring system.

The fourth scenario falls in the middle of the Summer of 2022. In this period, Italy observed a recrudescence of the pandemic. Omicron is the primary strain responsible for this peak. In this latter case, we consider the period from June 14$^{\text{th}}$, 2022 to July 14$^{\text{th}}$, 2022

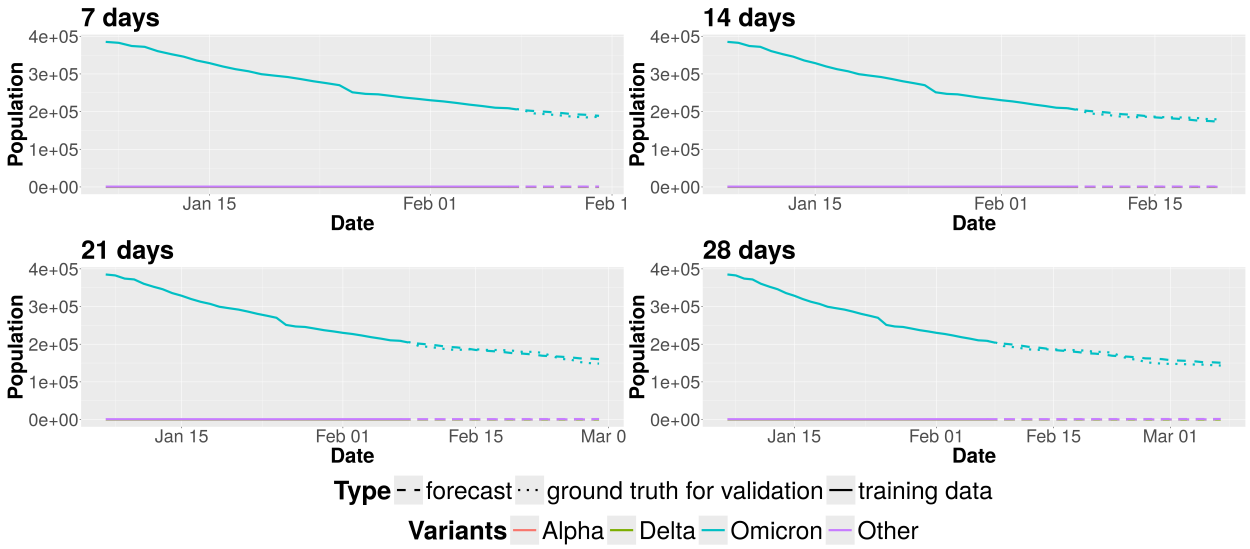


Figure 3.9: Evolution of the $I_{Omicron}(\hat{t})$ compartment using Sybil on the sixth scenario in Italy in which we forecast starting from February 7$^{\text{th}}$, 2023 (the dashed line shows the prediction, while solid and dotted lines represent the training data and the ground-truth values extracted from the surveillance data, respectively).

as training data, and we forecast on July-August 2022 (starting from July 15$^{\text{th}}$). Again, there is a slope change in the Omicron trajectory—although not as intense as in the second scenario—and Sybil captures it very well—see Figure 3.7. In this particular scenario, Sybil performs surprisingly well, and the forecast accuracy does not degrade when moving forward in the future.

In the fifth scenario, we used the period from May 30$^{\text{th}}$, 2020 to June 30$^{\text{th}}$, 2020 as training data, and we forecasted on July 2020 (starting from July 1$^{\text{st}}$). In this scenario, only one active variant (the *Other* variant) exists. In this period, we do not have a drastic change in the slope of the trajectory, and in Figure 3.8 we can see that Sybil is able to capture this less sudden change of slope, while Prophet alone is not able to do so.

Finally, in the sixth and last scenario, we used the period from January 7$^{\text{th}}$, 2023 to February 7$^{\text{th}}$, 2023 as training data, and we forecasted on February-March 2023 (starting from February 15$^{\text{th}}$). Here there are two active variants: Omicron and the *Other* variant. Here the trajectory is decreasing, and we capture almost perfectly the future trend—Figure 3.9.

3.2.2 Errors

Tables 3.1 and 3.2 show the RMSE with the standard deviation between the ground truth used for validation and the obtained forecast with Sybil and the plain use of different state-of-the-art approaches—i.e., Prophet [257], ARIMA / SARIMA [61], Neural Prophet [263], LSTM [131] and GRU [79] neural networks, and EpiNow2 [235]—for all the scenarios considered in Italy. In particular, for all the considered state-of-the-art approaches we used the same period as

Table 3.1: RMSE with standard deviation between the ground truth used for validation and the obtained forecast with Sybil and the plain use of different state-of-the-art approaches—i.e., Prophet [257], ARIMA / SARIMA [61], Neural Prophet [263], LSTM [131] and GRU [79] neural networks, and EpiNow2 [235]—for the first three scenarios considered in Italy. For each scenario and forecasting window (one-, two- three-, and four-week), values in bold represent minimum values, while underlined values represent values close to minimum values. For the third scenario, we do not report the errors for the two-, three-, and four-week forecast because they are too high—same for LSTM and GRU in the second scenario for the two-, three-, and four-week forecast and four-week forecast, respectively.

Scenario	Approach	1 week *mean (± std)*	2 weeks *mean (± std)*	3 weeks *mean (± std)*	4 weeks *mean (± std)*
1st	Sybil	1381 (± 747)	**4141** (± 2425)	**7149** (± 4120)	**13197** (± 8408)
	Prophet	3893 (± 2361)	10851 (± 6476)	19615 (± 11653)	32593 (± 20125)
	ARIMA	**778** (± 722)	4144 (± 3253)	9736 (± 6885)	19717 (± 13896)
	SARIMA	**778** (± 722)	4144 (± 3253)	9736 (± 6885)	19717 (± 13896)
	Neural Prophet	4035 (± 2364)	11111 (± 6553)	19983 (± 11791)	33063 (± 20317)
	LSTM	4623 (± 2106)	11411 (± 6192)	19943 (± 11262)	32955 (± 19890)
	GRU	2114 (± 1302)	6643 (± 4047)	12339 (± 7448)	21848 (± 13922)
	EpiNow2	5112 (± 3364)	106799 (± 990)	104885 (± 3054)	100485 (± 8815)
2nd	Sybil	96883 (± 34583)	**99589** (± 63486)	**190196** (± 186429)	**380341** (± 326962)
	Prophet	300266 (± 133638)	676381 (± 356182)	1226726 (± 703819)	1946285 (± 1158953)
	ARIMA	196298 (± 106323)	502508 (± 290262)	984971 (± 602757)	1636012 (± 1022713)
	SARIMA	196298 (± 106323)	502508 (± 290262)	984971 (± 602757)	1636012 (± 1022713)
	Neural Prophet	299890 (± 130724)	675966 (± 354417)	1225983 (± 702251)	1945198 (± 1157424)
	LSTM	857350 (± 447682)	-	-	-
	GRU	259534 (± 134814)	768775 (± 470723)	2362125 (± 1714411)	-
	EpiNow2	**64997** (± 63032)	2417232 (± 161043)	2429904 (± 147212)	2322897 (± 245505)
3rd (Dec 18th, 2021)	Sybil	89604 (± 55154)	-	-	-
	Prophet	79959 (± 35214)	-	-	-
	ARIMA	30728 (± 21829)	-	-	-
	SARIMA	30728 (± 21829)	-	-	-
	Neural Prophet	54956 (± 29974)	-	-	-
	LSTM	217615 (± 174845)	-	-	-
	GRU	**7030** (± 3204)	-	-	-
	EpiNow2	58149 (± 41276)	-	-	-
3rd (Dec 27th, 2021)	Sybil	**68624** (± 57142)	-	-	-
	Prophet	239562 (± 117411)	-	-	-
	ARIMA	200101 (± 107520)	-	-	-
	SARIMA	200101 (± 107520)	-	-	-
	Neural Prophet	233592 (± 114532)	-	-	-
	LSTM	316040 (± 270008)	-	-	-
	GRU	124546 (± 61369)	-	-	-
	EpiNow2	213874 (± 114639)	-	-	-

training data (the previous month)—with LSTM and GRU we used 10 values to predict the next one. Moreover, the predictions provided by Neural Prophet, LSTM and GRU are averaged over 10 different runs.

These tables confirm that Sybil almost always performs better than the other approaches for the two-, three-, and four-week forecast, while for the one-week forecast, Sybil provides forecasts that are close to the predictions made by the best approach—except in the 3rd scenario (Dec 18th, 2021). In this last case, Sybil needs more data to calibrate the predictions and after nine days—3rd scenario (Dec 27th, 2021)—we can see that the predictions have improved and Sybil performs better with respect to the other approaches.

When there are small changes or no changes in the trajectory, Sybil performs well together with other approaches. In the 5th scenario we can see that LSTM or GRU provide the best

Table 3.2: RMSE with standard deviation between the ground truth used for validation and the obtained forecast with Sybil and the plain use of different state-of-the-art approaches—i.e., Prophet [257], ARIMA / SARIMA [61], Neural Prophet [263], LSTM [131] and GRU [79] neural networks, and EpiNow2 [235]—for the last three scenarios considered in Italy. For each scenario and forecasting window (one-, two- three-, and four-week), values in bold represent minimum values, while underlined values represent values close to minimum values. For LSTM in the fourth scenario, for the four-week forecast, we do not report the error because it is too high.

Scenario	Approach	1 week *mean (± std)*	2 weeks *mean (± std)*	3 weeks *mean (± std)*	4 weeks *mean (± std)*
4^{th}	Sybil	44008 (± 27278)	**65583** (± 32874)	**70533** (± 32036)	**62499** (± 41414)
	Prophet	125139 (± 60796	298396 (± 162565)	518613 (± 292240)	761070 (± 431838)
	ARIMA	126859 (± 63014)	302769 (± 165795)	525680 (± 296710)	770852 (± 437619)
	SARIMA	126859 (± 63014)	302769 (± 165795)	525680 (± 296710)	770852 (± 437619)
	Neural Prophet	126996 (± 61562)	301660 (± 164008)	523304 (± 294393)	767204 (± 434716)
	LSTM	195751 (± 88381)	515099 (± 294551)	1324873 (± 907346)	-
	GRU	119809 (± 58668)	299894 (± 168192)	604343 (± 371511)	1419235 (± 1023522)
	EpiNow2	**26944** (± 23758)	1424314 (± 33177)	1366564 (± 93969)	1293287 (± 165282)
5^{th}	Sybil	809 (± 414)	1073 (± 454)	1384 (± 575)	1900 (± 895)
	Prophet	1771 (± 828)	3213 (± 1533)	4954 (± 2490)	7101 (± 3758)
	ARIMA	1940 (± 960)	3564 (± 1739)	5495 (± 2786)	7833 (± 4147)
	SARIMA	1940 (± 960)	3564 (± 1739)	5495 (± 2786)	7833 (± 4147)
	Neural Prophet	1858 (± 807)	3369 (± 1564)	5175 (± 2561)	7383 (± 3865)
	LSTM	**191** (± 155)	604 (± 383)	1002 (± 583)	1120 (± 570)
	GRU	295 (± 218)	**290** (± 287)	**531** (± 439)	**567** (± 415)
	EpiNow2	921 (± 496)	14062 (± 786)	13537 (± 999)	13278 (± 975)
6^{th}	Sybil	5580 (± 2866)	**5411** (± 5411)	**6424** (± 6247)	**6639** (± 5967)
	Prophet	**2412** (± 2071)	12141 (± 9513)	14725 (± 9523)	20230 (± 12190)
	ARIMA	7878 (± 5946)	27206 (± 17457)	37705 (± 20764)	51414 (± 27714)
	SARIMA	7878 (± 5946)	27206 (± 17457)	37705 (± 20764)	51414 (± 27714)
	Neural Prophet	2702 (± 2491)	12566 (± 9758)	15351 (± 9835)	21071 (± 12610)
	LSTM	10982 (± 3967)	12046 (± 3368)	23276 (± 12768)	31589 (± 16799)
	GRU	9149 (± 3467)	9157 (± 2671)	19841 (± 11774)	27705 (± 15683)
	EpiNow2	7201 (± 2898)	187451 (± 6533)	178719 (± 15232)	171106 (± 19200)

prediction, but Sybil predictions are also good. In the 6^{th} scenario Sybil performs better, except in the one-week forecast where the predictions are close to the best ones. Finally, for the third scenario, we do not report the errors for the two-, three-, and four-week forecast because they are too high—same for LSTM and GRU in the second scenario for the two-, three-, and four-week forecast and four-week forecast, respectively, and for LSTM in the fourth scenario for the four-week forecast.

3.2.3 Austria

Figure 3.10-*a)* shows the evolution of the $I(\tilde{t})$ compartment from February 2020 to May 2023 in Austria. Figures 3.10-*b)*, 3.10-*c)* and 3.10-*d)* show the evolution of infection, recovery and fatality rates respectively. Figures 3.10-*e)*, 3.10-*f)* and 3.10-*g)* show the daily proportion of each variant, the evolution of the $I_v(\tilde{t})$ compartments and the evolution of infection rates for each variant, respectively.

From the previous scenarios, we can observe that Sybil provides a robust, easily portable approach to different use cases without requiring modifications. For instance, in Figure 3.10-*f)*

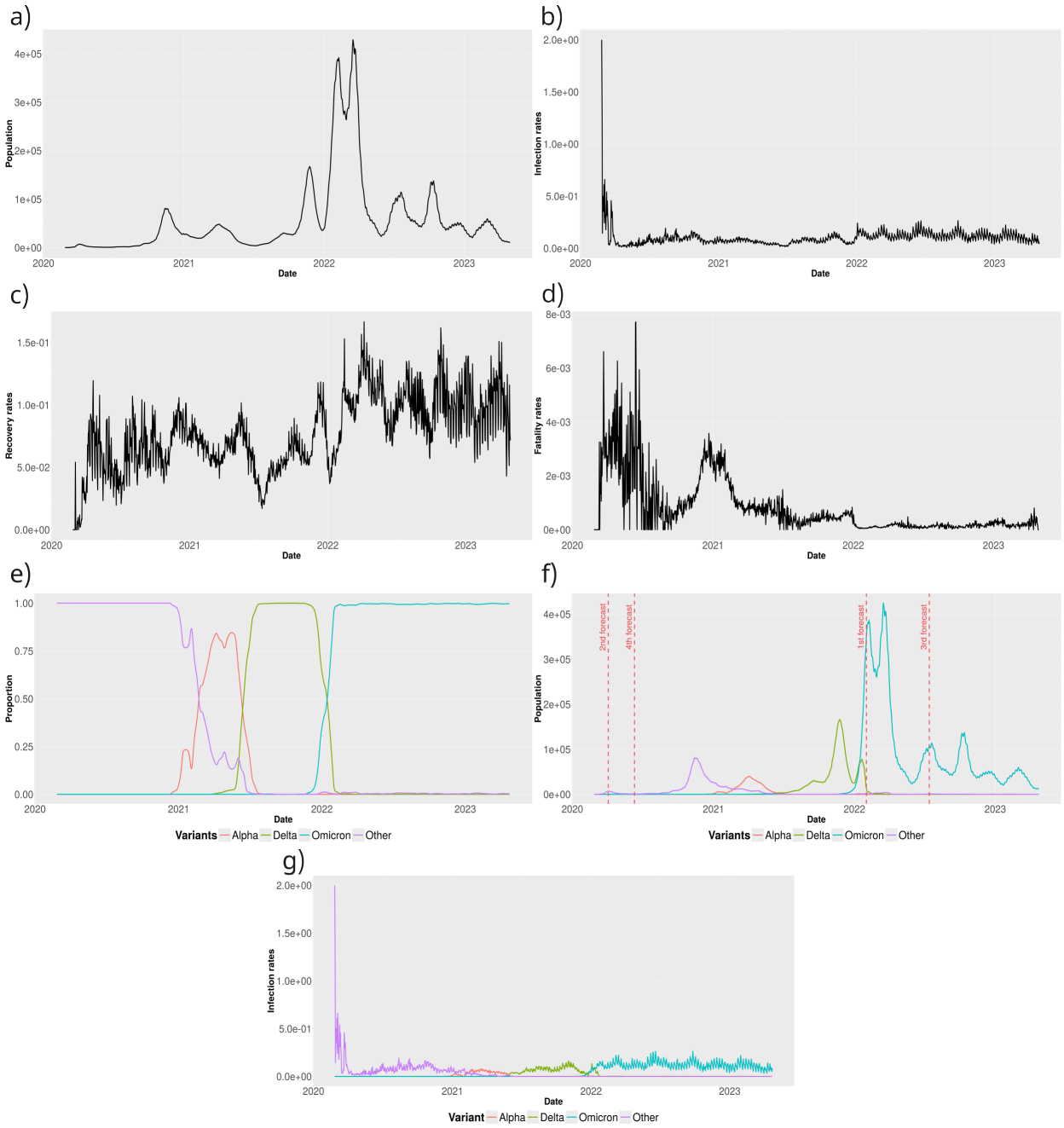


Figure 3.10: The figure shows the evolution of *a)* $I(\tilde{t})$ compartment, *b)* infection rates $\beta(\tilde{t})$, *c)* recovery rates $\gamma(\tilde{t})$, and *d)* fatality rates $\lambda(\tilde{t})$, *e)* the proportion of the aggregated variants, *f)* the evolution of $I_v(\tilde{t})$ compartments with the four considered forecast scenarios—vertical dashed lines mark the selected dates to test Sybil's forecasting—, and *g)* the evolution of infection rates $\beta_v(\tilde{t})$ from February 2020 to May 2023 in Austria.

we apply Sybil to surveillance data from Austria considering four different scenarios. This figure shows the daily infections for the different SARS-CoV-2 strains active in the period from February 2020 to May 2023. As in the case of Italy, lineages are grouped into four main virus strains—Alpha, Delta, Omicron, and the *Other* variant.

Figure 3.11 refers to the first scenario in Austria, where data from January 1st, 2022 to

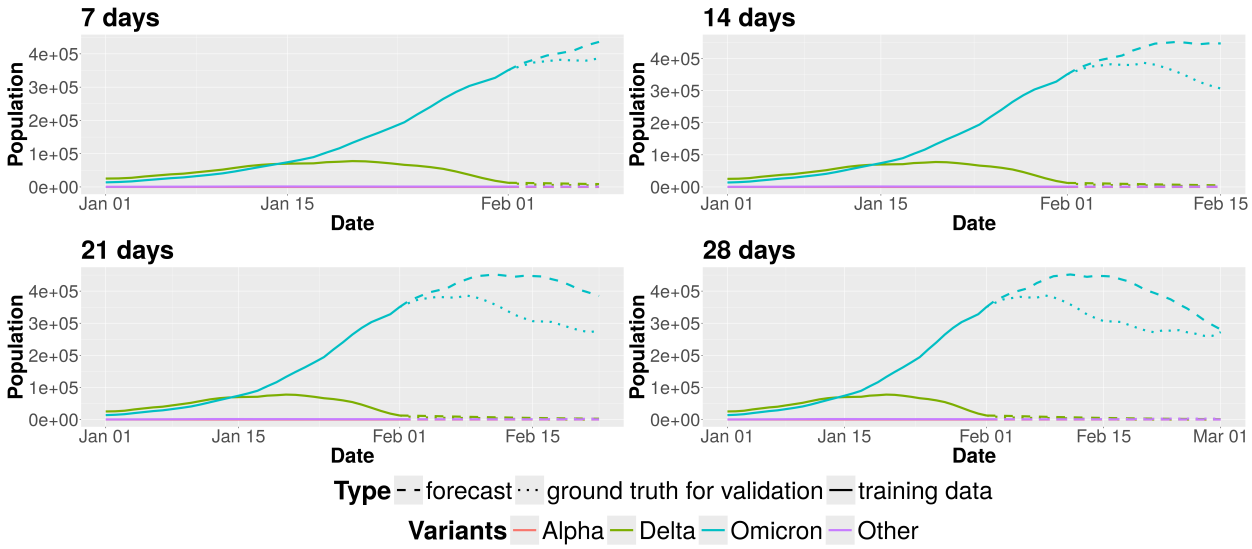


Figure 3.11: Evolution of infections using Sybil on the first scenario considered in Austria in which we forecast starting from February 1st, 2022 (the dashed line shows the prediction, while solid and dotted lines represent the training data and the ground-truth values extracted from the surveillance data, respectively). The RMSEs between the ground-truth and the Sybil forecast are 30171 (1 week), 80197 (2 weeks), 100774 (3 weeks) and 93937 (4 weeks). The RMSEs between the ground-truth and the Prophet forecast are 63197 (1 week), 170881 (2 weeks), 278884 (3 weeks), and 373726 (4 weeks).

February 1st, 2022 are employed to train the models, and we compare forecasts obtained for February 2022 (starting from February 2nd). During this period, there were two active variants: Omicron and the *Other* variant, as in the second scenario considered for Italy. The initial point chosen for the forecasts is again placed in the surroundings of the highest peak so that correctly predicting the future evolution of the infection is much more challenging. Nonetheless, Sybil shows high accuracy in predicting the future trajectory of the infection; despite a slightly different evolution, Sybil correctly predicts the diffusion slowdown, the successive decreasing phase. Eventually, prediction and surveillance data meet at the end of the four-week-long forecast.

Figure 3.12-*(a)* and 3.12-*(b)* refer to the second scenario in Austria, in which we are positioned on the first pandemic peak caused by the *Other* variant. Here, we consider the period from March 3rd, 2020 to April 3rd, 2020 as training data, and we forecast the daily infections for the period April-May 2020 (starting from April 4th). Also in this scenario, Sybil's predictions demonstrate remarkable accuracy in foreseeing the infection's future trajectory, anticipating the deceleration in diffusion and the subsequent decline. Figure 3.12-*(b)* shows the comparison between Sybil's and Prophet's predictions.

Figure 3.13 refers to the third scenario in Austria, in which we are positioned slightly before the Omicron pandemic peak occurred in the Summer of 2022. Here, we consider the period from June 14th, 2022 to July 14rd, 2022 as training data, and we forecast the daily infections for the period July-August 2022 (starting from July 15th). From these figures, we

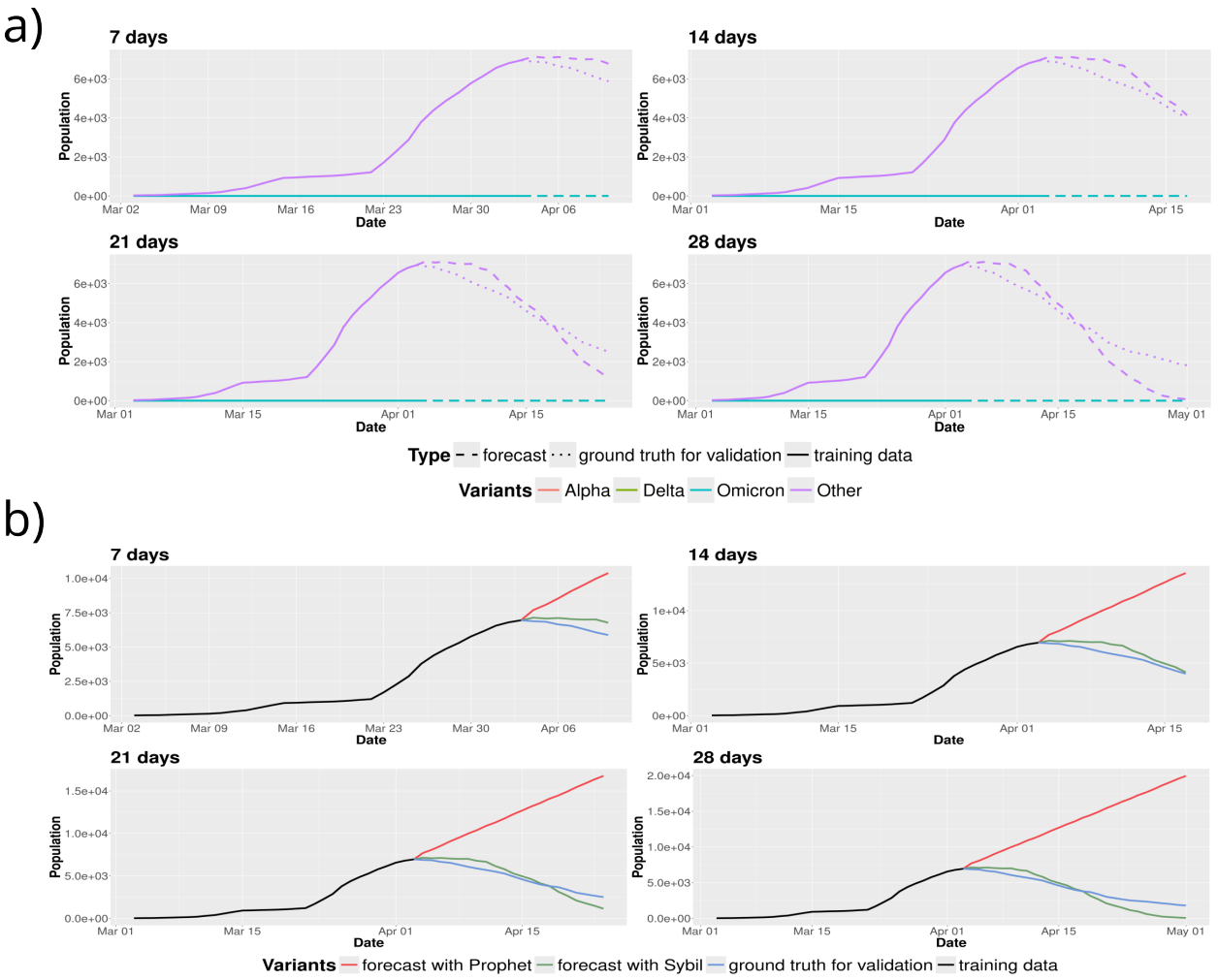


Figure 3.12: The figures show *(a)* the evolution of the $I_{Other}(\tilde{t})$ compartment using Sybil on the second scenario in Austria in which we forecast starting from April 3rd, 2020 (the dashed line shows the prediction, while solid and dotted lines represent the training data and the ground-truth values extracted from the surveillance data, respectively) and *(b)* the comparison between Sybil (green line) and Prophet (red line) on the number of infections for the *Other* variant using the same period as training data (black line) comparing and contrasting the predictions against the surveillance data for the period spanning the forecasting window (blue line). The RMSEs between the ground-truth and the Sybil forecast are 581 (1 week), 518 (2 weeks), 733 (3 weeks) and 1101 (4 weeks). The RMSEs between the ground-truth and the Prophet forecast are 2900 (1 week), 5698 (2 weeks), 8549 (3 weeks), and 11120 (4 weeks).

can see how well Sybil is able to capture the weekly seasonality presents in the data, which is one of Sybil's strengths.

Finally, Figure 3.14 shows the fourth and last scenario in Austria, in which we are positioned in a flat period after the first pandemic peak of the *Other* variant. Here, in the short-term forecast, the prediction and ground truth are close and relatively stable, showing a slight decline in cases. As the forecast window extends, there is a divergence between the forecast and actual data. Notably, after three and four weeks, the ground truth trends upward, while the forecast initially trends downward before showing an uptick.

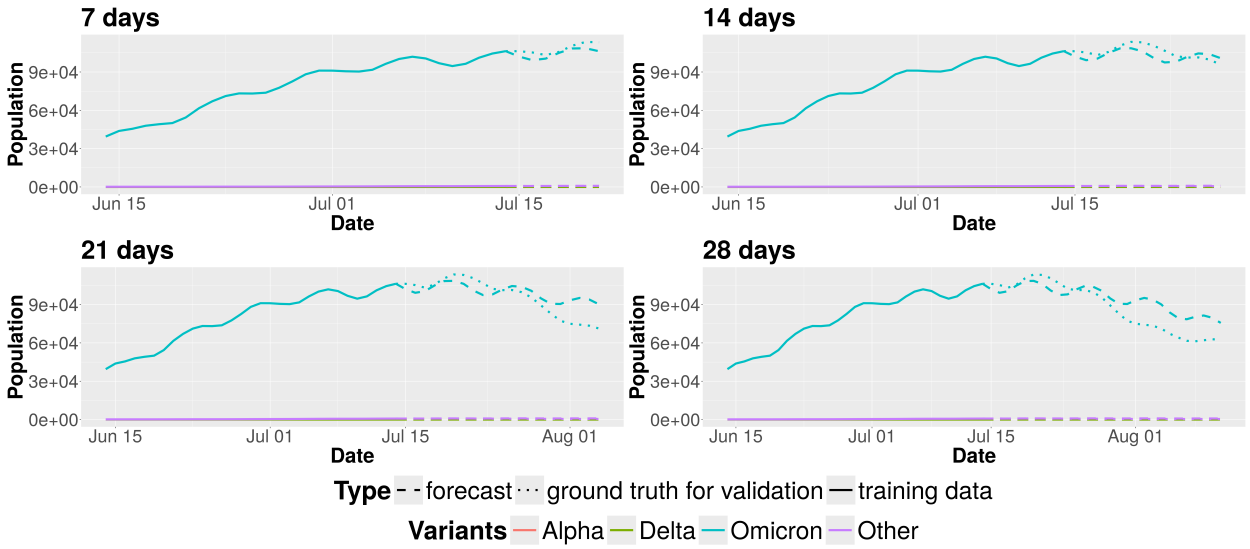


Figure 3.13: Evolution of the $I_{Omicron}(\tilde{t})$ compartment using Sybil on the third scenario in Austria in which we forecast starting from July 14$^{\text{th}}$, 2022 (the dashed line shows the prediction, while solid and dotted lines represent the training data and the ground-truth values extracted from the surveillance data, respectively).

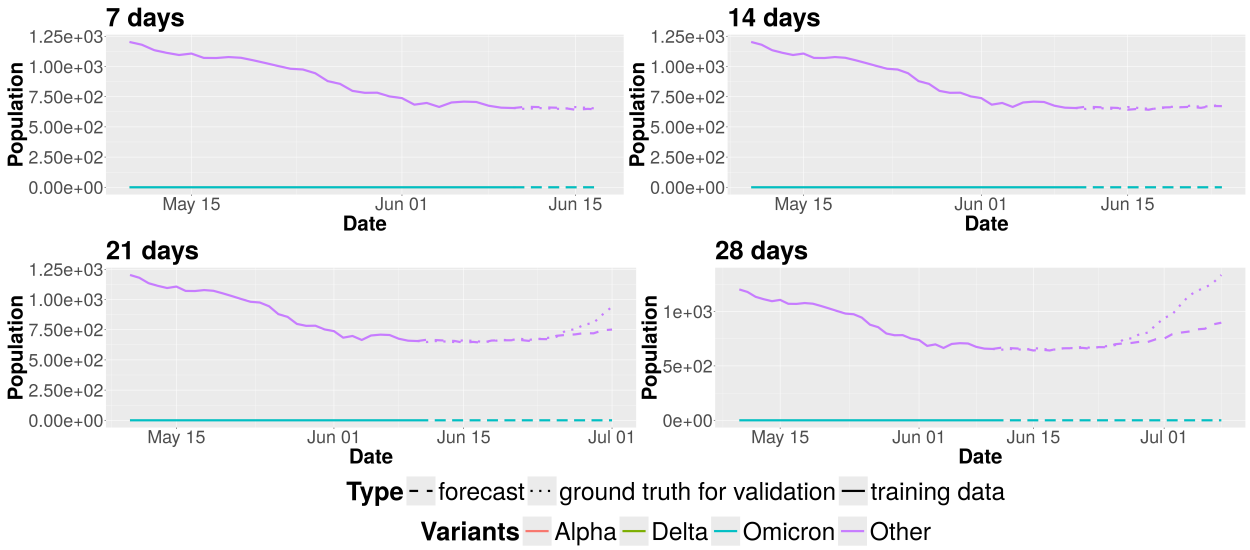


Figure 3.14: Evolution of the $I_{Other}(\tilde{t})$ compartment using Sybil on the fourth scenario in Austria in which we forecast starting from June 10$^{\text{th}}$, 2020 (the dashed line shows the prediction, while solid and dotted lines represent the training data and the ground-truth values extracted from the surveillance data, respectively).

3.2.4 Errors

Table 3.3 shows the RMSE with standard deviation between the ground truth used for validation and the obtained forecast with Sybil and the plain use of different state-of-the-art approaches—i.e., Prophet [257], ARIMA / SARIMA [61], Neural Prophet [263], LSTM [131] and GRU [79] neural networks, and EpiNow2 [235]—for all the scenarios considered in Austria. Again, for all the considered state-of-the-art approaches we used the same period as training

data (the previous month)—with LSTM and GRU we used 10 values to predict the next one. Moreover, the predictions provided by Neural Prophet, LSTM and GRU are averaged over 10 different runs. For the first scenario, we do not report the errors for the three- and four-week forecast for LSTM and for the four-week forecast for GRU because they are too high.

Table 3.3: RMSE with standard deviation between the ground truth used for validation and the obtained forecast with Sybil and the plain use of different state-of-the-art approaches—i.e., Prophet [257], ARIMA / SARIMA [61], Neural Prophet [263], LSTM [131] and GRU [79] neural networks, and EpiNow2 [235]—for the four scenarios considered in Austria. For each scenario and forecasting window (one-, two- three-, and four-week), values in bold represent minimum values, while underlined values represent values close to minimum values. For LSTM and GRU in the first scenario for three-, and four-week forecasts and four-week fforecasts respectively, we do not report the errors because they are too high.

Scenario	Approach	1 week *mean (± std)*	2 weeks *mean (± std)*	3 weeks *mean (± std)*	4 weeks *mean (± std)*
1st	Sybil	30171 (± 15880)	**80197** (± 45647)	**100774** (± 48972)	**93937** (± 45837)
	Prophet	63197 (± 31025)	170881 (± 98196)	278844 (± 155191)	373726 (± 199555)
	ARIMA	75981 (± 39985)	197430 (± 113176)	319520 (± 177087)	428617 (± 228470)
	SARIMA	75981 (± 39985)	197430 (± 113176)	319520 (± 177087)	428617 (± 228470)
	Neural Prophet	62845 (± 30308)	170465 (± 97719)	278293 (± 154718)	373008 (± 199027)
	LSTM	130610 (± 69354)	568862 (± 394725)	-	-
	GRU	57175 (± 32272)	188224 (± 118717)	481200 (± 334214)	-
	EpiNow2	**15423** (± 4989)	361539 (± 25603)	338909 (± 40724)	323021 (± 45713)
2nd	Sybil	**581** (± 244)	**518** (± 267)	**733** (± 733)	**1101** (± 998)
	Prophet	2900 (± 1268)	5698 (± 2757)	8549 (± 4211)	11120 (± 5404)
	ARIMA	1279 (± 654)	2845 (± 1522)	4455 (± 2353)	5785 (± 2929)
	SARIMA	1279 (± 654)	2845 (± 1522)	4455 (± 2353)	5785 (± 2929)
	Neural Prophet	2922 (± 1188)	5667 (± 2672)	8460 (± 4105)	10975 (± 5273)
	LSTM	4950 (± 2260)	18709 (± 12706)	268471 (± 235953)	5911799 (± 5422730)
	GRU	2503 (± 1111)	5289 (± 2702)	10039 (± 5915)	25807 (± 18965)
	EpiNow2	8706 (± 3973)	5752 (± 917)	5039 (± 1441)	4492 (± 1716)
3rd	Sybil	4383 (± 2052)	**5237** (± 4589)	**10463** (± 9874)	**12605** (± 10515)
	Prophet	3830 (± 969)	7230 (± 7104)	21879 (± 17776)	33265 (± 23837)
	ARIMA	8689 (± 2818)	23719 (± 13307)	46740 (± 28039)	66184 (± 38009)
	SARIMA	8689 (± 2818)	23719 (± 13307)	46740 (± 28039)	66184 (± 38009)
	Neural Prophet	4027 (± 1149)	7145 (± 7055)	21636 (± 17674)	32931 (± 23698)
	LSTM	3245 (± 2870)	7702 (± 6778)	22690 (± 17181)	33687 (± 22817)
	GRU	4434 (± 3173)	5931 (± 5847)	19115 (± 15538)	28769 (± 20578)
	EpiNow2	**2696** (± 2695)	105529 (± 5070)	97197 (± 14127)	89894 (± 18865)
4th	Sybil	13 (± 13)	**11** (± 11)	**57** (± 51)	**176** (± 144)
	Prophet	27 (± 16)	63 (± 36)	168 (± 115)	373 (± 267)
	ARIMA	78 (± 37)	159 (± 80)	304 (± 179)	550 (± 352)
	SARIMA	78 (± 37)	159 (± 80)	304 (± 179)	550 (± 352)
	Neural Prophet	19 (± 17)	46 (± 32)	146 (± 108)	346 (± 257)
	LSTM	**7** (± 7)	17 (± 13)	100 (± 82)	281 (± 221)
	GRU	23 (± 12)	42 (± 21)	124 (± 88)	306 (± 226)
	EpiNow2	11 (± 9)	661 (± 11)	712 (± 80)	852 (± 219)

3.2.5 Italy with eleven variants (V = 11)

This section shows some results obtained using a different aggregation of lineages in Italy (V = 11). In particular, we focused the attention on showing how the forecast changes moving Sybil's training window.

Figures 3.15-*a)*, 3.15-*b)* and 3.15-*c)* show the proportion of each variant, the evolution of the $I_v(\tilde{t})$ compartments and the evolution of the infection rates for each variant v from

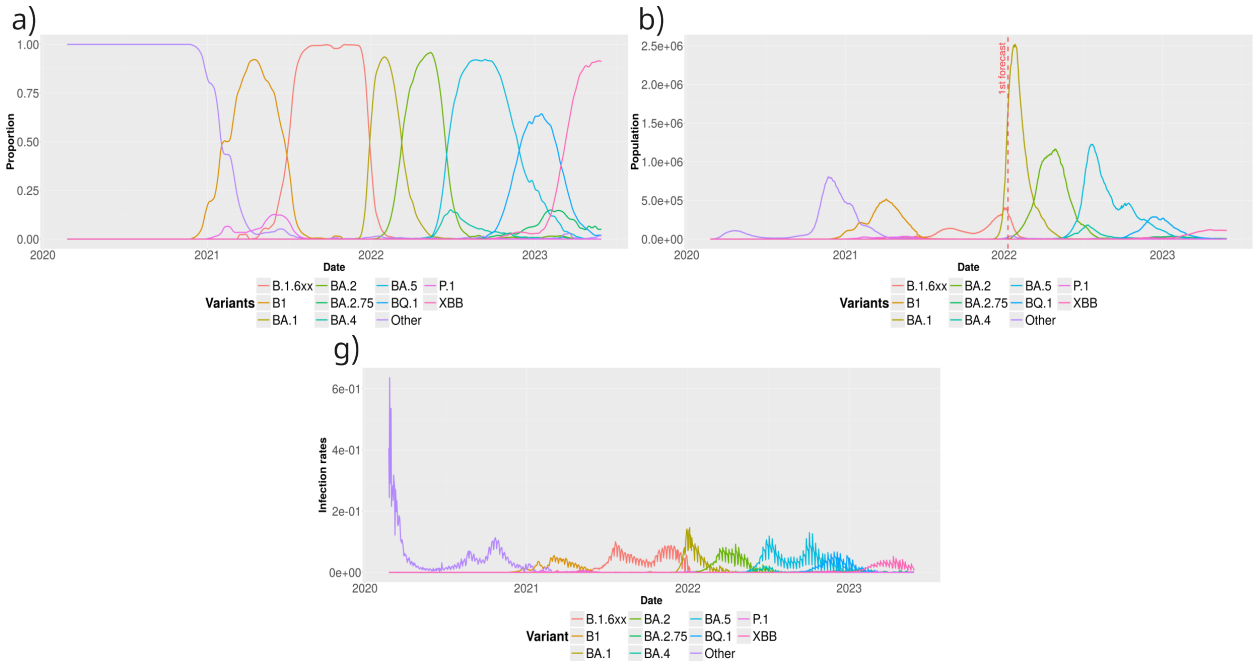


Figure 3.15: The figure shows *a)* the proportion of the aggregated variants, *b)* the evolution of the $I_v(\tilde{t})$ compartments, and *c)* the evolution of the infection rates $\beta_v(\tilde{t})$ from February 2020 to May 2023 in Italy.

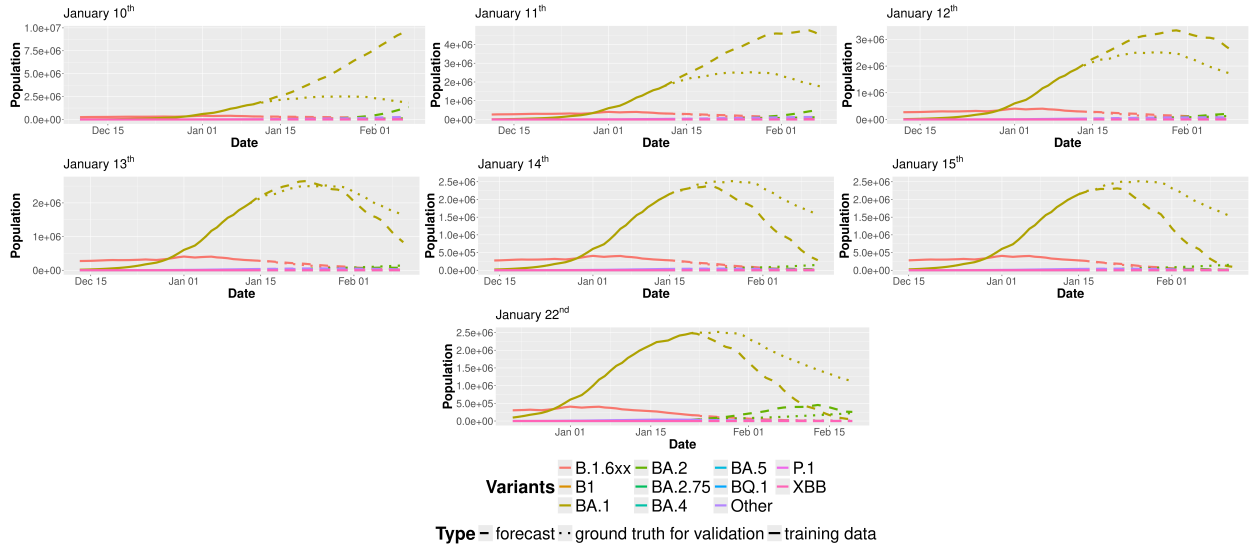


Figure 3.16: Evolution of the $I_v(\tilde{t})$ compartments in Italy using Sybil starting the forecast from January 10^{th}, 11^{th}, 12^{th}, 13^{th}, 14^{th}, 15^{th}, and January 22^{th} 2022 and moving the training window by one day (the dashed line shows the prediction, while solid and dotted lines represent the training data and the ground-truth values extracted from the surveillance data, respectively). All plots refer to a forecast four weeks into the future.

February 2020 to May 2023 in Italy, respectively. With this configuration, we considered again the main peak of infections. Figure 3.16 shows how the forecast changes moving the training window from January 10^{th}, 2022 to January 15^{th}, 2022 by one day and how Sybil can capture the change of slope of the BA.1 variant—Figure 3.17-*a)* shows how Sybil is able to capture the infection rates. We can also see a new emerging variant (BA.2). In particular, the

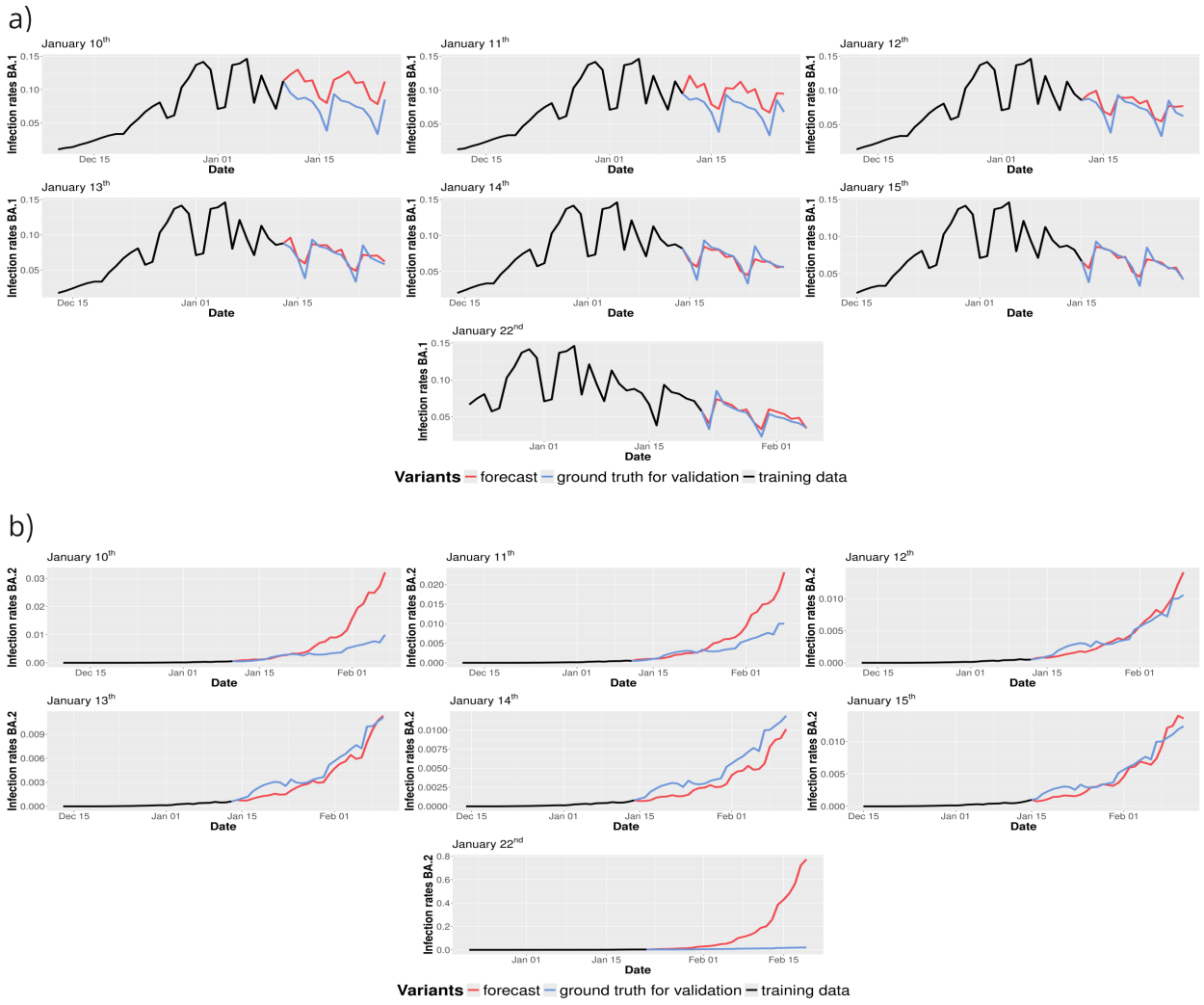


Figure 3.17: The figure shows the evolution of the infection rates *a)* $\beta_{BA.1}(\tilde{t})$ and *b)* $\beta_{BA.2}(\tilde{t})$ in Italy starting the forecast from January 10^{th}, 11^{th}, 12^{th}, 13^{th}, 14^{th} and 15^{th}, 2022 and moving the training window by one day (the red line shows the prediction, while black and blue lines represent the training data and the ground-truth values extracted from the surveillance data, respectively). All plots refer to a forecast four weeks into the future.

first row of Figure 3.16 shows that the model initially sees this new variant, while the second row shows that it would appear that we can't follow the trend. However, Figure 3.17-*b)* shows that we have obtained good predictions on the infection rates of the BA.2 variant. This behavior is because, when there is a new emerging variant—or, more generally, in periods in which we have a trajectory that is initially growing –, we have small values for the rates and Sybil sees exponential growth. Indeed, from a quantitative point of view, we may not be able to follow precisely the trajectory, but from a qualitative point of view, we are able to capture the new emerging variant/outbreak. Figure 3.16 shows that after one week we capture well the trajectory of the same variant (especially after one and four weeks). We point out that the model will only see a new emerging variant if it is already present in the training data.

3.2.6 Other European countries

In this section we show some results obtained considering other two European countries, specifically Belgium and France. The surveillance data used are available in the COVID-19 R library [121, 120] (for the results presented we used a snapshot dated May 2nd, 2024). Information about active virus strains comes from the ECDC, specifically, the one reported in [1, 148] (for the results presented we used a snapshot dated July 25th, 2023).

Compared to data on Italy and Austria, data on Belgium and France were reported worse, especially with regard to recoveries. In particular, the trajectories of infects are slightly different with respect to the official ones related to the actual number of infects—see Figures 3.18-*a)* and 3.19-*a)*—and despite of this, Sybil still provides good predictions—see Figures 3.18-*b)* and 3.19-*b)*. This confirms the fact that our framework is robust, but it is also true that the quality and the availability of data are important (Sybil needs good reported official data on daily new cases and the daily new fatalities).

Belgium

For Belgium, using the COVID-19 data available in the R library [121, 120], we observe that recovery data is only available for the period from March 2020 to November 2020. This limited recovery data, which is affected by inconsistent reporting, contributes to a trajectory for infection counts that resembles a cumulative trend. This effect is clearly visible in Figure 3.18-*a)*, where the population affected by various COVID-19 variants appears to increase continuously, without any significant reductions that would normally reflect recoveries. Figure 3.18-*b)* shows the results obtained with the considered forecast scenario.

France

For France, using the COVID-19 data from the R library [121, 120], recovery data is available from March 2020 to August 2021. In Figure 3.19-*a)*, we observe that the infection trajectory deviates somewhat from the official reports, likely due to inconsistencies in recovery reporting or data adjustments. Figure 3.19-*b)* shows the results obtained with the considered forecast scenario.

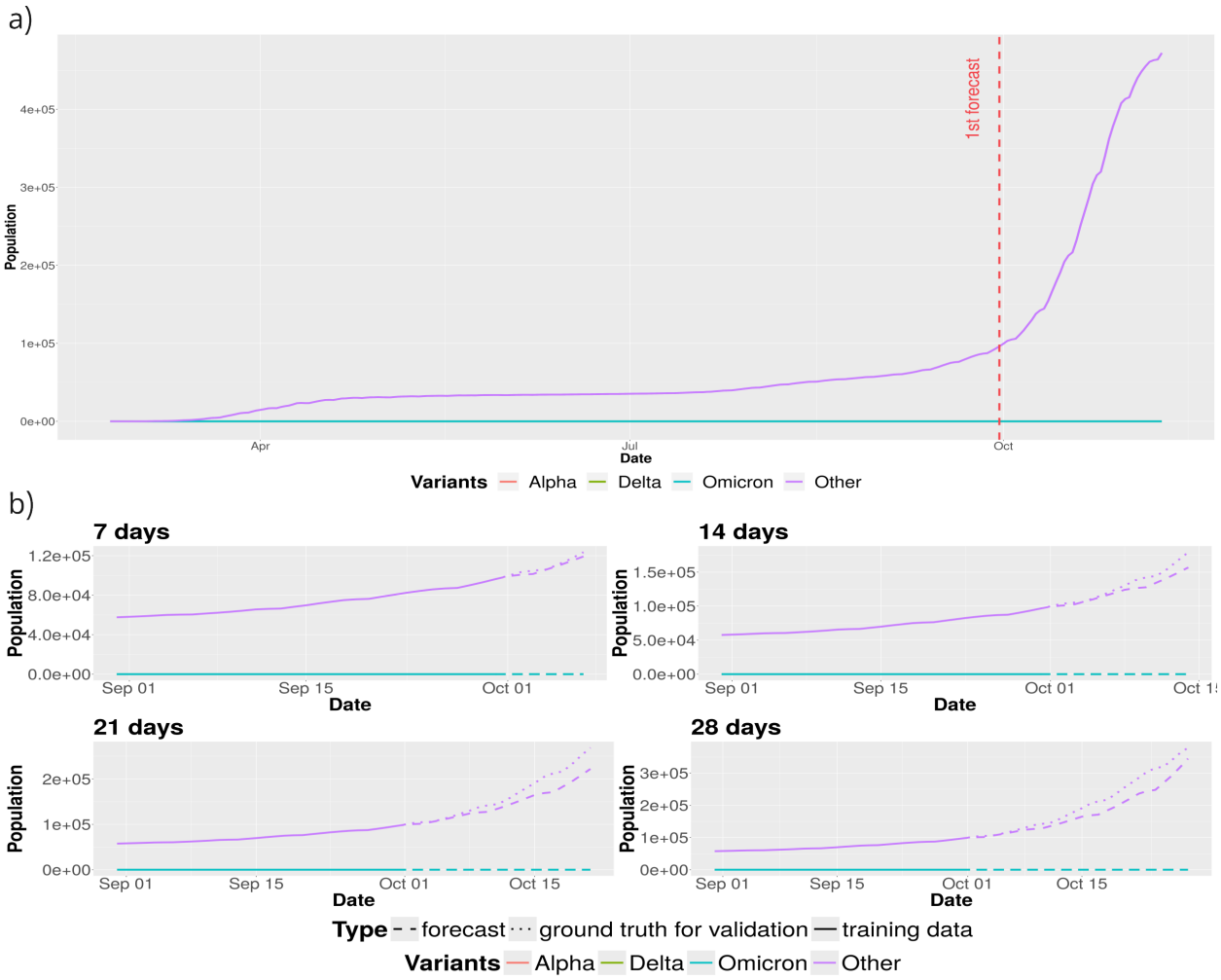


Figure 3.18: The figure shows *a)* the evolution of $I_v(\tilde{t})$ compartments from February 2020 to November 2020 in Belgium with the considered forecast scenario and *b)* the evolution of the $I_{Other}(\tilde{t})$ compartment in Belgium using Sybil starting the forecast from September 30$^{\text{th}}$, 2020 (the dashed line shows the prediction, while solid and dotted lines represent the training data and the ground-truth values extracted from the surveillance data, respectively).

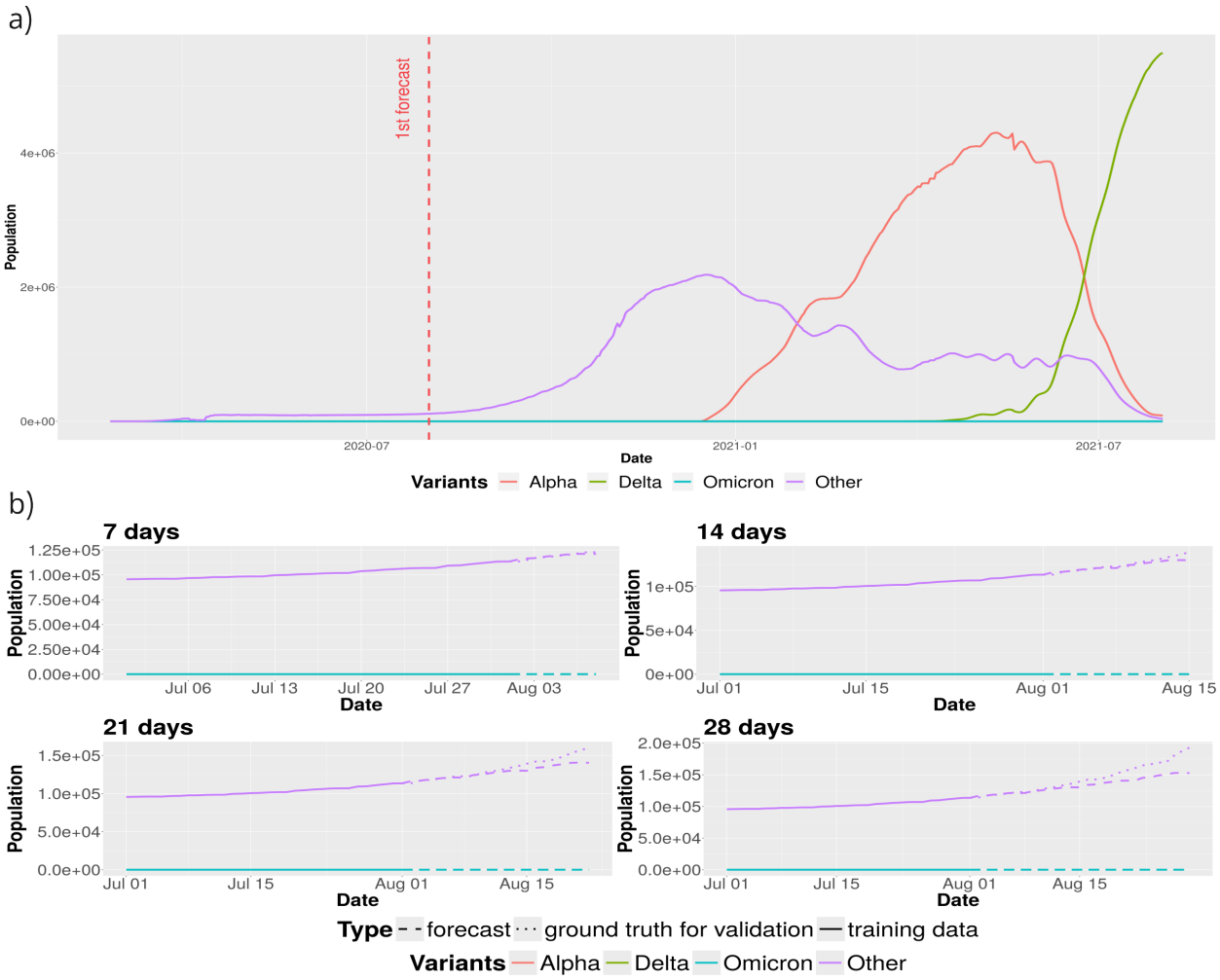


Figure 3.19: The figure shows *a)* the evolution of $I_v(\tilde{t})$ compartments from March 2020 to August 2021 in France with the considered forecast scenario and *b)* the evolution of the $I_{Other}(\tilde{t})$ compartment in France using Sybil starting the forecast from August 1st, 2020 (the dashed line shows the prediction, while solid and dotted lines represent the training data and the ground-truth values extracted from the surveillance data, respectively).

3.2.7 Fixed recovery rate

Due to the limited availability of data on recoveries, we tried to use a fixed recovery rate with mean $\frac{1}{\gamma}$ equal to 14 days [65]—same for each variant. However, another possible extension could be to estimate a different recovery rate that depends on variants. For countries with relatively good reported official data on recoveries—i.e. Italy and Austria—, the results obtained are comparable to the ones obtained with time-dependent recovery rates, while for countries with bad data on recoveries, the use of a fixed recovery rate has improved the results. In the following subsections, we show some results obtained using data on Italy, Austria, Belgium, France, Germany, and Sweden.

Italy

Figure 3.20-*a)* shows the three scenarios that we used to test Sybil using a fixed recovery rate–the three selected scenarios are the same as the first three in Figure 3.2-*f)*. Figures 3.20-*b)*, 3.20-*c)* and 3.21 show the predictions obtained in the three scenarios considered. From these figures we can see that the use of a fixed recovery rate does not alter the course of infection too much, at most it anticipates it or delays it by a few days.

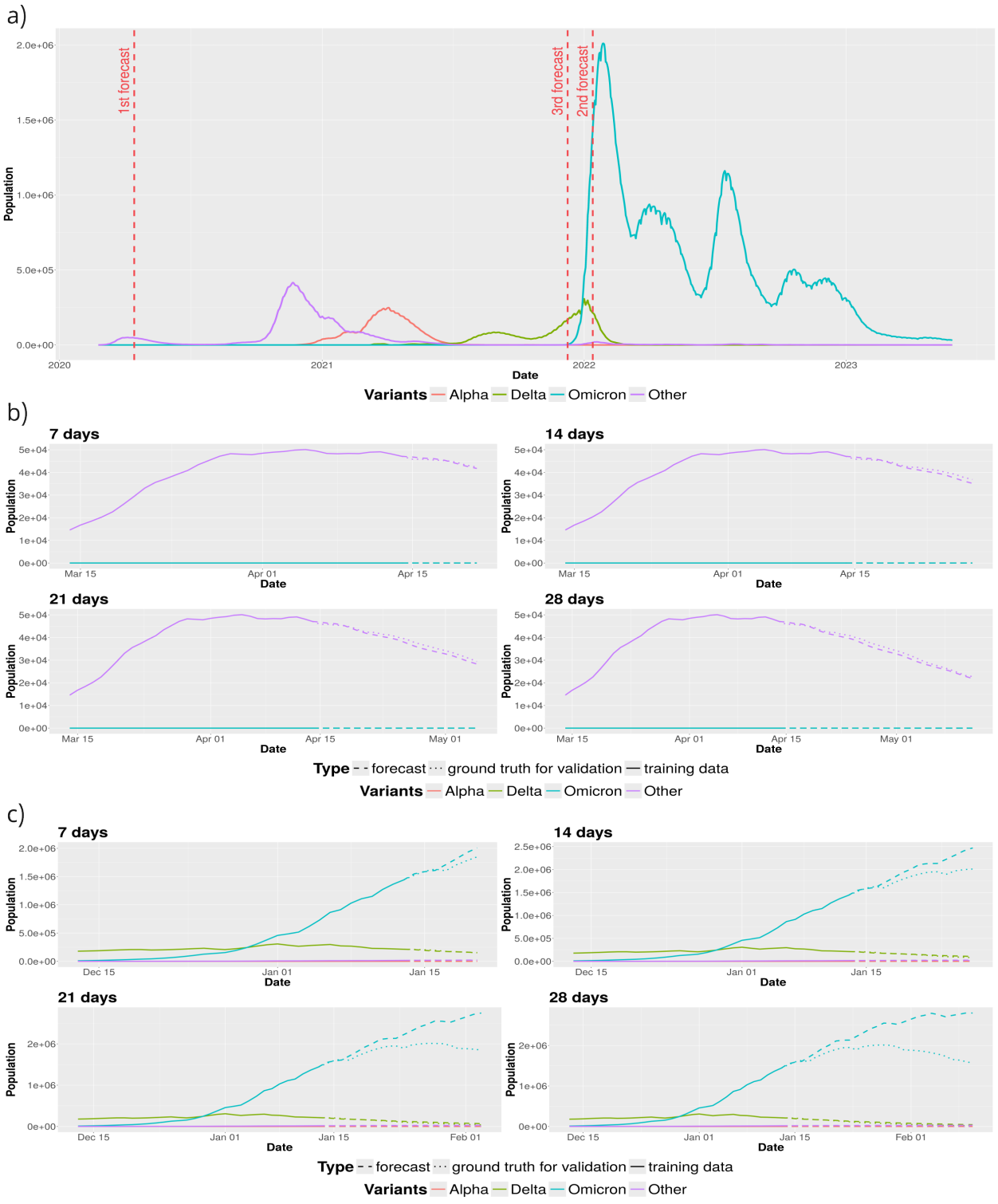


Figure 3.20: The figure shows *a)* the evolution of $I_v(\tilde{t})$ compartments from February 2020 to May 2023 in Italy with the three considered forecast scenarios, *b)* the evolution of the $I_{Other}(\tilde{t})$, and *c)* of $I_v(\tilde{t})$ compartments using Sybil with a fixed recovery rate on the first and second scenarios in Italy starting the forecast from April 14$^{\text{th}}$, 2020 and January 13$^{\text{th}}$, 2022, respectively (the dashed line shows the prediction, while solid and dotted lines represent the training data and the ground-truth values extracted from the surveillance data, respectively).

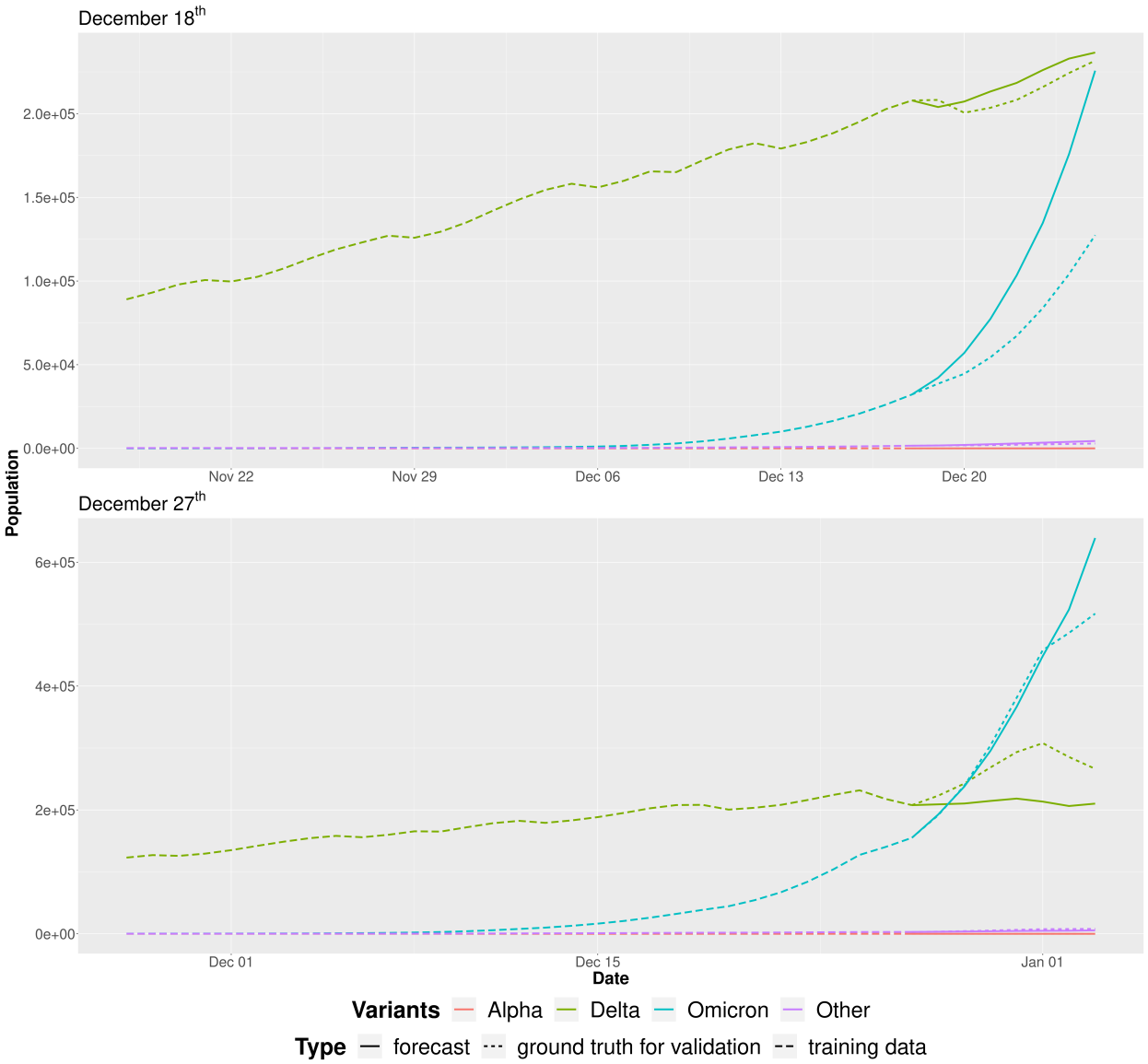


Figure 3.21: Evolution of $I_v(\tilde{t})$ compartments using Sybil on the third scenario in Italy with a fixed recovery rate. In the first plot, we forecast starting from December 18$^{\text{th}}$, 2021, while in the second one, we forecast starting from December 27$^{\text{th}}$, 2021. Both the plots refer to a forecast one week into the future. The dashed line shows the prediction, while solid and dotted lines represent the training data and the ground-truth values extracted from the surveillance data, respectively.

Austria

Figure 3.22-*a)* shows the scenario that we used to test Sybil using a fixed recovery rate—the selected scenario is the same as the first one in Figure 3.10-*f)*. Figure 3.22-*b)* shows the predictions obtained in the scenario considered.

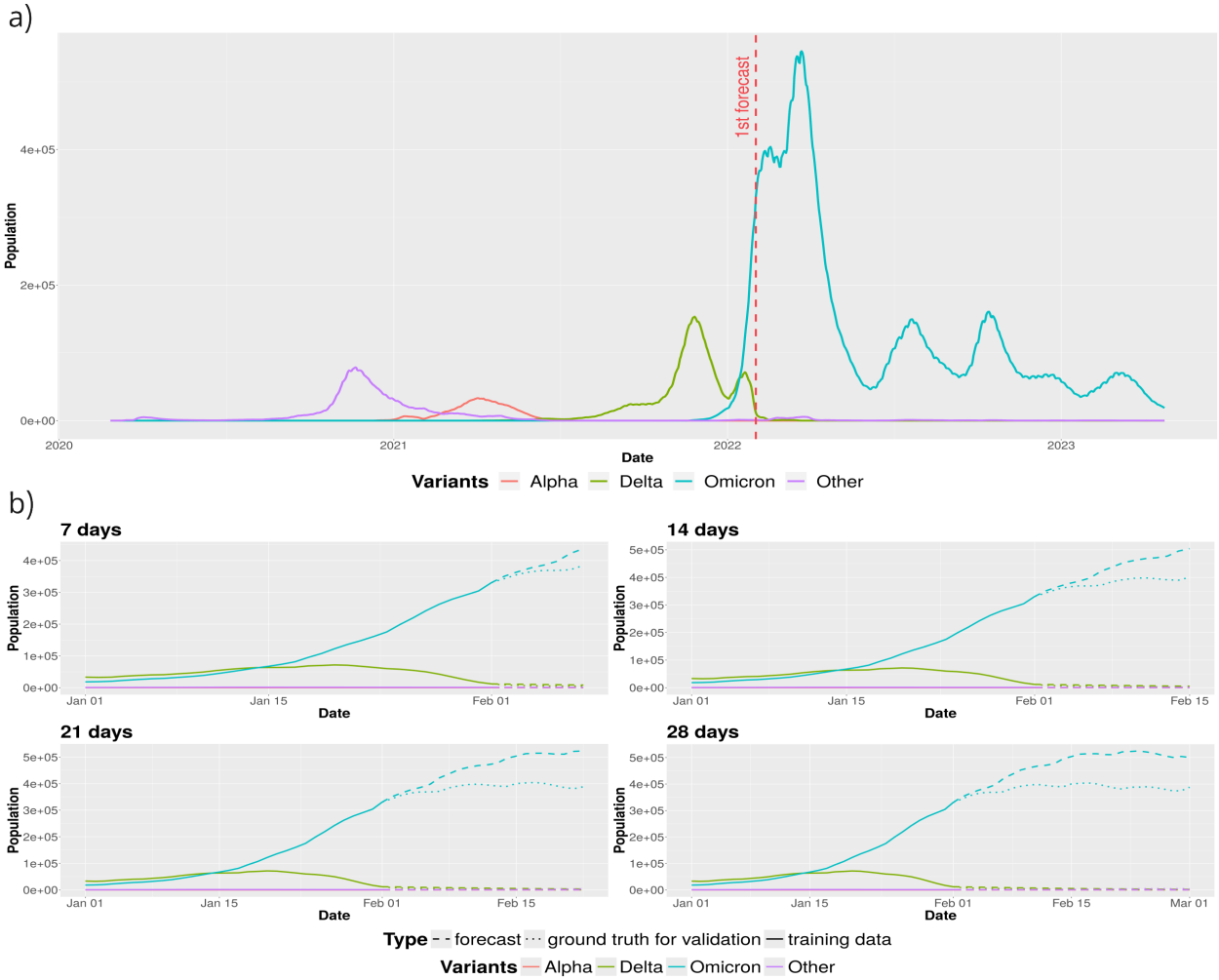


Figure 3.22: The figure shows *a)* the evolution of $I_v(\tilde{t})$ compartments from February 2020 to May 2023 in Austria with the considered forecast scenario and *b)* the evolution of the $I_{Omicron}(\tilde{t})$ compartment in Austria using Sybil starting the forecast from February 2$^{\text{nd}}$, 2022 (the dashed line shows the prediction, while solid and dotted lines represent the training data and the ground-truth values extracted from the surveillance data, respectively).

Belgium

Figure 3.23-*a)* shows the scenario that we used to test Sybil using a fixed recovery rate. Figure 3.23-*b)* shows the predictions obtained in the scenario considered.

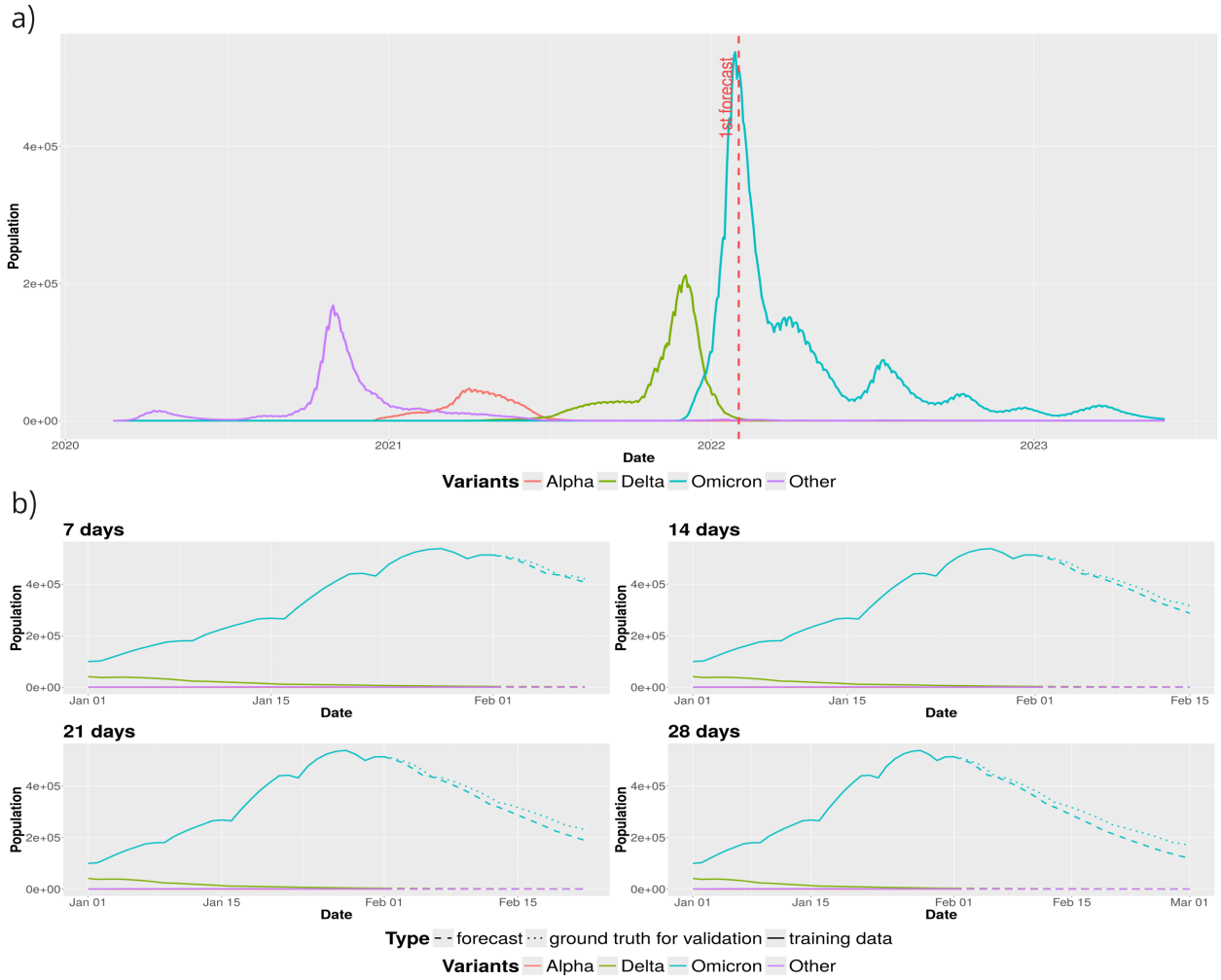


Figure 3.23: The figure shows *a)* the evolution of $I_v(\tilde{t})$ compartments from February 2020 to May 2023 in Belgium with the considered forecast scenario and *b)* the evolution of the $I_{Omicron}(\tilde{t})$ compartment in Belgium using Sybil. We forecast starting from February 2$^{\text{nd}}$, 2022 (the dashed line shows the prediction, while solid and dotted lines represent the training data and the ground-truth values extracted from the surveillance data, respectively) .

France

Figure 3.24-*a)* shows the scenario that we used to test Sybil using a fixed recovery rate. Figure 3.24-*b)* shows the predictions obtained in the scenario considered.

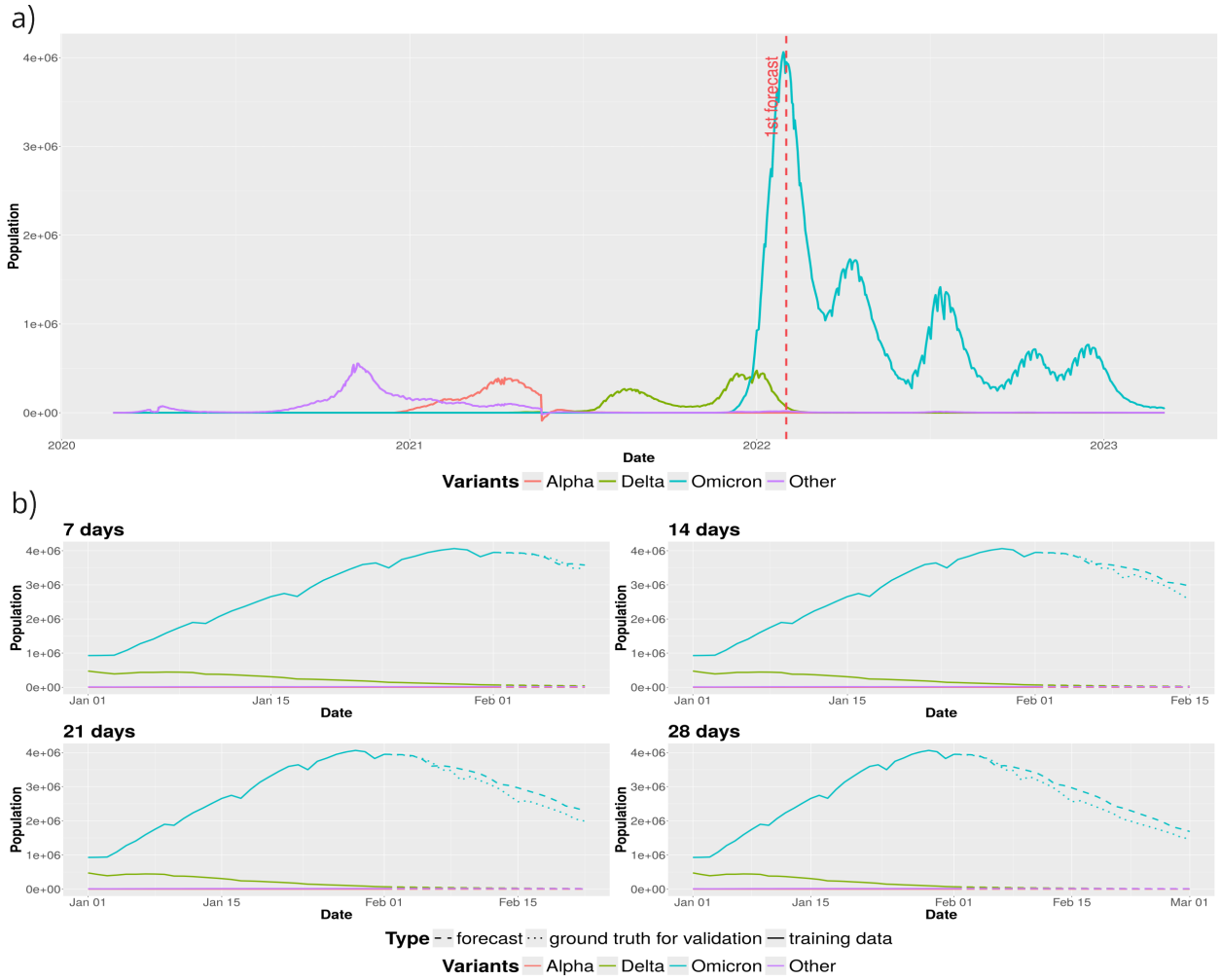


Figure 3.24: The figure shows *a)* the evolution of $I_v(\tilde{t})$ compartments from February 2020 to May 2023 in France with the considered forecast scenario and *b)* the evolution of the $I_{Omicron}(\tilde{t})$ compartment in France using Sybil. We forecast starting from February 2$^{\text{nd}}$, 2022 (the dashed line shows the prediction, while solid and dotted lines represent the training data and the ground-truth values extracted from the surveillance data, respectively).

Germany

Figure 3.25-*a)* shows the scenario that we used to test Sybil using a fixed recovery rate. Figure 3.25-*b)* shows the predictions obtained in the scenario considered.

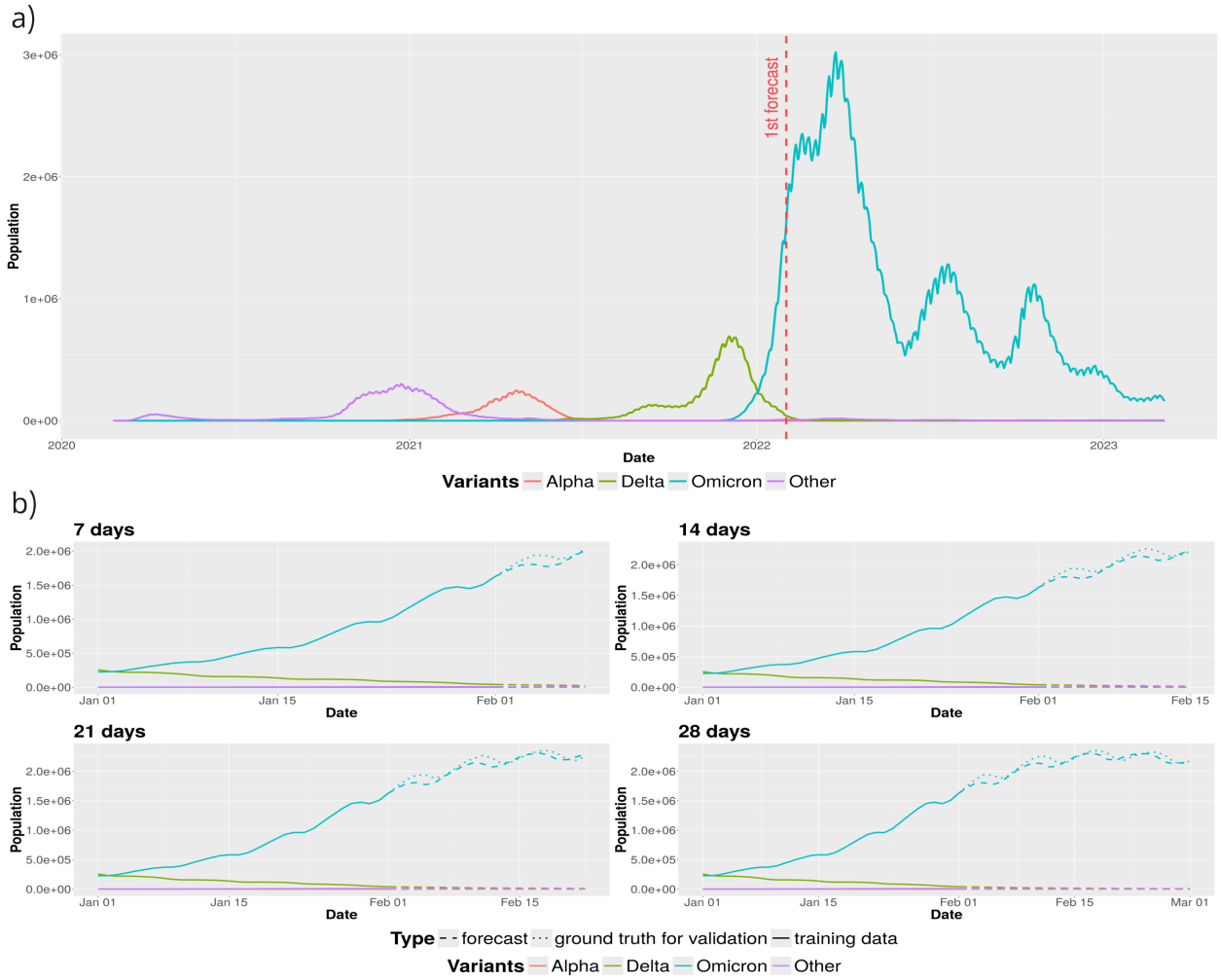


Figure 3.25: The figure shows *a)* the evolution of $I_v(\tilde{t})$ compartments from February 2020 to May 2023 in Germany with the considered forecast scenario, and *b)* the evolution of the $I_{Omicron}(\tilde{t})$ compartment in Germany using Sybil. We forecast starting from February 2$^{\text{nd}}$, 2022 (the dashed line shows the prediction, while solid and dotted lines represent the training data and the ground-truth values extracted from the surveillance data, respectively).

Sweden

Figure 3.26-*a)* shows the scenario that we used to test Sybil using a fixed recovery rate. Figure 3.26-*b)* shows the predictions obtained in the scenario considered.

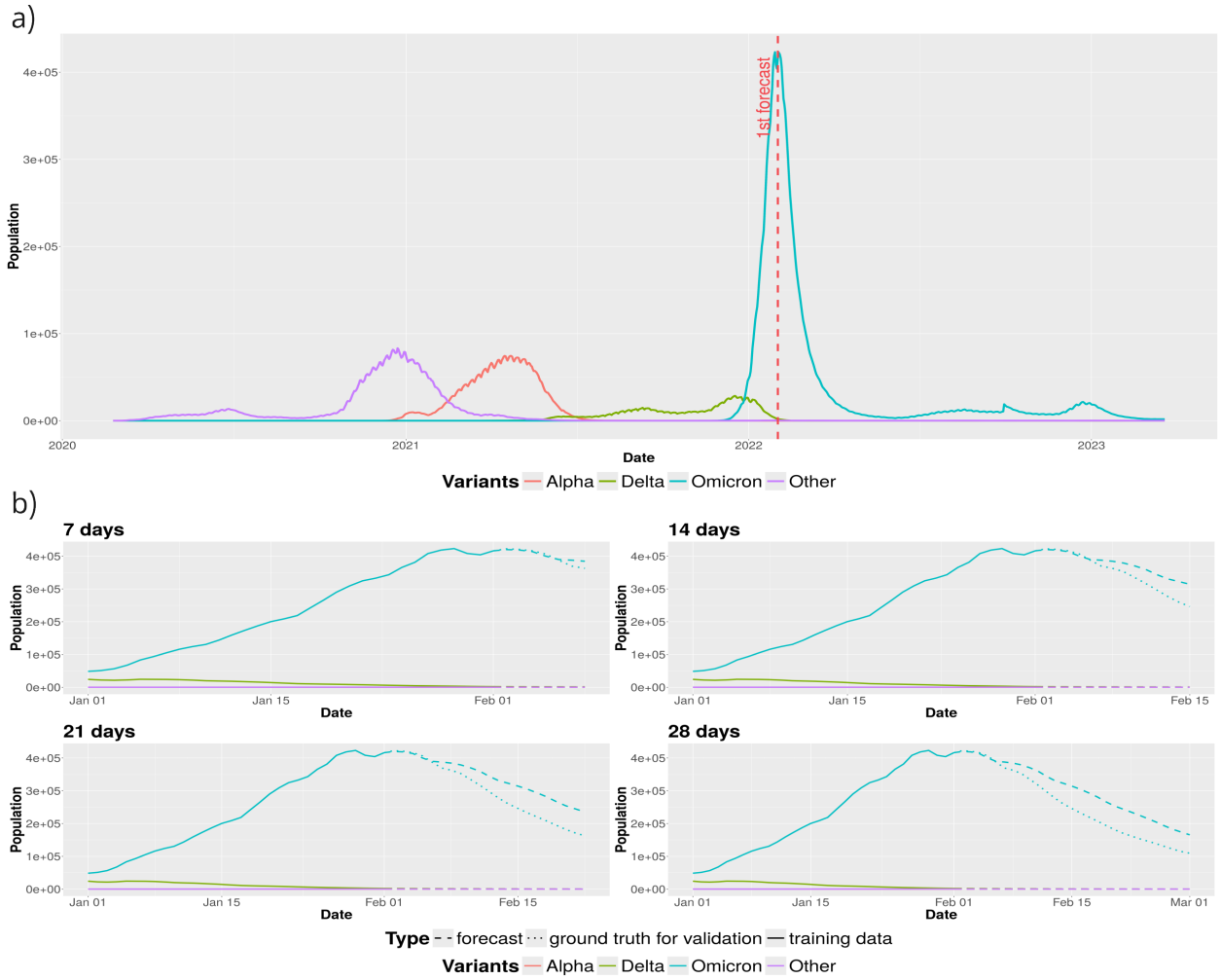


Figure 3.26: The figure shows *a)* the evolution of $I_v(\tilde{t})$ compartments from February 2020 to May 2023 in Sweden with the considered forecast scenario, and *b)* the evolution of the $I_{Omicron}(\tilde{t})$ compartment in Sweden using Sybil. We forecast starting from February 2$^{\text{nd}}$, 2022 (the dashed line shows the prediction, while solid and dotted lines represent the training data and the ground-truth values extracted from the surveillance data, respectively).

3.2.8 USA data

In this section, we present results obtained considering three U.S. states—namely, Colorado, New York and West Virginia—with three objectives: *i)* to compare Sybil with the work of Watson et al. [274], *ii)* to demonstrate that Sybil can be used for predictions at different levels—country, state (for U.S. states), regional, and city levels—, and *iii)* to show that Sybil can be used with various data sources—in this case, we used an alternative source for variants data, but it is also possible to use a different source for COVID-19 data—like cases and fatalities. The surveillance data used are available in the COVID-19 R library [121, 120] (for the results presented we used a snapshot dated June 13$^{\text{th}}$, 2024). Information about active virus strains comes from the CDC [71] (for the results presented we used a snapshot dated June 13$^{\text{th}}$, 2024). In particular, we used data related to the specific Health and Human Services (HHS) U.S. region for each considered state—i.e., Denver (Region 8) for the state of Colorado, New York (Region 2) for the state of New York and Philadelphia (Region 3) for the state of West Virginia. Due to the unavailability of data on recoveries for the USA, we used a fixed recovery rate with mean $\frac{1}{\gamma}$ equal to 14 days [65]—same for each variant.

To compare Sybil to the work of Watson et al., we chose the period from December 1$^{\text{st}}$, 2020 to December 31$^{\text{st}}$, 2020 as training data—both for Sybil and the framework by Watson et al.—, and we forecast the daily infections in January 2021 (starting from January 1$^{\text{st}}$). We chose this period because *i)* this is the final training date for the scenarios presented in the study of Watson et al., *ii)* we are near to a peak of the *Other* variant—before or after the peak—, and *iii)* the data used by Watson et al. only extend until March 7$^{\text{th}}$, 2021.

Figures 3.27-*a)*, 3.28-*a)* and 3.29-*a)* show the evolution of $I_v(\tilde{t})$ compartments from February 2020 to March 2023 in the considered U.S. states, while figures 3.27-*b)*, 3.28-*b)* and 3.29-*b)* confirm that Sybil provides good predictions—except for the state of New York after three and four weeks. Finally, Table 3.4 shows that Sybil provides similar or better prediction with respect to the work of Watson et al., especially for two-, three-, and four-week forecasts.

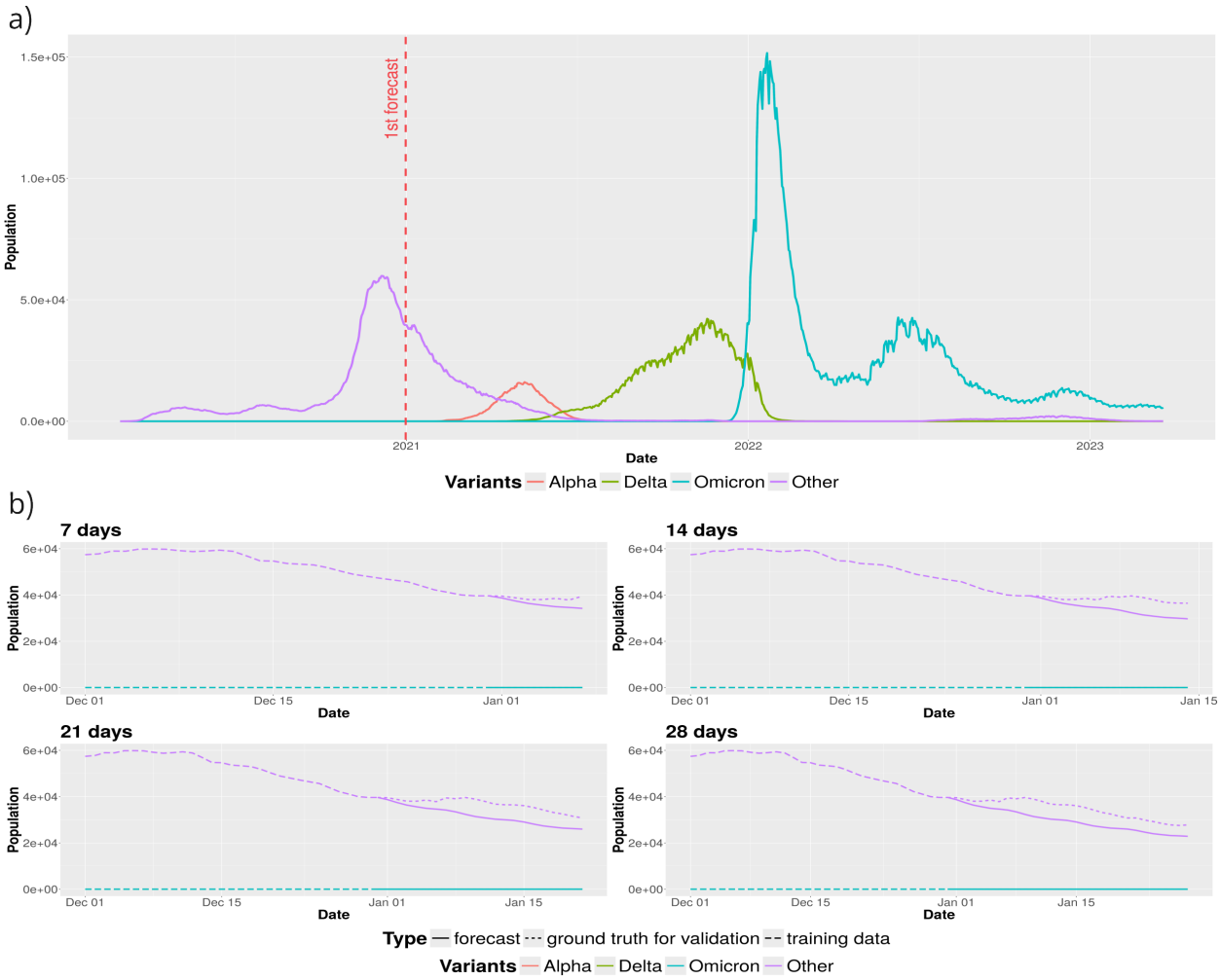


Figure 3.27: The figure shows *a)* the evolution of $I_v(\tilde{t})$ compartments from February 2020 to March 2023 in the state of Colorado, USA, with the considered forecast scenario, and *b)* the evolution of the $I_{Other}(\tilde{t})$ compartment in the state of Colorado, USA, using Sybil starting the forecast from January 1st, 2021 (the dashed line shows the prediction, while solid and dotted lines represent the training data and the ground-truth values extracted from the surveillance data, respectively)..

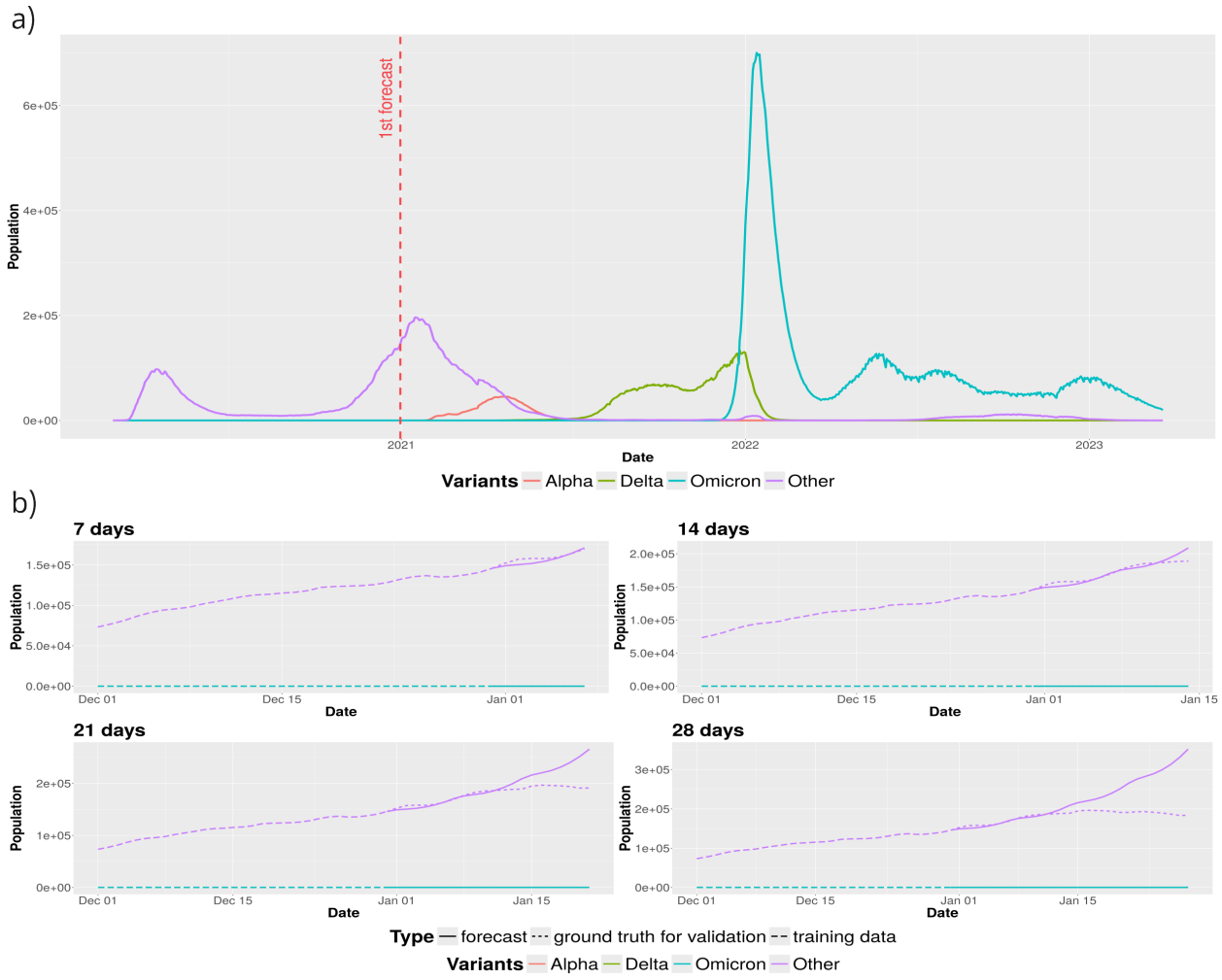


Figure 3.28: The figure shows *a)* the evolution of $I_v(\tilde{t})$ compartments from February 2020 to March 2023 in the state of New York, USA, with the considered forecast scenario and *b)* the evolution of the $I_{Other}(\tilde{t})$ compartment in the state of New York, USA, using Sybil starting the forecast from January 1$^{\text{st}}$, 2021 (the dashed line shows the prediction, while solid and dotted lines represent the training data and the ground-truth values extracted from the surveillance data, respectively).

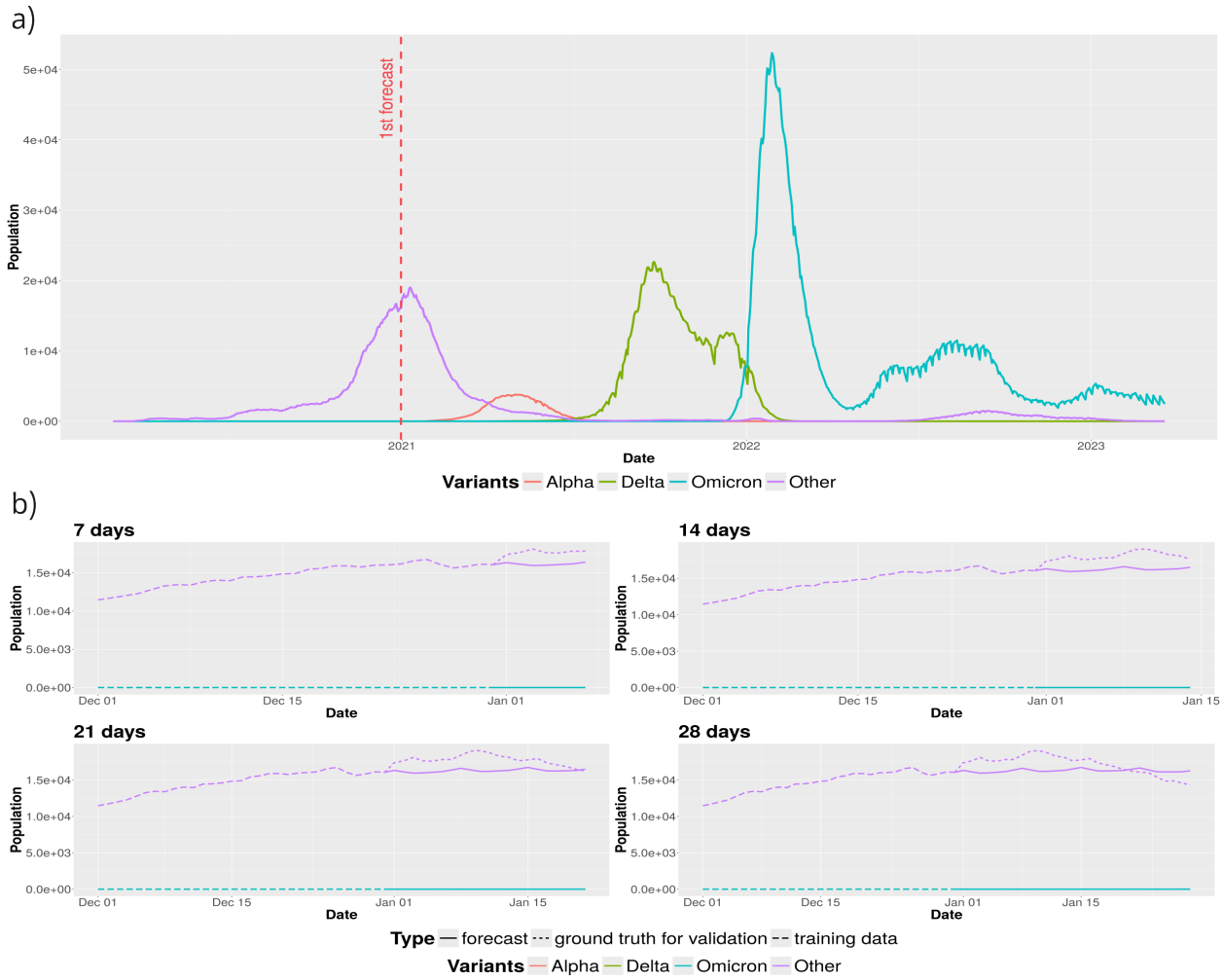


Figure 3.29: The figure shows *a)* the evolution of $I_v(\tilde{t})$ compartments from February 2020 to March 2023 in the state of West Virginia, USA, with the considered forecast scenario, and *b)* the evolution of the $I_{Other}(\tilde{t})$ compartment in the state of West Virginia, USA, using Sybil starting the forecast from January 1st, 2021 (the dashed line shows the prediction, while solid and dotted lines represent the training data and the ground-truth values extracted from the surveillance data, respectively).

Table 3.4: RMSE with standard deviation between the ground truth used for validation and the obtained forecast with Sybil and the framework by Watson et al. [274] for the three considered U.S. states. For each scenario and forecasting window (one-, two- three-, and four-week), values in bold represent minimum values, while underlined values represent values close to minimum values.

Approach	1 week *mean (± std)*	2 weeks *mean (± std)*	3 weeks *mean (± std)*	4 weeks *mean (± std)*
Colorado				
Sybil	<u>2958</u> (± 1404)	**5202** (± 2317)	**5562** (± 2079)	**5395** (± 1813)
Watson et al.	**2399** (± 1217)	<u>5231</u> (± 3026)	10104 (± 6110)	14690 (± 8598)
New York				
Sybil	**3948** (± 2873)	**6860** (± 6851)	**26540** (± 22617)	**64364** (± 50734)
Watson et al.	15526 (± 8771)	27622 (± 14368)	47854 (± 26877)	<u>74932</u> (± 44059)
West Virginia				
Sybil	1622 (± 304)	<u>1945</u> (± 525)	**1682** (± 735)	**1559** (± 1280)
Watson et al.	**731** (± 207)	**1532** (± 1405)	3746 (± 2896)	6329 (± 4474)

3.2.9 Continuous monitoring

To set up a continuous monitoring system we have to obtain good predictions also in periods in which there is a new emerging variant or a new exploding outbreak. In Section 3.2.1 we showed an example in which we forecast just before the exploding of the largest outbreak in Italy and, in Section 3.2.5, we showed another example using 11 different aggregations of lineages. In particular, considering again the case with 4 aggregation of lineages, Figure 3.30 shows two scenarios in which we forecast just before a new outbreak—of the new emerging Omicron variant (first scenario) and of the Alpha variant (second scenario), respectively.

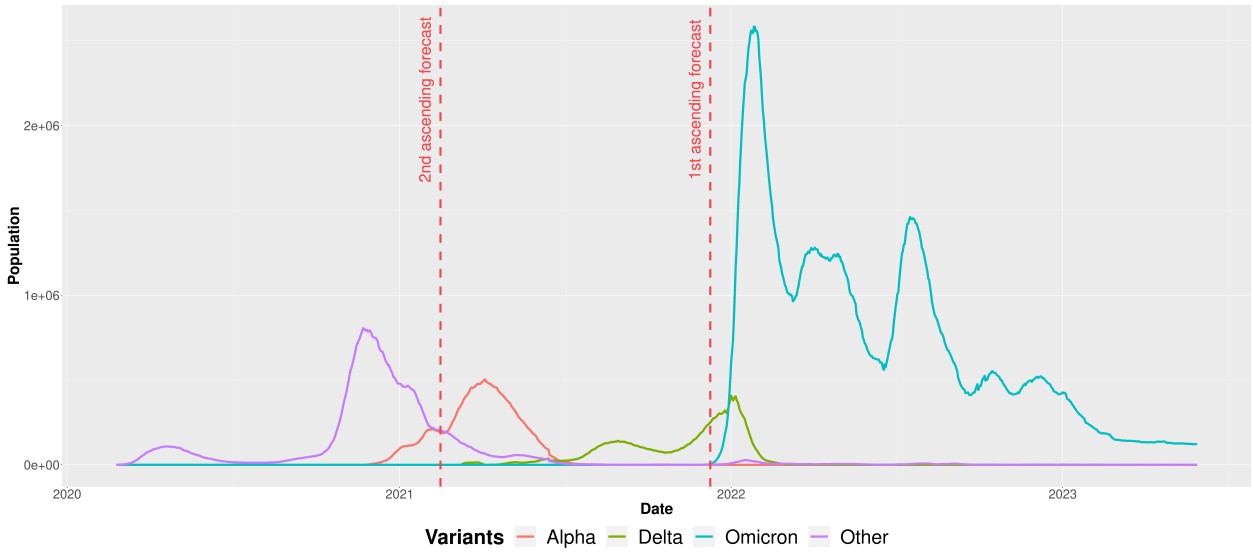


Figure 3.30: Evolution of $I_v(\tilde{t})$ compartment from February 2020 to May 2023 in Italy with the two considered ascending forecast scenarios.

Figures 3.31 and 3.32 show an extension of the first ascending scenario presented in Section 3.2.1. In particular, these figures show how the one-week forecast changes moving the training window from December 9$^{\text{th}}$, 2021 to January 11$^{\text{th}}$, 2022 by three days. In this case, the upward trajectory is very steep, and Sybil finds it harder to calibrate with respect to the previous cases. Here, the Omicron variant is a new emerging variant and Sybil initially foresees a more aggressive exponential growth. Looking at Figure 3.31 we can state that Sybil nailed the qualitative prediction, but needed more data to calibrate it. We can also see that Sybil captures well the prediction on the other active variant—the Delta variant.

In the second ascending scenario there are three active variants: Alpha, Delta and the *Other* variant. Figure 3.33 shows how the one-week forecast changes moving the training window from February 15$^{\text{th}}$, 2021 to March 20$^{\text{th}}$, 2021 by three days. In particular, the Alpha variant is ascending while the *Other* variant is descending. The first three rows show that Sybil initially sees an exponential growth for the Alpha variant—see also Figure 3.34 for the

predictions on the infection rates. This exponential growth increases as the forecasted period increases. Again, we can state that Sybil nailed the qualitative prediction, but needs more data to calibrate the quantitative prediction—see the last two rows of Figure 3.33.

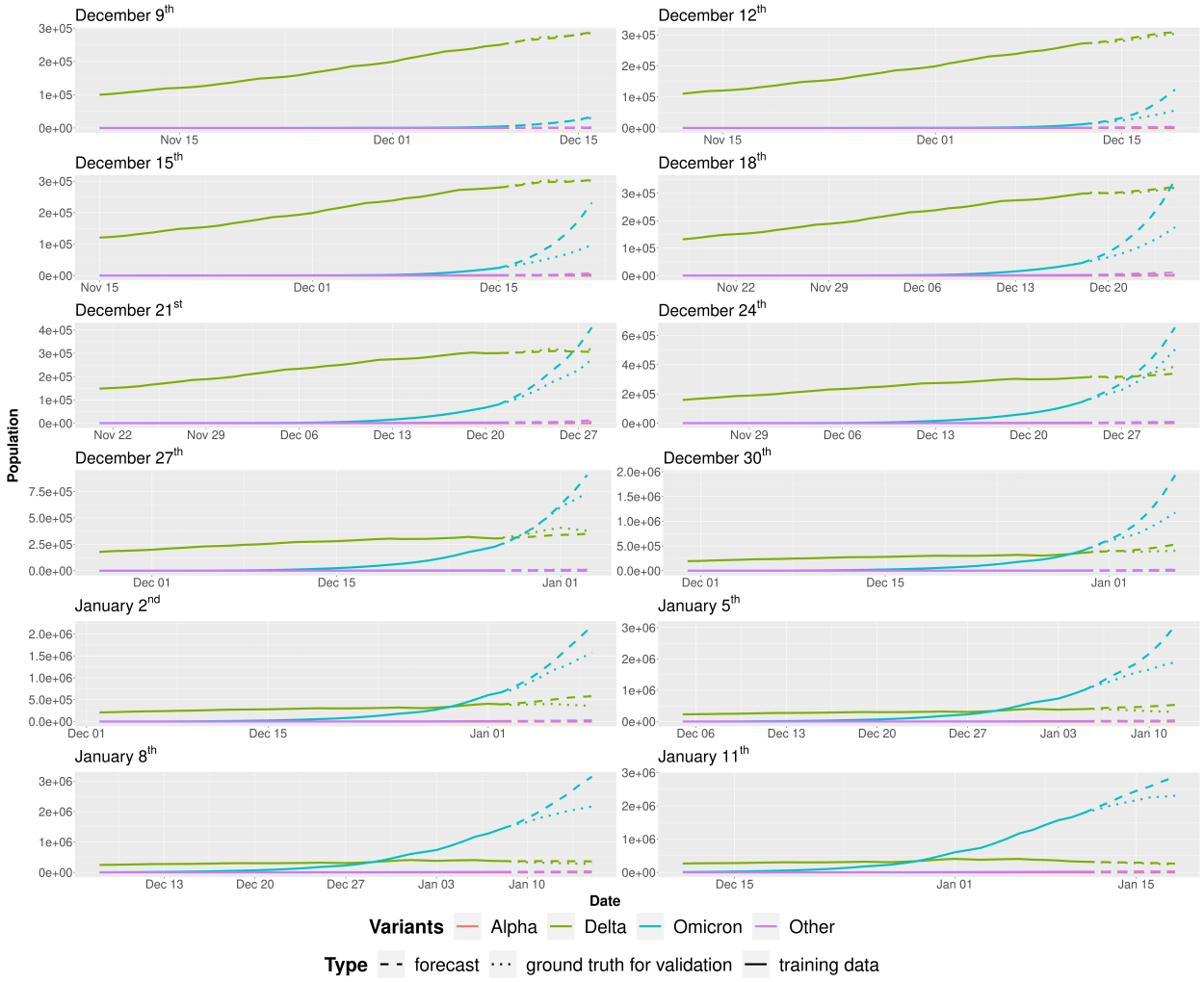


Figure 3.31: Evolution of $I_v(\tilde{t})$ compartments using Sybil in the first scenario in Italy starting the forecast from December 9^{th}, 12^{th}, 15^{th}, 18^{th}, 21^{st}, 24^{th}, 27^{th}, 30^{th}, 2021 and from January 2^{nd}, 5^{th}, 8^{th} and 11^{th}, 2022 moving the training window by three days (the dashed line shows the prediction, while solid and dotted lines represent the training data and the ground-truth values extracted from the surveillance data, respectively). All plots refer to a forecast one week into the future.

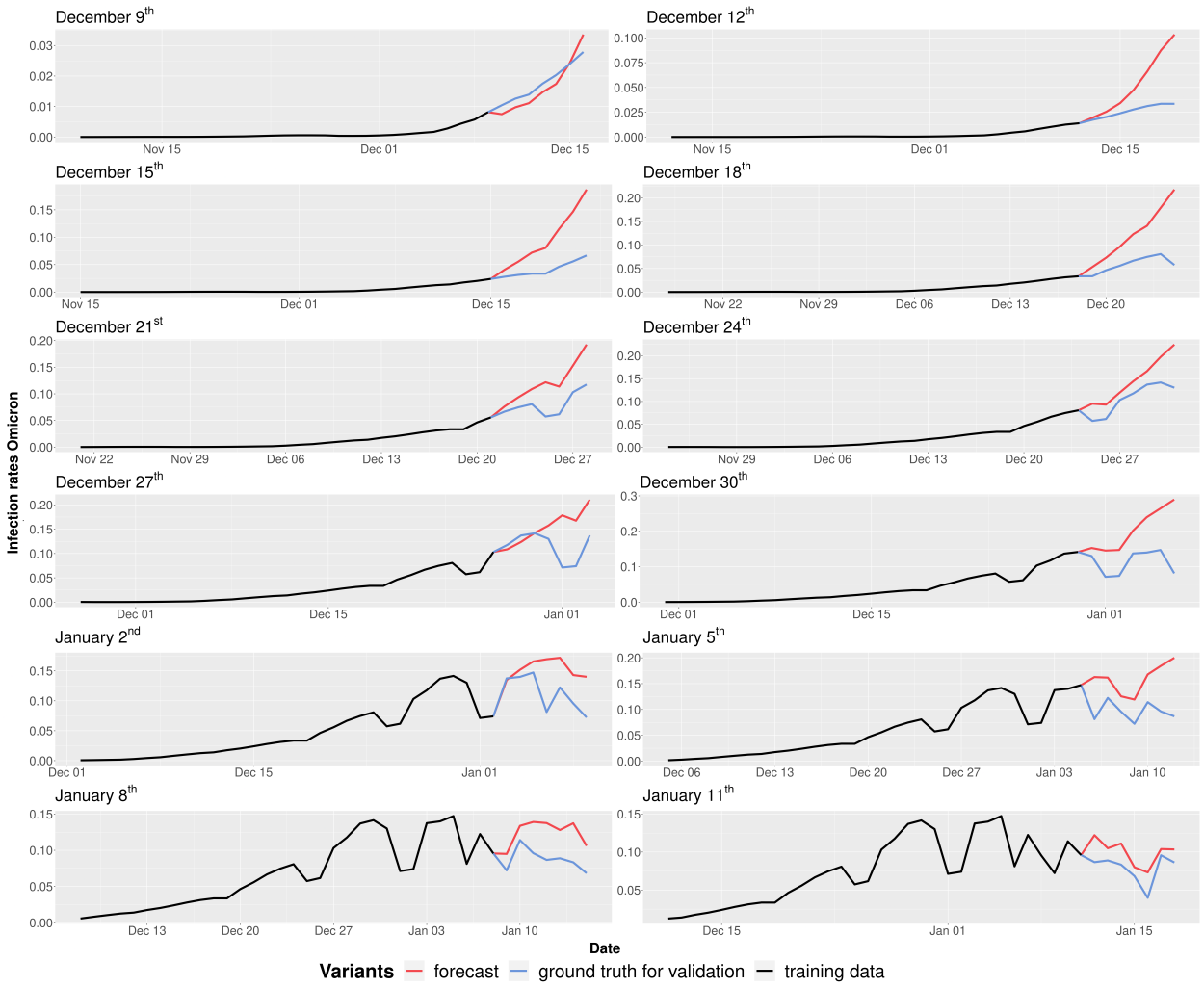


Figure 3.32: Evolution of infection rates $\beta_{Omicron}(\tilde{t})$ using Sybil in the first scenario in Italy starting the forecast from December 9^{th}, 12^{th}, 15^{th}, 18^{th}, 21^{st}, 24^{th}, 27^{th}, 30^{th}, 2021 and from January 2^{nd}, 5^{th}, 8^{th} and 11^{th}, 2022 moving the training window by three days (the dashed line shows the prediction, while solid and dotted lines represent the training data and the ground-truth values extracted from the surveillance data, respectively). All plots refer to a forecast one week into the future.

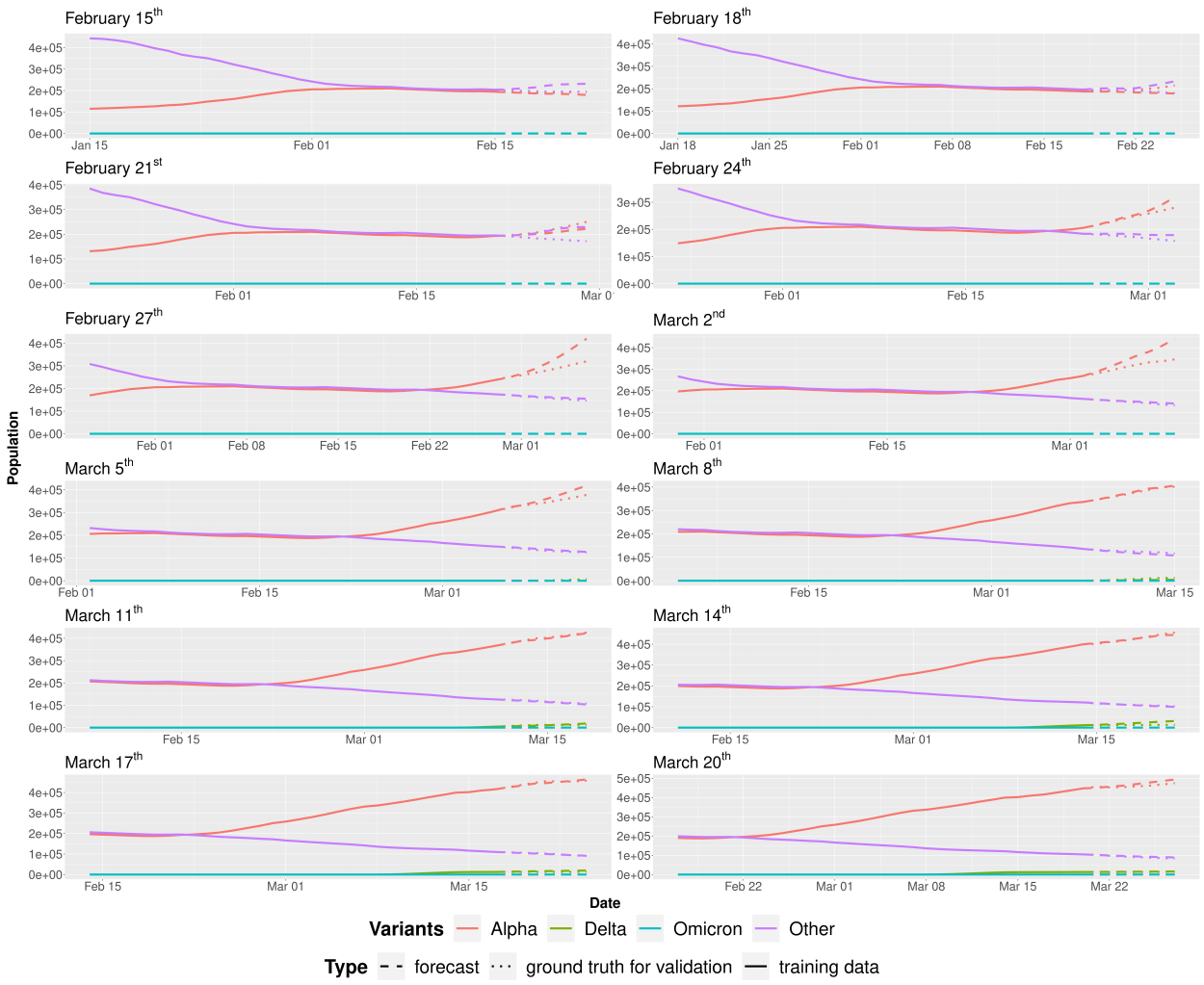


Figure 3.33: Evolution of $I_v(\tilde{t})$ compartments using Sybil in the second scenario in Italy starting the forecast from February 15^{th}, 18^{th}, 21^{st}, 24^{th}, 27^{th}, March 2^{nd}, 5^{th}, 8^{th}, 11^{th}, 14^{th}, 17^{th} and 20^{th}, 2021 moving the training window by three days (the dashed line shows the prediction, while solid and dotted lines represent the training data and the ground-truth values extracted from the surveillance data, respectively). All plots refer to a forecast one week into the future.

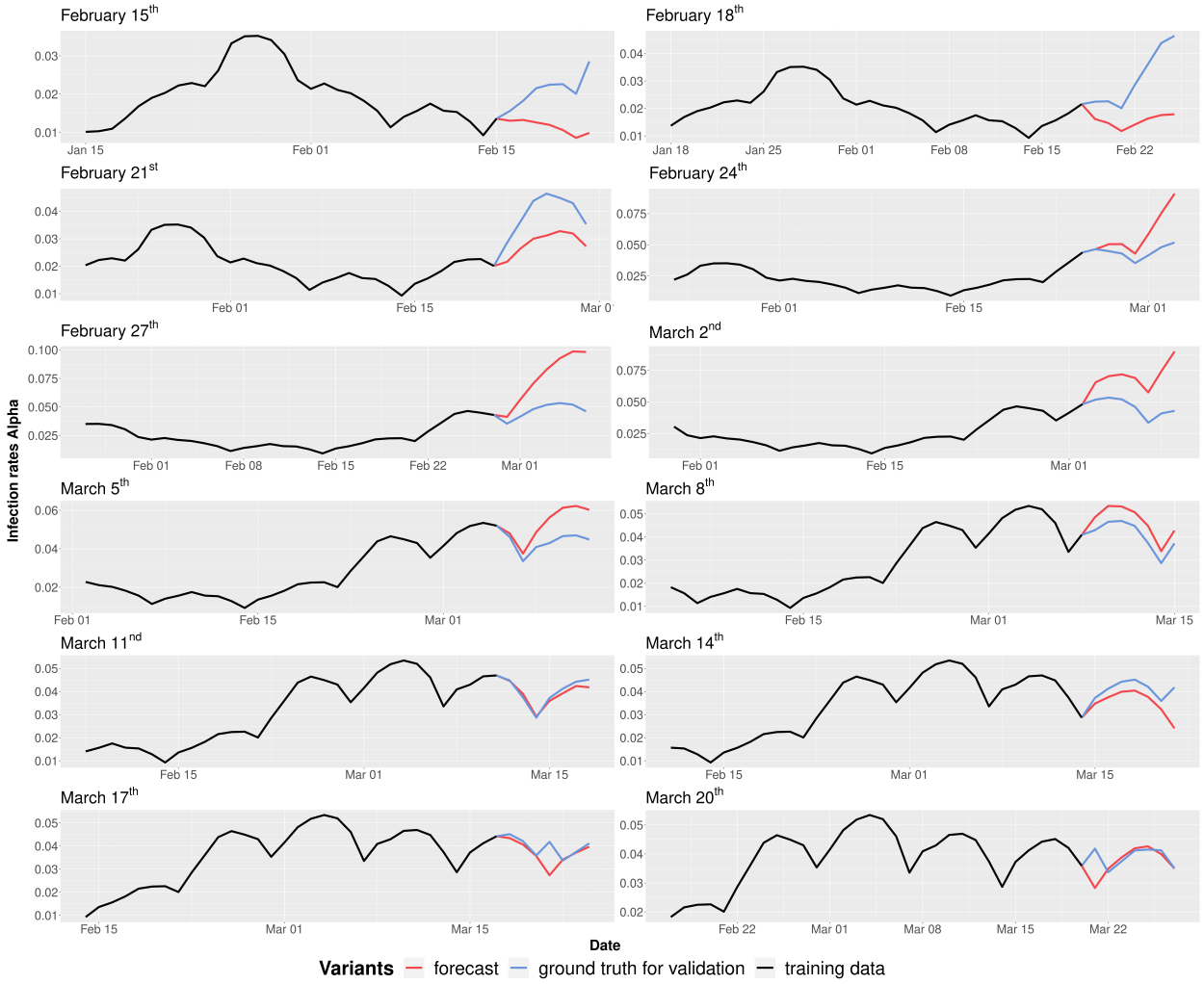


Figure 3.34: Evolution of infection rates $\beta_{Alpha}(\tilde{t})$ in the second scenario in Italy starting the forecast from February 15^{th}, 18^{th}, 21^{st}, 24^{th}, 27^{th}, March 2^{nd}, 5^{th}, 8^{th}, 11^{th}, 14^{th}, 17^{th} and 20^{th}, 2021 moving the training window by three days (the dashed line shows the prediction, while solid and dotted lines represent the training data and the ground-truth values extracted from the surveillance data, respectively). All plots refer to a forecast one week into the future.

3.3 Discussion

ML approaches have already shown their power when dealing with complexity, thanks to their excellent capabilities in exploiting non-trivial correlations often inaccessible with other tools. Despite the many advantages of employing ML approaches, there are also some caveats worth noting. First, most of these approaches require a tremendous amount of data to learn from, and data available from the surveillance might not be enough for most ML approaches. Sybil faces this issue from two different standpoints; it employs Prophet, a hybrid approach, using ML techniques in combination with simulations—refer to Section 3.1.2 for further details. On the other hand, Sybil does not face the challenge of forecasting the virus spread as a single task. First, it tries to predict the rates and then computes the future dynamics using a compartmental model. Therefore, Sybil, combining Prophet and compartmental models, exploits the two approaches in the respective field of application—where they perform the best—and strengthens the individual weaknesses. Specifically, by providing the compartmental model with parameters extracted from the real data or forecasts, there is no need to tune the model and estimate the missing parameters. Estimating the model's parameters is a resource-demanding and time-consuming task. Also, it is an activity tightly related to the specific situation or scenario under evaluation. This makes the model almost impossible to apply in a different setting without estimating the parameters again. Hence, Prophet makes Sybil's approach easily deployable in new scenarios without requiring additional tasks, but the data must comply with the daily requirements. Likewise, the use of compartmental models in predicting the future trend of the infection makes results more straightforward to explain, for instance, to policy-makers, who often do not trust predictions without a robust interpretation.

In Section 3.2, we presented Sybil's forecasts of different lengths for several European states and different periods spanning from February 2020 to May 2023. To showcase Sybil's capabilities, all the forecasting periods cover important changes in the first derivative of the number of daily infections. Then, Sybil's predictions are contrasted against the surveillance data and the plain application of Prophet. Results show the superiority of Sybil's approach on the plain forecast of the number of daily infections; in particular, Sybil outperforms the plain application of Prophet when primary changes in virus diffusion happen. For instance, Figure 3.5-*(b)* clearly shows the added value of the Sybil approach: it predicts with great accuracy the peak of the current wave while Prophet alone fails to predict the decreasing phase of the infection peak. Moreover, some experiments made choosing forecasting periods far from peaks show that similar and accurate forecasting can be obtained using both approaches, and new variants rising or new outbreaks can be predicted accurately using Sybil to set up a

continuous monitoring system.

Furthermore, we can assert that Sybil almost always outperforms the other considered state-of-the-art approaches—namely, Prophet [257], ARIMA / SARIMA [61], Neural Prophet [263], LSTM [131], and GRU [79] neural networks, EpiNow2 [235], as well as the framework of Watson et al. [274]—particularly for the two-, three-, and four-weeks forecast periods. In contrast, for the one-week forecast, Sybil provides predictions that are closely aligned with those of the best-performing approach. Additionally, when there are minor changes or no changes in the trajectory, Sybil's performance remains robust, effectively matching the performance of the other methods considered.

The presented methodology is nearly a key element of a continuous monitoring system, but it heavily depends on accurate data collection. In particular, in [121, 120] there are few nations with a complete time series (with a daily step) for the data we used (i.e., cumulative number of cases, cumulative number of deceases, cumulative number of recoveries). For this reason, we extended this methodology to include the use of a fixed recovery rate—same or different for each variant—to dispense with data on recoveries which are often not available and add a pre-processing step using a technique (e.g., splines) to fill missing data and to move from a weekly to a daily step in case of availability of data with a weekly step—without alter too much the information present in the time series. Finally, here we focus on country-level predictions (and state-level predictions for the USA), but Sybil can also be used at regional and city levels.

3.4 Conclusion

The global COVID-19 pandemic has brought to light the urgent necessity for sophisticated tools capable of monitoring and predicting the trajectory of infections within the population. This chapter introduces a cutting-edge framework designed for continuous monitoring and forecasting, seamlessly integrating machine learning-based predictive models with compartmental models.

Sybil distinguishes itself by delivering forecasts that are not only reliable and replicable but also readily explainable, as evidenced by thorough experimental validation. The adaptability of this innovative approach is clearly demonstrated through its successful application to diverse surveillance data from several European countries—specifically Italy, Austria, Belgium, France, Germany, and Sweden—and some U.S. states. In addition, the ability to provide accurate predictions even in the presence of a peak or a new emerging outbreak and the integration of variants are of vital help to policy-makers. Finally, by utilizing only data from the previous month, Sybil significantly reduces the demand for training data, enhancing computational efficiency and enabling effective operation with limited data.

By integrating data-centric and analytic approaches, Sybil effectively addresses the inherent limitations of each method. This amalgamation not only enhances the tool's ability to forecast the evolution of COVID-19 but also positions Sybil as a versatile instrument for predicting the trajectory of various other diseases, thus broadening its scope and impact in the field of infectious disease modeling by helping policy-makers in quickly reacting to a new emerging disease.

Future work could involve developing an R package that generalizes the underlying compartmental model, allowing users to easily apply Sybil in diverse contexts beyond infectious diseases. This would be possible as long as the model is represented by a system of ordinary differential equations that meet specific conditions, such as the absence of cyclic dependencies among the rates the user intends to estimate.

3.5 Availability of data and software code

COVID-19 data used in this study are available in the COVID19 R library [121, 120] and in the ECDC and CDC variants data [1, 148, 71]. The framework developed for the analyses presented in this work is available at `https://github.com/daniele-baccega/sybil-forecasting`.

3.5.1 How to reproduce the results

Requirements

You need to have docker installed on your computer, for more info see this document: `https://docs.docker.com/engine/installation/`. Ensure your user has the right to run docker (without the use of sudo). To create the docker group and add your user in a Unix system:

- Create the docker group:

```
$ sudo groupadd docker
```

- Add your user to the docker group:

```
$ sudo usermod -aG docker $USER
```

- Log out and log back in so that your group membership is re-evaluated.

Reproduce

To reproduce the results clone the repository (`https://github.com/daniele-baccega/sybil-forecasting.git`) and run:

```
$ cd sybil-forecasting
$ ./reproduce.sh
```

To reproduce the results with a fixed recovery rate—see Section 3.2.7—run (inside the $sybil - forecasting$ directory):

```
$ git checkout sybil-v2.0
$ ./reproduce.sh
```

To reproduce the results on the U.S. states—see Section 3.2.8—run (inside the $sybil - forecasting$ directory):

```
$ git checkout sybil-v2.5
$ ./reproduce.sh
```

Chapter 4

Social compliance with NPIs, mobility patterns, and reproduction number: lessons from COVID-19 in Europe

In this chapter, we explore the interplay between country-imposed restrictions, people's mobility, and the effective reproduction number R_t. As a key metric for monitoring a pandemic, the R_t is crucial for predicting its evolution, although it is challenging to measure directly. Conversely, mobility and NPIs are more accessible parameters, where policies implemented by decision-makers easily influence mobility, and they play a significant role in shaping the evolution of the pandemic. In particular, we aim to provide insights on how more easily observable variables (such as NPIs and people's mobility) can serve as surrogates for policy-makers, as these variables influence harder-to-observe metrics (like the R_t).

Utilizing restrictions and mobility data from major platforms [32, 104, 151, 176, 123], we explore the interplay between policy changes and movement trends, providing insights into disease transmission dynamics. In particular, our analyses:

i) demonstrates a strong correlation between the considered restrictions and mobility variables, contrasted by a weaker correlation between mobility and infection rates;

ii) reveals cultural patterns in mobility, highlighting how official restrictions and self-imposed measures shape behavior differently across diverse countries;

iii) examines the contrast between Northern and Southern European nations in terms of social compliance and interaction patterns using clustering techniques like K-means [125] and CONNECTOR [203];

iv) estimates missing mobility data based on restriction indicators and infer R_t values from mobility trends using XGBoost regression models [77].

Specifically, our analyses focuses on several European countries: Norway, Sweden, Finland, Ireland, the United Kingdom, Belgium, France, Germany, Austria, Italy, Greece, Spain, and Portugal.

As highlighted in Chapter 1, NPIs allow for reduction of the transmission through changes in behaviors, environments, and community interactions. One of the most immediate and observable effects of NPIs has been their impact on human mobility. Mobility, which refers to the movement of individuals within and between geographic areas, plays a crucial role in the transmission dynamics of infectious diseases. The relationship between NPIs and mobility is complex and multifaceted. On the one hand, stringent measures such as nationwide lock-downs and curfews have led to dramatic declines in mobility, as evidenced by data from mobility tracking tools and transportation statistics [136, 22]. On the other hand, the relaxation of these measures has often been followed by a resurgence in mobility, sometimes correlating with subsequent waves of infections. Moreover, variations in compliance and enforcement, socio-economic factors, and public trust in government policies have influenced the effectiveness of NPIs and the extent of mobility reductions [15, 81].

During the COVID-19 pandemic, changes in mobility patterns have been both a direct consequence of policy measures and a reflection of individual behavioral responses to the perceived risk of infection. The reduction in mobility, especially through lock-downs and travel bans, aimed to decrease the opportunities for person-to-person transmission, thereby slowing the spread of the virus and reducing the burden on healthcare systems [31, 246, 47]. The implementation of NPIs has varied significantly across different regions and periods, reflecting diverse epidemiological, political, and socio-economical contexts [204].

The differences among countries in their approach to NPIs and their impact on mobility are striking. For instance, countries with robust healthcare systems and high levels of public trust in government, such as Germany, saw more consistent compliance with mobility restrictions, resulting in more effective control of the effective reproduction number R_t. In contrast, countries with fragmented healthcare systems and lower public trust, such as the USA and Brazil, experienced greater challenges in enforcing NPIs and saw less dramatic declines in mobility [46, 72].

Social compliance, shaped by trust in institutions and adherence to societal norms, rules, and regulations, played a crucial role in the effectiveness of NPIs [267, 196]. In countries where individuals more readily adhered to collective public health goals, NPIs proved to be more effective in reducing disease spread. Cultural factors, such as societal norms, values, and attitudes toward government authority, further influenced these outcomes.

For example, in East Asian countries, where there is a strong tradition of community responsibility and compliance with public health measures, NPIs were more effective in reducing mobility compared to Western nations, where individual freedoms are more heavily emphasized [164]. Additionally, social interactions, representing the dynamic exchanges between individuals or groups that collectively shape relationships, cultural norms, and societal structures, contributed significantly to the effectiveness of these measures—where social networks facilitated compliance or, conversely, hindered adherence in regions with less social cohesion. Economic disparities also played a role, as lower-income countries and communities often faced greater difficulties in adhering to lock-down measures due to the necessity of continuing economic activities. These variations underscore the importance of context-specific strategies that consider both social compliance and the socio-economic landscape in managing mobility and R_t during a pandemic [185].

Thus, mobility data plays a critical role in understanding the dynamics of infectious disease transmission. Still, it often contains significant gaps due to limited geographic coverage, inconsistent data collection practices, and varying levels of technology adoption. Similarly, R_t data can be impacted by irregular reporting, testing backlogs, and inconsistent public health monitoring across different regions. These data limitations hinder the evaluation of policy impacts and the development of reliable predictive models. To overcome these challenges, it is essential to employ advanced methods that address these gaps, allowing for more accurate insights into how NPIs influence mobility patterns and, ultimately, the spread of infectious diseases. This leads to better-informed policy decisions and more effective public health interventions.

In this chapter, we begin by presenting the data used in our study, which includes government-imposed restrictions, various categories of people's mobility, and COVID-19 data. We then explain the pre-processing steps and provide preliminary analysis that highlight a strong correlation among the considered variables (Section 4.1). Next, we present the obtained results, focusing on patterns, clustering, and regression models (Section 4.3). Finally, we discuss the findings (Section 4.4), draw conclusions (Section 4.5), and describe how to replicate the results (Section 4.7).

4.1 Data preparation

This section details the data employed in the study (Section 4.1.1), providing a comprehensive account of the resources used to achieve the research objectives, including the pre-processing (Section 4.1.2) and some preliminary analysis (Section 4.1.3).

4.1.1 Data

We utilized several types of COVID-19 data, including official surveillance data from the COVID-19 Data Hub [121, 120], estimates of active cases computed from the UMD Global CTIS's survey [32, 231, 24, 234], mobility data from UMD Global CTIS's survey [32, 104, 151], information on restrictions from the Oxford Covid-19 Government Response Tracker (OxCGRT) [123], and Google mobility data [176]. All datasets used in this work are open-source and publicly available. Data from UMD Global CTIS's survey are available in an aggregated form—see Section 4.7 for more details.

The COVID-19 Data Hub is the primary source for surveillance data [121, 120]. Specifically, for each country, we used the cumulative number of fatalities to calculate daily new fatalities and the total population. While the Data Hub offers a comprehensive range of information, including vaccinations, testing numbers, hospitalizations, and intensive care unit occupancy, these additional data are not relevant to our current study. Official COVID-19 mortality data should be generally more reliable than case data due to the consistent and rigorous reporting standards applied to fatalities, unlike the variability and potential under-reporting often seen in case counts—also due to asymptomatic infections. For these reasons, we decided to use estimates of active cases from prior studies [32, 231, 24, 234], which were based on direct symptom responses indicative of COVID-like illness (CLI) proposed in UMD Global CTIS—i.e., fever along with cough, shortness of breath, or difficulty breathing—, along with the cumulative number of fatalities, rather than relying on surveillance data for officially reported cases.

Since the spring of 2020, Facebook has been collaborating in the context of their program Data for Good [6] with the University of Maryland (UMD) promoting surveys created by this institution [32, 104, 151]. The University of Maryland Social Data Science Center Global COVID-19 Trends and Impact Survey (UMD Global CTIS), in partnership with Facebook [104], has been recording daily responses from invited Facebook users on topics related to the COVID-19 pandemic, including RT-PCR results, symptoms, and vaccinations, from over 114 countries. We have obtained access to the responses through an agreement with the UMD and Facebook—see Section 4.6. The survey questions are available on the

UMD Global CTIS website [258]. In particular, from this platform, we used aggregated mobility information considering the following questions—the labels correspond to those used in the survey:

- In the last 24 hours, have you done any of the following?
  - *C0_1*: Gone to work outside the place where you are currently staying.
  - *C0_2*: Gone to a market, grocery store, or pharmacy.
  - *C0_3*: Gone to a restaurant, cafe, or shopping center.
  - *C0_4*: Spent time with someone who isn't currently staying with you.
  - *C0_5*: Attended a public event with more than 10 people.
  - *C0_6*: Used public transit.
- *C5*: In the past 7 days, how often did you wear a mask when in public?

Starting from May 20$^{\text{th}}$, 2021, all these questions—except *C5*—were modified, leading to a significant change in their meaning and values. For example, question *C0_1* changed from *In the last 24 hours, have you done any of the following? Gone to work outside the place where you are currently staying* to *In the past 24 hours, have you done any of the following? Gone to work or school* ***indoors****, outside the place where you are currently staying.* The change in values is likely attributable to the use of the word ***indoors*** in the questions. For this reason, we focused the attention on data from April 29$^{\text{th}}$, 2020 to May, 19$^{\text{th}}$, 2021. We will refer to this data as Facebook data.

To combat the COVID-19 pandemic, countries worldwide have implemented stringent policies (NPIs), such as stay-at-home lock-downs, school and workplace closures, event cancellations, and public transport restrictions, to slow virus spread by enforcing physical distancing. The Oxford Covid-19 Government Response Tracker (OxCGRT) [123] provides information on which pandemic response measures were enacted by governments and when. The OxCGRT furnishes accessible data on 24 policy indicators, accompanied by a categorization of government responses into four groups: *i)* containment and closure policies (C)—i.e., schools and workplaces closure, cancel public events, restrictions on gathering, close public transport, stay at home requirement, restrictions on internal movement, and on international travel—, *ii)* economic policies (E)—i.e., income support, debt/contract relief for households, fiscal measures, and giving international support—, *iii)* health system policies (H)—i.e., public information campaign, testing policy, contact tracing, emergency investment in healthcare, investment in COVID-19 vaccines, facial coverings, vaccination policy, and protection of elderly people—, and *iv)* vaccination policies (V)—i.e., vaccine prioritization,

vaccine eligibility/availability, vaccine financial support, vaccine requirement/mandate, and many other indices. They also provide four indices that aggregate the information into a single number with each index reporting a number between 0 and 100, reflecting the level of government response across specific dimensions: *i)* overall government response index (includes all indicators, i.e. C, E, H, and V), *ii)* containment and health index (includes all C and H indicators), *iii)* stringency index (includes all C indicators, plus H1, which records public information campaigns), and *iv)* economic support index (includes all E indicators). In this study, we used the whole dataset on restrictions, focusing in some cases on the stringency index.

Using anonymized data from apps like Google Maps, Google's COVID-19 Community Mobility Reports [176] provides a dataset showing the percentage of changes in people's mobility throughout the pandemic. This dataset measures daily visits to specific locations (e.g., workplaces, homes, grocery stores, parks, transit stations) and compares them to baseline days preceding the outbreak, introducing a significant bias. This baseline represents typical values calculated as the median over the five weeks from January 3 to February 6, 2020, accounting for routine differences between weekdays and weekends. In particular, we used data on workplaces (time spent at work), retail and recreation (includes places like restaurants, cafes, shopping centers, theme parks, museums, libraries, and movie theaters), grocery and pharmacy stores (includes places like grocery markets, food warehouses, farmers markets, specialty food shops, drug stores, and pharmacies), transit stations (includes public transport hubs such as subway, bus, and train stations) and residential (time spent at home). We confirmed a strong correlation between Facebook and Google mobility data (see Section 4.1.3 for details).

4.1.2 Pre-processing

In this section, we describe the data pre-processing steps taken to clean, transform, and prepare the raw data for effective analysis and modeling. In particular:

1. We normalized in [0, 1] the mobility variables considered from Facebook data.

2. We normalized the stringency index and the mobility variables considered from Google data.

3. We extracted the infection rates of COVID-19—similar to R_t—using Sybil [39], a machine learning and variant-aware compartmental model framework introduced in Chapter 3.

The first step involves the pre-processing of the considered UMD Global CTIS's survey questions—we took the number of people that answered *Yes* to the first six questions (see Section 4.1.1), divided by the daily number of responses, obtaining values in [0, 1]—to represent different mobility categories: *i) C0_1* as *Workplaces* mobility, *ii) C0_2* as *Grocery* mobility, *iii) C0_3 + C0_5* as *Recreation* mobility, *iv) 1 - C0_4* as the complement of *Residential* mobility—we took the complement to ensure uniformity in the direction of all mobility trajectories—, and *v) C0_6* as *Stations* mobility.

Moreover, we processed question *C5* in the following way: for each answer, there are five possible binary options after applying one-hot encoding[1]. Specifically, *i) ATT* indicates that the person used a mask all the time, *ii) MOTT* most of the time, *iii) SOTT* some of the time, *iv) ALOTT* a little of the time, and *v) NOTT* none of the time. We computed a weighted average—normalized in the range [0, 1]—using the following fraction:

$$masks = \frac{1 \cdot ATT + 0.75 \cdot MOTT + 0.5 \cdot SOTT + 0.25 \cdot ALOTT + 0 \cdot NOTT}{ATT + MOTT + SOTT + ALOTT + NOTT} \tag{4.1}$$

In particular, for each option, we applied a decreasing weight to the values, reflecting how frequently individuals used masks in public. Higher values indicate a greater proportion of people who consistently wore masks in public over the past week. We then divided the weighted sum by the total number of responses, resulting in values within the range $[0, 1]$ that estimate the proportion of people who wore a mask in the past seven days. This variable was named *Masks*.

In the second step, we normalized the Google mobility variables—*Workplaces*, *Stations*, *Recreation*, *Residential*, and *Grocery*—and the *Stringency* index by dividing them by 100, as both are expressed as percentages. Google mobility data ranges from -1 to N, where -1 indicates no mobility compared to the baseline, and N represents the total population increase. The maximum possible increase occurs when only one person moves at time *t*, followed by the entire population moving at time *t + 1*. Conversely, the maximum possible decrease occurs when some individuals move at time *t*, followed by no one moving at time *t + 1*. However, these extreme scenarios do not appear in the data. As illustrated in Figure 4.1, the five Google mobility variables we analyzed range between -0.75 and 1. Using Google mobility data, we were unable to compute the complement of *Residential* mobility because the values are relative to the baseline introduced above.

[1]The possible answers are: *All the time*, *Most of the time*, *Some of the time*, *A little of the time*, and *None of the time*. For example, if a person answered *Some of the time*, we assigned the following vector to that response: [0, 0, 1, 0, 0]. By applying this approach to all responses and summing the vectors, we can determine the number of people who selected each specific answer.

In the third and last step, we extracted the *Infection Rates*—similar to R_t—using Sybil [39]. As introduced in Chapter 3, Sybil is an integrated machine learning and variant-aware compartmental model framework designed to enhance prediction accuracy and explainability. Sybil leverages the relative stability of disease characteristic indices to make future projections and utilizes a straightforward, widely recognized analytical model to outline the infection dynamics. In particular, Sybil combines a simple compartmental model—in this case, a Susceptible-Infected-Recovered-Deceased-Susceptible (SIRDS) compartmental model with reinfections—with a machine learning-based predictive model—Prophet [257]—to forecast the future progression of infection, even in the presence of multiple virus strains. At its core, Sybil features an analytical model with dual functionality. In the first stage of Sybil's operation, this analytical model derives critical parameter values from the data. These parameter values, particularly the reproduction number over time, R_t, are then used as training data for Sybil's ML component. Based on this training, the ML component predicts future values for the key parameters, which are subsequently fed back into the analytical model. Using these future parameter values, Sybil calculates the future evolution of daily infections using the analytical model.

We extended Sybil by *i)* including a pre-processing step using a technique—i.e., splines [93], piece-wise-defined mathematical functions that use multiple polynomial segments to create a smooth and continuous curve—to fill missing data, and *ii)* adding the possibility to work with weekly data in case of unavailability of daily data—without significantly changing the information in the time series. In particular, we used an average number of days necessary to recover from the infection $\frac{1}{\gamma}$ equal to 14 days [65] and an average number of end-of-immunization days $\frac{1}{\nu}$ equal to 180 days [277]. In this study, we focus solely on extracting infection rates from COVID-19 data, rather than modeling variants or using Sybil to make predictions. As mentioned in Section 4.1.1, for each country considered, we used official surveillance data [121, 120] on the cumulative number of deaths and the total population, as well as estimates of the fraction of active cases [32, 231, 24, 234]—using a spline function to smooth these values.

Equation 4.2 shows the discretized version of the system of ordinary differential equations of the SIRDS compartmental model, with reinfections, used by Sybil. In this model, the population is divided into four categories: *i)* the Susceptible (S) population consists of healthy individuals who have not yet been infected by the disease, *ii)* the Infected (I) population comprises individuals who can transmit the disease to others, *iii)* the Recovered (R) population includes those who have recovered from the disease, and *iv)* the Deceased (D) population includes those who have died from it.

The model dynamics involve susceptible individuals becoming infected through contact

with infected people. Infected individuals either recover and move to the recovered group or die of the disease, moving to the deceased group. Recovered people become again susceptible after $\frac{1}{\nu}$ days on average. The model is governed by time-dependent rates—infection rates $\beta(\tilde{t})$, the recovery rates $\gamma(\tilde{t})$, and the fatality rates $\lambda(\tilde{t})$—meaning that they may vary at each time step, with the time step corresponding to one day, with the only exception of the end-of-immunization rate ν.

$$\begin{aligned} S(\tilde{t}+1) &= S(\tilde{t}) - \beta(\tilde{t})\frac{S(\tilde{t})I(\tilde{t})}{N} + \nu R(\tilde{t}) \\ I(\tilde{t}+1) &= I(\tilde{t}) + \beta(\tilde{t})\frac{S(\tilde{t})I(\tilde{t})}{N} - \gamma(\tilde{t})I(\tilde{t}) - \lambda(\tilde{t})I(\tilde{t}) \\ R(\tilde{t}+1) &= R(\tilde{t}) + \gamma(\tilde{t})I(\tilde{t}) - \nu R(\tilde{t}) \\ D(\tilde{t}+1) &= D(\tilde{t}) + \lambda(\tilde{t})I(\tilde{t}) \end{aligned} \tag{4.2}$$

We primarily rely on Facebook data from April 29$^{\text{th}}$, 2020 to May, 19$^{\text{th}}$, 2021 for analyzing cultural patterns and clustering, as we focus on the first year of COVID-19, which saw a greater number of government-imposed restrictions. Additionally, we use Google data from April 2020 to March 2022 for XGBoost regression models, incorporating mask usage data from Facebook to enhance our analyses. We considered the following European countries: Finland, Norway, Sweden, Belgium, Ireland, the United Kingdom, France, Germany, Austria, Italy, Greece, Spain, and Portugal.

4.1.3 Preliminary data analysis

Figure 4.1 illustrates the temporal evolution of each variable from April 2020 to March 2022, including Google mobility variables—*Workplaces*, *Stations*, *Recreation*, *Residential*, and *Grocery*—, Facebook's *Masks* usage, the complement of the OxCGRT *Stringency* index, and *Infection Rates* from Sybil. For each day, a boxplot aggregates values across the considered countries, showing the variability, trends, and seasonality over time. A clear and consistent correlation, whether positive or negative, can be observed across some of these variables. For example, a decrease in mobility at *Recreation* and *Stations* tends to correspond with a decrease in the complement of the *Stringency* index, reflecting the impact of stricter government-imposed restrictions aimed at reducing mobility. This suggests that as policies become more stringent, people are less likely to engage in activities that require them to travel or gather in public spaces. On the other hand, *Residential* mobility generally exhibits a negative correlation with these variables, indicating that as mobility restrictions increase, individuals tend to spend more time at home, possibly due to enforced lock-downs or heightened fear of exposure.

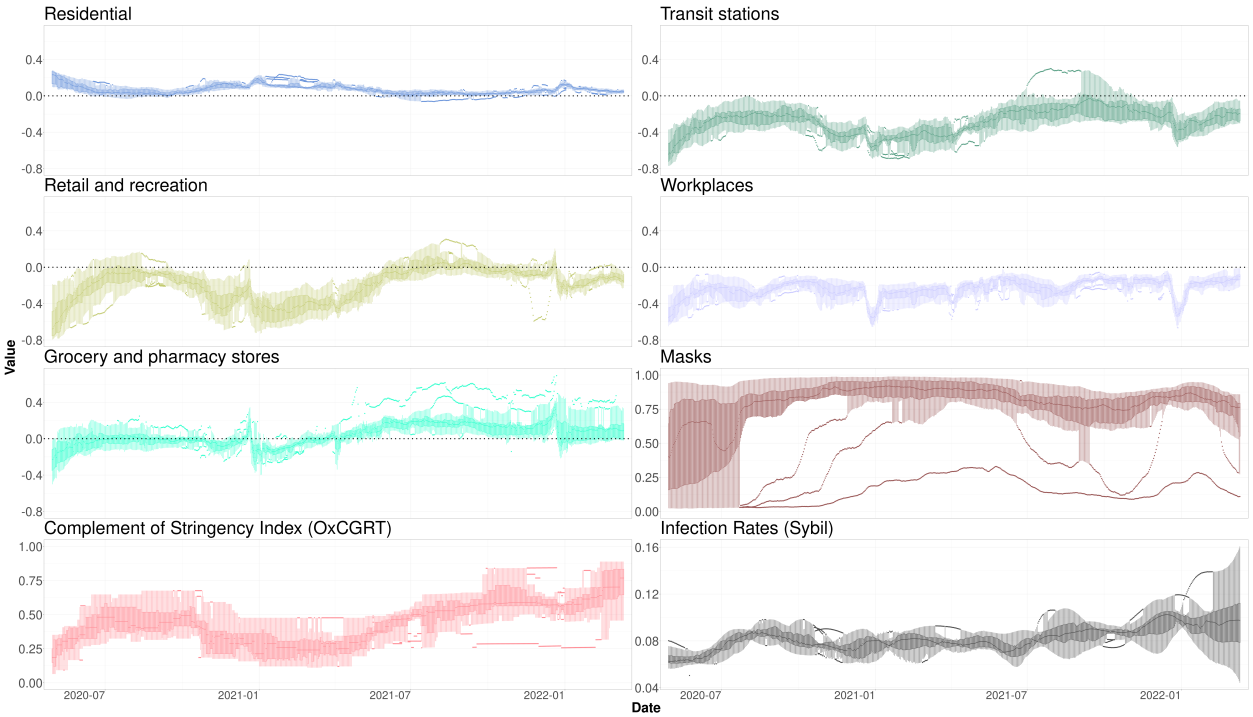


Figure 4.1: Temporal evolution of each variable—Google mobility variables (*Workplaces*, *Stations*, *Recreation*, *Residential*, and *Grocery*), the Facebook *Masks* variable, the complement of the OxCGRT *Stringency* index, and *Infection Rates* from Sybil—displayed as daily boxplots summarizing data across all considered countries from April 2020 to March 2022. For each variable, the median is displayed along with two bands: the darker band represents data from the first quartile (Q1) to the third quartile (Q3), while the lighter band includes data from $Q1 - 1.5 \cdot IQR$ to $Q3 + 1.5 \cdot IQR$, where IQR stay for inter-quartile range. In addition, outliers can be easily identified.

Similarly, the complement of the *Stringency* index appears to follow changes in *Infection Rates*, suggesting that reductions in NPIs may align with increases in *Infection Rates*, whereas stricter measures may coincide with a decline in *Infection Rates*. Lastly, *Grocery* and *Workplaces* mobilities appear to show weaker correlations with the other variables, indicating that changes in these particular types of mobility are less directly influenced by the complement of the *Stringency* index or *Infection Rates*. This may be due to the essential nature of grocery shopping and workplace attendance—strongly influenced by seasonality—, which might be less affected by the broader social distancing measures.

To address the first point outlined in the introductory section of this chapter, we utilized the Spearman [249] correlation index. Table 4.1 shows the correlation between the considered variables and averaged over all the selected countries. The complement of the *Stringency* index has moderately high correlations with most mobility variables, indicating that stricter policies tend to decrease people's mobility. Generally, there is quite a high correlation among the mobility variables—including the Facebook *Masks* variable, except for *Grocery* and *Workplaces* where the correlation is lower. *Infection rates* have weaker correlations with most variables, showing a slightly higher correlation with the complement of the *Stringency*

	Compl. of stringency index	*Grocery*	*Recreation*	*Workplaces*	*Stations*	*Residential*	*Masks*	*Inf. rates*
Compl. of stringency index	1	0.52	0.69	0.52	0.68	-0.64	-0.52	0.48
Grocery	0.52	1	0.80	0.50	0.76	-0.69	-0.33	0.31
Recreation	0.69	0.80	1	0.48	0.91	-0.89	-0.54	0.39
Workplaces	0.52	0.50	0.48	1	0.61	-0.60	-0.26	0.28
Stations	0.68	0.76	0.91	0.61	1	-0.92	-0.59	0.37
Residential	-0.64	-0.69	-0.89	-0.60	-0.92	1	0.54	-0.32
Masks	-0.52	-0.33	-0.54	-0.26	-0.59	0.54	1	-0.12
Inf. rates	0.48	0.31	0.39	0.28	0.37	-0.32	-0.12	1

Table 4.1: Spearman [249] correlation among selected variables—including Google mobility indicators (*Workplaces*, *Stations*, *Recreation*, *Residential*, and *Grocery*), the Facebook *Masks* variable, the complement of the OxCGRT *Stringency* index, and *Infection Rates* from Sybil—averaged across all considered countries.

index (0.48). This suggests that fluctuations in *Infection Rates* correspond to changes in the complement of the *Stringency* index and are affected to a lesser extent by variations in mobility. The correlation is not particularly strong due to the delay between an increase in mobility and the subsequent rise in *Infection Rates*, as well as the lag between stricter restrictions and a decline in *Infection Rates*. In contrast, this delay is absent among mobility variables, and the lag between mobility and the complement of the *Stringency* index remains minimal.

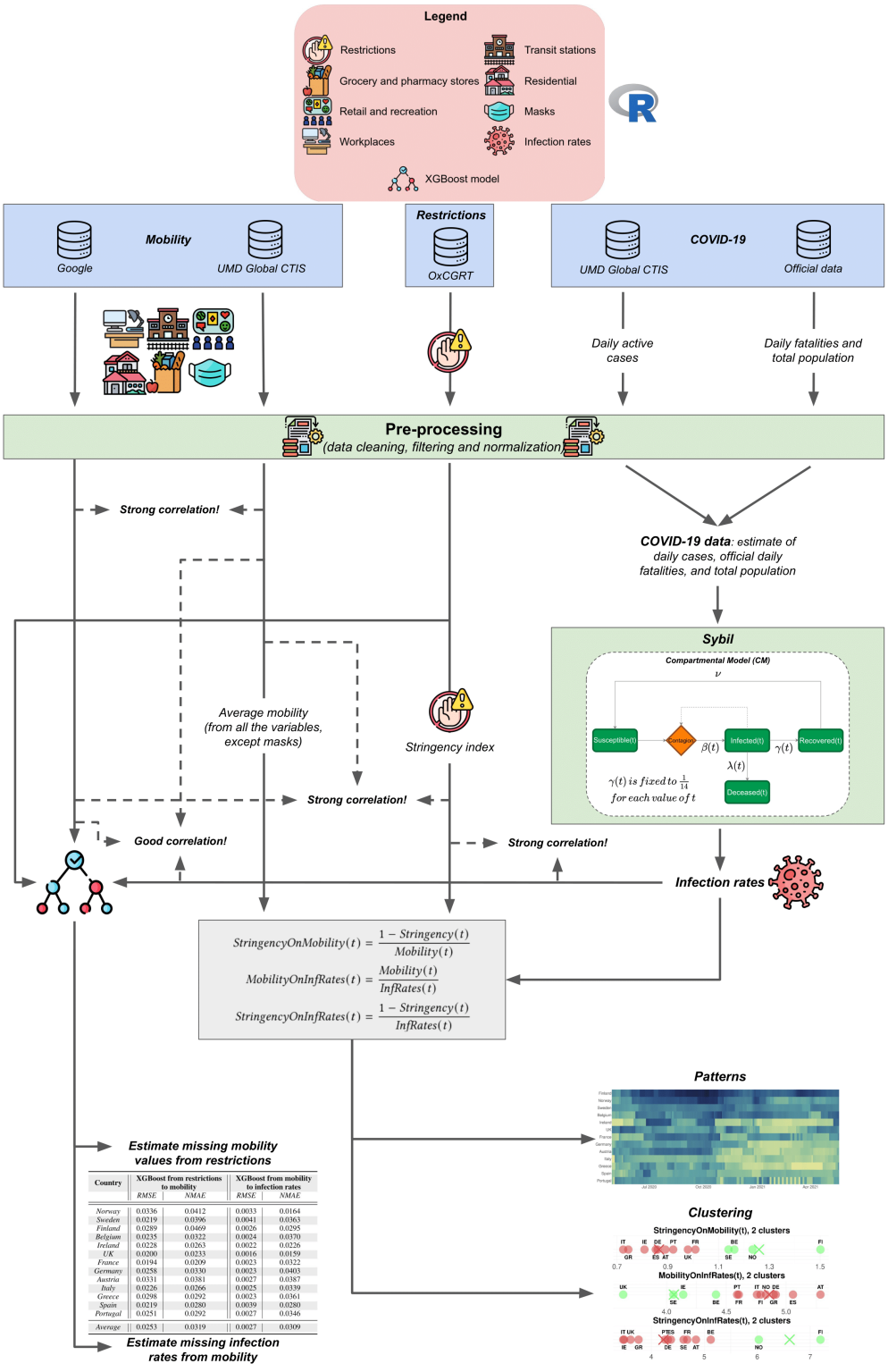
Finally, we computed the Spearman correlation index among Google and Facebook mobility variables, averaged across all chosen countries, in the period from April 29$^{\text{th}}$, 2020 to May, 19$^{\text{th}}$, 2021—we could not consider the whole period due to questions change (refer to Section 4.1.1 for more details). In particular, we obtained the following values, concluding that the correlation is high, especially for *Recreation*, *Residential*, and *Stations*: 0.70 for *Grocery*, 0.95 for *Recreation*, 0.75 for *Workplaces*, 0.83 for *Stations*, and -0.83 for *Residential*—data on mask usage is only available in the Facebook dataset.

4.2 Our approach

Figure 4.2 presents a schematic overview of our approach. Starting with the datasets described in Section 4.1.1, we *pre-processed the data* (Section 4.1.2) to generate cleaned, filtered, and normalized mobility variables: *Workplaces*, *Stations*, *Recreation*, *Residential*, and *Grocery* (sourced from both Google and Facebook) and *Masks* (from Facebook). Additionally, we *incorporated restriction indices* and *Infection Rates* derived using Sybil [39]. Next, we *calculated three key ratios*, detailed later, based on average Facebook mobility (excluding masks), the stringency index, and infection rates. These ratios allowed us to identify cultural patterns by examining how official restrictions and self-imposed limitations (in the absence of formal regulations) influenced mobility and R_t across the studied countries (Section 4.3.1). Next, we *applied clustering techniques*, specifically K-means [125] and CONNECTOR [203], to analyze differences in social compliance with restrictions between Northern and Southern European countries, also highlighting no disparities in social interactions (Section 4.3.2). Finally, we *employed XGBoost regression models* [77] to estimate missing mobility data from restriction indicators and infer missing R_t values based on mobility patterns (Section 4.3.3).

XGBoost [77] is a powerful machine learning algorithm known for its high accuracy, particularly in classification and regression tasks. It efficiently handles missing data and uses regularization (L1 and L2) to reduce overfitting, making it robust in complex scenarios. Additionally, it is fast and scalable, leveraging parallel processing to handle large datasets, and provides clear feature importance for interpretability. However, XGBoost can be computationally intensive, requiring significant memory and processing power for large datasets. It also needs careful hyperparameter tuning for optimal performance and may overfit noisy data if not properly regularized. Despite its complexity, it offers strong results when properly configured. XGBoost is an advanced implementation of the gradient boosting algorithm designed for efficiency and performance, which builds an ensemble of decision trees sequentially to optimize predictive accuracy. By focusing on minimizing errors of previous trees through gradient descent and incorporating regularization techniques to prevent over-fitting, XGBoost achieves high accuracy and scalability.

On the other hand, K-means [125] is a popular clustering (unsupervised machine learning) algorithm to cluster similar data points into a predefined number of clusters. It is simple, fast, and efficient, especially for large datasets. The algorithm works by iteratively assigning data points to the nearest cluster center and then recalculating the centers until convergence. However, K-means has some limitations. It assumes clusters are spherical and evenly sized, which may not always be the case in real-world data. It is also sensitive to the initial selection of cluster centers, which can affect the final result. Additionally, K-means struggle with



Country	XGBoost from restrictions to mobility		XGBoost from mobility to infection rates	
	RMSE	NMAE	RMSE	NMAE
Norway	0.0336	0.0412	0.0033	0.0164
Sweden	0.0219	0.0396	0.0041	0.0363
Finland	0.0289	0.0469	0.0026	0.0295
Belgium	0.0235	0.0322	0.0024	0.0370
Ireland	0.0228	0.0263	0.0022	0.0226
UK	0.0200	0.0233	0.0016	0.0159
France	0.0194	0.0209	0.0023	0.0322
Germany	0.0258	0.0330	0.0023	0.0403
Austria	0.0331	0.0381	0.0027	0.0387
Italy	0.0226	0.0266	0.0025	0.0339
Greece	0.0298	0.0292	0.0023	0.0361
Spain	0.0219	0.0280	0.0039	0.0280
Portugal	0.0251	0.0292	0.0027	0.0346
Average	0.0253	0.0319	0.0027	0.0309

Figure 4.2: Schematic representation of our study.

identifying complex or non-convex cluster shapes and can be influenced by outliers. To assess the quality of clustering, the Silhouette score [239] is often used. This score measures how

similar a data point is to its own cluster compared to other clusters, with a range from -1 to 1. A higher score indicates better-defined clusters, while values near 0 or negative suggest poor clustering or potential misclassification.

Finally, CONNECTOR [203] is an R package for unsupervised analysis of longitudinal data, using a model-based approach to cluster functional data, particularly effective for sparse and irregular curves. It employs spline functions to create smooth curves for accurate modeling and clustering. The package provides multiple visual inspections of the data, specific indexes, and graphics to set the two model's free parameters (i.e., the dimension of the spline basis vector and the optimal number of clusters). Finally, different visualization plots might be exploited to show the longitudinal clustering result and its cluster stability.

Concretely, we used Facebook mobility data in the first period—from April 29$^{\text{th}}$, 2020 to May, 19$^{\text{th}}$, 2021—to address the second and third points outlined in the introductory section of this chapter. We selected this initial period because the pandemic's onset saw stricter restrictions in place, and individuals were generally more compliant with these measures. Additionally, we computed the following ratios for each country to identify cultural patterns:

$$\begin{aligned} StringencyOnMobility(t) &= \frac{1 - StrIdx(t)}{AvgMob(t)} \\ MobilityOnInfRates(t) &= \frac{AvgMob(t)}{InfRates(t)} \\ StringencyOnInfRates(t) &= \frac{1 - StrIdx(t)}{InfRates(t)} \\ &= \frac{1 - StrIdx(t)}{AvgMob(t)} \cdot \frac{AvgMob(t)}{InfRates(t)} \\ &= StringencyOnMobility(t) \cdot MobilityOnInfRates(t) \end{aligned} \tag{4.3}$$

where *AvgMob(t)* is the average mobility computed on the considered variables—*Workplaces*, *Stations*, *Recreation*, *Residential*, and *Grocery*—, without considering *Masks*, at time *t*, *InfRates(t)* is the *Infection Rates* at time *t*, and *StrIdx(t)* is the *Stringency* index at time *t*. *AvgMob(t)* and *InfRates(t)* are always greater than 0. Based on these time series, we identified patterns and applied clustering techniques using K-means [125] and CONNECTOR [203].

Specifically, the *StringencyOnMobility(t)* ratio emphasizes the relationship between restrictions and mobility, primarily reflecting aspects of social compliance with restrictions. The *MobilityOnInfRates(t)* ratio illustrates the connection between mobility and infection rates, mainly capturing factors related to social interactions. Finally, the *StringencyOnInfRates(t)* ratio highlights the relationship between restrictions and infection rates, influenced by both social compliance with restrictions and social interactions.

To address the last point outlined in the introductory section of this chapter, we employed XGBoost regression models [77]. Specifically, we used 5-fold cross-validation (80% training and 20% testing) to train XGBoost models for each country. This approach can be viewed as estimating missing data, where the 20% of discarded data represents unobserved values used for model testing, while the models aim to estimate the missing values. These models were designed to estimate mobility variables—namely, *Workplaces*, *Stations*, *Recreation*, *Residential*, and *Grocery* (from Google), as well as *Masks* (from Facebook)—at time t, using the whole restriction indices from OxCGRT at the same time point. Additionally, we trained a separate XGBoost model to estimate *Infection Rates* (extracted from Sybil) at time t, using the mobility variables from that time.

4.3 Results

Here, we present the key results obtained using the methods outlined in Section 4.2. Specifically, Sections 4.3.1 and 4.3.2 detail the patterns and clusters identified by analyzing the ratios introduced in Section 4.2, while in Section 4.3.3, we discuss the results of the XGBoost regression models.

4.3.1 Patterns

Figure 4.3 illustrates the ratios calculated using Equation 4.3 for each considered country. Figure 4.3-*a)*, showing the *StringencyOnMobility(t)* ratio, highlights the relationship between restrictions and mobility. Higher values of this ratio indicate that minimal restrictions (low stringency) are sufficient to limit mobility or keep people at home. This often reflects the influence of external factors such as fear of infection or strong social compliance with NPIs. As a result, fewer interventions are needed to achieve the same mobility reduction. Notably, countries like Norway, Sweden, Finland, and Belgium exhibit consistently high values, visible as dark bands in the heatmap.

Figure 4.3-*b)*, which illustrates the *MobilityOnInfRates(t)* ratio, illustrates the relationship between mobility and infection rates. Higher values indicate that average mobility is increasing faster than the infection rate, or that the infection rate is decreasing relative to mobility. For example, in June and July 2020, most countries showed higher values due to increased mobility and lower infection rates following the first pandemic wave. However, the United Kingdom was an exception, with mobility remaining lower while infection rates began to rise again. Two other notable points occurred around December 2020 (Christmas) and April 2021 (Easter), when seasonality played a crucial role. In December 2020, a sharp decline in mobility caused the ratio to drop—except in Greece, where the infection peak occurred later and the decline in mobility was slower. After this decline, the ratio increased as mobility gradually rose, though more slowly in the United Kingdom. In April 2021, a similar pattern emerged, but with a faster rise in mobility. In general, Figure 4.3-*b)* reveals no significant distinctions among countries in the impact of mobility on infection rates, with few exceptions such as the United Kingdom, Ireland, and Sweden where lighter bands are noticeable, particularly after Christmas 2020. This suggests a broadly similar relationship across populations. The lack of clear differentiation indicates that, despite regional and cultural variations, mobility patterns and their influence on infection rates were largely consistent during the observed period, driven primarily by social interactions.

Figure 4.3-*c)* shows the *StringencyOnInfRates(t)* ratio, which highlights the relationship

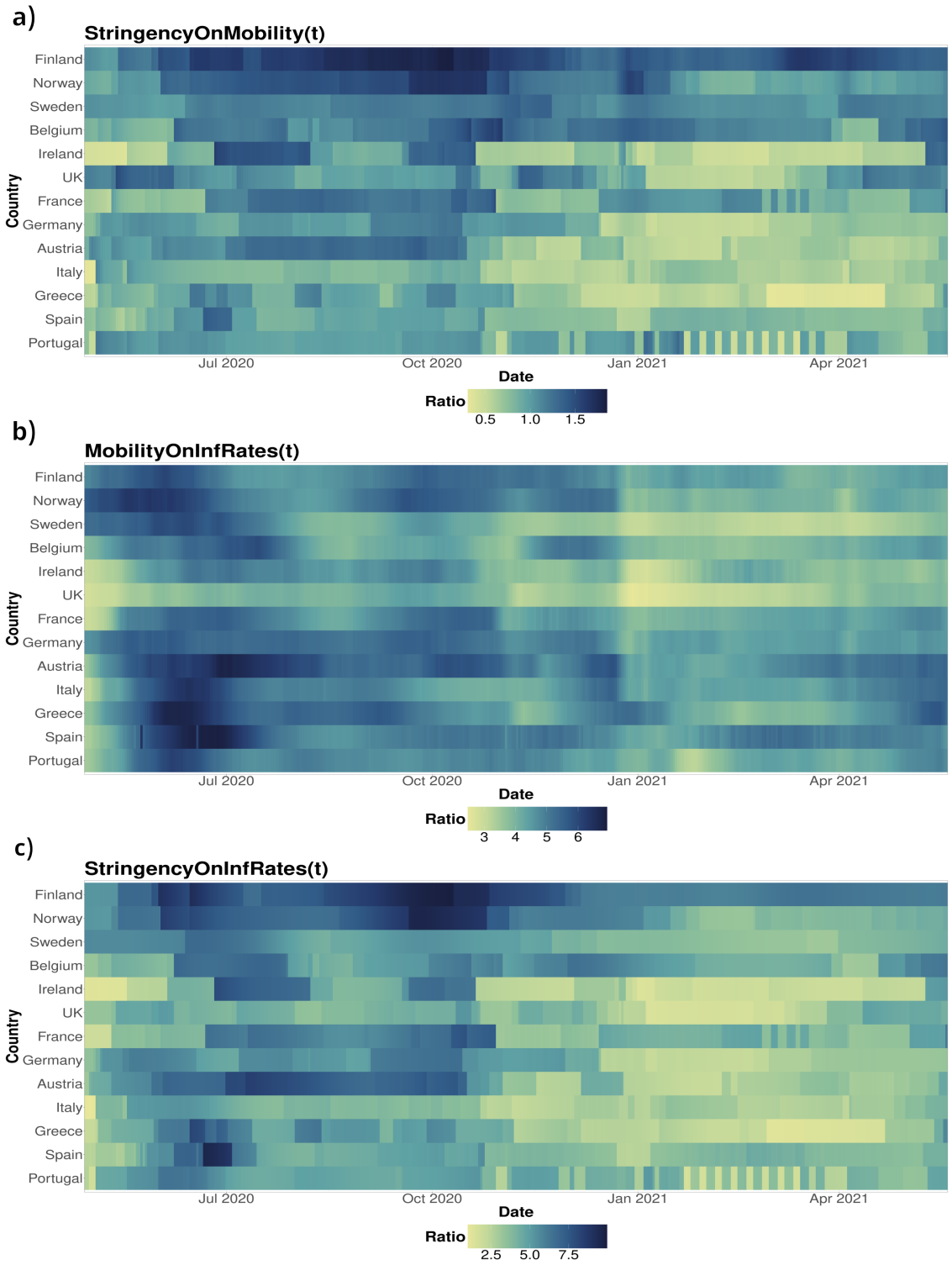


Figure 4.3: Heatmaps display country ratios: *StringencyOnMobility*, *MobilityOnInfRates*, and *StringencyOnInfRates*, with darker colors for higher values.

between restrictions and infection rates. Higher values of this ratio signify that the complement of the stringency index is high (i.e., restrictions are minimal) while infection rates remain

relatively low. This suggests scenarios where either limited restrictions are sufficient to manage the infection rate or where infection rates are inherently low, resulting in a higher ratio. In this case, multiple factors come into play, such as effective natural mitigation mechanisms, high societal compliance, the nature of social interactions, and other external influences that collectively help control the spread of infections despite minimal interventions. Additionally, Finland exhibits different patterns starting from the end of 2020, largely because restrictions remained unchanged for several months.

4.3.2 Clustering

In this section, we cluster countries based on the ratios defined in Section 4.2, employing two different clustering techniques: K-means [125] and CONNECTOR [203].

K-means

In this section, we apply K-means clustering [125] to group countries based on the three ratios described in Section 4.2. Specifically, we perform clustering using the average values of these ratios computed over the entire period for each country. Additionally, we explore clustering by incorporating both the average values and their corresponding standard deviations to assess the impact of variability.

Figure 4.4 shows the average Silhouette score [239], computed using K-means and evaluated with both the average alone and the combination of the average and standard deviation, for the three considered ratios. Specifically, the optimal number of clusters appears to be two in all cases. However, when considering the average, three clusters for the $StringencyOnInfRates$ ratio also seem to be a reasonable choice, while five clusters are appropriate for the $MobilityOnInfRates$ and $StringencyOnMobility$ ratios. Additionally, when considering both the average and standard deviation, three clusters also appear to be a reasonable choice for the $StringencyOnMobility$ ratio and six for the $MobilityOnInfRates$ and $StringencyOnMobility$ ratios.

Figure 4.5 illustrates the clusters derived using only the average values. Focusing on the $StringencyOnMobility$ ratio, the two plots in the first row show the results for two and five clusters, respectively. These clusters reveal a clear separation between Northern and Southern European countries. Notably, in both cases, the three analyzed Northern European countries—Finland, Norway, and Sweden—form a distinct cluster, separate from the other nations. When considering five clusters, Finland is placed in its own separate cluster. Additionally, in both cluster configurations, Belgium is grouped with the Northern countries, likely due to the fact that fewer restrictions were implemented there as well.

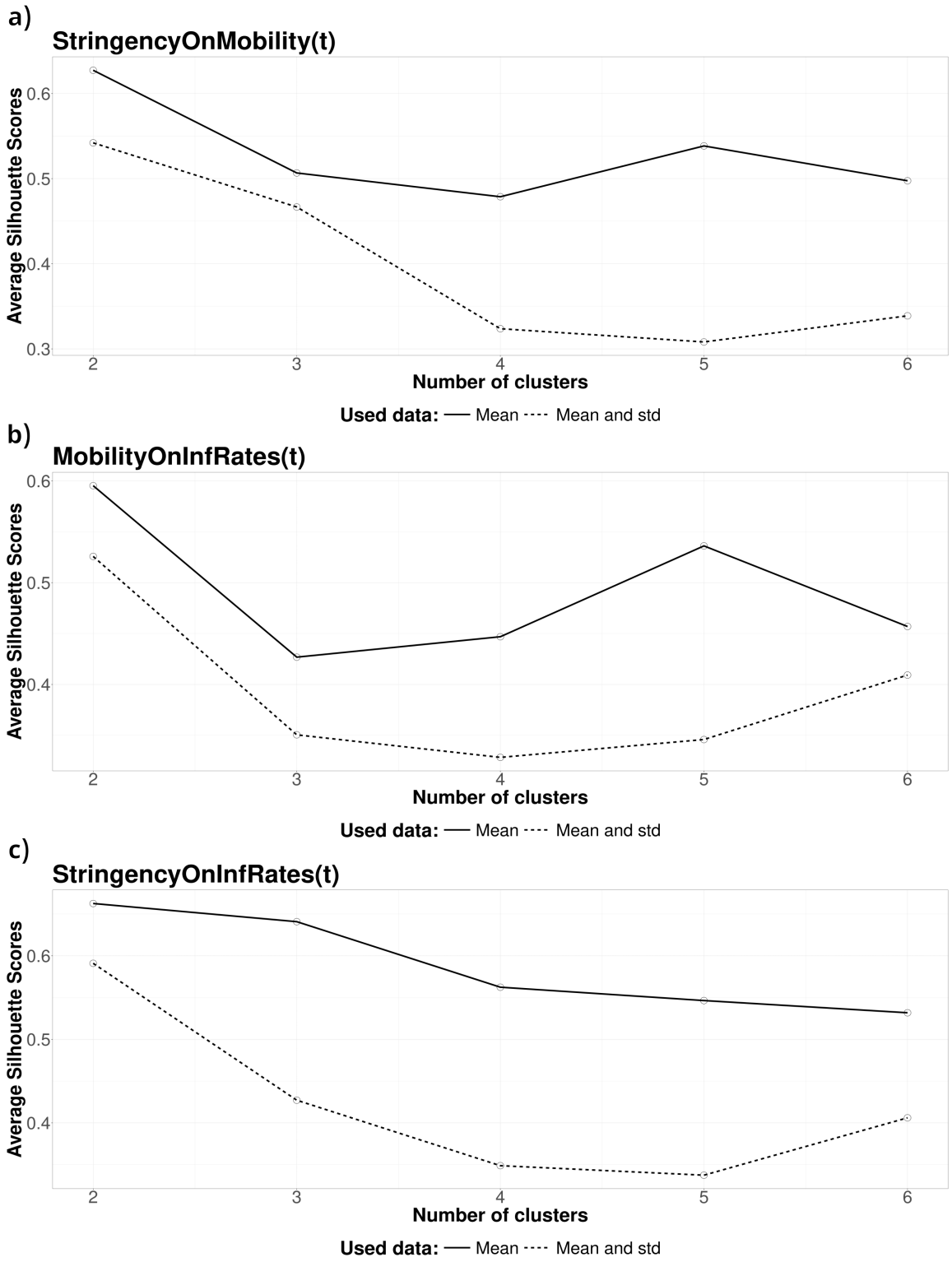


Figure 4.4: Silhouette score [239] computed using K-means—evaluated both with the average alone and with the combination of the average and standard deviation—for the three considered ratios: $MobilityOnInfRates$, $StringencyOnInfRates$, and $StringencyOnMobility$.

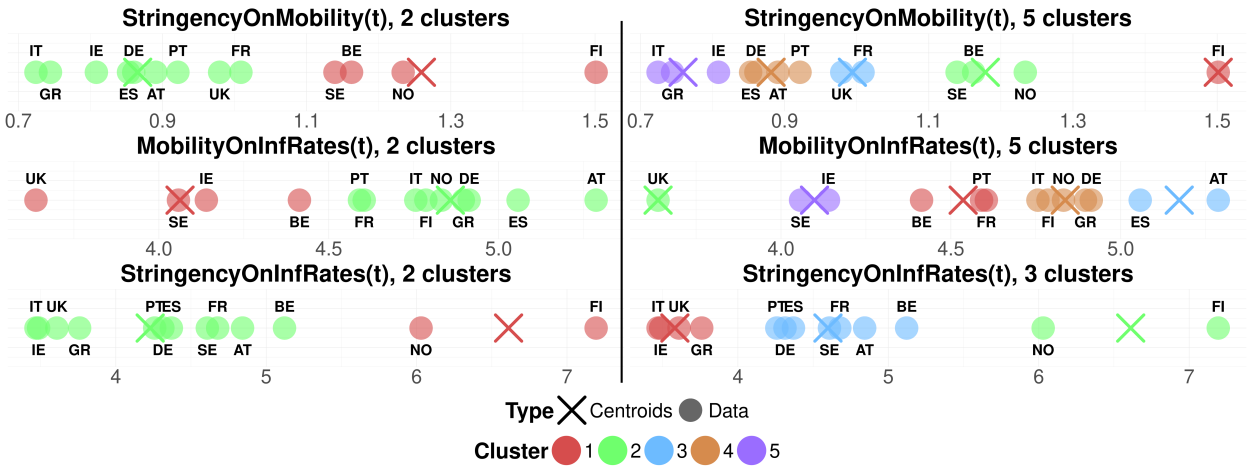


Figure 4.5: Clusters derived using the average values for the *StringencyOnMobility*, *MobilityOnInfRates*, and *StringencyOnInfRates* ratios with K-means [125].

As noted in Section 4.3.1, in these countries, less stringent measures were effective in controlling infections, largely due to high levels of social compliance. In Northern European countries, cultural norms, behavioral tendencies, and strong trust in public health guidance reduce the need for coercive measures, as people are more likely to voluntarily adhere to recommendations. Moreover, we observe that countries like Italy, Greece, Spain, and Portugal—typically considered Southern European countries—cluster more to the left, alongside several intermediate European countries like Ireland, Germany, France, Austria, and the United Kingdom. This suggests that these countries needed more stringent restrictions to encourage people to stay at home. The three clusters on the left in the second plot reflect progressively increasing levels of restrictions, from right to left.

Considering the *MobilityOnInfRates* ratio—the two plots in the second row in Figure 4.5, which show two and five clusters, respectively—we observe no clear distinction between Northern and Southern countries. Instead, countries appear mixed, as noted in Section 4.3.1. This can be attributed to the fact that, when examining the effect of mobility on infection rates, we are primarily considering the role of social interactions instead of social compliance. Unlike measures that target mobility alone, the influence of social interactions encompasses a wide range of factors, including cultural norms and community behaviors. In countries with higher social interactions, mobility may have less of an impact on infection rates, while in others, even small changes in mobility can lead to significant fluctuations in infections. This mix of factors contributes to the lack of a clear distinction and highlights the complexity of how mobility and social behaviors interact in different contexts.

Finally, when examining the *StringencyOnInfRates* ratio—the two plots in the last row in Figure 4.5, which show two and three clusters, respectively—we observe that Finland

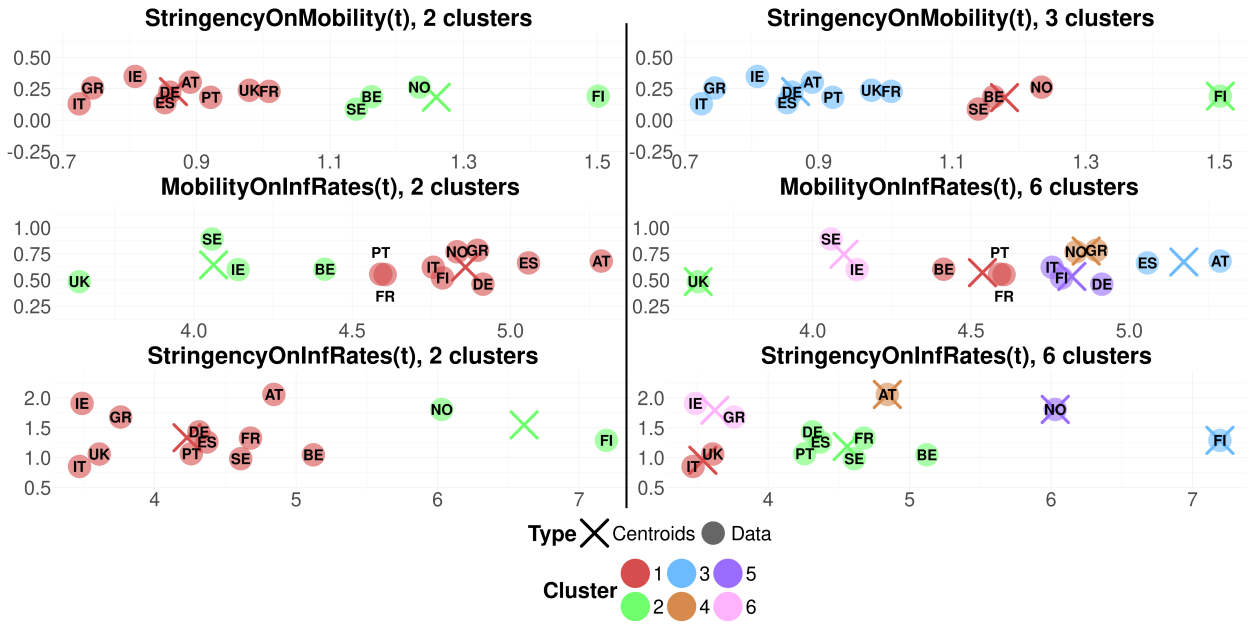


Figure 4.6: Clusters derived using both average values and standard deviations for the *StringencyOnMobility*, *MobilityOnInfRates*, and *StringencyOnInfRates* ratios with K-means [125]. Colors may differ with respect to Figure 4.5.

and Norway remain distinctly separated, while Sweden does not. In this case, both social compliance and social interactions are considered, as we are directly linking restrictions to infection rates. This suggests that the clusters observed here are more similar to those seen in the first and second plots than to those in the third and fourth plots, as is indeed the case.

Figure 4.6 illustrates the clusters derived using both the average values and the standard deviation. These clusters align closely with those shown in Figure 4.5. For the *StringencyOnMobility* ratio with three clusters, a clear separation is observed among Finland and Norway, Belgium, and Sweden, consistent with the pattern noted in the second plot of Figure 4.5. For the *MobilityOnInfRates* ratio with six clusters, we observed the same general separation as with five clusters when considering only the average; however, the cluster containing Norway, Greece, Germany, Italy, and Finland is further split into two distinct clusters. Similarly, for the *StringencyOnMobility* ratio with six clusters, Finland, Norway, and Austria form three separate singleton clusters. Meanwhile, Ireland and Greece are grouped into a fourth cluster, Italy and the United Kingdom constitute a fifth cluster, and all remaining countries are included in the final cluster.

CONNECTOR

In this section, we cluster countries using an alternative clustering technique, CONNECTOR [203]. Figure 4.7 shows the dimension of the spline basis vector, determined by the *cross-log-likelihood*—see [203] for more details. This dimension can be chosen as the one

corresponding to the largest cross-validated likelihood, as proposed in [139]. In particular, we chose 8 for the $StringencyOnMobility$ ratio and 11 for the $MobilityOnInfRates$ and $StringencyOnInfRates$ ratios.

Figure 4.8 illustrates the cluster separation measure, known as fDB, as defined in [203]. This measure is an adaptation of the well-known Davies–Bouldin (DB) index [92] for functional data. The optimal cluster dimension is selected by maximizing this value. Additionally, the figure presents the total tightness, a metric used to quantify dispersion within clusters. As the number of clusters increases, total tightness decreases, approaching zero when the number of fitted clusters matches the number of sampled curves. For this analysis, we select two, five, and six clusters for each ratio.

Figure 4.9 illustrates the clustering results for the $StringencyOnMobility$ ratio, with configurations of two, five, and six clusters. For the two-cluster configuration, the results replicate the separation between Northern and Southern countries achieved using K-means. Specifically, Finland, Norway, Sweden, and Belgium form one cluster, while the remaining countries are grouped into the other cluster. In the five-cluster configuration, Finland forms its own cluster, while the United Kingdom constitutes another separate cluster. Italy, Spain, and Portugal are grouped into a third cluster. Norway, Sweden, Belgium, and France are assigned to a fourth cluster, while Ireland, Germany, Austria, and Greece comprise the fifth cluster. In this case, while the clusters differ slightly from those obtained using K-means, the separation between Northern and Southern countries remains evident. Finally, in the six-cluster configuration, the fourth cluster is divided into two subgroups: Germany and Austria form one cluster, while Ireland and Greece constitute the other.

Figure 4.10 illustrates the clustering results for the $MobilityOnInfRates$ ratio with two, five, and six cluster configurations. In the two-cluster configuration, Sweden, the United Kingdom, and Ireland are grouped in the first cluster, while the remaining countries form the second cluster. This clustering yields similar results to those obtained using K-means, where Belgium was also included in the first cluster. In the five-cluster configuration, Sweden is placed in the first cluster, Finland, Norway, and Germany in the second, Ireland and the United Kingdom in the third, Austria and Spain in the fourth, and Italy, Greece, Belgium, France, and Portugal in the fifth. In this case, these clusters differ slightly from the ones obtained using K-means. Finally, for the six-cluster configuration, the fifth cluster is split into two subgroups: Italy and Greece form one cluster, while Belgium, France, and Portugal constitute the other. These six clusters are slightly different with respect to the ones obtained using K-means considering the average and the standard deviation in the previous section.

Figure 4.11 illustrates the clustering results for the $StringencyOnInfRates$ ratio, with configurations of two, five, and six clusters. In the two-cluster configuration, Finland and

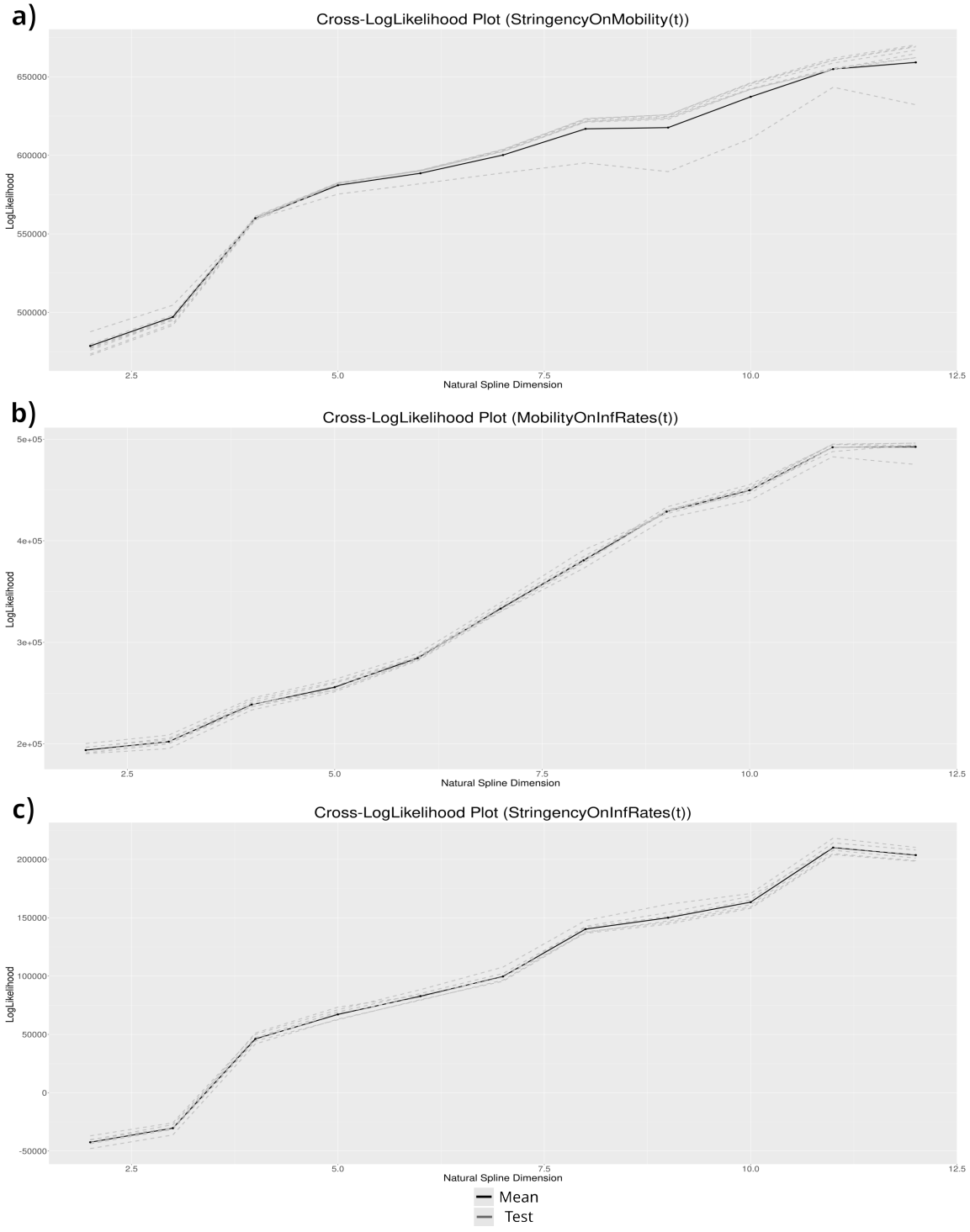


Figure 4.7: The dimension of the spline basis vector, determined by the *cross-log-likelihood* [203], is shown for each ratio.

Norway form one cluster, while all other countries belong to the second cluster. As with this ratio, we observe the same separation as obtained using K-means. In the five-cluster configuration, Finland and Norway are grouped in one cluster, Ireland forms the second

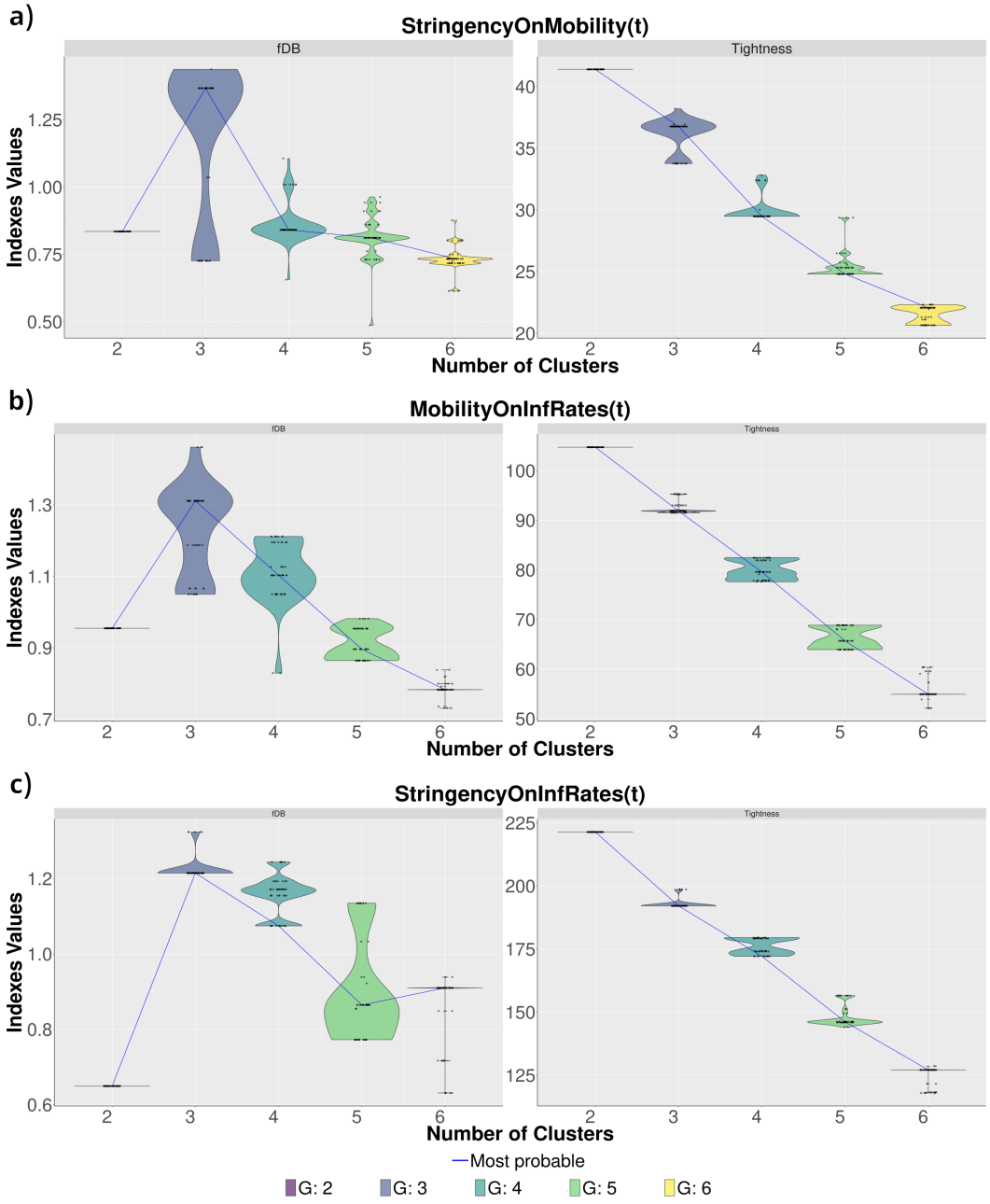


Figure 4.8: Cluster separation measure (fDB) and total tightness [203] are shown for each ratio. The optimal number of clusters is chosen to minimize the fDB while balancing total tightness.

cluster, Austria is in the third, Sweden, Belgium, France, and Spain are in the fourth, and the United Kingdom, Germany, Italy, Greece, and Portugal are placed in the fifth cluster. Finally, for the six-cluster configuration, Finland and Norway are grouped in one cluster, Ireland forms the second cluster, Austria is in the third, Belgium and France are in the fourth, the United Kingdom, Italy, and Greece are in the fifth, and Sweden, Germany, Spain, and Portugal are in the sixth. In this case, as well, these six clusters are slightly different with

respect to the ones obtained using K-means.

To summarize, in Section 4.3.1 we identified patterns analyzing the interplay among restrictions, mobility, and infection rates, while in Section 4.3.2 we analyzed country clustering based on three key ratios using K-means [125], considering both average values and their standard deviations, and CONNECTOR [203]. Northern European countries, like Finland, Norway, and Sweden formed distinct clusters when social compliance plays a crucial role in the ratios, reflecting high compliance and effective infection control with fewer restrictions. Southern and intermediate European countries required stricter measures, highlighting differences in compliance levels. The $MobilityOnInfRates$ ratio showed mixed clustering, driven by social interactions rather than compliance. The $StringencyOnInfRates$ ratio, linking restrictions to infection rates, revealed clusters similar to those based on mobility ratios, emphasizing the combined role of compliance and interactions. These findings underscore the importance of tailored strategies in managing public health measures, potentially by grouping countries into clusters, as we have identified.

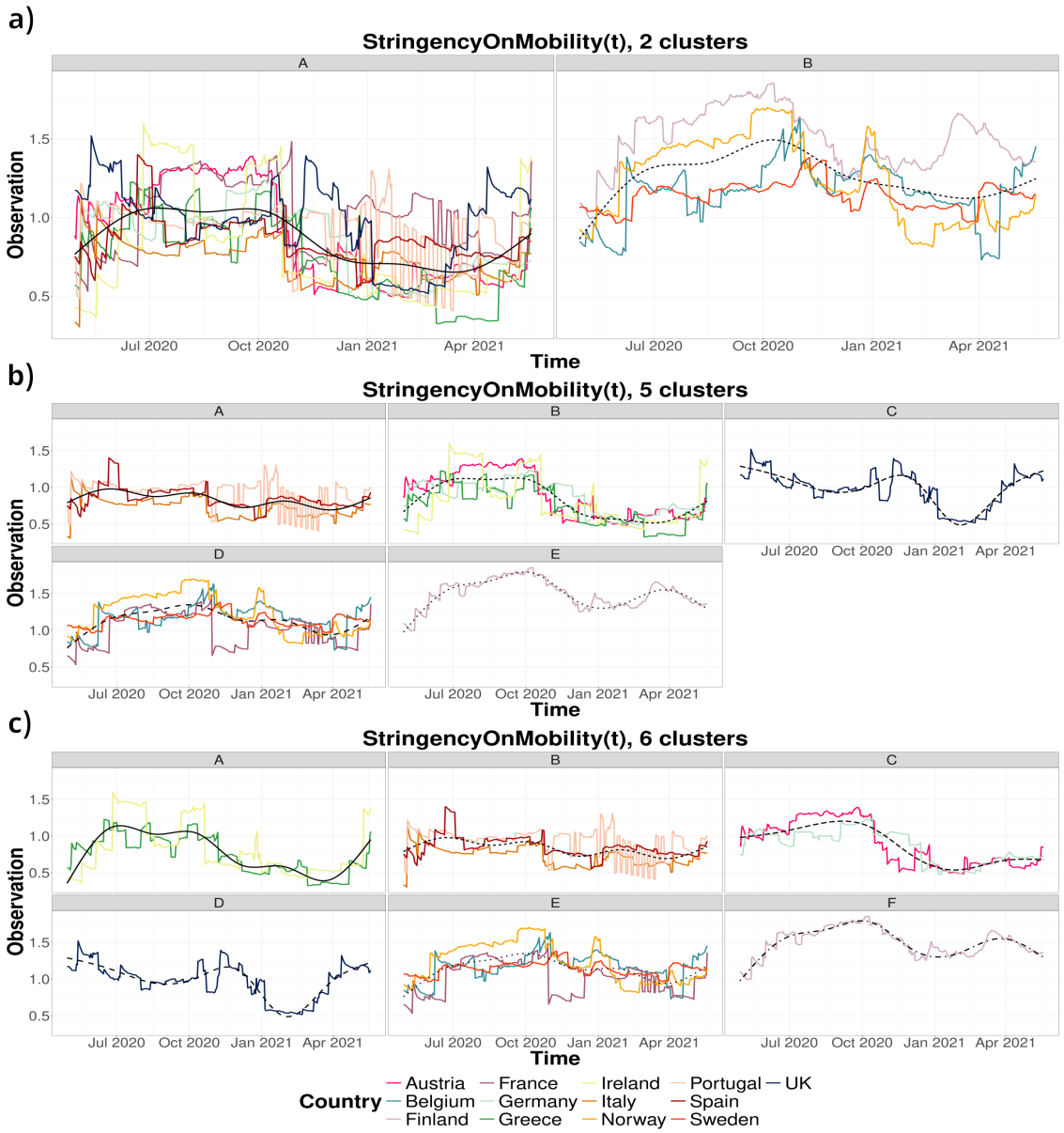


Figure 4.9: Two, five, and six clusters for the *StringencyOnMobility* ratio.

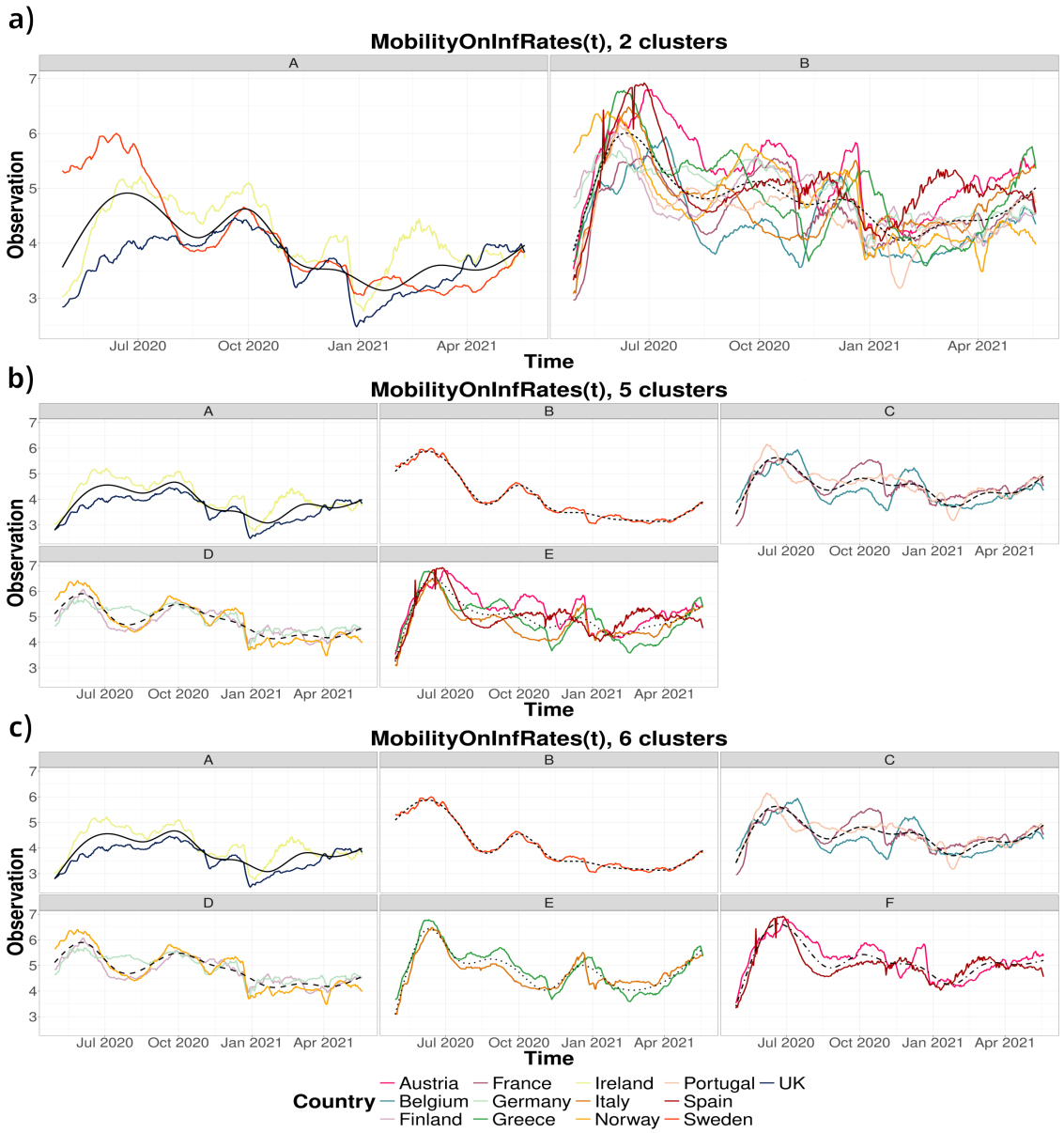


Figure 4.10: Two, five, and six clusters for the $MobilityOnInfRates$ ratio.

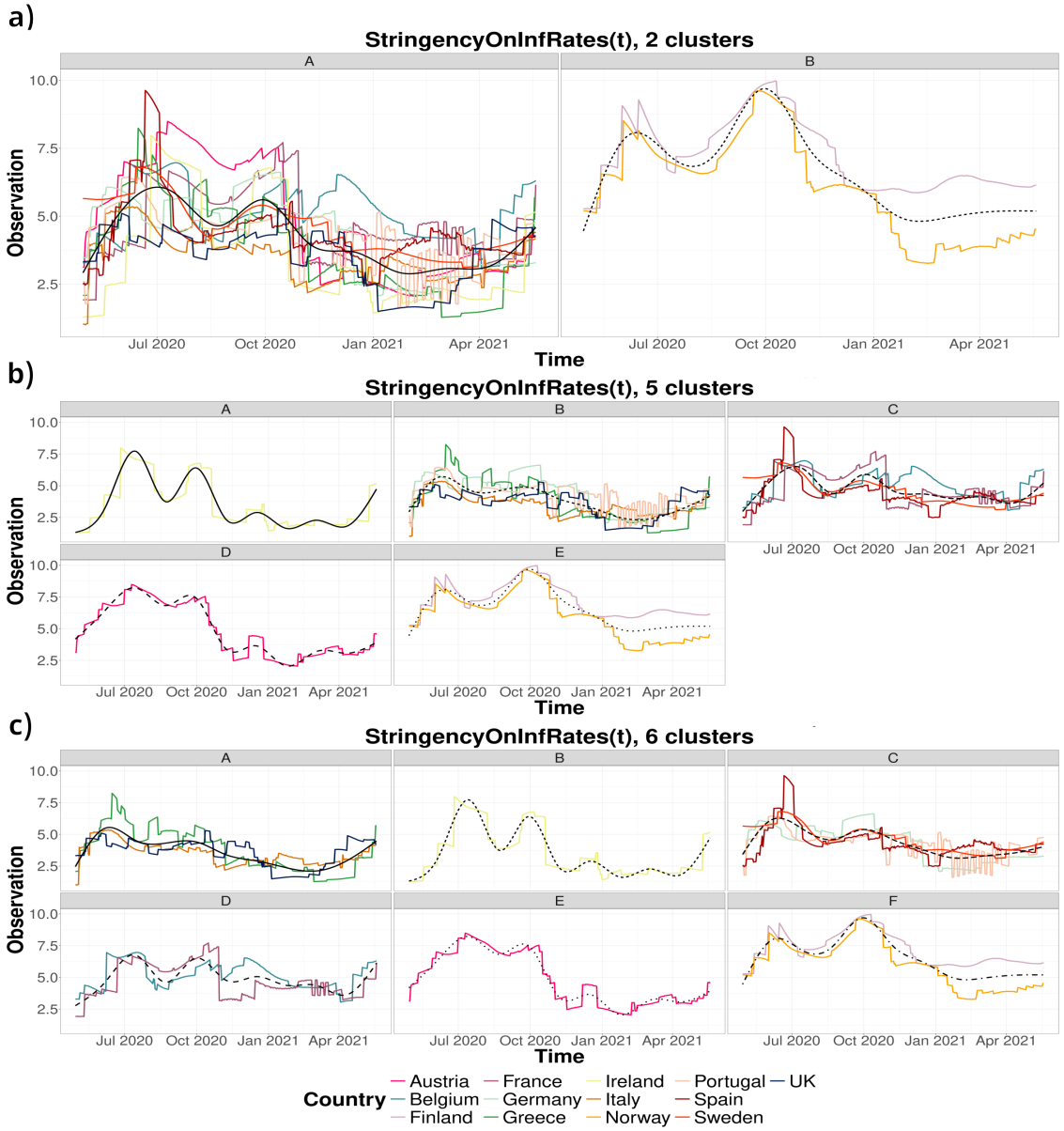


Figure 4.11: Two, five, and six clusters for the *StringencyOnInfRates* ratio.

4.3.3 Regression Models

Table 4.2 presents the average root mean squared error (RMSE) and normalized mean absolute error (NMAE) between the test set (ground truth, 20% of the data) and the corresponding estimates, averaged across the five considered folds, for the XGBoost [77] regression models. These models were used to estimate missing mobility data based on restrictions—we averaged the errors over the six mobility variables—and missing R_t data based on mobility patterns for each country. The reported errors highlight the models' strong performance in accurately estimating missing values.

The estimated values can be used to fill in gaps in time series data, allowing us to uncover patterns, simulate various scenarios, or test hypotheses regarding the relationships between variables. Furthermore, the completed time series can be utilized with other models, such as neural networks for forecasting, decision trees for classification, or reinforcement learning models for optimizing policy interventions.

Country	**XGBoost from restrictions to mobility**		**XGBoost from mobility to infection rates**	
	RMSE	*NMAE*	*RMSE*	*NMAE*
Norway	0.0336	0.0412	0.0033	0.0164
Sweden	0.0219	0.0396	0.0041	0.0363
Finland	0.0289	0.0469	0.0026	0.0295
Belgium	0.0235	0.0322	0.0024	0.0370
Ireland	0.0228	0.0263	0.0022	0.0226
United Kingdom	0.0200	0.0233	0.0016	0.0159
France	0.0194	0.0209	0.0023	0.0322
Germany	0.0258	0.0330	0.0023	0.0403
Austria	0.0331	0.0381	0.0027	0.0387
Italy	0.0226	0.0266	0.0025	0.0339
Greece	0.0298	0.0292	0.0023	0.0361
Spain	0.0219	0.0280	0.0039	0.0280
Portugal	0.0251	0.0292	0.0027	0.0346
Average	0.0253	0.0319	0.0027	0.0309

Table 4.2: Average RMSE and NMAE across five folds for XGBoost models estimating missing mobility data from restrictions—the errors are averaged over the six variables—and missing R_t data from mobility patterns for each country.

4.4 Discussion

	Findings
Correlation	A strong correlation is observed between the analyzed restrictions and mobility variables, in contrast to the weaker correlation identified between mobility and infection rates.
Patterns	Clear cultural patterns emerge when examining the impact of the stringency index on mobility and infection rates and the effect of mobility on infection rates.
Clustering	There is a clear separation between Northern and Southern European countries when social compliance is the key factor, but this distinction disappears when social interactions are the main focus.
Estimations	XGBoost models perform very well in estimating missing mobility data from restrictions and missing infection rates from mobility.

Table 4.3: Summary of key findings.

Table 4.3 shows a summary of key findings. Our analyses highlight a strong *correlation* between government stringency measures and mobility reductions, as evidenced by the high correlation values observed in Figure 4.1 and Table 4.1. This indicates that restrictions are highly effective in modifying public behavior, with an almost immediate response in mobility patterns, largely due to social compliance with public health measures. However, the relationship between mobility and R_t is weaker, suggesting a delayed and more complex interaction between changes in movement and its influence on infection rates. This lag can be attributed to additional factors, such as social interactions, incubation periods, pre-existing transmission dynamics, and the delay between an increase in mobility and the subsequent rise in R_t.

Analyzing the ratios defined in Section 4.2, distinct *cultural patterns* emerge in the interplay between restrictions, mobility, and infection rates. These patterns enable us to cluster countries accordingly. The *clustering analyses* in Section 4.3.2 reveals a clear separation between Northern and Southern European countries when social compliance plays a major role. Northern European countries, such as Finland, Norway, and Sweden, demonstrated higher compliance with restrictions, leading to significant mobility reductions with less stringent measures. In contrast, Southern European countries required stricter measures to achieve similar levels of compliance. This distinction reflects cultural and societal differences, including public trust in authorities, social norms, and economic conditions. When social interactions become the dominant factor, the clustering distinction between

Northern and Southern countries disappears. This finding underscores the universal role of interpersonal behavior in driving disease transmission, irrespective of regional or cultural differences. It highlights the importance of addressing social interaction patterns alongside mobility restrictions in designing effective public health strategies.

Finally, regression models enable us to accurately estimate *missing* mobility data from restrictions and missing infection rates from mobility. For instance, these estimated values can be used to complete time series data, which can then be employed to identify patterns in the data, simulate different scenarios, or test hypotheses about the relationships between variables. Additionally, the completed time series can be applied to other models, such as neural networks for forecasting, decision trees for classification tasks, or reinforcement learning models for optimizing policy interventions.

The findings have significant implications for policy-making in managing infectious diseases. First, the rapid and robust response of mobility to restrictions underscores the importance of timely and decisive government interventions. Policymakers should prioritize measures that effectively reduce high-risk mobility patterns, such as public gatherings and inter-regional travel. Second, the observed cultural and regional variations suggest the need for context-specific strategies. Northern European countries may benefit from less coercive measures, harnessing their higher levels of public trust and compliance. In contrast, Southern European countries may require more stringent interventions to achieve similar outcomes. Such tailored approaches can optimize the balance between restriction effectiveness and societal acceptance. Finally, the weaker correlation between mobility and R_t emphasizes the importance of complementing mobility-focused measures with other interventions, such as improved testing, contact tracing, and vaccination campaigns. Addressing factors beyond mobility can accelerate the reduction in transmission rates.

4.5 Conclusions

In this study, we utilized diverse data sources to investigate the relationship between government-imposed restrictions, mobility patterns, and the effective reproduction number R_t of COVID-19. Due to its unprecedented global impact and the extensive data collected during the pandemic, COVID-19 served as an ideal case study, offering valuable insights into the dynamics of infectious disease control in a highly interconnected world.

We focused on several European countries—specifically Finland, Norway, Sweden, Belgium, Ireland, the United Kingdom, France, Germany, Austria, Italy, Greece, Spain, and Portugal—to identify cultural patterns and cluster these nations using two state-of-the-art approaches: K-means [125] and CONNECTOR [203]. Moreover, we employed XGBoost [77] to estimate missing mobility data from the whole restriction dataset and to infer missing R_t values from mobility data—an innovative approach that represents a novel contribution to the existing literature. Our findings emphasize the importance of tailoring public health strategies to account for cultural and regional differences, particularly in compliance with NPIs. Moreover, the methodological approach adopted in this study—combining machine learning techniques with epidemiological insights—provides a robust framework for analyzing mobility and infection dynamics in future pandemics.

Despite the valuable insights provided by this study, several limitations should be noted. First, while the focus on European countries offers important cross-cultural comparisons, the findings may not be fully generalizable to regions with different socio-cultural or economic contexts. Second, we trained separate XGBoost models for each country, which limits the generalization of the models. Additionally, due to issues with changing survey questions on platforms like Facebook, the study primarily focuses on the initial phase of the COVID-19 pandemic and may not fully account for evolving pandemic dynamics, such as the emergence of new variants, which could impact the relationship between mobility patterns and infection rates. Finally, we attempted to predict future mobility trends based on restrictions and future R_t values based on mobility, achieving relatively good results. However, other time series techniques, such as Prophet [257], yielded more accurate predictions.

Future work may include: *i)* expanding the analysis to include additional countries and regions, particularly those with differing socio-cultural and economic contexts, would enhance the generalization of the results, *ii)* training general XGBoost models to estimate mobility from restrictions and infection rates from mobility data, and *iii)* extending the current approach to other infectious diseases, public health scenarios, or more broadly, to different contexts. This would provide a deeper understanding of how easily observable variables can influence more challenging-to-observe variables.

4.6 Ethical declaration

The Ethics Board (IRB) of IMDEA Networks Institute gave ethical approval for this work on 2021/07/05. IMDEA Networks has signed Data Use Agreements with Facebook, Carnegie Mellon University (CMU) and the University of Maryland (UMD) to access their data, specifically UMD project 1587016-3 entitled C-SPEC: Symptom Survey: COVID-19 and CMU project STUDY2020_00000162 entitled ILI Community-Surveillance Study. The data used in this study was collected by the University of Maryland Social Data Science Center Global COVID-19 Trends and Impact Survey (UMD Global CTIS) in partnership with Facebook. Informed consent has been obtained from all participants in this survey by this institution. All the methods in this study have been carried out in accordance with relevant of ethics and privacy guidelines and regulations.

4.7 Availability of data and software code

All datasets used in this work are open-source and publicly available. In particular, we used official surveillance data from the COVID-19 Data Hub [121, 120], mobility data from UMD Global CTIS's survey [32, 104, 151], estimates of active cases computed from the UMD Global CTIS's survey [32, 231, 24, 234], information on restrictions from the Oxford Covid-19, and Google mobility data [176]. The UMD Global CTIS's survey data (in aggregated form) are openly accessible at `https://github.com/GCGImdea/coronasurveys/`. The microdata of the CTIS survey from which the aggregated data was obtained cannot be shared, as per the Data Use Agreements signed with Facebook, Carnegie Mellon University (CMU), and the University of Maryland (UMD). Thus, the authors do not have permission to share the microdata of the CTIS survey.

4.7.1 How to reproduce the results

Requirements

You need to have docker installed on your computer, for more info see this document: `https://docs.docker.com/engine/installation/`. Ensure your user has the right to run docker (without the use of sudo). To create the docker group and add your user in a Unix system:

- Create the docker group:

```
$ sudo groupadd docker
```

- Add your user to the docker group:

```
$ sudo usermod -aG docker $USER
```

- Log out and log back in so that your group membership is re-evaluated.

Reproduce

To reproduce the results clone the repository (`https://github.com/daniele-baccega/mobility-study`) and run:

```
$ cd mobility-study
$ ./reproduce.sh
```

Extended data analysis on NPIs, mobility, infection rates and regression models

Additional details on some of the results presented in this chapter are provided in the Supplementary Material of `https://www.medrxiv.org/content/10.1101/2025.01.10.25320334v1`.

Chapter 5

Forge4Flame: an intuitive dashboard for designing GPU Agent-Based Models to simulate the spread of infectious diseases

This chapter presents the first results of the TrustAlert project [212]. Before discussing the work, we provide a brief introduction to the project and its main objectives. The goal of the project—firstly introduced in [86]—is to advance the preparedness and response capabilities of territorial public health agencies in emergencies. To achieve this goal, the project aims to establish a comprehensive, integrated platform, to provide early warnings, monitoring, and forecasting tools to public health response agencies and local healthcare services for anticipating medical needs. We can distinguish three main core objectives—Figure 1.3:

i) to *early detect* the approaching of a health crisis, i.e. spreading of cases due to certain pathologies;

ii) to *assess the impact* of the health crisis at the territorial level;

iii) to *forecast scenarios* at macro- and micro-level.

We have been most involved in the third sub-project (forecast scenarios) and, in particular, we proposed a new simulation paradigm integrating micro- and macro-simulations to achieve higher robustness and trustability with respect to the single simulation paradigms. Our proposal for this simulation paradigm is sketched in Figure 5.1 where circles represent models with different levels of detail, the arrows the data exchange among the models, and the dashed

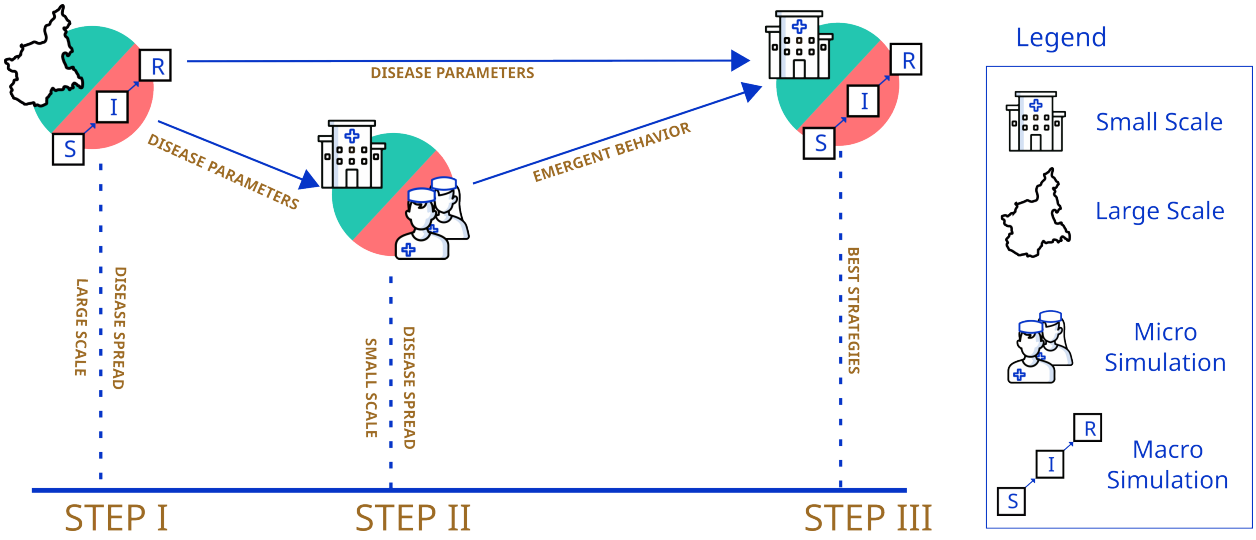


Figure 5.1: Diagram representation of the novel modeling paradigm.

lines the main outputs provided by each model that will be integrated into the developed smart dashboard.

The proposed approach combines three different models: a first model for macro-simulations at large scale (namely MaS-LaS), a second model for micro-simulations at small scale (namely MiS-SmS), and a third model for macro-simulations at small scale (namely MaS-SmS). These models will be exploited to derive *i)* a more trustable forecast on the spread of infectious disease at large and small scales, and *ii)* a set of possible interventions to reduce the impact of a health emergency in a (public) environment—e.g., schools, public offices, universities, health facilities, etc.. The validation of the proposed solution will be performed with retrospective data from COVID-19 provided by Cottolengo Hospital Health Unit.

In detail, initially in Step I the MaS-LaS model describing the macroscopic behavior of infectious disease at the national, regional, and provincial levels (i.e., large scale level) is created and trained to reproduce the observed infection disease spread and forecast its evolution. Then, the outcomes obtained by simulating such a model will be exploited in Step II to better parameterize the MiS-SmS and MaS-SmS models in terms, for instance, of incoming infected individuals and the ability of a disease to spread (i.e. R_0 or R_t). Thus, the MiS-SmS taking as input the future evolution of the disease estimated through the MaS-LaS model will simulate its spread in a specific (public) environment from a microscopic point of view. The micro-simulation of this model based on an agent-based model will provide the emergent behaviors for the specific environments. Still, evaluating different containment strategies using genetic algorithms or reinforcement learning could not be explored because each simulation run is too computationally expensive. To deal with this aspect in Step

III the MaS-SmS is introduced to describe the spread of the infectious disease in the same specific (public) environment but from a macroscopic point of view and taking as input those emergent behaviors reported by MiS-SmS in Step I. This allows us to explore different containment strategies using genetic algorithms or reinforcement learning directly on this model for macro-behavior at a small scale whose solution is less computationally demanding.

While beginning to develop the agent-based model environment for the Cottolengo Hospital, we realized that this task is proving to be particularly challenging and represents one of the main limitations of current agent-based modeling tools—as introduced in Chapter 1. These tools typically require a manual setup, where users must manually configure different environmental components, which can be time-consuming and technically demanding. Users need to define the spatial structure of the environment by hand, including the dimensions and layout of rooms, the placement of walls, doors, and windows, as well as furniture and other obstacles that may influence the agents' movements or interactions. For example, simulating an office environment requires detailed specifications of cubicles, meeting rooms, hallways, and other spatial features that affect how agents navigate the space. Similarly, modeling a hospital would require mapping out waiting areas, wards, operating rooms, and corridors, each with specific rules for agent behavior. As a result, the manual and non-practical process of environment creation can limit the flexibility and accessibility of ABMs, particularly for users who may not have specialized knowledge of modeling techniques and the simulated domain. This creates a significant barrier to using ABMs for large-scale, real-world applications, where the environment's structure plays a crucial role in agent interactions and outcomes.

We were already aware of this, as we encountered similar difficulties in previous works [41, 40], where we faced the same problem using NetLogo [285]. For the agent-based model of the Cottolengo Hospital, we selected FLAME GPU 2 [223, 222] for the reasons outlined in Section 1.4, where the environment definition is even more complex compared to NetLogo. This highlights the necessity of a dedicated tool to facilitate this task.

To address this limitation, we implemented—in synergistic collaboration with the FLAME GPU 2 [222, 223] developers—***Forge4Flame (F4F)***, a user-friendly dashboard—developed in R Shiny—designed to simplify the definition of an agent-based model for FLAME GPU 2 agent-based models, automatically generating the necessary code. This tool: *i)* simplifies the modeling process, making it accessible to users without prior coding experience and enabling them to define the model's environment—by drawing it through a graphical interface—, the agents interacting within it, the disease model, and other components relevant to an agent-based simulation of disease spread—like the definition of *what-if* scenarios based on the implementation of NPIs and vaccination strategies; *ii)* facilitates effective cooperation among diverse research fields, improving the efficiency of disease model development and analysis; as

already said, this work arises from an Italian project [212] aimed at developing an integrated modeling environment to evaluate the potential impacts of health emergencies at various levels of detail and to support healthcare decision-making regarding interventions to mitigate the effects of such emergencies; *iii)* enhances the model development and result evaluation by leveraging Slurm [288] for efficient distributed computing on High-Performance Computing (HPC) infrastructure, thereby improving the scalability of FLAME GPU 2; and *iv)* integrates macro-level population models with micro-level simulations, enabling comprehensive analysis of disease dynamics; notably, it explores the integration of ABMs with a compartmental model used to study COVID-19 diffusion in the Piedmont region of Italy [202], incorporating the prevalence of the virus in the broader population extracted using the macro-level model into the micro-level model. Finally, we emphasize that **this tool does not have an equivalent counterpart in the existing literature**.

To validate the tool, we used COVID-19 as a case study replicating findings from previous studies [41, 40]—depicted in Chapter 2—that model the COVID-19 emergency in an Italian middle school, demonstrating Forge4Flame's ability to accelerate both the modeling phase and the evaluation of results. Moreover, leveraging the combination of *F4F* with a macro-level model [202], we explore how the spread of the disease would have been affected in the school environment if an alert had been issued in February 2020, to demonstrate the utility of an early warning system in slowing the spread of infection.

In this chapter, we start by outlining the background of FLAME GPU 2 [223, 222], Docker [181], and Slurm [288] (Section 5.1). Next, we introduce the main outcome of this work, *F4F*, along with the architecture used to run it (Section 5.2). We then present a comparison of computation times and results between FLAME GPU 2 and NetLogo [285], using a case study published in [41, 40], and present the results obtained from combining macro- and micro-level models (Section 5.3). Finally, we provide a discussion (Section 5.4) and draw conclusions regarding the key findings (Section 5.5).

5.1 Background

This section provides an overview of the key components and tools used in our work. Specifically, Section 5.1.1 focuses on FLAME GPU 2 [223, 222], Section 5.1.2 introduces Docker [181], and Section 5.1.3 highlights Slurm [288].

5.1.1 FLAME GPU 2

FLAME GPU 2 [223, 222] is an agent-based modeling and simulation framework designed to simplify GPU programming while offering substantial performance improvements over CPU-based agent-based model frameworks. It employs a state-based representation, enabling users to define agents as state machines with internal memory updated by functions. This approach supports parallel execution, contrasting with traditional serial simulations that process agents sequentially and require agent shuffling to prevent bias.

In FLAME GPU 2, conflict resolution is necessary when agents compete for resources. This is typically handled by assigning agents a priority, often determined randomly, to resolve contention during selection. Models in FLAME GPU 2 consist of agents with defined variables, states, and functions. These functions govern state transitions by updating agents' memory based on specified mathematical operations. The execution order is controlled using layers, where functions within the same layer run in parallel, subject to constraints that prevent race conditions caused by shared resource conflicts.

Agents communicate indirectly through global message lists, enabling information exchange without direct interaction. Additionally, environment variables, called macro-environment properties, provide a lightweight communication mechanism by allowing agents to access and update shared global values atomically. However, since the execution order of atomic operations is determined by the GPU, results may vary across runs due to the non-deterministic nature of parallel execution.

As highlighted in Chapter 1, ABMs can include a description of the spatial domain. Specifically, constructing a 3D environment in FLAME GPU 2 involves defining a bounded Cartesian space where agents are positioned and interact. Each agent's location is represented by 3D coordinates (x, y, z), and their behavior, such as movement or sensing, is influenced by these positions. Spatial partitioning techniques, such as uniform grids or hashing, are employed to efficiently manage interactions and proximity queries in the 3D space. Agents can move and interact under defined rules and physics-based dynamics, like forces for attraction or repulsion, can further govern movement. Communication between agents is facilitated through spatially filtered message passing, ensuring that interactions are localized and computationally

efficient.

Creating a closed environment like a school or a hospital can be highly challenging due to the need to model complex spatial structures and agent interactions. The environment requires careful design to represent rooms, hallways, and other features, often using grid-based layouts or mathematical boundaries to define walls and restricted areas. Agent navigation within this space is another challenge, requiring rules or algorithms like pathfinding to ensure agents follow logical paths and avoid obstacles.

Communication between agents must also account for the confined structure, with localized interactions limited to nearby zones. Efficient message passing and filtering are crucial for maintaining computational performance in such a setting. The simulation's complexity increases further with dynamic changes, such as agents moving between rooms or doors opening and closing. Additionally, the presence of heterogeneous agent roles—like students, teachers, doctors, and patients—introduces further challenges to the system.

Visualization presents another layer of difficulty, as it needs to accurately depict the closed environment and agent movements in a confined 3D space. While FLAME GPU 2's parallel processing capabilities help manage large-scale simulations, achieving a realistic and functional closed environment demands advanced programming skills and a deep understanding of the framework's features.

5.1.2 Docker

Docker [181] is an open-source platform for developing, deploying, and managing containerized applications. Docker allows to package and run applications in loosely isolated environments called containers. These containers provide isolation and security, allowing multiple containers to run simultaneously on a single host. Lightweight and self-contained, containers include everything needed to run an application, eliminating dependencies on the host's configuration. Containers can also be shared easily, ensuring consistent behavior and functionality across different environments for all users. This approach enhances reproducibility, scalability, and efficiency in software development and operations.

The Docker system includes tools like Docker Engine [2], a runtime for managing containers, Docker Hub [3], a cloud-based repository for sharing container images, and Docker Compose [140], which simplifies multi-container application deployment. Docker also integrates with orchestration frameworks like Kubernetes [220], reinforcing its position as a foundational technology in modern software deployment workflows.

Specifically, Docker Compose [140] is a tool that simplifies managing multi-container applications by defining services, networks, and volumes in a single YAML file. It simplifies

the setup, deployment, and scaling of containerized applications, enabling developers to start or stop all services with simple commands. Ideal for micro-services, it ensures consistency across environments and supports scaling for testing.

5.1.3 Slurm

Slurm [288] is a robust, fault-tolerant, and highly scalable job scheduling system designed to manage workloads on large-scale Linux clusters. It is widely utilized in High-Performance Computing (HPC) environments to efficiently allocate and manage resources. Slurm performs several critical functions that ensure efficient resource utilization and job management across heterogeneous cluster environments. It has three main functions: *i)* it allocates exclusive and/or non-exclusive access to resources for some duration of time, *ii)* it provides a framework for starting, executing, and monitoring works and *iii)* it arbitrates contention for resources by managing a queue of pending works. In addition to its core functions, Slurm is highly extensible, offering a variety of optional plug-ins and modules to enhance its capabilities. Notable plug-ins provide features such as accounting, advanced reservations, backfill scheduling, and more, enabling tailored solutions for diverse workload management needs.

5.2 Dashboard

In this section, we present the primary outcome of this work, the *F4F* dashboard (Section 5.2.1), along with the architecture employed to conduct experiments (Section 5.2.2), combining Docker [181], Docker Compose [140], and Slurm [288] for HPC systems.

5.2.1 Forge4Flame (F4F)

To address the primary challenge of ABMs—i.e., the difficulty in constructing realistic environments—, we developed ***Forge4Flame (F4F)***, a user-friendly dashboard designed to streamline the creation of FLAME GPU 2 [223, 222] ABMs automatically generating the necessary code. With its intuitive interface, users can easily define key components, including the model's environment, the agents interacting within it, the disease model, and other essential elements for agent-based simulations.

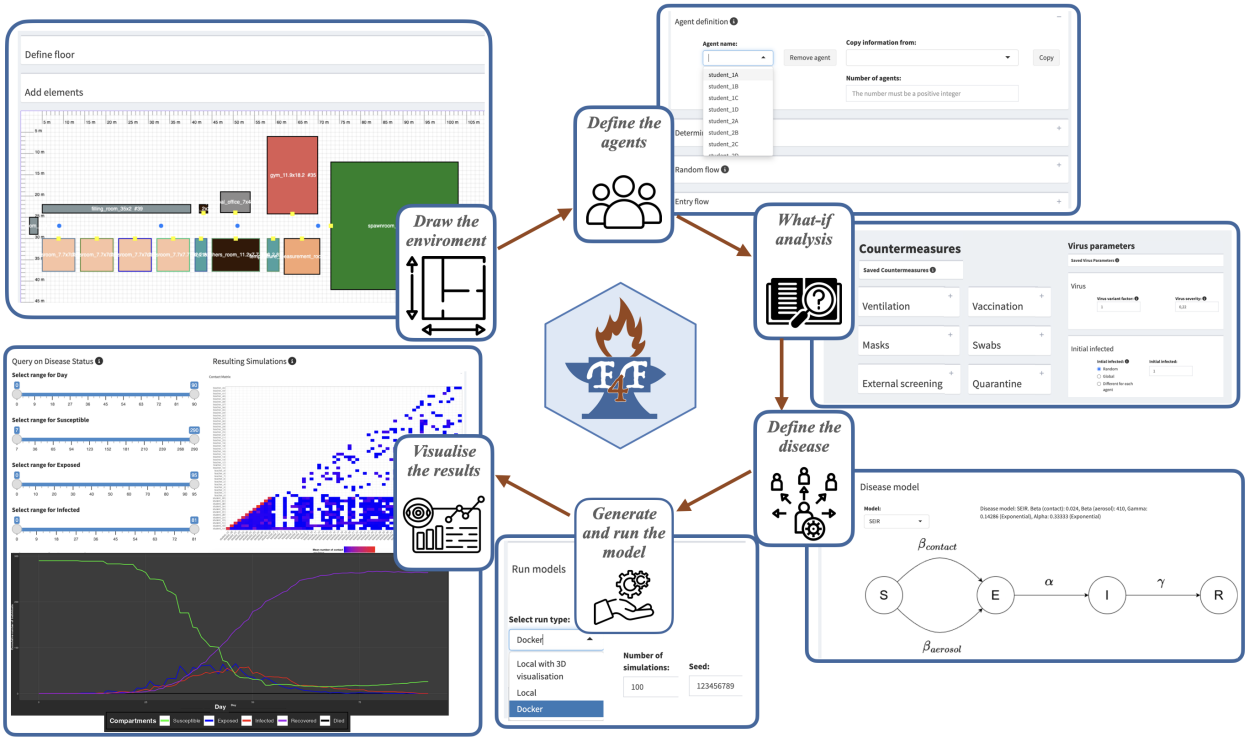


Figure 5.2: *F4F* schematic overview: the six main pages of the application that can be exploited to build an ABM model, accompanied by illustrative screenshots. The workflow is characterized by the six key steps: drawing the environment, defining agents, setting optional features to run what-if analysis, setting disease parameters, running simulations, and visualizing results.

Figure 5.2 shows the *F4F* workflow, offering a concise overview of the tool and featuring a menu on the left-hand side that organizes the primary functionalities of *F4F*, each tailored to simplify various aspects of model creation. Specifically, Figure 5.2 represents a workflow for building an ABM using the six key functionalities. In detail, an *Home* page serves as

the starting point. The *Canvas* page allows users to design the model's environment by arranging rooms through a drag-and-drop interface. In the *Rooms* page, users can specify new room types, while the *Agents* page enables the creation of new agent types and their movement patterns within the simulation. The *Resources* page allows users to define the capacity of agents permitted in each room. The application also includes a dedicated *Infection* page for specifying disease models and a *What-If* analysis tool for exploring hypothetical scenarios. Users can configure initial model parameters in the *Configuration* page, adjust canvas dimensions, and manage model files in *Settings*. The *Run* page allows users to execute a defined model. Lastly, *F4F* provides a *Post-processing* feature for examining simulation logs visually, enhancing the ability to interpret results.

In particular, the main strengths of the tool are outlined as follows:

- *User-friendly interface*: *F4F* is highly intuitive and user-friendly, simplifying the process of creating complex agent-based models. The drag-and-drop functionality allows users to easily define the model's environment quickly and efficiently.

- *Comprehensive model definition*: *F4F* enables the definition of complete ABMs, encompassing the model's environment, the agents interacting within it along with their movement flows, resource limitations, the infectious disease model, hypothetical scenario analysis, and configuration parameters.

- *Automated code generation*: *F4F* automatically generates the necessary code for running agent-based model simulations using FLAME GPU 2 [223, 222]. This ensures that users, regardless of their coding expertise, can effectively utilize the application.

- *Scenario exploration*: the *What-If* page allows users to explore various scenarios, analyzing the effectiveness of different NPIs and vaccination strategies.

- *Visualization features*: the *2D visualisation* page offers an intuitive way to visualize simulation results, particularly in situations where FLAME GPU 2's 3D visualization capability is not available.

- *Detailed user guide and application examples*: a comprehensive guide detailing all the features available in *F4F*, along with a couple of model construction examples, is provided and will be continuously updated as new features are introduced.

- *Ample room for future improvements and new features*: *F4F* offers a wide range of features, with additional functionalities planned for future releases, allowing for continuous enhancement and expansion.

A brief overview of how to define a model is now provided. In addition, a comprehensive user guide with more details, along with a couple of application examples illustrating the practical use and versatility of the tool, is provided (see Section 5.6 for more details). To build an agent-based model using *F4F*, the user must first define new room types on the *Rooms* page by specifying their dimensions. Once the rooms are defined, the user can add them to the canvas on the *Canvas* page, thereby defining the model's environment. For each room added, the user can specify the location of the door, assign the room to a particular area, remove previously added rooms, clear floors, add specific points to assist agent movements and perform other relevant actions. Next, the user must define agent types on the *Agents* page. On this page, the user can create agent classes, copy information from previously defined agents, establish the initial number of agents for each type, and define determined, random, and entry flows. Then, on the *Resources* page, the user can limit access to each room for a specific number of agents (either generally or for each agent type) while also specifying an alternative room to be used if resources are unavailable in the primary room. On the *Infection* page, the user must select the infectious disease model to be used for simulating the infection, ensuring that the virus parameters are correctly chosen according to the modeled virus. On the *What-If* page, the user can enable countermeasures (such as room air ventilation, vaccination, mask usage, screening, and quarantine policies) and define other virus-related parameters, including the virus severity, the modeled virus variant, the initial number of infected individuals, and the potential for contagion outside the environment. On the *Configuration* page, the user can select various configuration parameters, such as the seed, the number of simulations to run, the number of days to simulate in each run, the starting day and hour, and the FLAME GPU 2 step size. On the *Settings* page, the user can download the defined model. Finally, on the *Run* and *Post-processing* pages, users can execute the model and load simulation results for post-processing and visualization.

5.2.2 Architecture

Figure 5.3 illustrates three possible approaches for running *F4F* and FLAME GPU 2. In all the cases, the user must create the model using *F4F* and link its output to FLAME GPU 2 through a specific button in the application. The output includes: *i)* a JSON file describing the model and *ii)* 3D objects used for visualization (which are automatically generated). A Python script will process the JSON file, generating the required code and FLAME GPU 2 variables. After the simulation is complete, the FLAME GPU 2 output can be uploaded back to *F4F* for post-processing, visualizing statistics, or viewing the simulation in 2D.

The first approach involves using Docker Compose [140] to manage the Docker containers

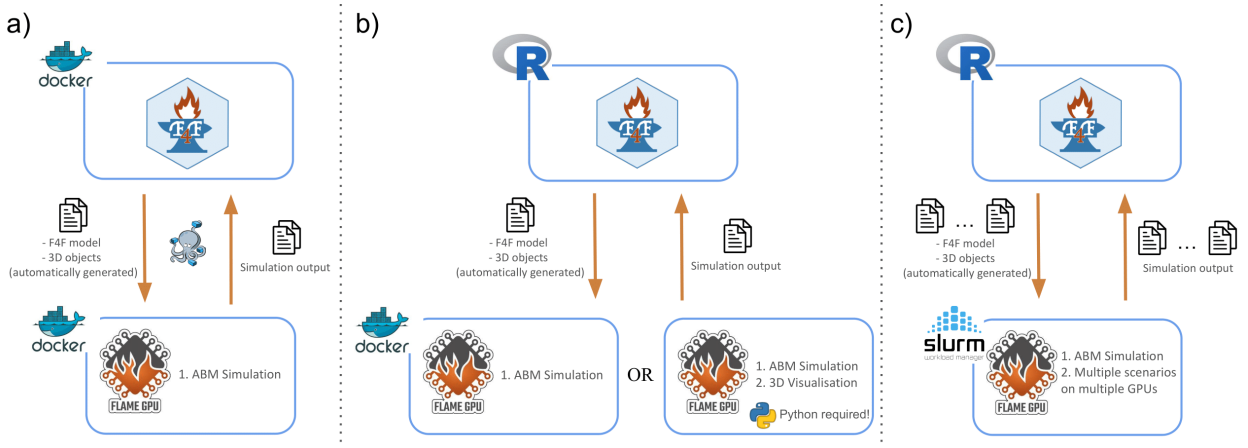


Figure 5.3: *F4F* architecture. It is possible to run *F4F* and FLAME GPU 2 *a)* together using Docker Compose [140], or separately using *F4F* without Docker and FLAME GPU 2 *b)* either with or without Docker or *c)* using Slurm [288] to run multiple scenarios on multiple GPUs without Docker.

of *F4F* and FLAME GPU 2 (see Figure 5.3-*a)*). This method simplifies the setup by eliminating the need for manual installation of packages and libraries. Once the model is defined in *F4F*, the user can run the simulation directly from the application using a dedicated button. Upon completion, the results can be visualized within the application itself. However, this approach does not support FLAME GPU 2's 3D visualization.

The second approach involves running the two components separately, *F4F* without Docker and FLAME GPU 2 either with or without it (see Figure 5.3-*b)*). If the user opts not to use Docker, they must manually install all required dependencies. The only way to enable FLAME GPU 2's 3D visualization is by running the application without using Docker.

The third and final approach exploits Slurm [288] (see Figure 5.3-*c)*) to run one or more scenarios on HPC systems exploiting a massive parallelization, running FLAME GPU 2 without Docker. Currently, *F4F* does not directly support running simulations on Slurm. Instead, users must execute them manually using the script provided in the GitHub repository (see Section 5.6 for more details).This functionality is currently being integrated into *F4F*.

Finally, the FLAME GPU 2 Docker image has been evaluated on the following GPUs: NVIDIA Tesla T4, NVIDIA GH200, and NVIDIA Tesla V100. For more information on how to install and run the tools and Docker containers, please refer to the user guide (Section 5.6).

5.3 Results

In this section, we compare FLAME GPU 2 [223, 222] and NetLogo [285] by reproducing a prior study [41, 40] (developed in NetLogo) using *F4F* (Section 5.3.1), focusing on computational cost and scalability. We then extend the comparison to stochastic processes using the ITP [209] across multiple runs performed with both FLAME GPU 2 and NetLogo (Section 5.3.2). Finally, we integrate macro- and micro-simulations—compartmental models and agent-based models, respectively—to demonstrate how early warnings can prevent widespread closures and mitigate the spread of infections (Section 5.3.3).

5.3.1 Case study: a general middle school environment

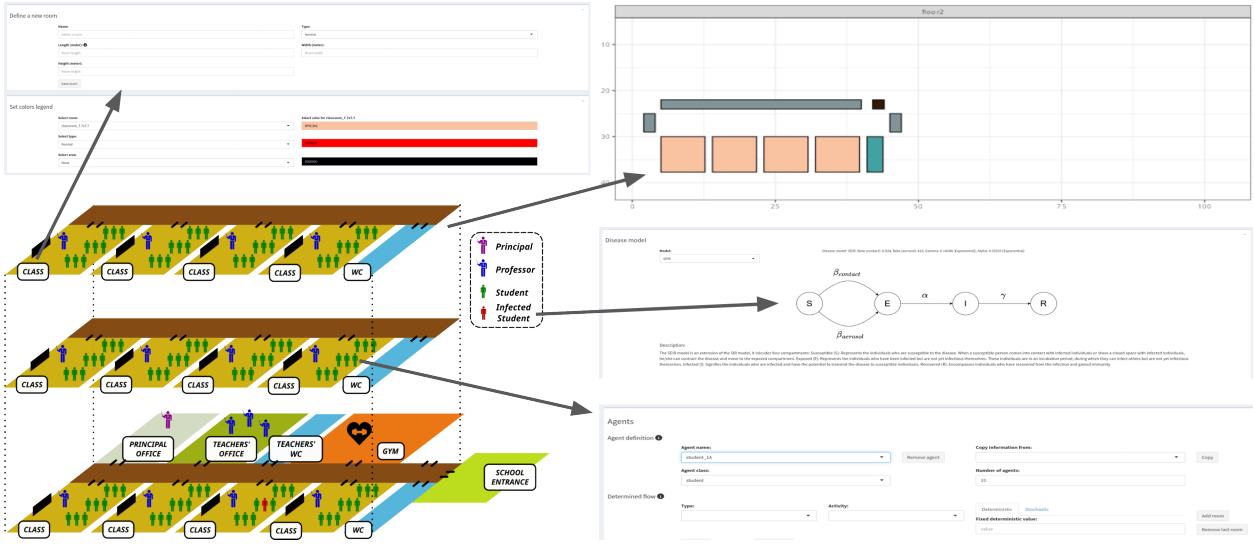


Figure 5.4: Graphical representation of a generic school described in Chapter 2 [41, 40]. The figure also illustrates the mapping between components of the school environment and the corresponding *F4F* pages: *Rooms* for defining the school's spaces, such as classrooms and bathrooms; *Agents* for specifying agents and their movements; *Infection* for modeling the disease dynamics; and *2D visualization* for visual representation of the simulation results.

In [41, 40] and in Chapter 2, we presented an agent-based model, developed in NetLogo [285], to simulate the dynamics of COVID-19 spread and evaluate the effectiveness of NPIs in a school setting (see Figure 5.4). The model integrates three key components:

- *Environment and population dynamics*: the school environment includes detailed spaces—e.g., classrooms, corridors, gym, offices—and a structured population of 290 individuals, comprising 240 students, 49 teachers, and one principal. Agents move and interact based on defined roles and schedules. Classrooms are organized with twenty-four desks, and a typical school day consists of six 50-minute classes with scheduled breaks.

Parameter	Value
Number of classrooms	12
Students per classroom	20
Initial infected	0
Vaccinated people	Teachers and students
Vaccine efficacy	0% / 50%
Fraction of vaccinated agents	0% / 30%
Average number of days of incubation [289, 194]	3
Average number of days of infection [66]	7
Duration of quarantine in days	0 / 10
Virus variant	Original
Ventilation	0 ACH / 3 ACH
Mask efficacy [200]	59% (surgical) / 90% (FFP2)
Mask type	None / Surgical

Parameter	Value
Fraction of population wearing mask	0% / 50%
Probability of going to the bathroom (per minute)	$8e^{-4}$
Probability of going to the principal office (per minute)	$1.4e^{-5}$
Screening policy	None
First probability of external screening (per day)	0 / 0.0075
Second probability of external screening (per day)	0 / 0.03
Probability of outside contagion	0.001
Contamination risk (c_r) [133]	0.024
Duration of the lessons in minutes	50
Simulated days for each run	60
Number of runs for each scenario	200

Table 5.1: Shared model parameters in FLAME GPU 2 and NetLogo for comparison.

- *Disease progression model*: the COVID-19 progression model uses a Susceptible-Exposed-Infected-Recovered (SEIR) compartmental model. Agents transition through states as follows: *i)* susceptible to exposed upon virus exposure, *ii)* exposed to infected after the incubation period, and *iii)* infected to recovered after the infectious period. Transmissions occur via close-range contact [133] or aerosol spread indoors [114].

- *Non-Pharmaceutical Interventions (NPIs)*: policies such as social distancing, environmental ventilation, mask usage, screening campaigns, and quarantine policies are modeled. The model allows for the evaluation of combined non-pharmaceutical intervention strategies to reduce transmission.

Additional features include the incorporation of COVID-19 variants—e.g., Alpha, Beta, Delta, Omicron BA.1, BA.2—, modeled with differing transmissibility parameters based on [200], the implementation of screening campaigns conducted outside the modeled environment—two different probabilities of external screening are applied to students every day: the probability of testing a student due to sporting activities that involve screening campaigns (e.g., playing soccer) and the probability of testing a student outside the school due to the presence of some symptoms—, and the possibility to get the infection outside the school. All these configurations can be defined using *F4F*.

Using *F4F*, we replicate the environment described in [41, 40] and conduct experiments to compare FLAME GPU 2 and NetLogo both in terms of computational times and obtained results—Table 5.1 summarizes the parameters used in both models.We examine two specific scenarios: one where no countermeasures are implemented, and another where all available countermeasures are applied, including room air ventilation, mask usage, vaccination, quarantine, and external screening—except for internal screening policies: this exclusion is due to

the specific policies used in [41, 40], which cannot be replicated through *F4F*. Specifically, we vaccinated 30% of students and teachers with a vaccine efficacy of 50%, ensured that 50% of individuals wore surgical masks, implemented room air ventilation at a rate of 3 ACH, applied a 10-day quarantine period, and set external screening probabilities at 0.0075 and 0.03, respectively. In both scenarios, the entire population is initially susceptible, and each susceptible agent has a daily probability of contracting the infection outside the school.

5.3.2 Comparison and validation: NetLogo vs. FLAME GPU 2

	Number of runs		
NetLogo [285]	1	100 (8 in parallel)	1.000 (300 in parallel)
Intel Xeon Processor Skylake, IBRS 8 cores	60m	15h 28m	-
Intel(R) Xeon (R) CPU E5-2640 v3 @ 2.60GHz 16 cores	1h 2m	13h 51m	-
13th Gen Intel (R) Core (TM) i9-13900H 20 cores	39m	9h 9m	-
FLAME GPU 2 [223, 222]	1	100 (8 in parallel)	1.000 (300 in parallel)
NVIDIA Tesla T4	1m 43s	55m	11h 12m
NVIDIA Tesla V100	1m 41s	55m	9h 14m
NVIDIA GH200	1m 2s	43m	7h 15m

Table 5.2: Comparison between NetLogo and FLAME GPU 2 on different CPUs and GPUs, respectively. The table clearly shows the performance gain when using GPUs instead of CPUs.

Table 5.2 compares NetLogo and FLAME GPU 2 in terms of execution time across different CPUs and GPUs. For a fair comparison, the computation time is reported for both a single run and 100 runs, with 8 runs executed in parallel, on both CPUs and GPUs. Additionally, the performance of the three GPUs was evaluated in an extreme scenario, running 1,000 executions with 300 in parallel. All GPU results were obtained using the high-performance computing infrastructures described in [21].

The table clearly demonstrates the performance gain when using GPUs, reducing the execution time from 39 minutes to 1 minute and 2 seconds (approximately 38 times faster) for a single run, and from 9 hours and 9 minutes to 43 minutes (approximately 13 times faster) for 100 runs, based on the best times recorded for both tools. Considering the worst-case times, GPUs reduced the execution time from 1 hour and 2 minutes to 1 minute and 43 seconds (approximately 36 times faster) for a single run, and from 15 hours and 28 minutes to 55 minutes (approximately 17 times faster) for 100 runs.

Focusing on GPUs, while no significant differences are observed with 1 or 100 runs, a clear performance increase emerges in the final scenario with 1.000 runs and 300 in parallel,

where the NVIDIA Tesla V100 outperforms the NVIDIA Tesla T4, and the NVIDIA GH200 surpasses both. This improvement is attributed to the higher computational power, memory bandwidth, and parallel processing capabilities of more advanced GPUs, which become increasingly relevant under intensive workloads.

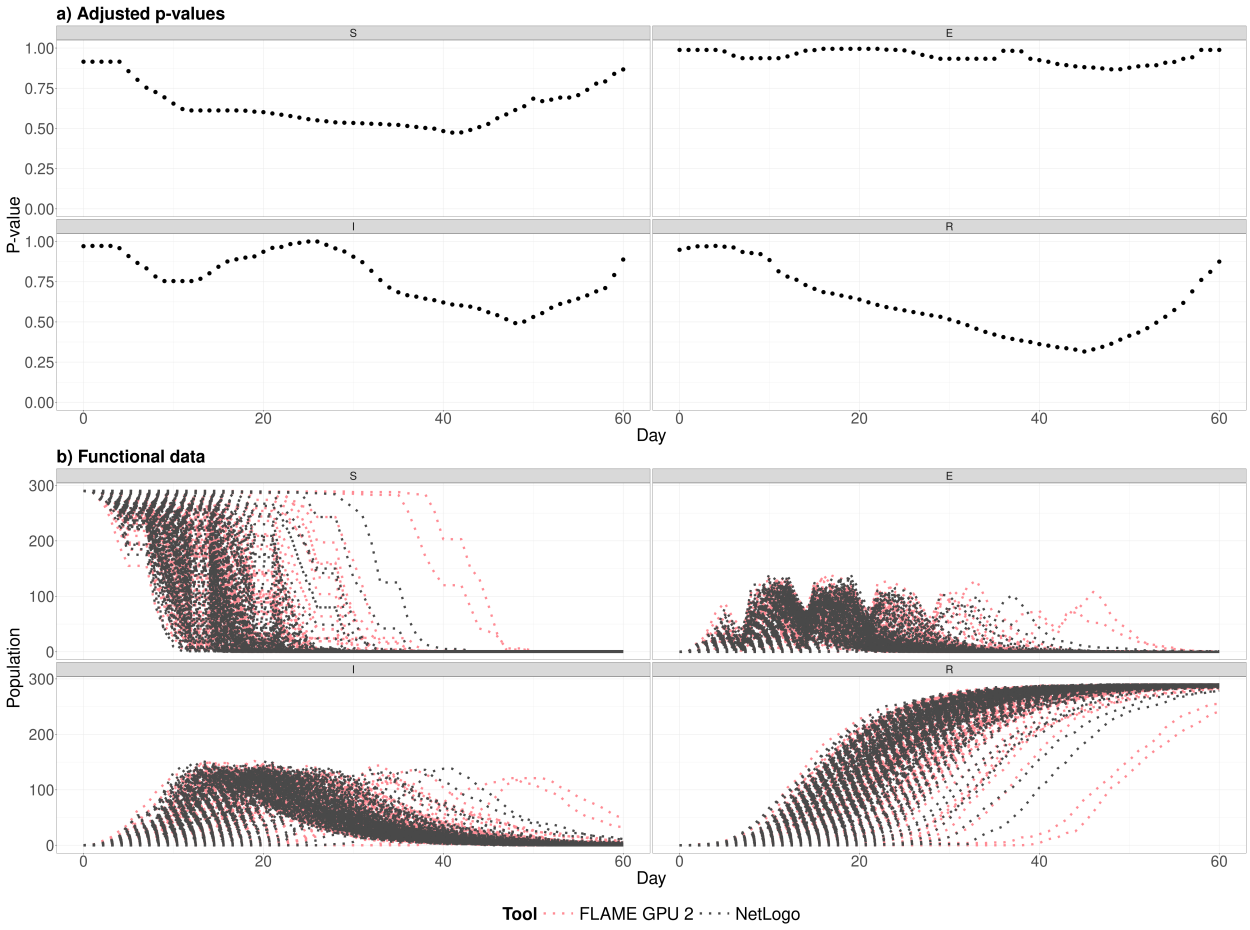


Figure 5.5: ITP [209] applied on the scenario without NPIs where a comparison among 200 traces from NetLogo and FLAME GPU 2 is performed.

To validate the results, we used the Interval Testing Procedure (ITP) [209]. This procedure allows for the comparison of two statistical processes by analyzing the confidence interval of the outputs generated at each time step. Under the null hypothesis that the two processes are equal, ITP determines whether the output intervals of the two processes overlap or differ significantly. If the intervals at the same time step from both processes do not overlap, it suggests a significant difference between them, rejecting the null hypothesis. Conversely, if the intervals overlap, the null hypothesis is not rejected, indicating that the two processes are statistically equivalent.

Figures 5.5 and 5.6 illustrate the output of the ITP for the two scenarios considered. Figures 5.5-*a)* and 5.6-*a)* illustrate that the p-values at each time step exceed 0.05—the

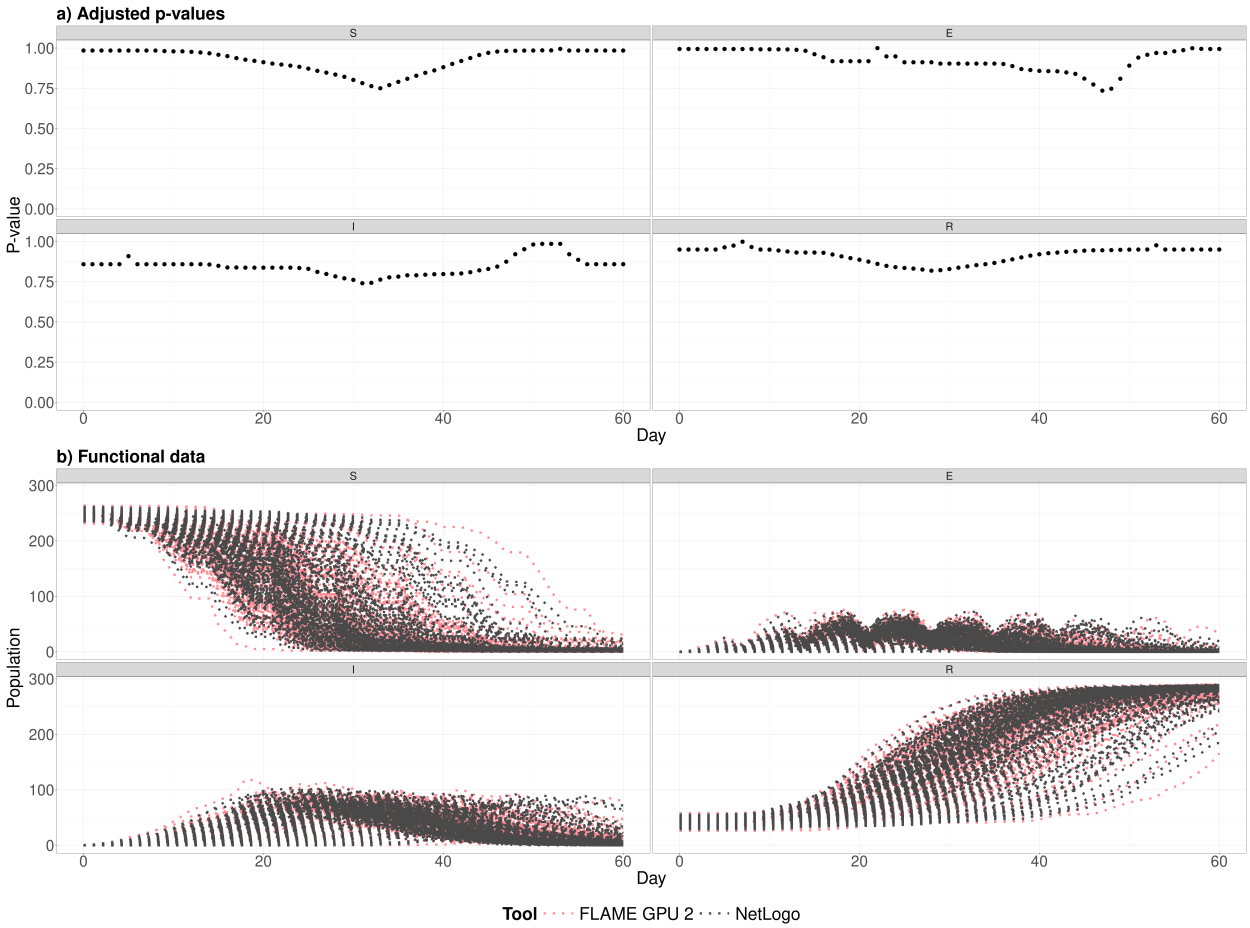


Figure 5.6: ITP [209] applied on the scenario with NPIs where a comparison among 200 traces from NetLogo and FLAME GPU 2 is performed.

selected significance level—, meaning we cannot reject the null hypothesis. Figures 5.5-*b)* and 5.6-*b)* show a comparison of NetLogo and FLAME GPU 2 traces, illustrating the evolution of the four compartments in the SEIR compartmental model.

5.3.3 From macro- to micro-simulation

Modeling a realistic probability of infection outside the agent-based model environment requires incorporating an additional step to account for external factors and interactions beyond the simulated environment. Specifically, our idea is to incorporate the prevalence of the virus in the broader population—whether at the national, regional, or city level—into the agent-based model. In the case of active cases, the user must provide the total population, enabling the computation of daily infection prevalence using Equation 5.1. According to this equation, the prevalence of the virus at time t is defined as the total number of infected individuals at time t (denoted as $I(t)$) divided by the total population (N).

$$prevalence(t) = \frac{I(t)}{N} \tag{5.1}$$

Within the context of a closed system, this approach is designed to model the introduction of an average daily number of infected individuals into the environment. Unlike a simple constant rate, this number follows a more complex trend. In our specific case, a compartmental model was utilized linking the simulated prevalence of the virus to *F4F*. Figure 5.7 shows a graphical schema of our approach, illustrating how infection prevalence is transferred from the large-scale macro-simulation compartmental model to the small-scale micro-simulation agent-based model.

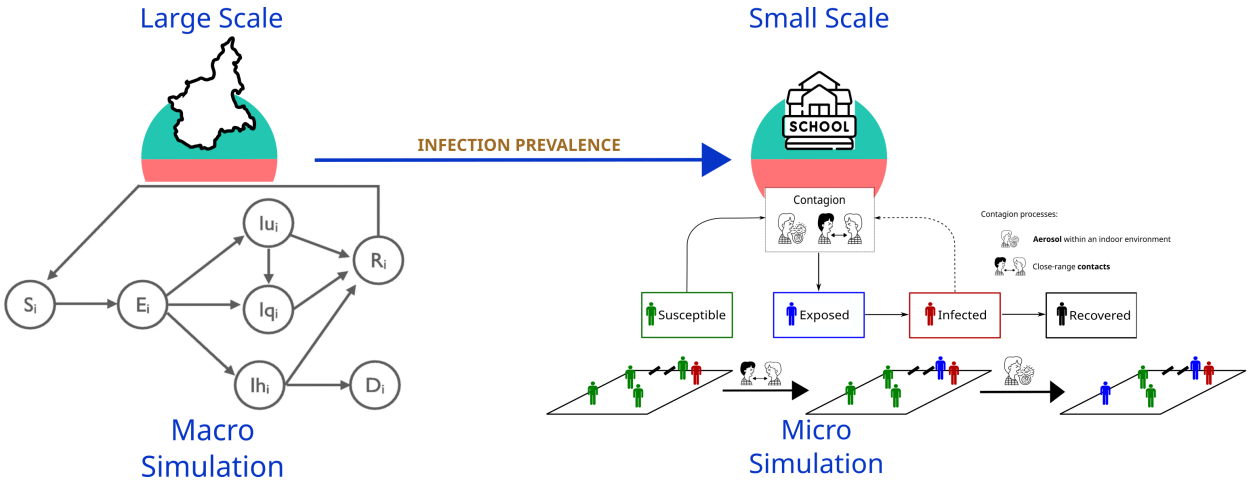


Figure 5.7: Graphical schema of the *from macro to micro* approach.

Specifically, we utilized the COVID-19 model introduced in [202], used to investigate different scenarios of the COVID-19 spread in Piedmont and the implementation of different infection-control measures and testing approaches. This model, formulated using the Petri Net formalism [206], is an extended version of the Susceptible-Exposed-Infected-Removed-Susceptible (SEIRS) compartmental model. It incorporates a population age structure and includes an additional state for deceased individuals (D), effectively making it a SEIRDS model with age stratification—see Figure 5.7, on the left, for a graphical representation of the compartmental model.

The infectious population is divided into three sub-groups: *i)* undetected infected (u), *ii)* quarantined infected (q), and *iii)* hospitalized infected (h). The model further stratifies individuals into three age groups—young (0-19 years), adults (20-69 years), and older adults (70+ years)—with uniform susceptibility and social activity patterns within each group. This structure allows for tracking age-specific disease incidence and social interactions.

Contacts are categorized into four types: *i)* within families, *ii)* at school, *iii)* at work, and

iv) other social interactions, according to [211]. The model simulates quarantine measures, the use of Personal Protective Equipment (PPE)—including masks, gloves, gowns, and eye protection designed to protect individuals from health risks and infections—and public restrictions imposed by the Italian government, which reduce social contact according to the severity of the restrictions.

Each age group is divided into seven compartments: susceptible (S_i), exposed (E_i), infected (I_{u_i}, I_{q_i}, I_{h_i}), recovered (R_i), and deceased (D_i). This detailed segmentation allows for effective modeling of quarantine and governmental measures by age group. The transition from susceptible (S_i) to exposed (E_i) is governed by the force of infection (FOI), a time- and age-dependent rate that includes four components: *i)* the infection rate, which depends on the age classes of both susceptible and infected individuals and the contact matrix; *ii)* the strength of governmental restrictions, modeled as a time-dependent step function; *iii)* compliance with governmental restrictions, which improves with higher disease severity; *iv)* and adherence to individual-level infection-control measures.

Experiments design and results

Various scenarios were analyzed to evaluate how the spread of COVID-19 in the school environment could have been influenced if an alert had been issued in February 2020 to slow the infection, focusing the attention on the Piedmont region of Italy.

Using the macro-level model described in the previous section, the daily prevalence of the virus in the Piedmont population was generated, using Equation 5.1. The average parameters defining the infection in the agent-based model align with those in the compartmental model and are as follows: *i)* a 5-day incubation period after infection, *ii)* a 5-day infectious period, and *iii)* an immunization period of one and a half years. The agent-based model also includes the possibility of individuals becoming susceptible again.

Scenario's name	**Day of the alert**	**Mask fraction**	**Mask type**
No Countermeasures	No alert	0%	None
Surgical 20%	February 23rd, 2020	20%	Surgical
FFP2 20%	February 23rd, 2020	20%	FFP2
Surgical 40%	February 23rd, 2020	40%	Surgical
FFP2 40%	February 23rd, 2020	40%	FFP2
Surgical 80%	February 23rd, 2020	80%	Surgical
FFP2 80%	February 23rd, 2020	80%	FFP2

Table 5.3: Scenarios considered both with the compartmental and the agent-based model. For each scenario, the macro-level model provides the prevalence in the broader population based on the specified parameters, which is then fed into *F4F*. Surgical and FFP2 masks have an inhalation and exhalation efficacy of 59% and 90%, respectively [200].

Table 5.3 outlines the scenarios considered both with the compartmental and the agent-based model. In particular, we included a baseline scenario with no PPE and no public restrictions. Next, we considered six scenarios in which a medical alert was issued on February 23$^{\text{rd}}$, 2020, varying the level of restrictions, particularly considering only the use of masks from that day. In these simulations, we calculated the spread of the disease assuming that *i)* 20%, *ii)* 40%, and *iii)* 80% of the population wore masks, with distinctions made between surgical masks (59% inhalation/exhalation efficacy) and FFP2 masks (90% inhalation/exhalation efficacy). For each scenario, the macro-level model calculates the prevalence in the broader population using the specified parameters, reflecting in this way the effects of government-imposed restrictions in the school. This information is subsequently fed into the dashboard—that is, into the agent-based model. Finally, we run 200 simulations of 90 days each, using February 19$^{\text{th}}$, 2020 as the starting date, for each scenario.

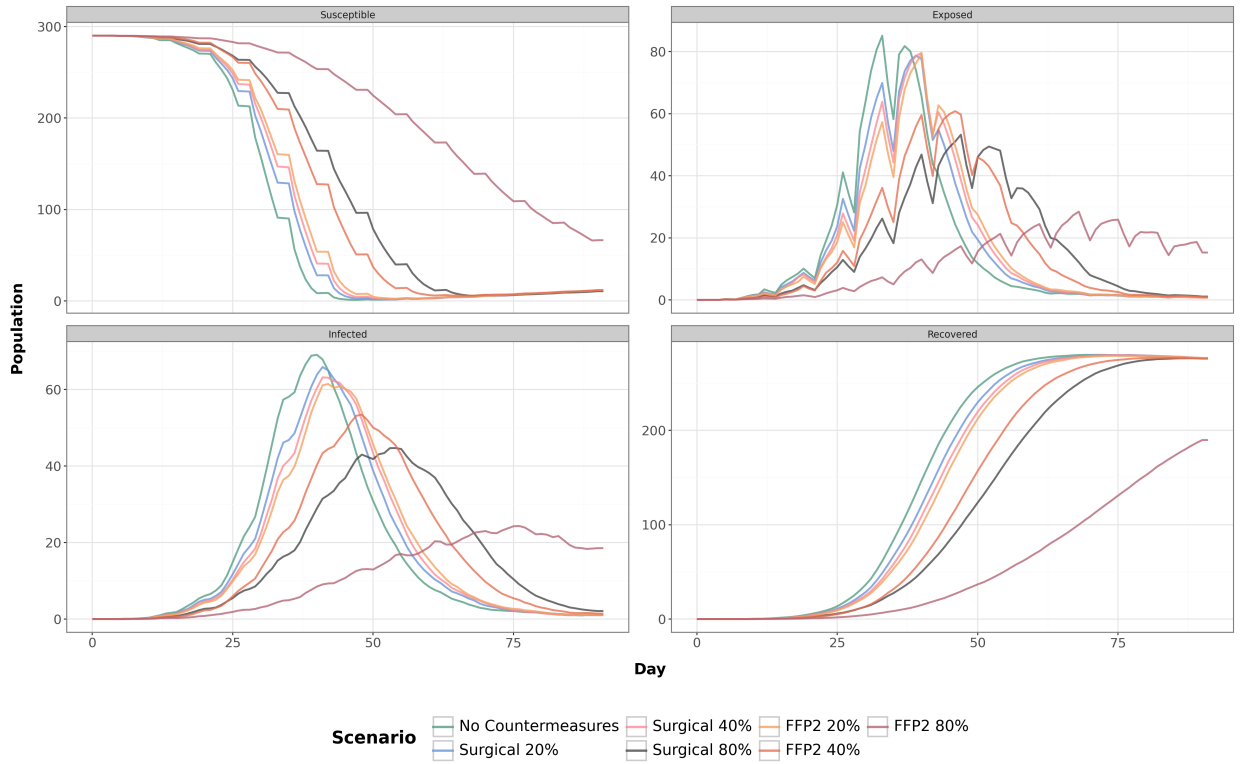


Figure 5.8: Trajectories of the susceptible, exposed, infected, and recovered populations across the seven scenarios considered.

Figure 5.8 shows the evolution of the four disease states in the micro-simulation across all scenarios. As expected, the virus spread rapidly within the school in the *No Countermeasures* scenario—it served as the baseline for comparison with subsequent scenarios. In contrast, the scenario *FFP2 80%* resulted in the slowest spread, with fewer infections as susceptible individuals were less likely to become exposed. Moreover, the results show that the three

scenarios *Surgical 20%*, *Surgical 40%*, and *FFP2 20%* produced similar outcomes, while both *FFP2 40%* and *Surgical 80%* slowed the infection spread.

Figure 5.9 shows the difference in the percentage of infected individuals on the 30th day of the simulation within the school across all scenarios, compared to the baseline. This figure confirms previous findings, where the *FFP2 80%* scenario had 90.53% fewer infections, making it the most effective, while the *Surgical 20%* scenario resulted in a 21.4% reduction in infections. We also observe that the *FFP2 20%* scenario led to a 39.34% reduction in infections, which is a positive outcome, considering the relatively low percentage of mask utilization.

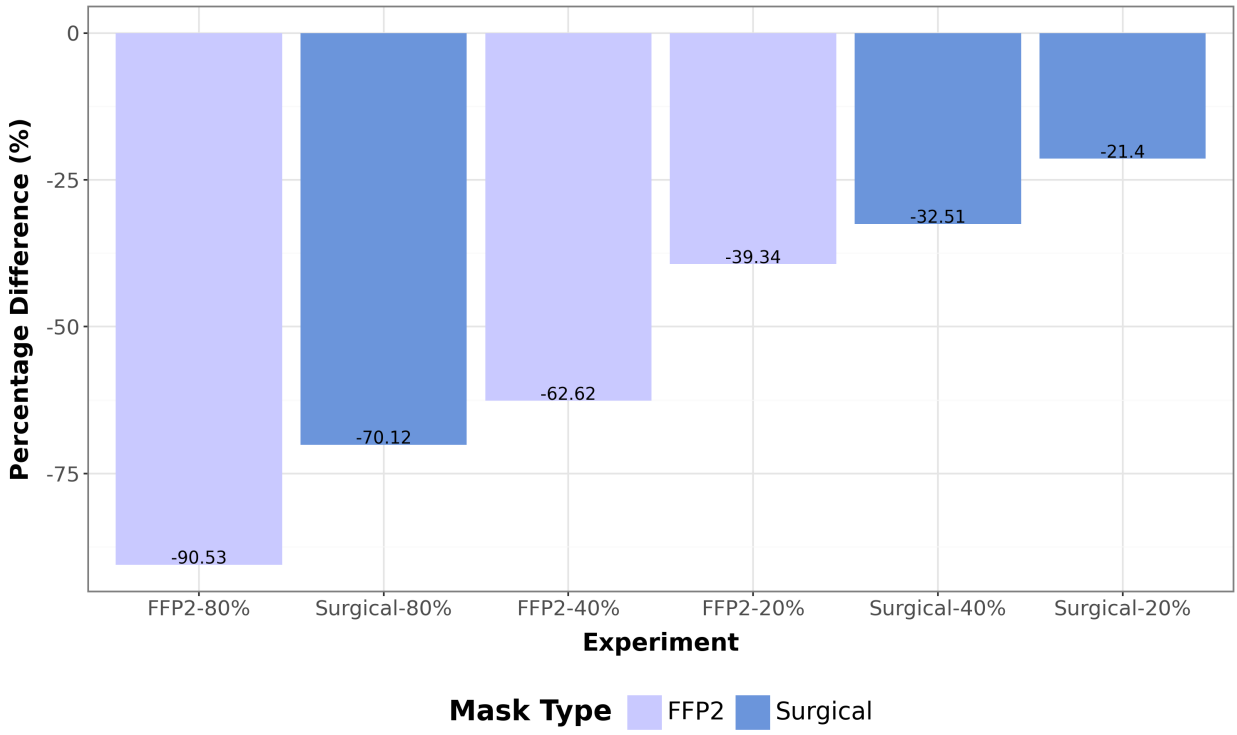


Figure 5.9: Percentage reduction in the infected population compared to the baseline, where no countermeasures were applied.

5.4 Discussion

Modeling and studying infectious disease spreading is a challenging task that requires domain knowledge and expertise. Agent-based models serve as a valuable tool for studying disease transmission, especially within closed environments. However, ABMs are complex and require advanced programming capabilities. In addition, the main limitation of ABMs, i.e. the challenge of constructing realistic well-defined close environments, makes it even more complicated. To solve this problem, we developed *Forge4Flame*, a user-friendly dashboard—developed in R Shiny—designed to simplify the definition of an agent-based model for FLAME GPU 2 agent-based models, automatically generating the necessary code, with the main aim of simplifying the modeling process, making it accessible to users without prior coding experience.

As previously mentioned, **this tool represents a novel contribution, with no equivalent counterpart currently available in the existing literature**. This originality positions *F4F* as a one-of-a-kind tool, addressing gaps in existing methodologies and providing unique capabilities for agent-based modeling. The integration of Docker ensures that the tool and simulations are both reproducible and portable, significantly enhancing usability and accessibility. In addition, we developed a script for running simulations on HPC systems using Slurm, which represents one of the key innovations of our work.

To validate the tool and compare execution times between NetLogo and FLAME GPU 2, we reproduced a model developed in previous studies [41, 40]—a generic middle school environment. Notably, the time required to construct the environment using *F4F* is significantly shorter compared to the time needed in NetLogo or when using FLAME GPU 2 without the tool. As demonstrated in Section 5.3.2, FLAME GPU 2 outperforms NetLogo in terms of execution time while providing statistically equivalent results.

Finally, to model a realistic probability of infection outside the agent-based model environment, we integrate a macro-level compartmental model with the micro-level model, which represents another key innovation of our work—as part of a novel modeling paradigm, as depicted in Figure 5.1. Specifically, we demonstrated how the prevalence of the virus in the broader population can be incorporated into the agent-based model using *F4F*, enabling the exploration of various what-if scenarios. This approach demonstrates a potential method for connecting large-scale epidemiological trends with detailed, localized transmission dynamics, offering insights into how broader regional patterns might influence disease spread in specific settings.

5.5 Conclusion

We developed *Forge4Flame (F4F)*, an intuitive dashboard designed to streamline the process of defining agent-based models for FLAME GPU 2. This tool solved the main limitation of ABMs, namely the difficulty of constructing a realistic environment, like a school or a hospital.

F4F simplifies the modeling process, automatically generating the required code, making it accessible even to users with no prior coding experience, promoting interdisciplinary collaboration, and enhancing the efficiency of disease model development, execution, and analysis. The tool also accelerates the workflow, using Docker and Slurm for efficient distributed computing on HPC systems, and integrates macro-level population models with micro-level individual simulations, enabling a detailed analysis of disease dynamics.

The comparison with NetLogo confirmed that GPUs are crucial for working with detailed ABMs, providing a significant performance boost when transitioning from CPUs to GPUs. These findings demonstrate the potential of *F4F* to enhance both the accessibility and scalability of agent-based modeling in various research fields.

The main limitations of the tool are as follows: *i)* it is specifically designed for modeling infectious disease spread in well-defined closed environments and *ii)* it does not currently allow users to define custom policies directly within *F4F*—instead, users must modify the generated files by writing code manually.

We have many ideas for expanding the tool, including the definition of an advanced section within the dashboard that would allow users with some coding experience to implement custom functions, such as personalized policies. Additionally, we are considering features like allowing probabilistic components in the determined flow. Imagine that an agent has a probability p of entering a room r at a certain point in their determined flow; currently, the user has to define two different agents: one with room r in the determined flow and one without it, while correctly adjusting the number of agents or the rates associated with them. Other potential enhancements include more visualization tools, such as heatmaps to show contact locations or virus concentration in rooms or simple visualizations of disease evolution, the possibility to define rooms without walls to model corridors or open spaces, the possibility to add obstacles inside rooms, and many more.

Finally, a more biologically-focused version of the tool could expand its applicability to other fields, such as modeling cellular interactions for cancer research, simulating immune system responses, or studying the spread of infections at the microscopic level. These adaptations would open new options for using the tool in a wide range of biological and medical research.

5.6 Availability of data and software code

Here are the links to the repositories: `https://github.com/qBioTurin/FORGE4FLAME` and `https://github.com/qBioTurin/FLAMEGPU-FORGE4FLAME/`. To install and run the tool and reproduce the results, see the user guide at `https://github.com/qBioTurin/FORGE4FLAME/blob/main/UserGuide.pdf`.

Chapter 6

Conclusion and Future Works

This thesis explores computational epidemiology, with a primary focus on forecasting scenarios using a combination of computational models and knowledge extraction techniques. Focusing on COVID-19 as a case study, the aim is to support policymakers in effectively managing public health emergencies. We began by providing an overview of the central topics of the thesis, particularly continuous monitoring and early warning systems, emphasizing their role in forecasting potential surges, managing containment efforts, and evaluating the effectiveness of public interventions. Specifically, we concentrated on pandemics as a case study, with a particular emphasis on computational epidemiology. We introduced the tools and techniques essential for enhancing the effectiveness of these systems, such as computational models and knowledge extraction methods—this includes the use of data analysis, particularly for big data, and artificial intelligence to improve predictions and decision-making—, while also discussing how these approaches can be effectively combined. Next, we presented the contributions of the thesis and examined the state of the art relevant to our research.

In Chapter 2, we evaluated the impact of various NPIs and vaccination strategies on reducing COVID-19 infections in a generic middle school environment using an agent-based modeling approach developed in NetLogo [285]. Initially, we considered a baseline setting where no interventions are implemented. That allowed us to understand better the effect of the vaccination alone and what happens when NPIs are introduced. Then, we compared the impact of wearing masks all the time and the implementation of a testing protocol, using two quarantine policies used in the Piedmont region of Italy in 2021 [208] and 2022. The computational model allowed us to pinpoint several useful takeaways for policy-makers. Overall, vaccination campaigns covering less than 40% of the susceptible population provide poor results in protecting from a major outbreak. We observed that the combination of different NPIs can drastically reduce virus transmission in the school environment after the introduction of an infected case. Our model shows that, in the absence

of containment measures, the introduction of a single infective case can lead to the infection of the majority of school students. However, the combination of mask-wearing and room air ventilation can drastically reduce, on average, viral transmission. In our simulation, after introducing one infected student, the use of these two NPIs decreased the cumulative number of infected students by 79%. We tested the effect of different screening policies under different participation levels. Our model suggests that screening can further reduce viral transmission through early identification and isolation of infective cases. In addition, even with low participation, screening can still have a non-negligible impact in terms of COVID-19 cases avoided [41].

Moreover, we showed that testing campaigns are effective as far as the testing process is faster than the virus diffusion. On the other hand, wearing masks properly is an effective countermeasure also when the infection speeds up. Eventually, the introduction of an infection external to the school allows us to consider different settings of infection prevalence. In particular, in a setting with a medium or high circulation of the virus, the results show that it is more challenging to contain the spread compared to a low virus circulation level, but a combination of masks and quarantine policies allows to reduce the final number of infections. We concluded that NPIs are useful for infection, especially in the early stages, and can help to keep activities open even in scenarios with higher circulation, but they must be supported by other public health instruments [40].

In Chapter 3, we introduced Sybil [39], an integrated framework designed for continuous monitoring and forecasting, effectively integrating machine learning-based predictive models with compartmental models. Sybil distinguishes itself by delivering forecasts that are not only reliable and replicable but also readily explainable, as evidenced by thorough experimental validation. The adaptability of this innovative approach is clearly demonstrated through its successful application to diverse surveillance data from several European countries—specifically Italy, Austria, Belgium, France, Germany, and Sweden—and some U.S. states. In addition, the ability to provide accurate predictions even in the presence of a peak or a new emerging outbreak and the integration of variants are of vital help to policy-makers. Finally, by utilizing only data from the previous month, Sybil significantly reduces the demand for training data, enhancing computational efficiency and enabling effective operation with limited data.

By integrating data-centric and analytic approaches, Sybil effectively addresses the inherent limitations of each method. This amalgamation not only enhances the tool's ability to forecast the evolution of COVID-19 but also positions Sybil as a versatile instrument for predicting the trajectory of various other diseases, thus broadening its scope and impact in the field of infectious disease modeling by helping policy-makers in quickly reacting to a new emerging

disease.

In Chapter 4, we examined the relationship between NPIs, mobility patterns, and the effective reproduction number R_t of COVID-19, using data from 13 European countries. We observed a strong correlation between NPIs and people's mobility, whereas the correlation between mobility and the R_t was weaker, likely due to the more intricate relationship between these variables. Analyzing some ratios, distinct cultural patterns emerge in the interplay between NPIs, mobility, and the R_t. By clustering countries using these ratios, we find a strong correlation between NPIs and mobility reductions, highlighting the effectiveness of measures in changing public behavior. However, the link between mobility and R_t is weaker, indicating a delayed impact due to factors like social interactions, incubation periods, and pre-existing transmission dynamics. Our analyses also reveal cultural differences, with Northern European countries showing higher compliance and reduced mobility even with less stringent measures. Southern European nations required stricter interventions to achieve similar results. When social interactions dominated, these regional distinctions faded, emphasizing the universal role of interpersonal behavior in transmission.

In addition, we employed XGBoost [77] models to estimate missing mobility data and infer missing R_t values from mobility patterns, offering a novel approach that makes a significant contribution to the existing literature. Our findings underscore the need for public health strategies to be tailored to cultural and regional differences, particularly in terms of compliance with NPIs. At the same time, the methodological framework combining machine learning with epidemiological insights provides a robust tool for analyzing mobility and infection dynamics in future pandemics.

Finally, in Chapter 5, we developed *Forge4Flame (F4F)*, an intuitive dashboard designed to simplify the definition of an agent-based model for FLAME GPU 2. By automatically generating the required code, this tool simplifies the modeling process, making it accessible to users without coding experience, allowing them to define the model's environment, the agents interacting within it, the disease model, and several NPIs through a graphical interface. It enhances interdisciplinary collaboration, improving disease model development and analysis, evaluating health emergencies, and supporting decision-making. The tool also leverages Docker and Slurm for efficient distributed computing on High-Performance Computing (HPC) systems, improving FLAME GPU 2 scalability.

Moreover, integrating macro-level population models with micro-level simulations, *F4F* enables detailed analysis of disease dynamics, such as reflecting the prevalence of the virus in the broader population into the micro-level model. We explored how an early warning alert in February 2020 could have reduced the disease spread in the school environment using a macro-level compartmental model used to study the COVID-19 diffusion in the

Piedmont region [202]. Finally, we validated the tool, replicating previous studies [41, 40] on the COVID-19 emergency in an Italian generic middle school. This demonstrated *F4F*'s ability to accelerate model development and result analysis. This tool simplifies the modeling and analysis of infectious disease spread, offering a unique and comprehensive solution for researchers.

6.1 Future works

This thesis suggests several potential directions for future works:

- Integrate Neural Ordinary Differential Equations (Neural ODEs) and/or Physics-Informed Neural Networks (PINNs) with compartmental models to enhance predictions of a system's future evolution, aiming to advance and refine Sybil's capabilities, while also applying it in new contexts to further increase its impact and versatility.
- In the context of the TrustAlert project, use *F4F* to model the Cottolengo Hospital in Turin, Italy, validating the proposed simulation paradigm introduced in Chapter 5 using retrospective COVID-19 data—Figure 5.1.
- Apply the integration of diverse model types to other domains, such as cybersecurity or city planning, enhancing their ability to provide insights into critical, less-observable aspects.

Appendices

Appendix A

Extended data analysis on NPIs, mobility, infection rates and regression models

In this chapter, we provide additional details on some results presented in Chapter 4. Specifically, we show more details on the preliminary data analysis conducted in Section 4.1.3 for each country (Section A.1) and show more details on the XGBoost [77] regression models outlined in Section 4.3.3 (Section A.2).

A.1 Extended data analysis

In this section, we show detailed plots for each of the 13 countries considered. Specifically, Figures A.1 and A.2 show the time evolution of each of the six considered Facebook mobility variables—*Grocery*, *Residential*, *Stations*, *Recreation*, *Workplaces*, and *Masks*—, the *Stringency* index, and the *Infection Rates* from April 2020 to May 2021. Figures A.5 and A.6 show the same information considering Google mobility variables instead of Facebook ones—except for the Facebook *Masks* variable—from April 2020 to March 2022. Figures A.3 and A.4 show the average Facebook mobility—computed on *Grocery*, *Residential*, *Stations*, *Recreation*, and *Workplaces*—, along with the complement of the *Stringency* index and *Infection Rates*. Figures A.7 and A.8 show the same information considering Google mobility variables instead of Facebook ones from April 2020 to March 2022.

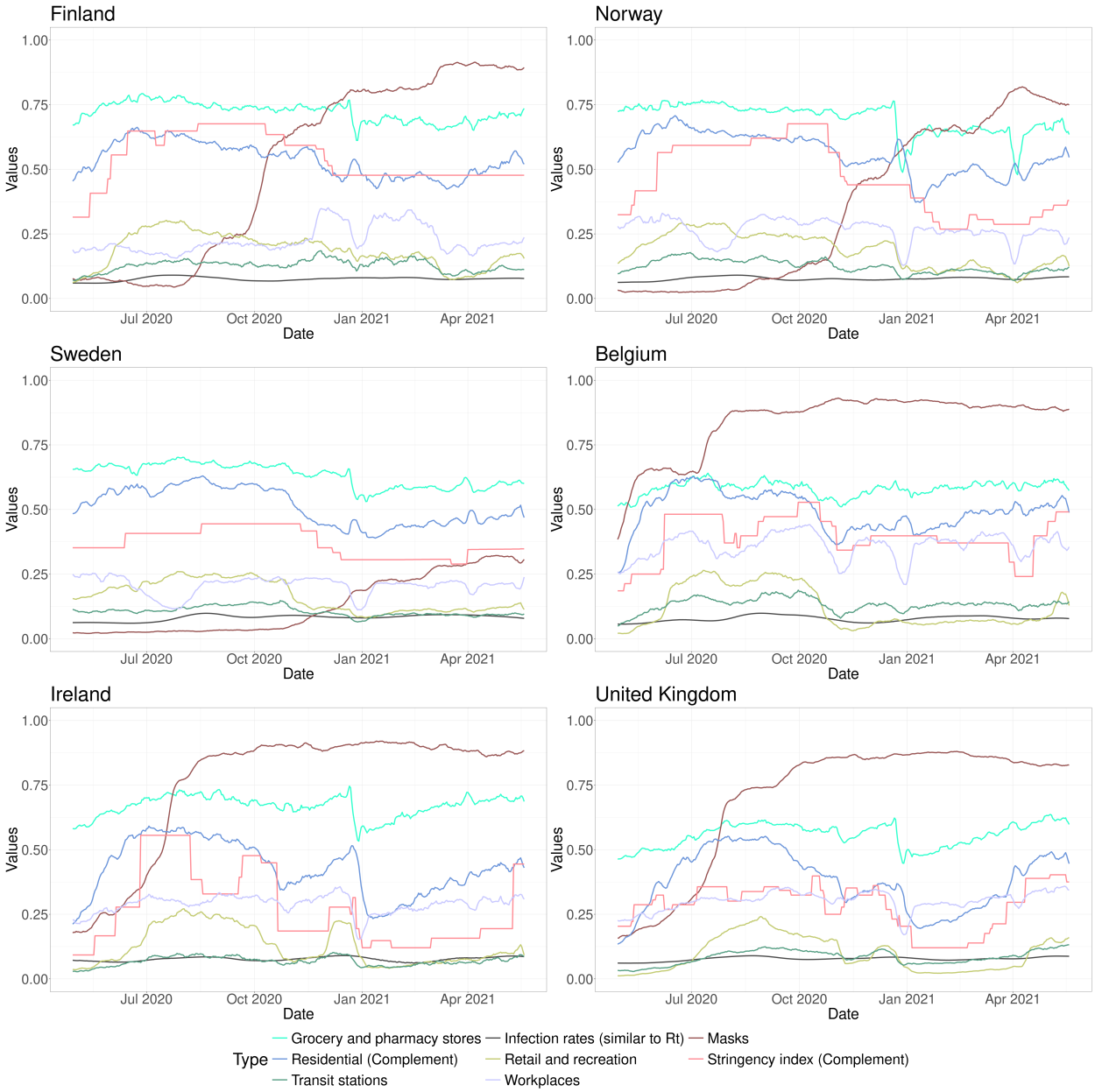


Figure A.1: Evolution of the considered Facebook mobility variables—*Grocery*, *Residential*, *Stations*, *Recreation*, *Workplaces*, and *Masks*—, alongside the complement of the *Stringency* index and *Infection Rates* for Finland, Norway, Sweden, Belgium, Ireland, and the United Kingdom from April 2020 to May 2021.

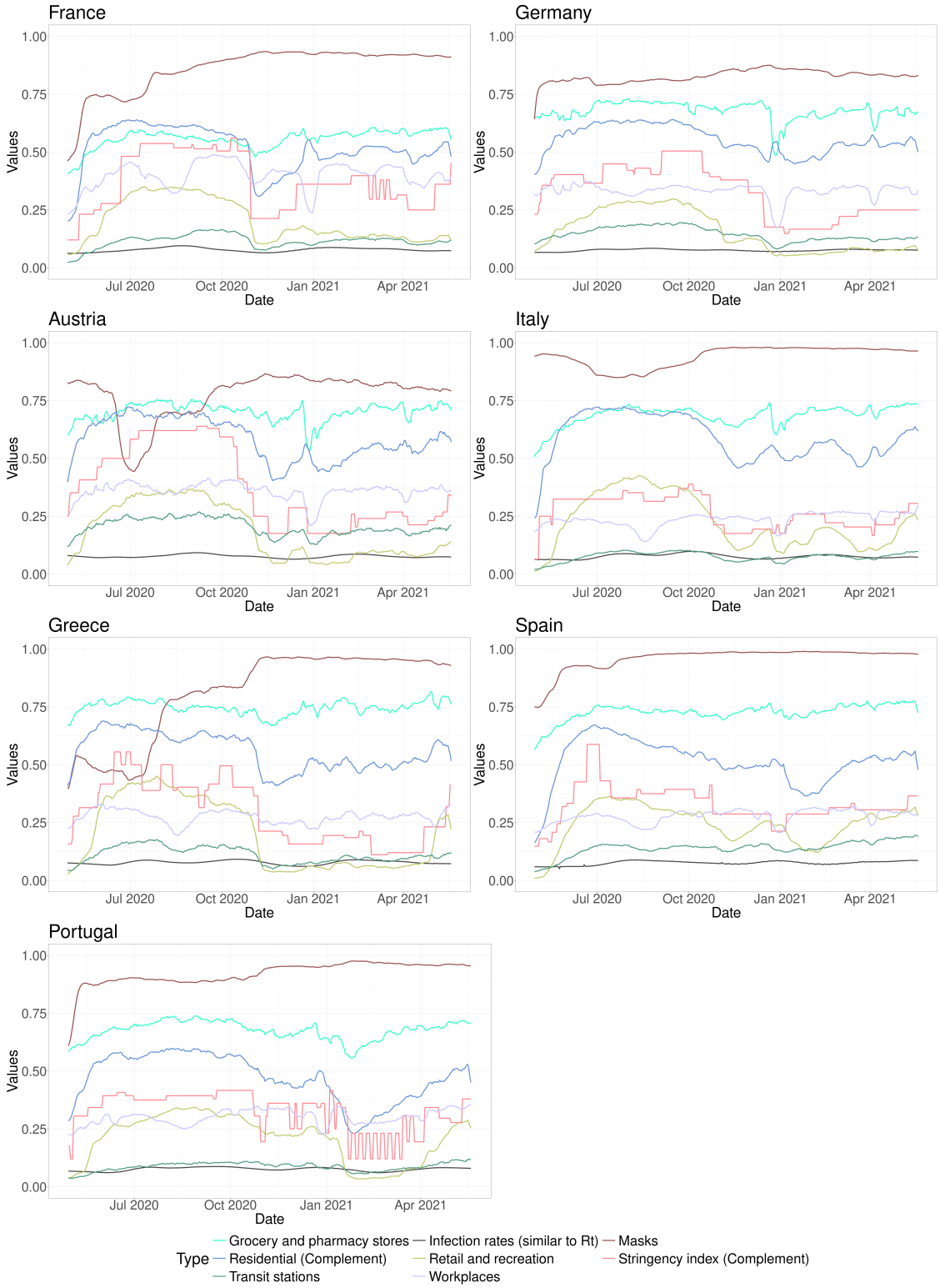


Figure A.2: Evolution of the considered Facebook mobility variables—*Grocery*, *Residential*, *Stations*, *Recreation*, *Workplaces*, and *Masks*—, alongside the complement of the *Stringency* index and *Infection Rates* for France, Germany, Austria, Italy, Greece, Spain, and Portugal from April 2020 to May 2021.

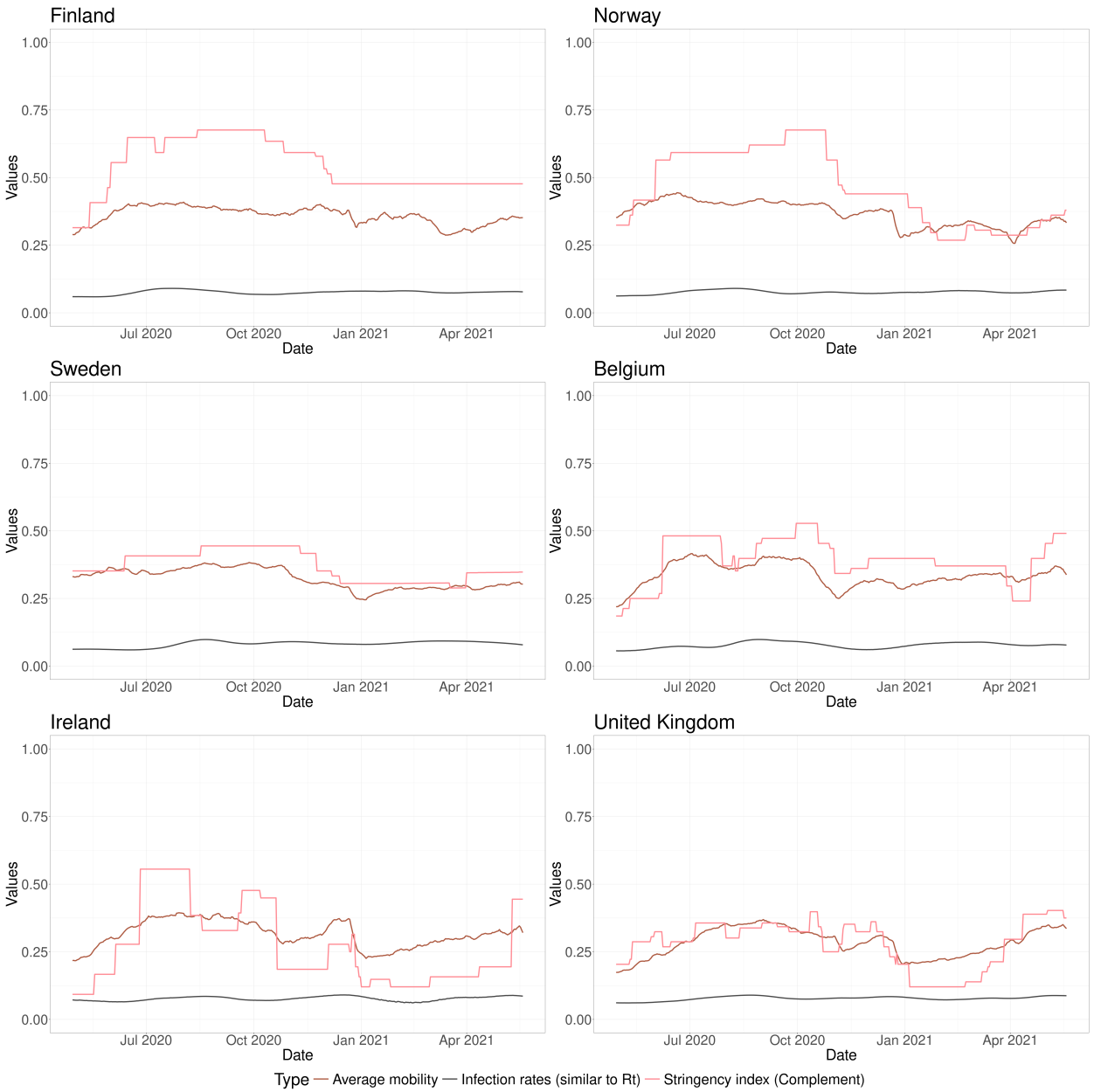


Figure A.3: Evolution of the average Facebook mobility—computed on *Grocery*, *Residential*, *Stations*, *Recreation*, and *Workplaces*—, alongside the complement of the *Stringency* index and *Infection Rates* for Finland, Norway, Sweden, Belgium, Ireland, and the United Kingdom from April 2020 to May 2021.

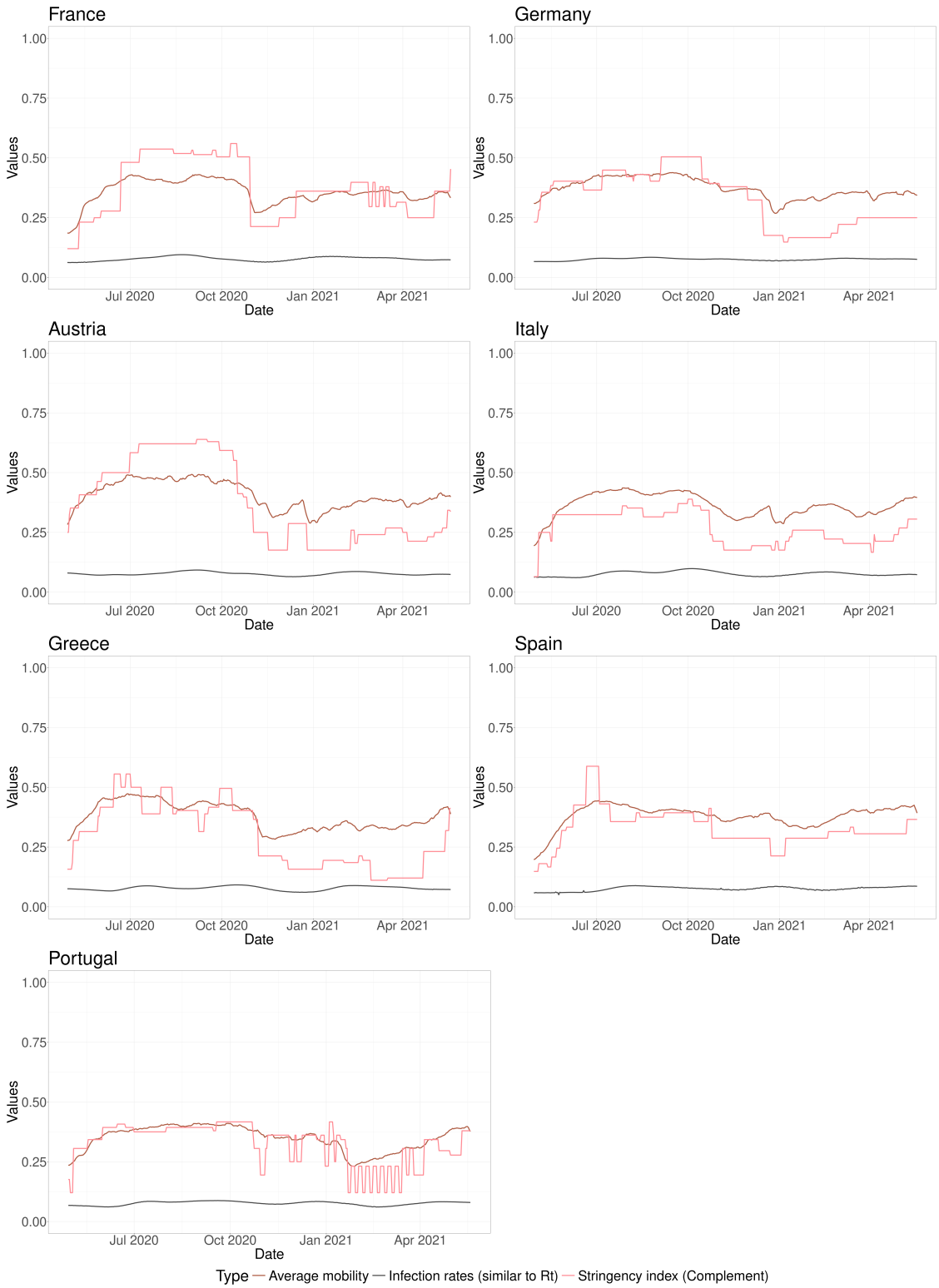


Figure A.4: Evolution of the average Facebook mobility—computed on *Grocery*, *Residential*, *Stations*, *Recreation*, and *Workplaces*—, alongside the complement of the *Stringency* index and *Infection Rates* for France, Germany, Austria, Italy, Greece, Spain, and Portugal from April 2020 to May 2021.

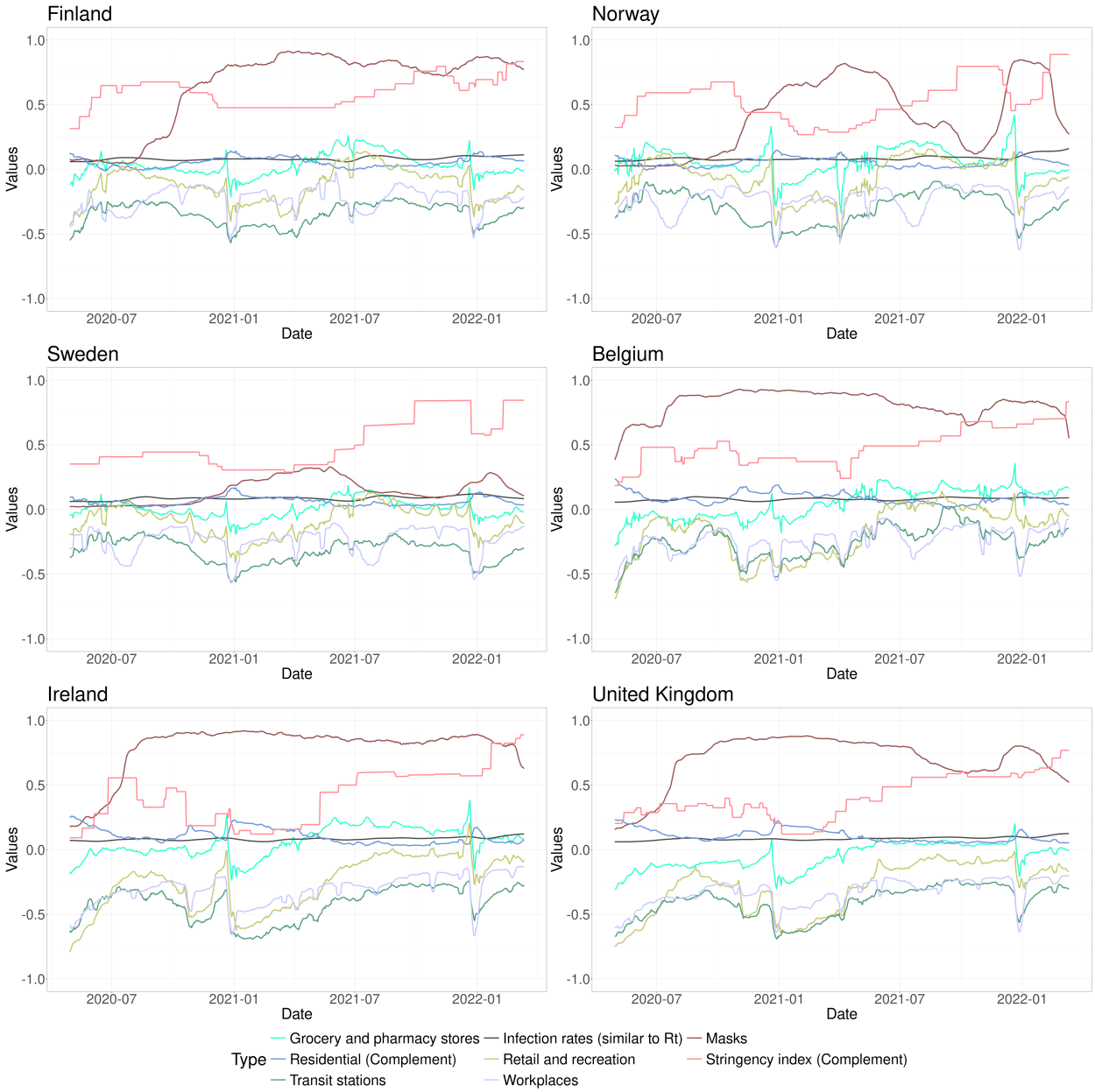


Figure A.5: Evolution of the considered Google mobility variables—*Grocery*, *Residential*, *Stations*, *Recreation*, *Workplaces*, and the Facebook *Masks*—, alongside the complement of the *Stringency* index and *Infection Rates* for Finland, Norway, Sweden, Belgium, Ireland, and the United Kingdom from April 2020 to March 2022.

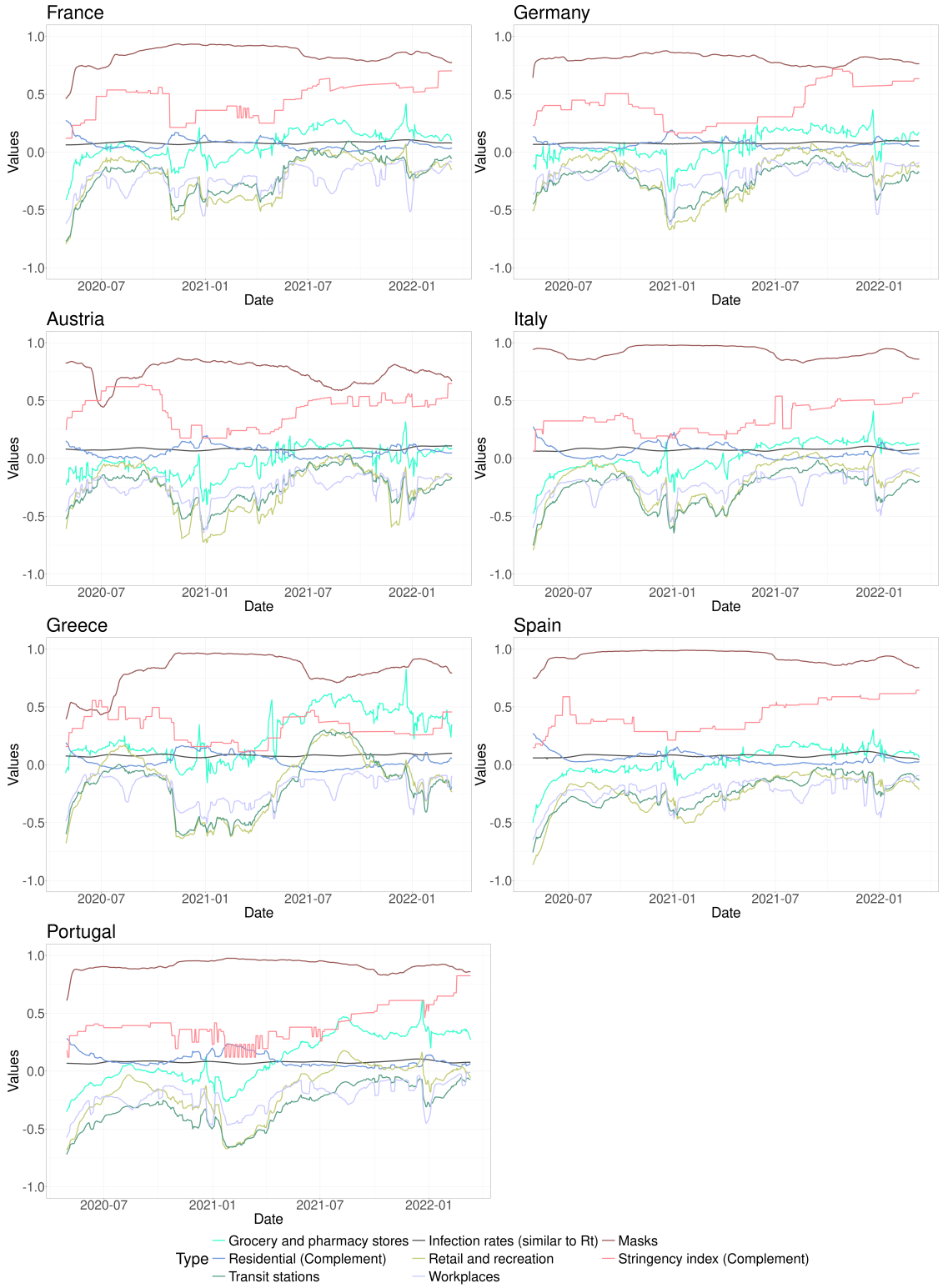


Figure A.6: Evolution of the considered Google mobility variables—*Grocery*, *Residential*, *Stations*, *Recreation*, *Workplaces*, and the Facebook *Masks*—, alongside the complement of the *Stringency* index and *Infection Rates* for France, Germany, Austria, Italy, Greece, Spain, and Portugal from April 2020 to March 2022.

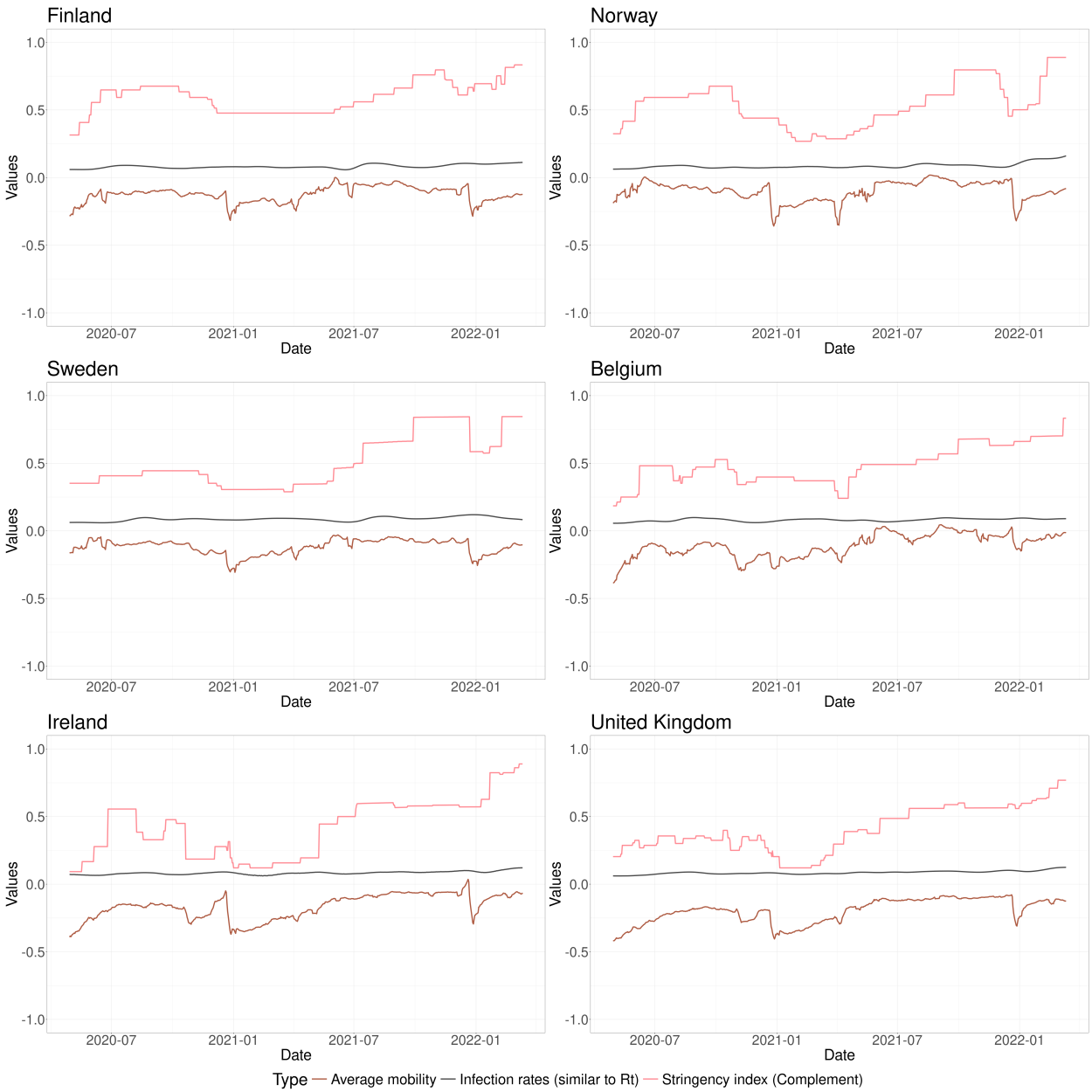


Figure A.7: Evolution of the average Google mobility—computed on *Grocery*, *Residential*, *Stations*, *Recreation*, and *Workplaces*—, alongside the complement of the *Stringency* index and *Infection Rates* for Finland, Norway, Sweden, Belgium, Ireland, and the United Kingdom from April 2020 to March 2022.

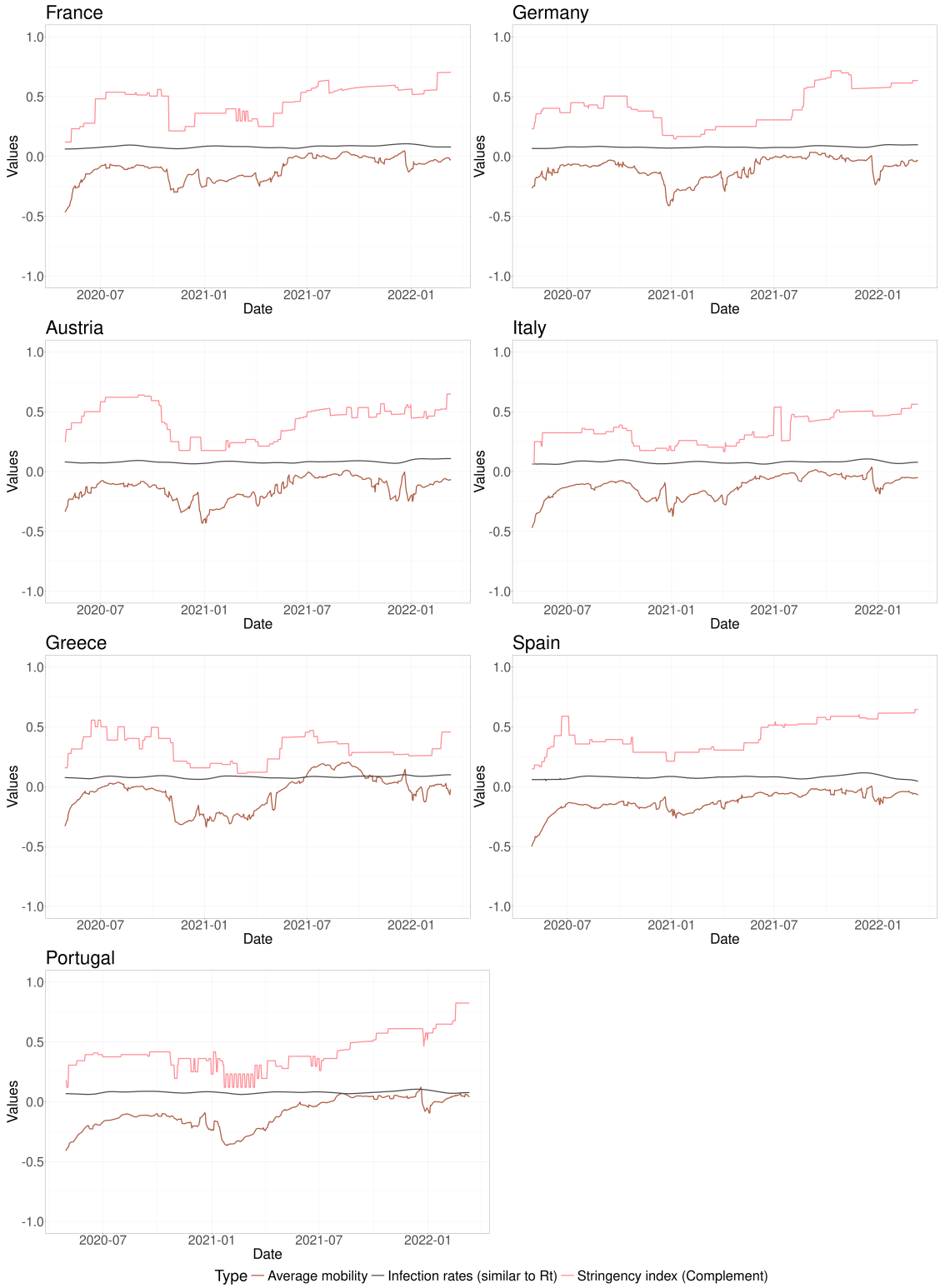


Figure A.8: Evolution of the average Google mobility—computed on *Grocery*, *Residential*, *Stations*, *Recreation*, and *Workplaces*—, alongside the complement of the *Stringency* index and *Infection Rates* for France, Germany, Austria, Italy, Greece, Spain, and Portugal from April 2020 to March 2022.

A.2 Extended results on regression models

In this section, we show more details on the XGBoost [77] regression models used in Chapter 4 to estimate missing mobility data based on NPIs and missing R_t data based on mobility patterns. In particular, we show some plots that illustrate a comparison between the ground truth and the predicted values for the test set (20% of the data), for each country. For brevity, we present results for the first fold only, noting that similar observations apply consistently across the remaining four folds.

Figures A.9 to A.21 show the predictions obtained using XGBoost regression models to estimate missing mobility data based on NPIs, for each mobility variable considered. Figures A.22 and A.23 show the predictions obtained using XGBoost regression models to estimate missing R_t values based on mobility patterns. These figures show that XGBoost models work very well in estimating missing data.

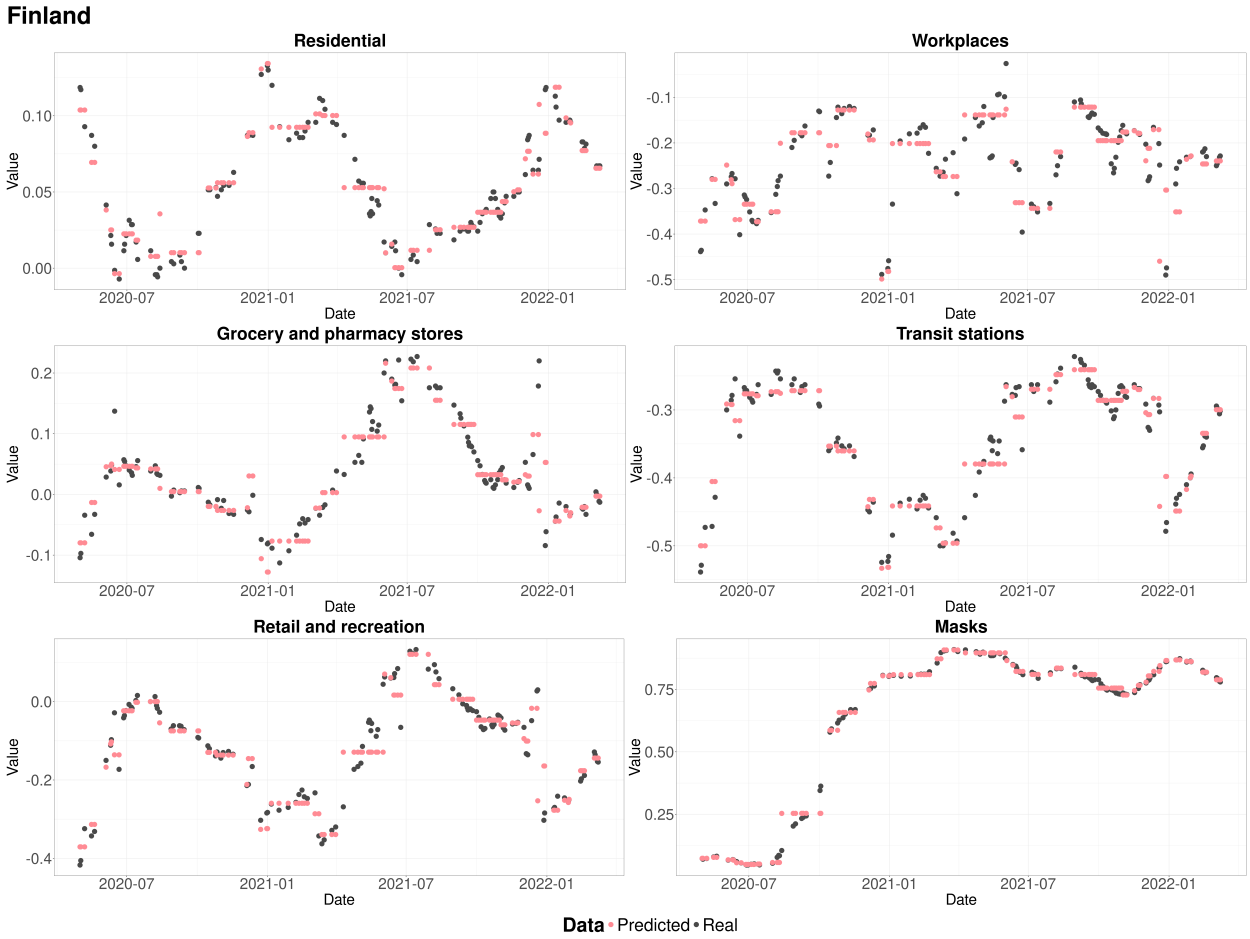


Figure A.9: Comparison between the ground truth and the predicted values for the test set (20% of the data, first fold) for the XGBoost model from NPIs to mobility in Finland.

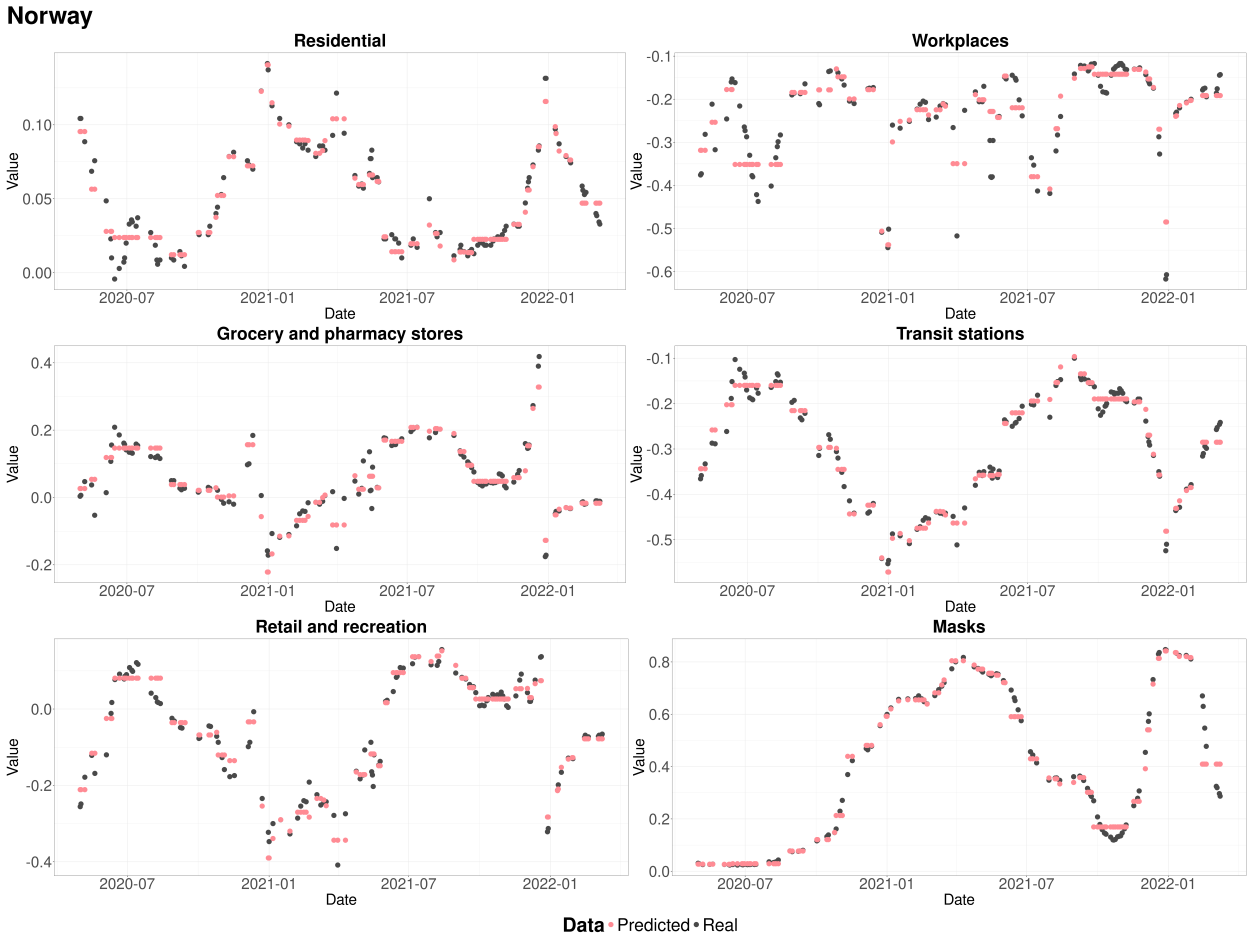


Figure A.10: Comparison between the ground truth and the predicted values for the test set (20% of the data, first fold) for the XGBoost model from NPIs to mobility in Norway.

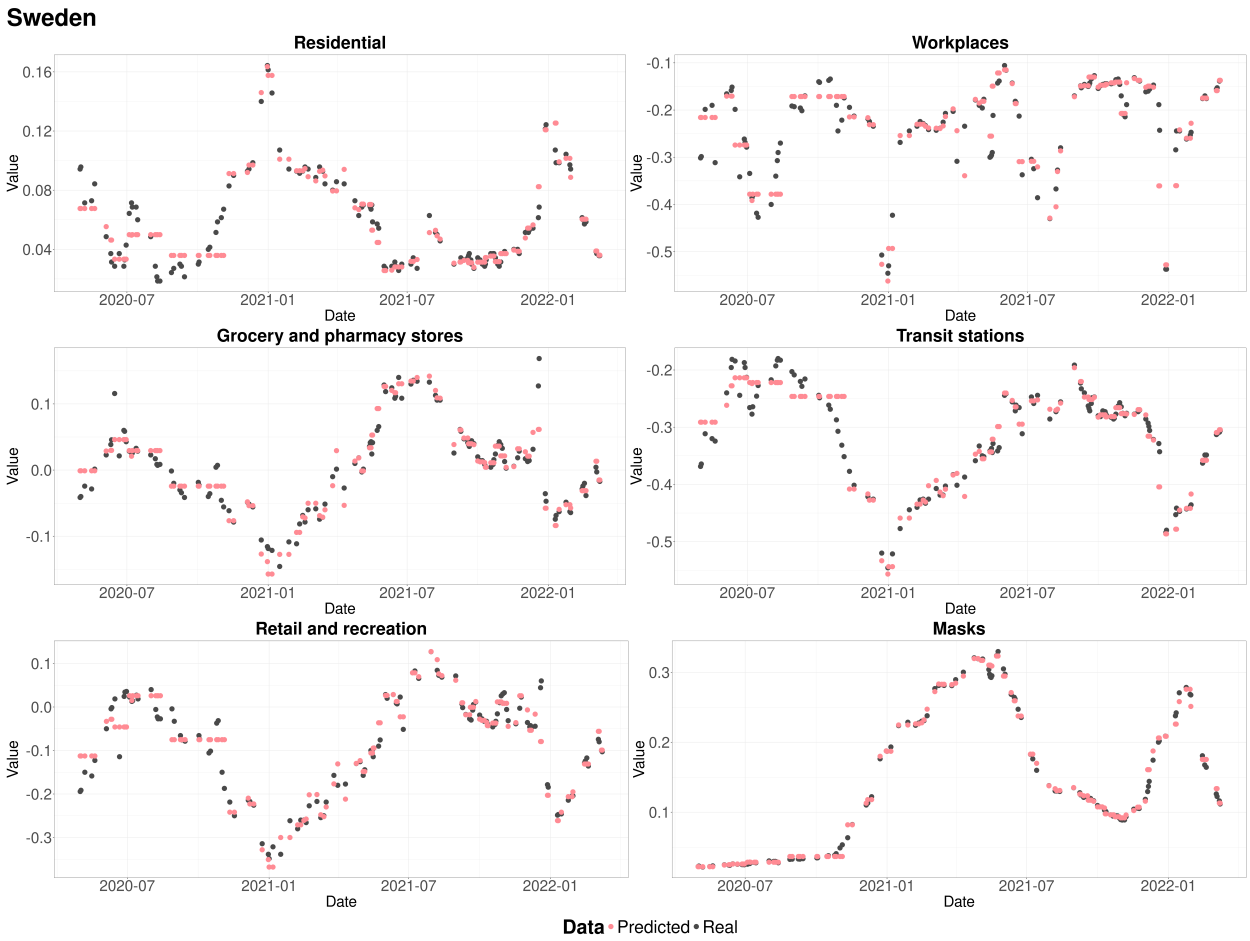


Figure A.11: Comparison between the ground truth and the predicted values for the test set (20% of the data, first fold) for the XGBoost model from NPIs to mobility in Sweden.

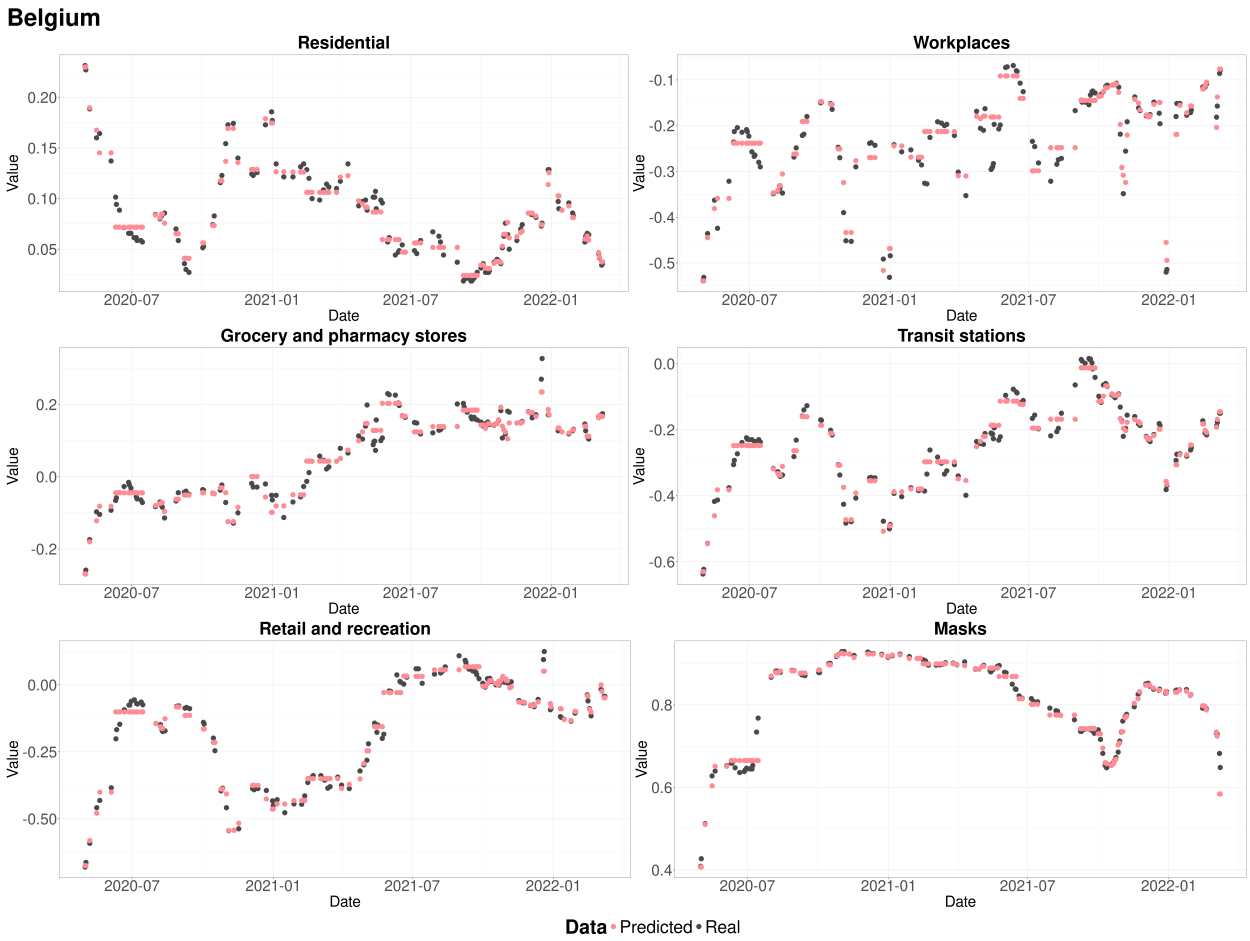


Figure A.12: Comparison between the ground truth and the predicted values for the test set (20% of the data, first fold) for the XGBoost model from NPIs to mobility in Belgium.

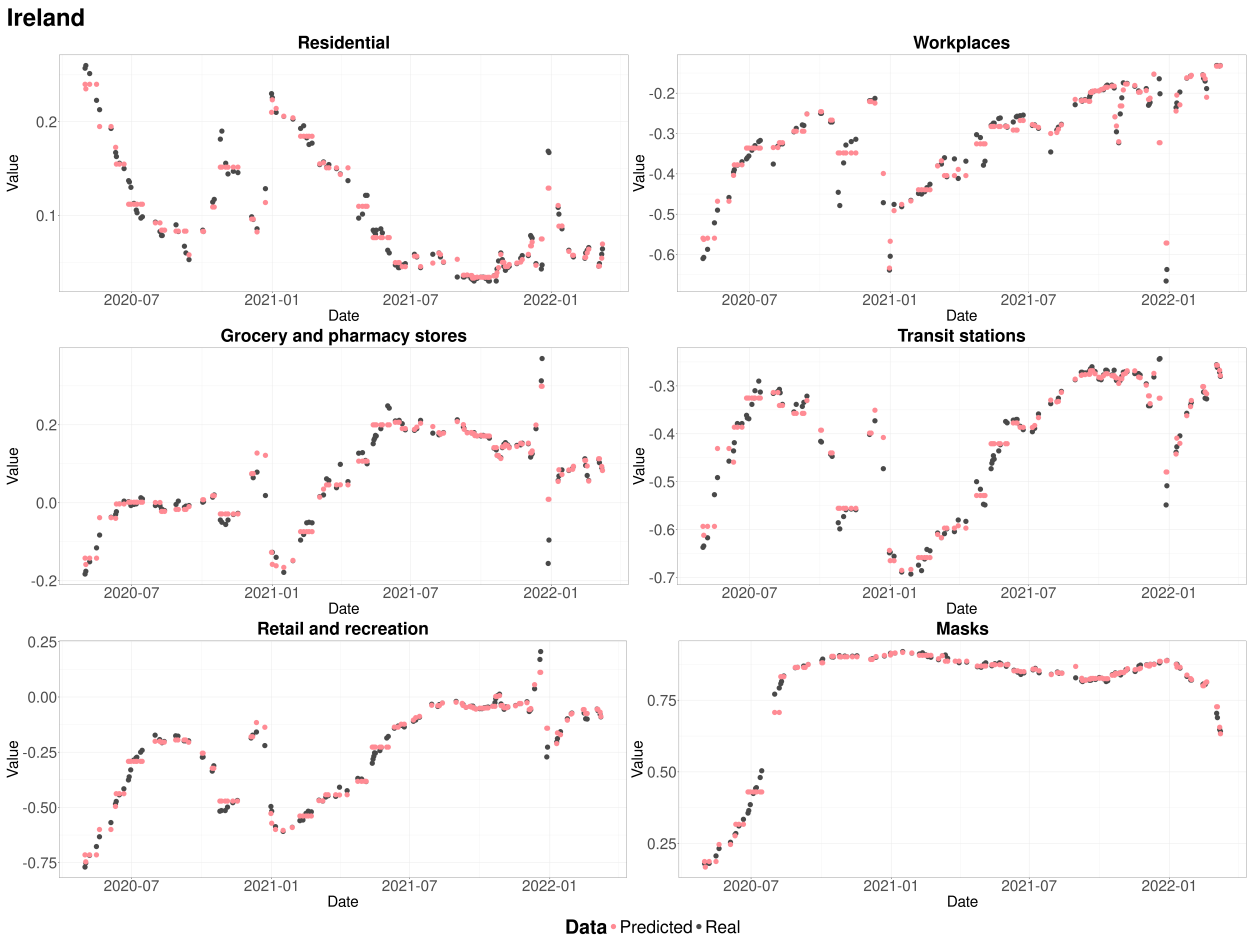


Figure A.13: Comparison between the ground truth and the predicted values for the test set (20% of the data, first fold) for the XGBoost model from NPIs to mobility in Ireland.

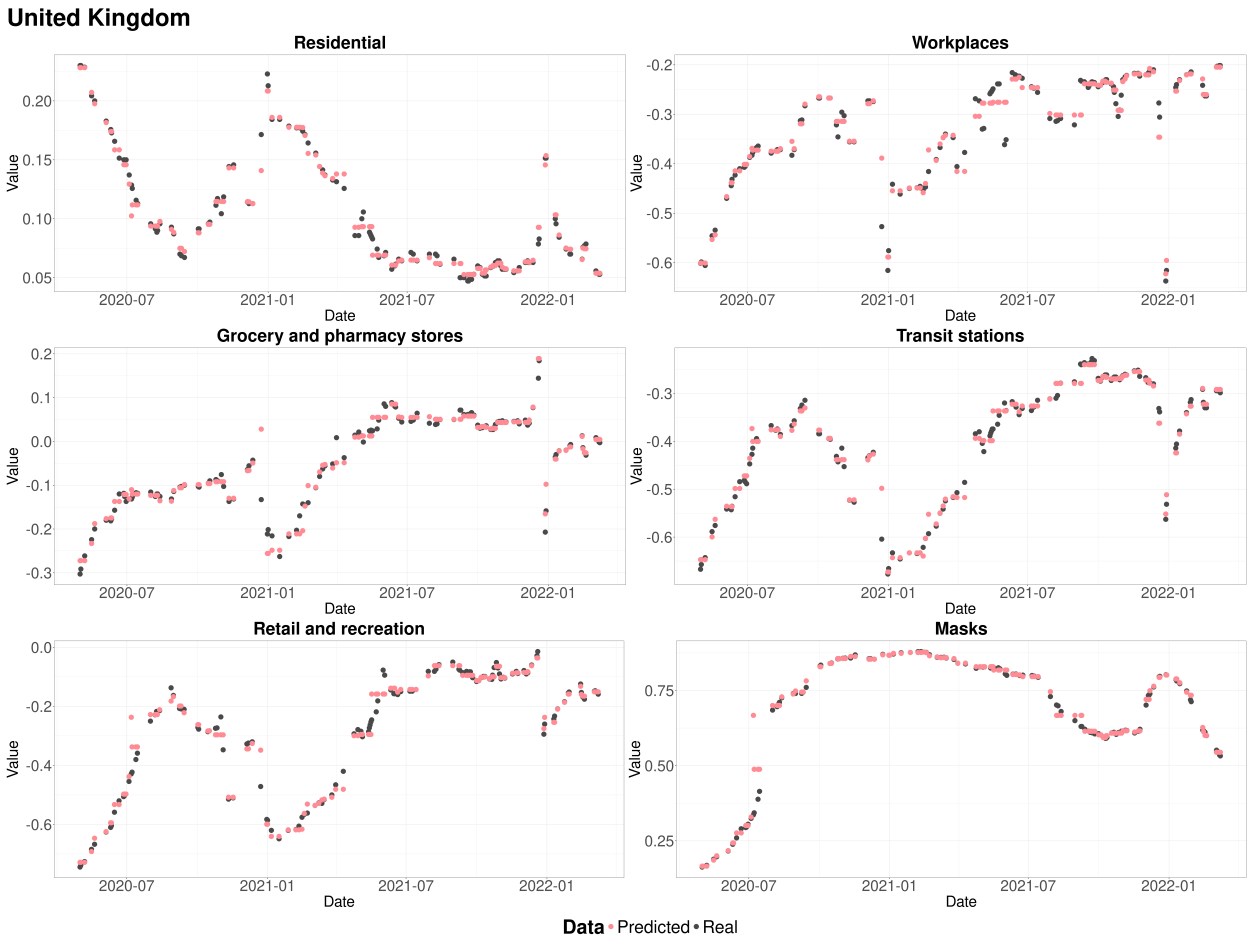


Figure A.14: Comparison between the ground truth and the predicted values for the test set (20% of the data, first fold) for the XGBoost model from NPIs to mobility in the United Kingdom.

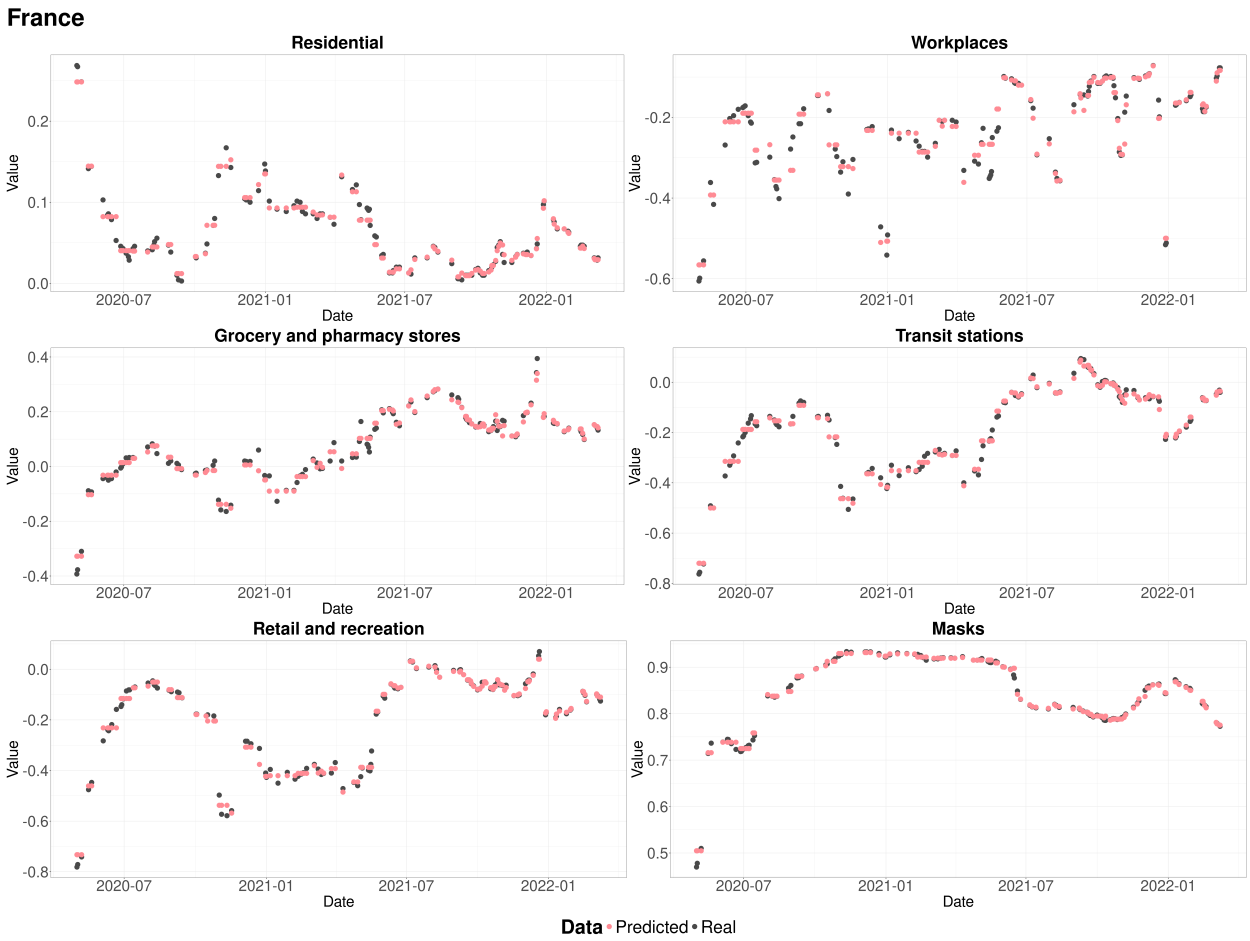


Figure A.15: Comparison between the ground truth and the predicted values for the test set (20% of the data, first fold) for the XGBoost model from NPIs to mobility in France.

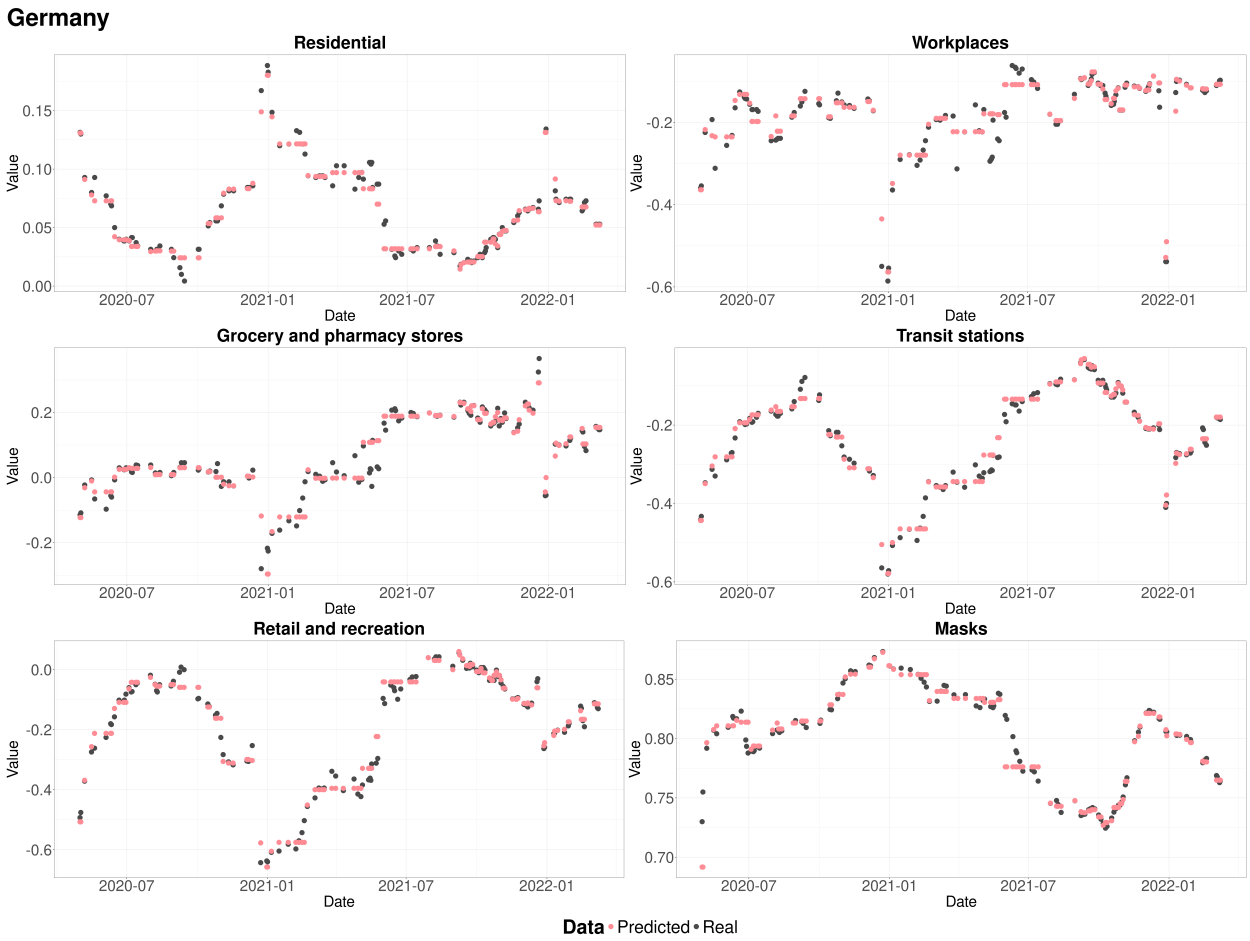


Figure A.16: Comparison between the ground truth and the predicted values for the test set (20% of the data, first fold) for the XGBoost model from NPIs to mobility in Germany.

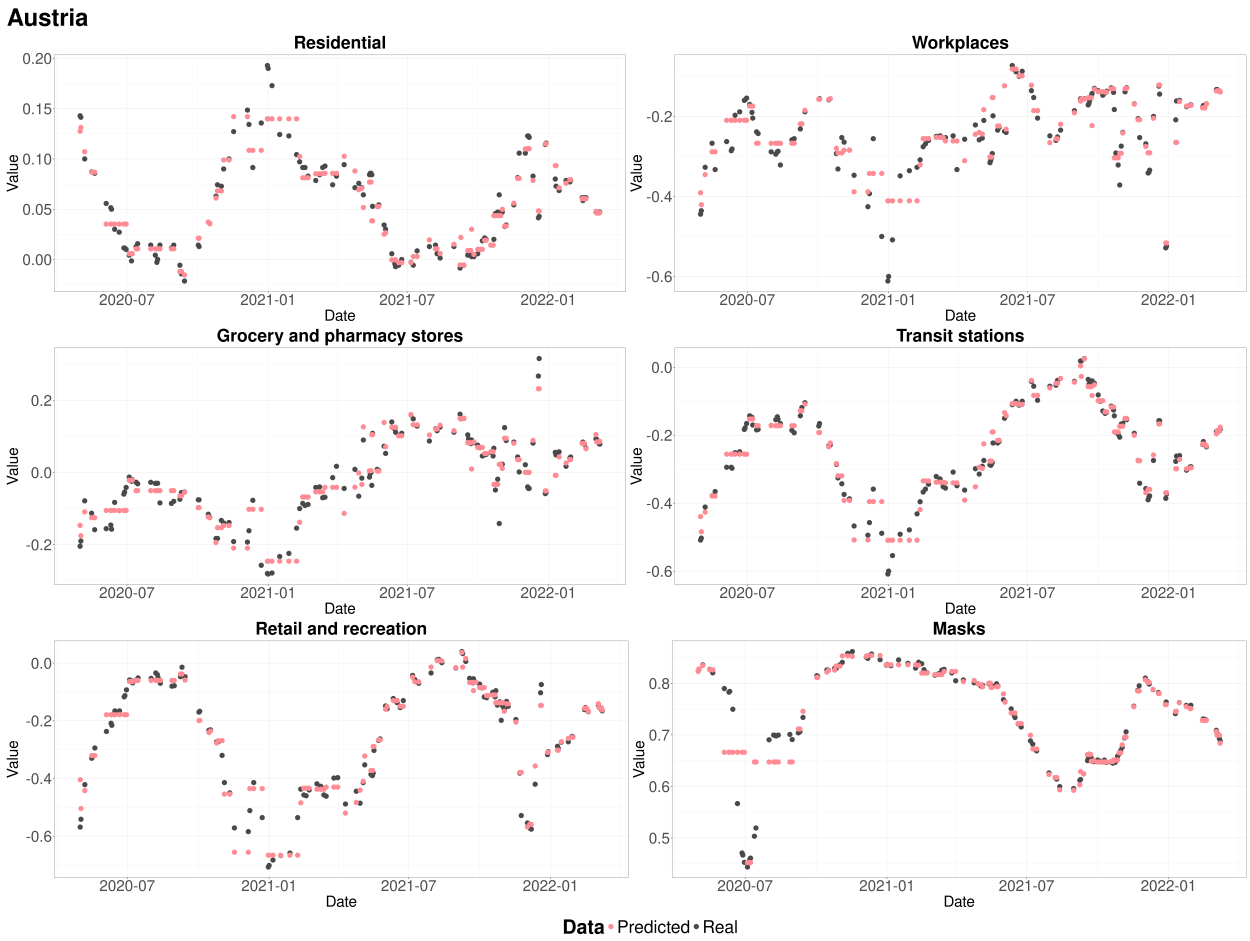


Figure A.17: Comparison between the ground truth and the predicted values for the test set (20% of the data, first fold) for the XGBoost model from NPIs to mobility in Austria.

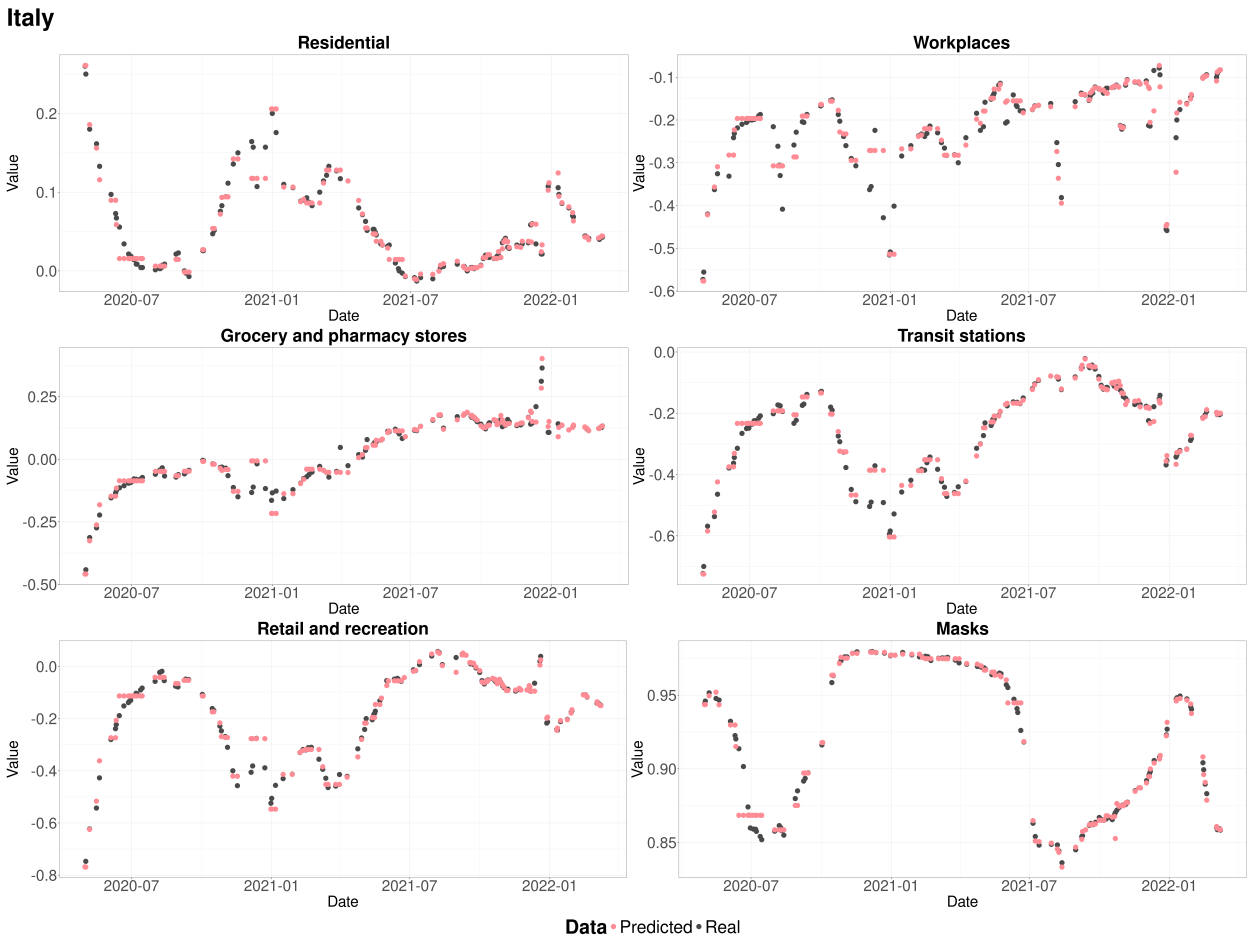


Figure A.18: Comparison between the ground truth and the predicted values for the test set (20% of the data, first fold) for the XGBoost model from NPIs to mobility in Italy.

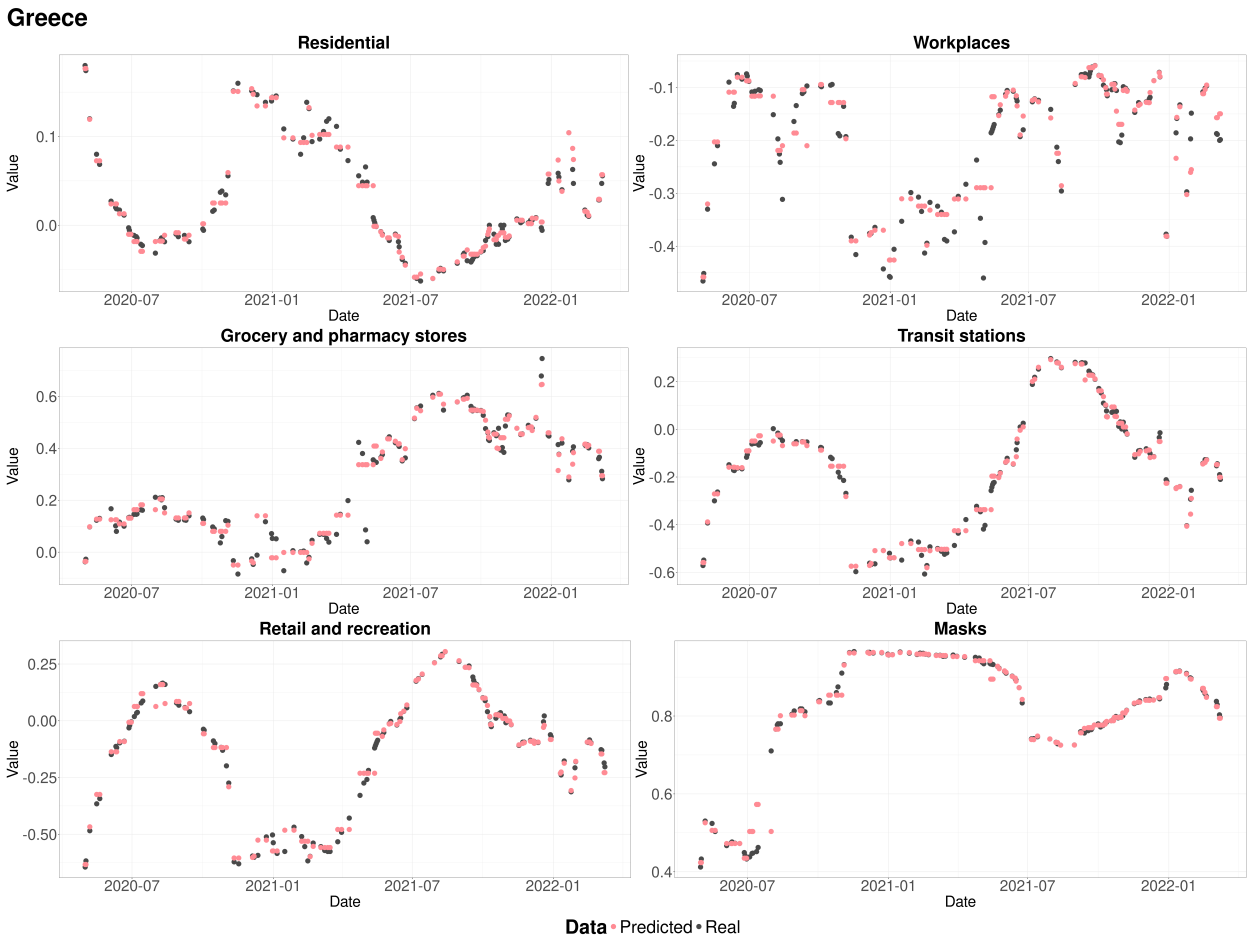


Figure A.19: Comparison between the ground truth and the predicted values for the test set (20% of the data, first fold) for the XGBoost model from NPIs to mobility in Greece.

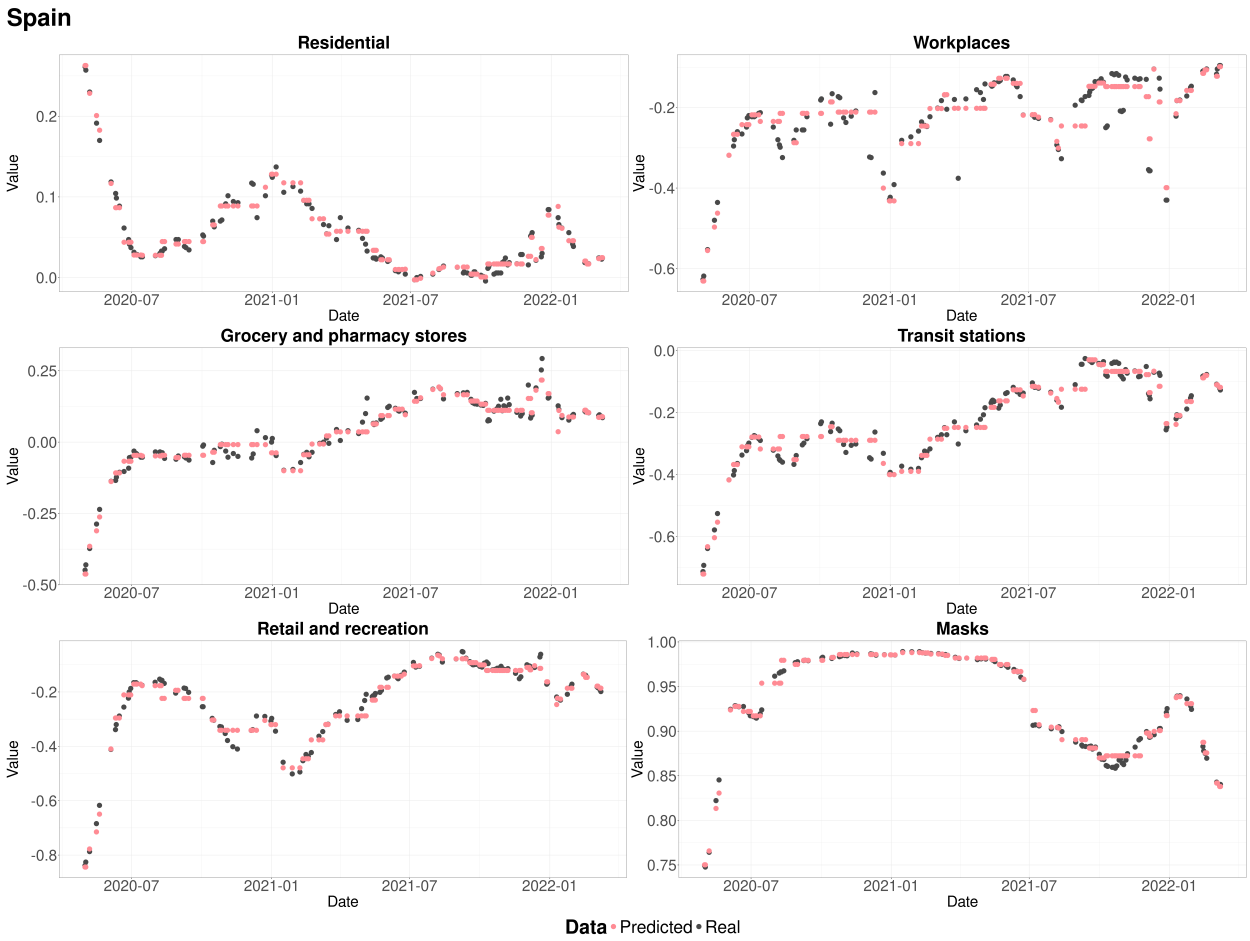


Figure A.20: Comparison between the ground truth and the predicted values for the test set (20% of the data, first fold) for the XGBoost model from NPIs to mobility in Spain.

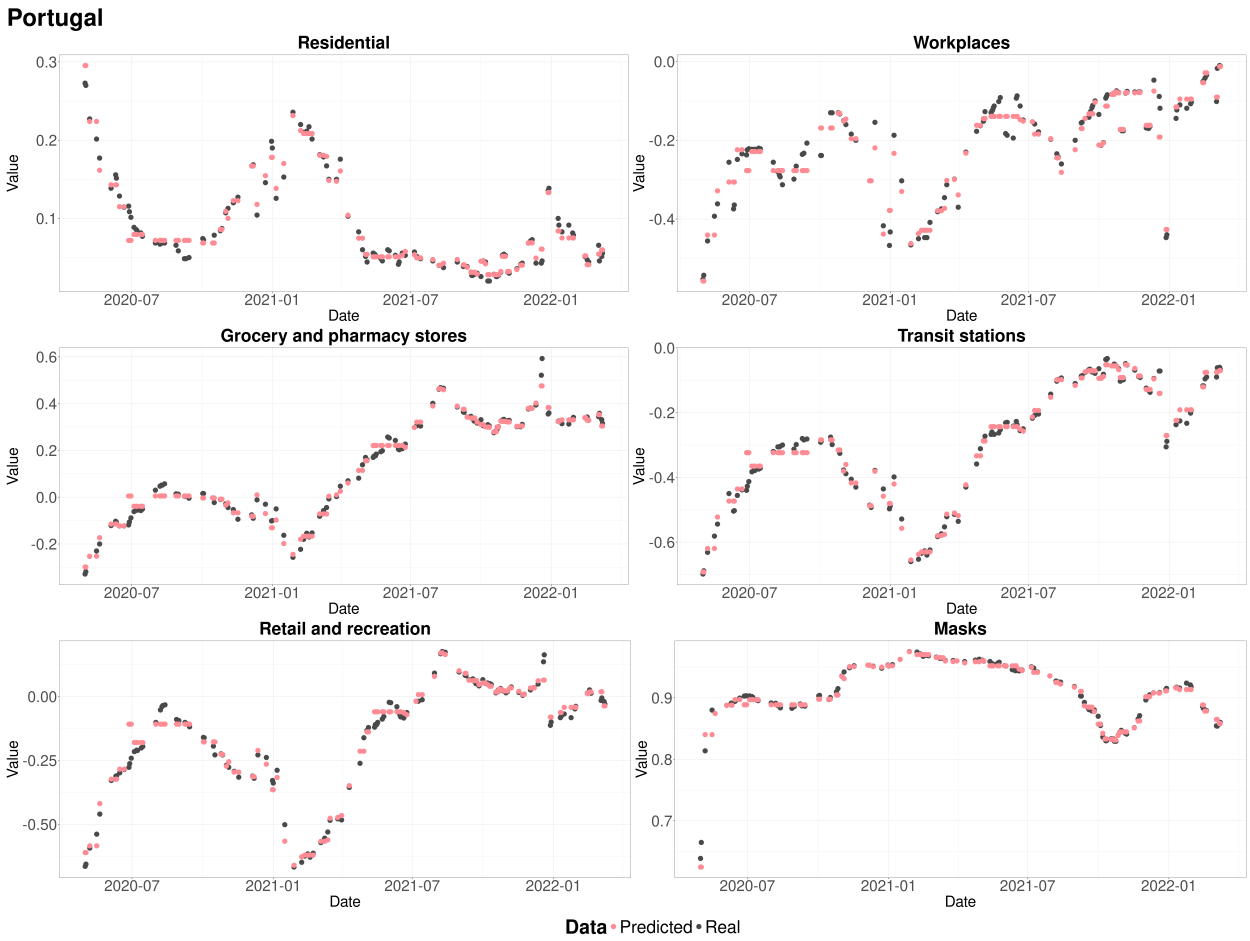


Figure A.21: Comparison between the ground truth and the predicted values for the test set (20% of the data, first fold) for the XGBoost model from NPIs to mobility in Portugal.

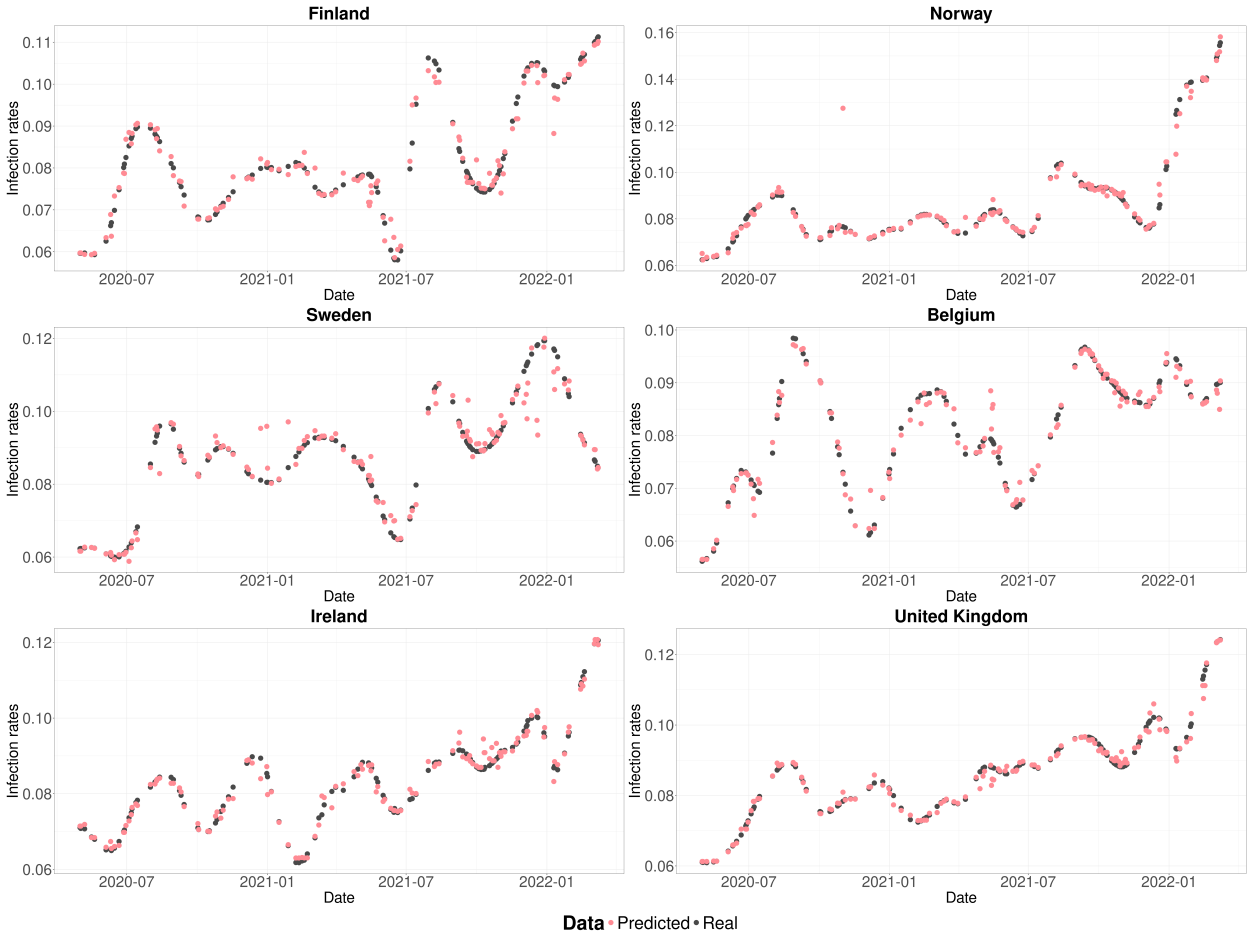


Figure A.22: Comparison between the ground truth and the predicted values for the test set (20% of the data, first fold) for the XGBoost model from mobility to infection rates in Finland, Norway, Sweden, Belgium, Ireland, and the United Kingdom.

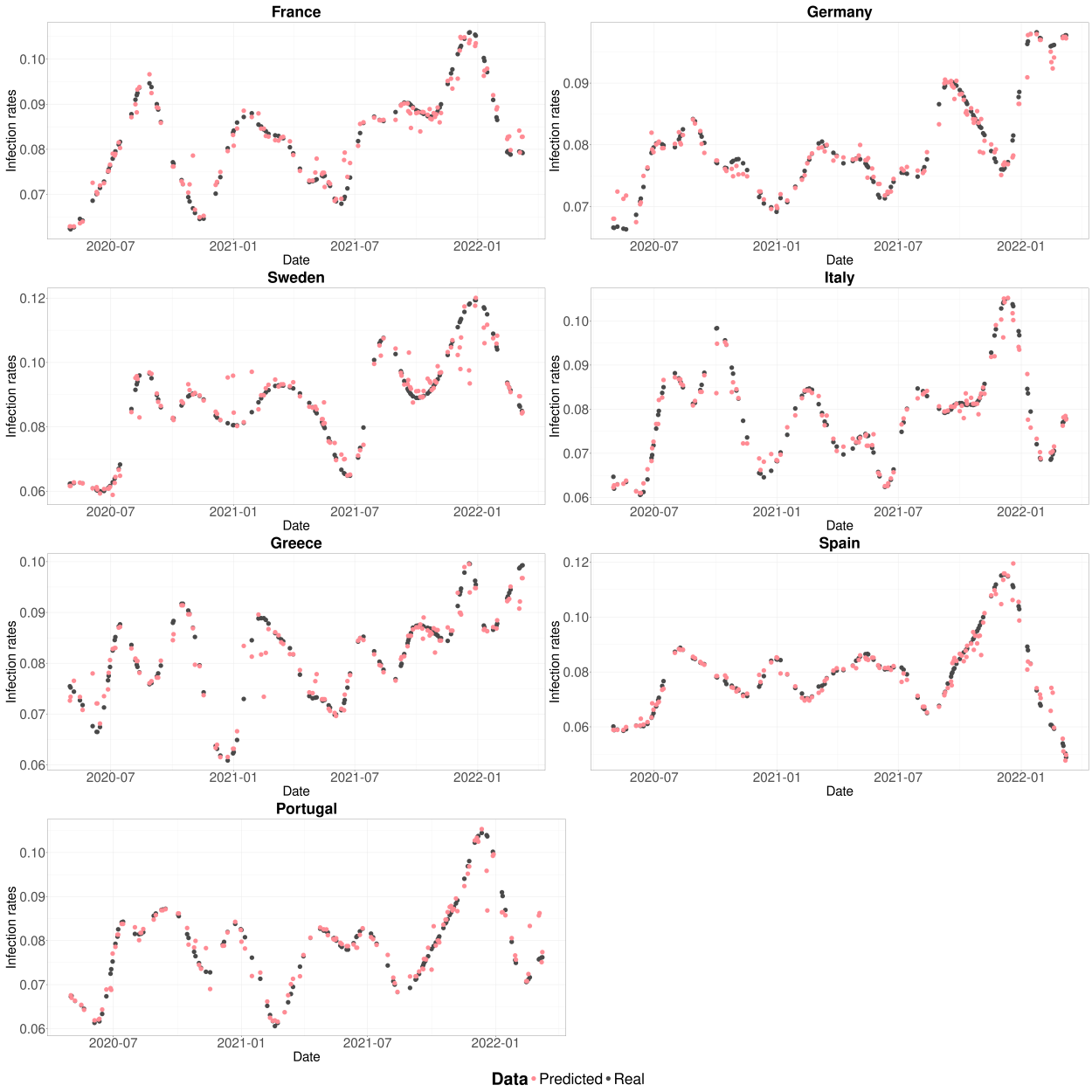


Figure A.23: Comparison between the ground truth and the predicted values for the test set (20% of the data, first fold) for the XGBoost model from mobility to infection rates in France, Germany, Austria, Italy, Greece, Spain, and Portugal.

Appendix B

Forge4Flame User Guide

This chapter serves as a user guide for ***Forge4Flame (F4F)***. Specifically, Section B.1 provides various instructions for downloading the repository, *F4F*, FLAME GPU 2, the Docker images, and Slurm, while Section B.2 explains how to run them in different ways. Section B.3 introduces the main features of the application, using an example as a guide. Section B.4 provides additional details about the generic middle school environment discussed in Chapter 5. Finally, Section B.5 briefly explains the integration between *F4F* and FLAME GPU 2.

B.1 How to download

This section provides instructions for downloading the repository, the *F4F* R package, FLAME GPU 2, and the Docker images and configuring the HPC environment for Slurm.

B.1.1 Repository

Cloning the repository is optional, especially for users intending to use the *F4F* package, dockers, or Slurm—see the following sections for more details. The GitHub repository is available at `https://github.com/qBioTurin/FORGE4FLAME`. To clone it, use the following Bash command:

```
git clone --recurse-submodules \
          https://github.com/qBioTurin/FORGE4FLAME.git
```

The main directory, `FORGE4FLAME`, contains the R Shiny application, while the `FORGE4FLAME/inst/FLAMEGPU-FORGE4FLAME` directory stores the FLAME GPU 2 template used by *F4F*.

B.1.2 Forge4Flame (F4F)

To install the *F4F* R package it is possible to use the R package *devtools* [282] or *remotes* [90], through these R commands:

```
if(!requireNamespace("devtools", quietly = TRUE))
    install.packages("devtools")

library(devtools)
install_github("https://github.com/qBioTurin/FORGE4FLAME")
```

or:

```
if(!requireNamespace("remotes", quietly = TRUE))
    install.packages("remotes")

library(remotes)
install_github("https://github.com/qBioTurin/FORGE4FLAME")
```

Dependencies

F4F is implemented in R Shiny [74]. It requires several CRAN packages to operate: *shinydashboard* [76], *shinyjs* [36], *jsonlite* [193], *dplyr* [281], *shinythemes* [75], *colourpicker* [34], *glue* [129], *readr* [283], *zip* [89], *sortable* [96], *shinyalert* [35], *shinybusy* [182], *shinyBS* [45], *stringr* [278], *ggplot2* [279], *tidyr* [280], *DT* [286], *shiny* [73], *shinyWidgets* [205], *shinyFiles* [**?**], and *htmltools* [**?**]. These packages are automatically installed.

Furthermore, it also requires the EBImage [198] and the EBImageExtra [195] packages. The user must install these packages manually using the following R commands:

```
if(!requireNamespace("BiocManager", quietly = TRUE))
    install.packages("BiocManager")

BiocManager::install("EBImage")
```

```
if(!requireNamespace("remotes", quietly = TRUE))
    install.packages("remotes")

library(remotes)
install_github("ornelles/EBImageExtra")
```

B.1.3 FLAME GPU 2

To install FLAME GPU 2 dependencies refer to the official documentation here.

B.1.4 Dockers

Users can download the Docker images for both *F4F* and FLAME GPU 2 to avoid any potential dependency-related issues. In this context, the user must have Docker installed on their computer. For more information, refer to this document: `https://docs.docker.com/engine/installation/`. Additionally, the user needs to ensure they have the necessary permissions to run Docker without using sudo. To create the Docker group and add the user on a Unix system, follow these steps:

- Create the docker group:

```
$ sudo groupadd docker
```

- Add user to the docker group:

```
$ sudo usermod -aG docker $USER
```

- Log out and log back in so that group membership is re-evaluated.

Additionally, the user must have the NVIDIA driver (installation instructions can be found here) and the NVIDIA container toolkit (installation instructions can be found here) installed. To download the Docker images, run the following Bash commands:

```
docker pull qbioturin/forge4flame
docker pull qbioturin/flamegpu2
```

An alternative method for downloading Docker images is via *F4F*, provided that the user has installed the R package (refer to Section B.1 for installation instructions):

```
library(FORGE4FLAME)
FORGE4FLAME::downloadContainers()
```

B.1.5 Slurm

To run FLAME GPU 2 simulations on an HPC system, the user must install Slurm on it (more information here). Generally, on HPC systems, Docker cannot be used due to privacy concerns. Therefore, the user must install all necessary FLAME GPU 2 dependencies on the user's system before running simulations (see Section B.1.3 for more details). However, on HPC4AI [21] (more information here), FLAME GPU 2 can also be executed using Docker, thanks to a tool that addresses privacy concerns.

B.2 How to run

This section describes how to run *F4F* and FLAME GPU 2, both with and without Docker, using the three approaches outlined in Figure **??** in Chapter 5.

B.2.1 F4F

Without Docker

To run *F4F* without using Docker, run the following R commands after installing the package (see Section B.1.2):

```
library(FORGE4FLAME)
FORGE4FLAME::FORGE4FLAME.run()
```

With Docker

To run it using Docker, execute the following Bash command (if running on a server, ensure that port 3838 is exposed and accessible via `http://<server-hostname>:3838`; if running locally, access at `http://localhost:3838/`):

```
docker run -p 3838:3838 qbioturin/forge4flame
```

In this case, there is no need to download the *F4F* repository or install the package. In addition, it is not possible to run a FLAME GPU 2 simulation using the *Run* button in the **Run** page—see Section B.3.9 for more details.

B.2.2 FLAME GPU 2

Without Docker

To run FLAME GPU 2 without using Docker, run the following Bash command (inside the `FORGE4FLAME/inst/FLAMEGPU-FORGE4FLAME` directory, after cloning the *F4F* repository):

```
# Executing a single run without using the FLAME GPU 2
# 3D visualization:

./abm.sh -expdir NameOfTheModel

############################################################

# Executing a single run using the FLAME GPU 2
# 3D visualization:

./abm.sh -expdir NameOfTheModel \
    -v ON

############################################################

# Executing n runs without using the FLAME GPU 2
# 3D visualization:

./abm_ensemble.sh -expdir NameOfTheModel

############################################################

# Visualize the helper:

./abm.sh -h
./abm_ensemble.sh -h
```

In particular, `NameOfTheModel` must be the name of the desired model and must correspond to a directory within `FORGE4FLAME/inst/FLAMEGPU-FORGE4FLAME/resources/f4f` that contains both a JSON file and an RDs file. If the user utilizes the *Link the model to FLAME GPU 2* button (see Section B.3.10 for more details), the directory with the provided

name will be automatically saved in the correct location. Otherwise, if the user downloads the model using the *Save the model* button, they must manually copy the unzipped directory into `FORGE4FLAME/inst/FLAMEGPU-FORGE4FLAME/resources/f4f`. Results will be saved in `FORGE4FLAME/inst/FLAMEGPU-FORGE4FLAME/results/NameOfTheModel`.

With Docker

To run FLAME GPU 2 using Docker, run the following Bash commands:

```
# Executing a single run without using the FLAME GPU 2
# 3D visualization:

docker run --user $UID:$UID \
   --rm \
   --gpus all \
   --runtime nvidia \
   -v /Absolute/Path/To/The/Directory/With/The/Model/
      NameOfTheModel:/home/docker/flamegpu2/FLAMEGPU-
      FORGE4FLAME/resources/f4f/NameOfTheModel \
   -v $(pwd):/home/docker/flamegpu2/FLAMEGPU-FORGE4FLAME/
      flamegpu2_results \
   qbioturin/flamegpu2 /usr/bin/bash -c "/home/docker/
      flamegpu2/FLAMEGPU-FORGE4FLAME/abm.sh -expdir
      NameOfTheModel"

#################################################################


# Executing n runs without using the FLAME GPU 2
# 3D visualization:

docker run --user $UID:$UID \
   --rm \
   --gpus all \
   --runtime nvidia \
   -v /Absolute/Path/To/The/Directory/With/The/Model/
      NameOfTheModel:/home/docker/flamegpu2/FLAMEGPU-
```

```
    FORGE4FLAME/resources/f4f/NameOfTheModel \
-v $(pwd):/home/docker/flamegpu2/FLAMEGPU-FORGE4FLAME/
    flamegpu2_results \
qbioturin/flamegpu2 /usr/bin/bash -c "/home/docker/
    flamegpu2/FLAMEGPU-FORGE4FLAME/abm_ensemble.sh -
    expdir NameOfTheModel"
```

In this case, there is no need to download the *F4F* repository or install the package. `/Absolute/Path/To/The/Directory/With/The/Model/NameOfTheModel` represents the **absolute** path to the local directory, named `NameOfTheModel`, which contains the model to run (the JSON and the RDs files). These files will be saved in a directory named `NameOfTheModel` inside the Docker (in `FLAMEGPU-FORGE4FLAME/resources/f4f`). Results will be saved in a directory named `results/NameOfTheModel` within the current directory. The user must replace the directory name containing the model with `NameOfTheModel`.

B.2.3 Docker Compose

To use the Docker Compose, the user must download the YAML file here. To start both *F4F* and FLAME GPU 2 containers, navigate to the directory containing the YAML file and run the following Bash command (if running on a server, ensure that port 3839 is exposed and accessible via `http://<server-hostname>:3839`; if running locally, access to `http://localhost:3839`):

```
docker compose up -d --build
```

To run a FLAME GPU 2 simulation using Docker Compose, the user must use the **Run** page of *F4F*—see Section B.2 for more details. Results will be saved in a directory named `results/NameOfTheModel` within the current directory, where `NameOfTheModel` is the name selected by the user when clicking on the *Run* button in the **Run** page—see Section B.3.9 for more details.

To stop the containers, run the following Bash command:

```
docker compose down
```

B.2.4 Slurm

To execute FLAME GPU 2 simulations on HPC systems using Slurm, assuming the necessary drivers are installed, the repository is cloned, and one or more reservations have been made on the HPC system, a Bash script like the one below can be used (this script, named `run_on_slurm.sh`, is included in the repository):

```
#!/bin/bash

# Define arrays
PARTITIONS=("MachineName1-booked", "MachineName2-booked",
   ..., "MachineNameN-booked")
RESERVATIONS=("Run1", "Run2", ..., "RunN")
GRES=("GPUName1", "GPUName2", ..., "GPUNameN")
MODEL_NAMES=("NameOfTheModel1", "NameOfTheModel2", ..., "
   NameOfTheModelN")

ARRAY_LENGTH=${#PARTITIONS[@]}

for ((i=0; i<ARRAY_LENGTH; i++)); do
    JOB_SCRIPT="job_script$i.sh"

    cat <<EOF > "$JOB_SCRIPT"
#!/bin/bash
#SBATCH --partition=${PARTITIONS[$i]}
#SBATCH --reservation=${RESERVATIONS[$i]}
#SBATCH -N 1
#SBATCH --gres=${GRES[$i]}

./abm_ensemble.sh -expdir ${MODEL_NAMES[$i]}
EOF

    # Submit the job and check if it was successful
    if sbatch "$JOB_SCRIPT"; then
        echo "Submitted: $JOB_SCRIPT"
    else
        echo "Failed to submit: $JOB_SCRIPT" >&2
    fi
done
```

In this script, the user can define a set of partitions, reservations, GPUs, and model names by specifying the four vectors:

- `PARTITIONS` specifies the names of the machines reserved.

- `RESERVATIONS` defines the reservation names chosen by the user at book time.
- `GRES` determines the name of the GPUs.
- `MODEL_NAMES` specifies the names of the models to run, which must correspond to the folder names within `FLAMEGPU-FORGE4FLAME/resources/f4f`.

Specifically, we run the experiments presented in Chapter 5 using Slurm on HPC4AI [21] (more information here).

B.2.5 How to reproduce the results

To reproduce the results presented in Chapter 5—comparing FLAME GPU 2 and NetLogo using the ITP and the alert scenarios—run the following Bash commands (note that this process may take time, especially for NetLogo):

```
git clone -b School --recurse-submodules \
    https://github.com/qBioTurin/FLAMEGPU-FORGE4FLAME.git

cd FLAMEGPU-FORGE4FLAME

# Reproduce results on comparison
# between FLAME GPU 2 and NetLogo:
./reproduce_comparison.sh

# Reproduce results on alert
# scenarios:
./reproduce_alert.sh
```

B.3 Forge4Flame application

In this section, we provide a detailed overview of all functionalities available in the *F4F* application, explaining each tab individually.

B.3.1 Homepage

Figure B.1 illustrates the **Homepage** of *F4F*, providing a clear overview of the application. The left-hand menu organizes *F4F*'s main functionalities, each designed to streamline various stages of model creation.

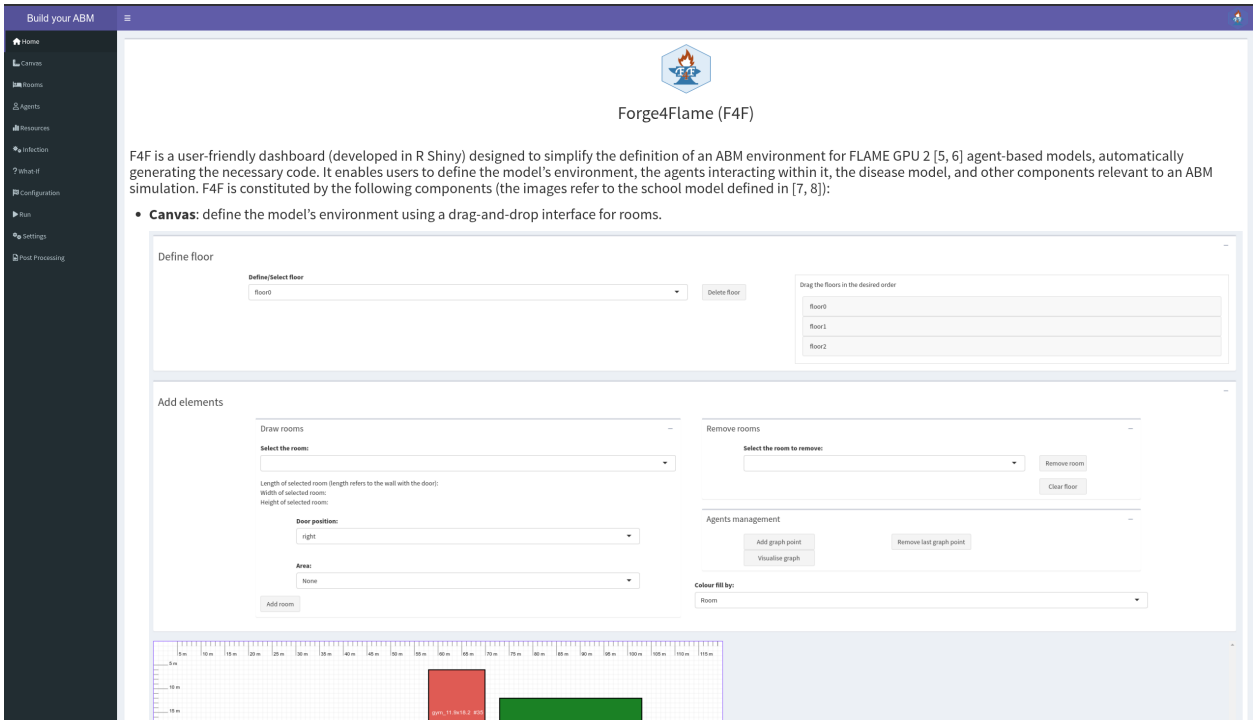


Figure B.1: The **Homepage** page.

B.3.2 Canvas

Figure B.2 shows the initial **Canvas** page. In the panel `Define floor`, a drop-down menu can be used to select a previously defined floor or to create a new floor giving it a name. The button *Delete floor* deletes the selected floor.

In the section *Draw rooms*, inside the panel `Add elements`, the user can add a previously defined type of room inside the canvas—by clicking on the button *Add room*—specifying the position of the door (none, bottom, left, top, right) and the area to which the room belongs (the user can define new areas or assign a room to no area). In the section *Remove rooms* the user can remove a room from the canvas by selecting it in the drop-down menu and clicking on the button *Remove room*. It is also possible to remove all the elements in the selected floor through the button *Clear floor*. The *Agent management* section is useful to create the graph of movements used by agents and to visualize it—see Section B.5 for more details on the graph of movement. The *Add graph point* button allows the user to insert a new node into the agents' movement graph. It is particularly useful to insert these graph points near the corners of the "corridors". Additionally, each room's door and the room's center are represented by a node in the graph. In addition, it is possible to remove the last inserted graph point by clicking on the *Remove last graph point* button and to visualize all the graph's edges by clicking on the *Visualise graph* button. The drop-down menu named *Colour fill by*

allows the user to color rooms based on names, types, or areas. Finally, under this panel, there is the canvas, initially empty. Figure B.3 shows the just created canvas associated with floor *floor0*—with the panel `Add elements` closed. Each square in the canvas represents 1x1 m^2.

Figure B.4 shows the canvas of floor *floor0* with 16 rooms and 4 graph points. In particular, in this case, the color of the room represents a previously defined room (that is, in the example light green represents the *room0*, dark blue the *room1*, gray the *fillingroom*, dark purple the *stairs*, and red the *spawnroom*). Moreover, Figure B.5 shows the canvas of floor *floor1* with 15 rooms and 4 graph points (here we have two more defined rooms: dark pink representing the *room2*, and dark green the *room3*).

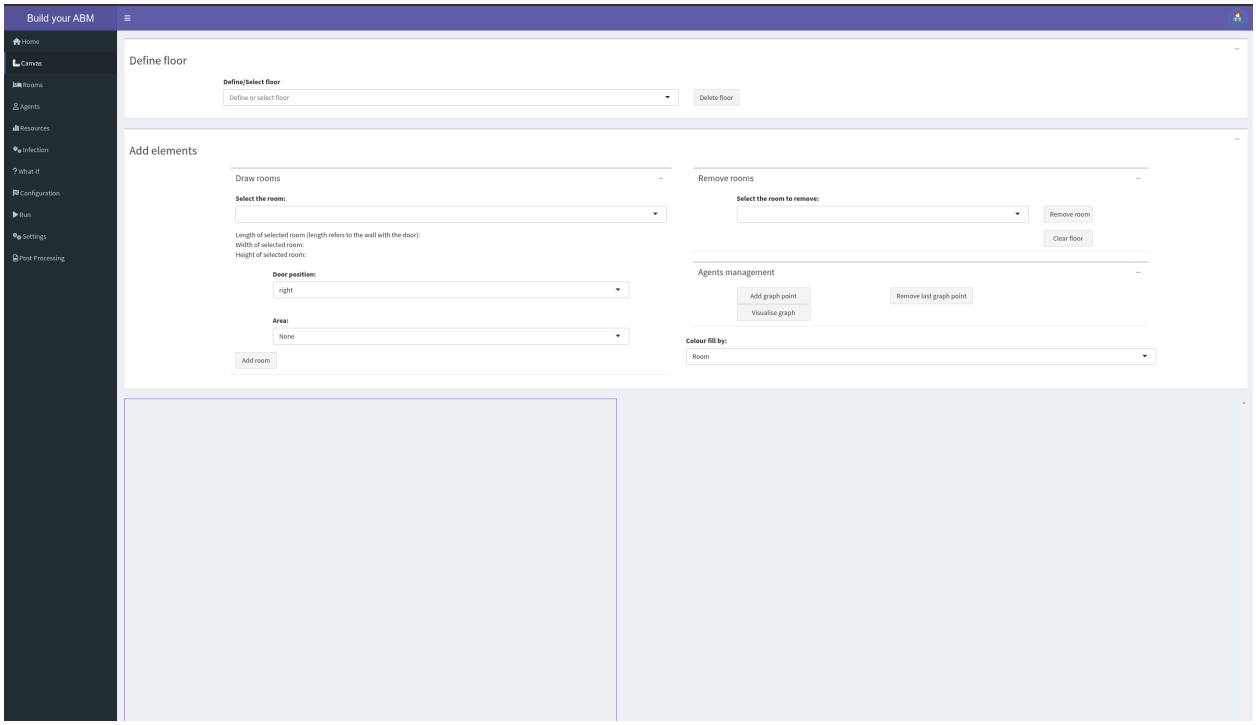


Figure B.2: The **Canvas** page.

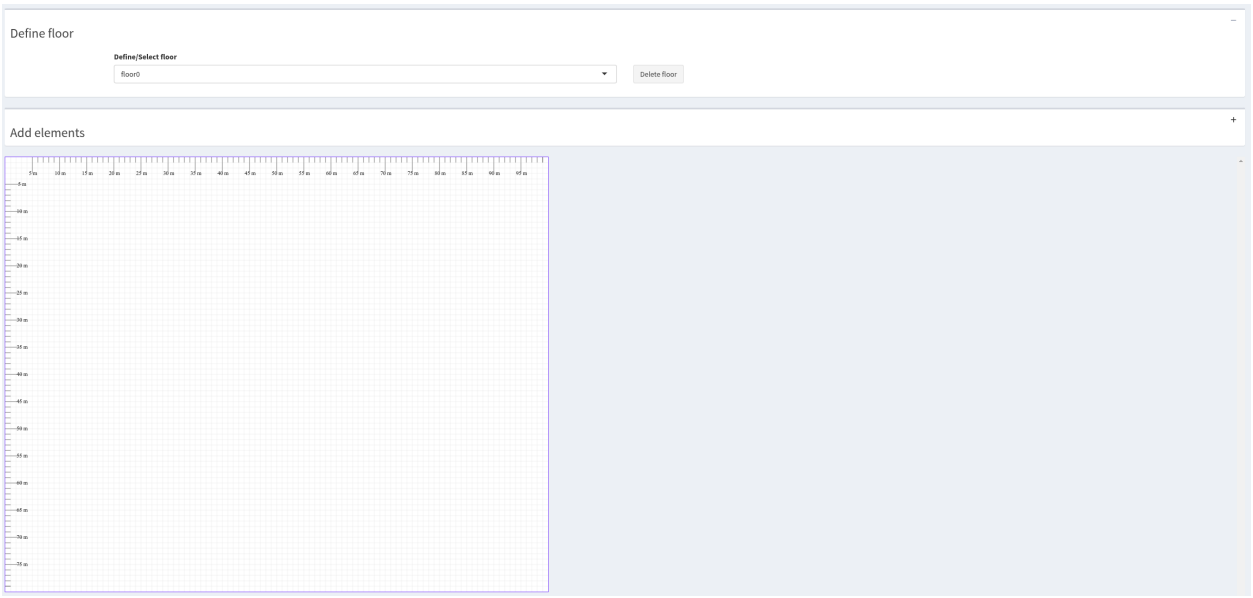


Figure B.3: The **Canvas** page with the empty canvas of the *floor0* visualized.

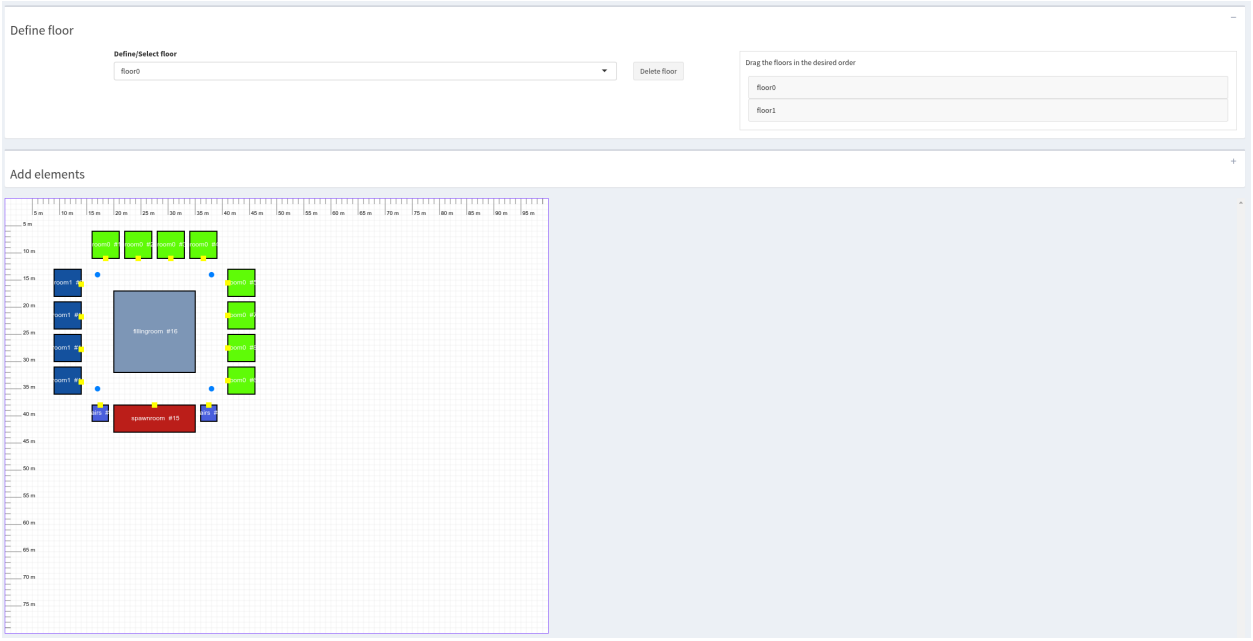


Figure B.4: The **Canvas** page with some rooms and graph points inside the canvas *floor0*. Rooms are colored by names.

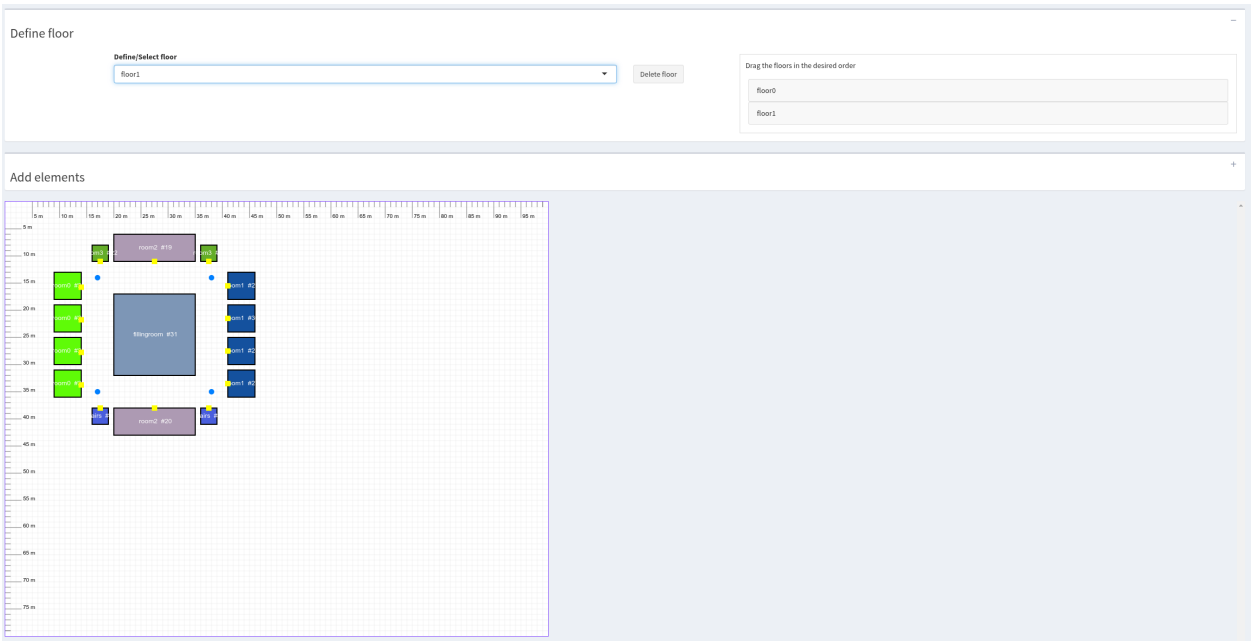


Figure B.5: The **Canvas** page with some rooms and graph points inside the canvas *floor1*. Rooms are colored by names.

B.3.3 Rooms

The **Rooms** page allows the user to define the rooms that will be placed inside the canvas. There are two panels inside this page, as shown in Figure B.6.

The first, `Define a new room`, allows creating the room from scratch and the second, `Set colour legend`, allows modification of the colors associated with each defined room. To define a new room the user must specify a *Name*—the name of a room is case insensitive, i.e. *room1* and *ROOM1* represent the same room—, the *Length*, the *Width*, and the *Height* of the room. Notice that the length refers to the wall with the door. The height of a room will not be visible in the canvas but will be displayed in the FLAME GPU 2 visualization. It is essential for accurately implementing the aerosol contagion process—see Section B.3.6 for more details.

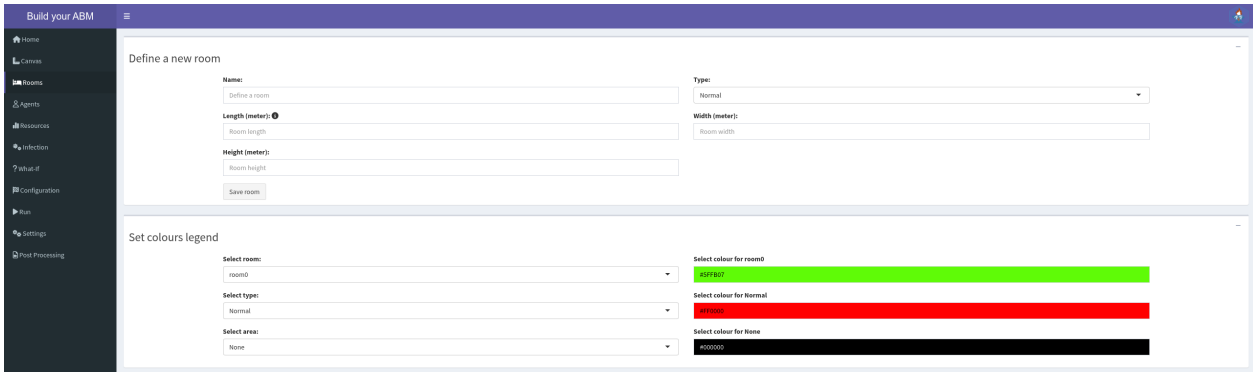


Figure B.6: The **Rooms** page.

Rooms can be grouped into clusters based on a specific *Type*. This allows for the definition of a particular room type, instructing the agents to visit every room of that type. For instance, the user may define the generic type **bathroom**, defining a *small_bathroom* and a *large_bathroom*, instructing the agents to visit the generic **bathroom**. An example of the definition of the room *small_bathroom* can be seen in Figure B.7. A set of default types is already defined inside our application:

- *Normal*: represents a generic room type.
- *Stairs*: room type used to move agents from one floor to another; stairs must be approximately in the same position on each floor.
- *Spawnroom*: room type from which agents enter and exit from the environment; currently, only one Spawnroom can be placed inside the canvas.
- *Fillingroom*: room type with no door; can be used to simulate the walls and rooms of no interest.
- *Waitingroom*: room type used by agents as a waiting area before entering another room when no resources are available.

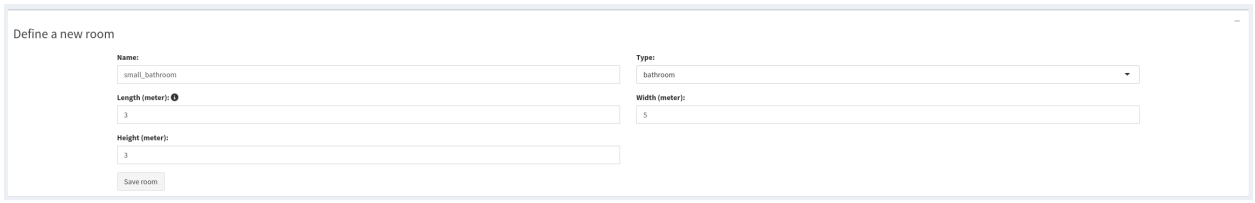


Figure B.7: Definition of a room called *small_bathroom* associated to a type called **bathroom**.

The panel `Set colour legend` allows the user to define the colors of individual rooms, rooms of a specific type, or rooms in a particular area, based on the chosen option from

the drop-down menu. The associated color will be displayed next to the selected item. To change the color, simply click on it. A color map will appear, enabling the user to select their preferred color, as shown in Figure B.8. These colors will be the ones the user sees on the canvas and in the FLAME GPU 2 simulation.

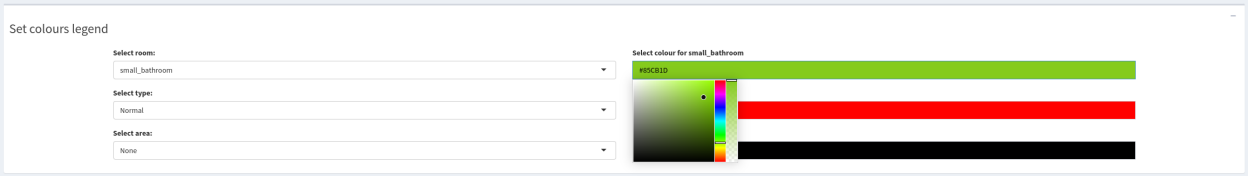


Figure B.8: Selection of a particular color for the room *small_bathroom*.

B.3.4 Agents

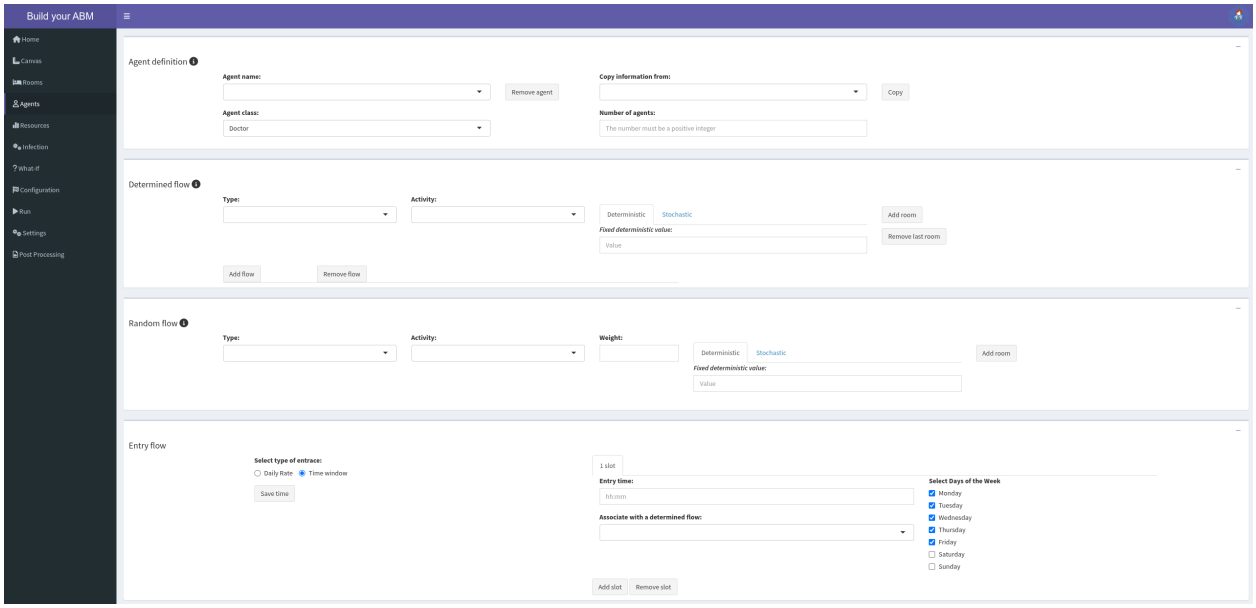


Figure B.9: The **Agents** page.

The **Agents** page serves to define the agents that will interact with the environment. It comprises several panels—see Figure B.9:

- In the `Agent definition` panel, the user can define all possible agent types. Each agent is defined by a name (*Agent name*) and can belong to a class (*Agent class*). A class is a category that groups different types of agents. Defining a broader class, such as *Doctor*, can be useful for organizing related agents like *Paediatrician*, *Neurologist*, *Surgeon*, and so on. Within this agent structure, the user can specify subclass-specific agents that belong to the same overarching group. For each specific agent, the user can

define details like determined and random flows and entry times, which will be covered in the following sections. The user can also copy all information concerning an agent using the drop-down menu *Copy information from*, which allows the user to select an existing agent and then modify it as needed. Additionally, the user can specify the number of agents to include in the model by entering the desired amount in *Number of agents*. An example of this process is shown in Figure B.10.

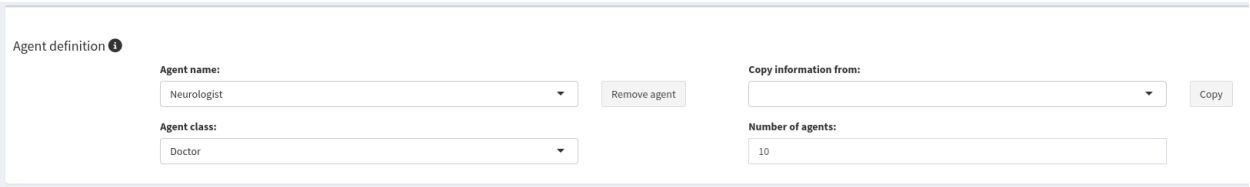


Figure B.10: An example of the definition of an agent named *Neurologist*, which belongs to the class *Doctor*, with a total of 10 agents of that type in the model.

- In the `Determined flow` panel, the user can define one or more flows that an agent must follow within the model. There are three key elements to configure:

  1. In the *Type* drop-down menu, the user can select the room to add to the flow. These rooms are those that have been added to the canvas—by type and area.
  2. In the *Activity* drop-down menu, the user can select the intensity of the activity that the agent performs in the selected room. This setting is crucial for calculating the virus concentration within a room: higher activity intensity leads to greater virus spread.
  3. The third and last element defines the duration of the task performed by the agent. The user can specify a deterministic time, which represents the exact number of minutes the agent spends on the task, or define the duration using a stochastic distribution—i.e., Exponential, Uniform, or Truncated Positive Normal. The distribution type can be selected from a drop-down menu, and its parameters can be entered below.

  Rooms are added to the flow using the *Add Room* button, and the entire flow will be displayed below. The first room in the flow will appear at the top, with the agent following the subsequent rooms in descending order. Both the first and last rooms in the flow **must** be the room with type *Spawnroom* and area *None* to allow the agent to enter and exit the environment—the time associated the last *Spawnroom* is **negligible**, it will be automatically computed based on the next entry time. The order of rooms can be adjusted by selecting and moving them up or down within the flow. Additionally,

the last room can be removed using the *Remove Last Room* button. Multiple flows can be defined for a single type of agent—except for agents with a daily rate, which can have only one flow, as explained below. An example of a defined flow is shown in Figure B.11.

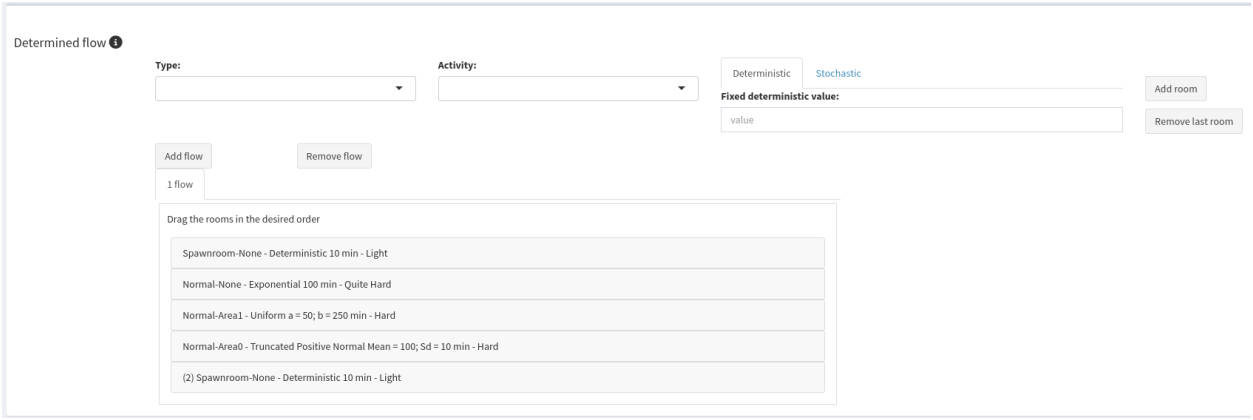


Figure B.11: An example of a determined flow. Imagine that the agent is scheduled to enter the environment at 10:00 A.M. Then, it will *i)* remain outside the environment until 10:10 A.M, while engaging in a light activity, *ii)* move to a room of type *Normal* with area *None*, where it will perform a quite hard activity for a duration determined by an exponential distribution with an average of 100 minutes, *iii)* proceed to a room of type *Normal* with area *Area1*, where it will perform a hard activity for a duration determined by a uniform distribution with $a = 50$ minutes and $b = 250$ minutes, *iv)* move to a room of type *Normal* with area *Area0*, where it will perform a hard activity for a duration determined by a truncated positive normal distribution with a mean of 100 minutes and standard deviation of 10 minutes, and finally *v)* leave the environment.

- In the `Random flow` panel, the user can add random events that interrupt the agent's determined flow, i.e. directing the agent to an additional room—a random event should occur infrequently and last for a short duration (few minutes). The elements are similar to those in the previous section, including the *Type* drop-down menu for selecting rooms—by type and area—, the *Activity* drop-down menu for selecting the activity type, and the definition of the time spent in that room. The *Weight* field allows the user to specify the probability of the event occurring by entering a number between 0 and 1. By default, the agent will follow the determined flow without any events. When a new event (e.g., going to the toilet) is added with a probability p, the probability of following the determined flow in each simulation step becomes $1 - p$. A random event should happen rarely and last only a few minutes. If an event lasts too long, it may be cut short due to the completion of a block in the determined flow. For example, consider an event e with a duration of t_e, while the agent is required to remain in its current position for t_{df}—where df represents the current block in the determined flow. If $t_e > t_{df}$, the agent will experience the event for t_{df} to maintain consistency with the determined flow. In each simulation step, at most one event can occur and the probabilities are to be considered **per simulation step**—for more details on the

simulation step, see Section B.3.8. An agent is restricted from triggering new events while actively handling an ongoing event to avoid excessively long or infinite event loops. New events can only be initiated once the current event has been completed. In addition, an event can be removed by simply clicking on it—an alert will appear asking for confirmation to delete or cancel the action. An example of a random flow is shown in Figure B.12.

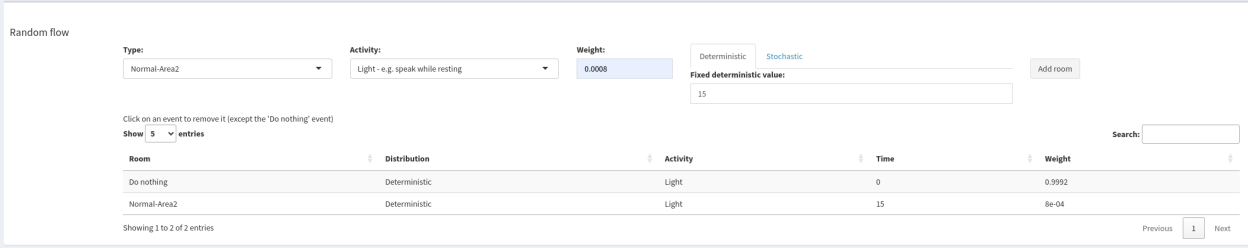


Figure B.12: An example of random flow. At each simulation step, the agent has a probability of 0.0008 to move to a room of type *Normal* with area *Area2* with light activity and will spend exactly 15 minutes there.

- The last panel is called `Entry flow`. It allows the user to define when the agent enters the environment and starts its determined flow. There are two ways to define an agent's entry flow:
  - By a **daily rate**: the user can define an entrance rate for the agent using either a deterministic value or a stochastic distribution, along with specifying an initial and a final generation time. Agents will be generated inside this time window. If a deterministic value is used, the exact number of agents specified will be generated, otherwise, the number of agents created will be determined according to the stochastic distribution parameters. Generated agents will follow the determined flow and leave the environment upon completion—these agents can have only one determined flow. Additionally, the user can set various entry rates for different time windows and specify particular days of the week for agent generation. See Figures B.13 and B.14 for an example of a definition of a daily rate entry flow.
  - By a **time window**: in this case, the number of agents is specified in the `Agent definition` panel—see Figure B.10. The user sets the entry time and the specific days of the week when the agent will enter the environment—it is possible to define multiple time slots. Additionally, the user can define different flows—see Figure B.11—associating one of them to each time slot. These agents will repeat the selected flow each week throughout the simulation. The total duration of a determined flow can only be estimated (especially if the user employs stochastic distributions). We check for possible overlaps among the entry flows using the

average of the selected distributions. In some cases, overlaps may occur. In these cases, when the agent finishes the current determined flow, it will start the next one with some delay. Our suggestion is to avoid placing different entry flows too close to each other. See Figure B.15 for an example of the definition of a time window entry flow.

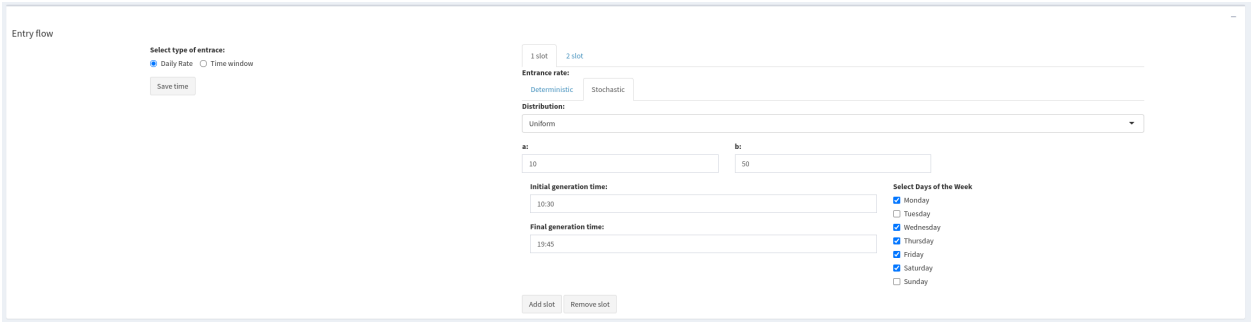


Figure B.13: An example of a daily rate entry flow (slot 1). Agents will be generated between 10:30 A.M. and 7:45 P.M. on every weekday except Tuesday and Sunday, with their number determined stochastically using a uniform distribution with $a = 10$ agents and $b = 50$ agents.

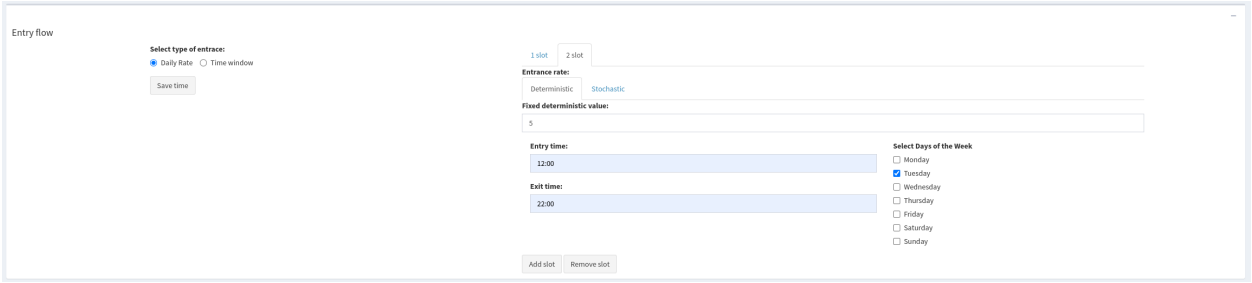


Figure B.14: An example of a daily rate entry flow (slot 2). Five agents will be generated between 12:00 A.M. and 10:00 P.M. on Tuesday and Sunday.

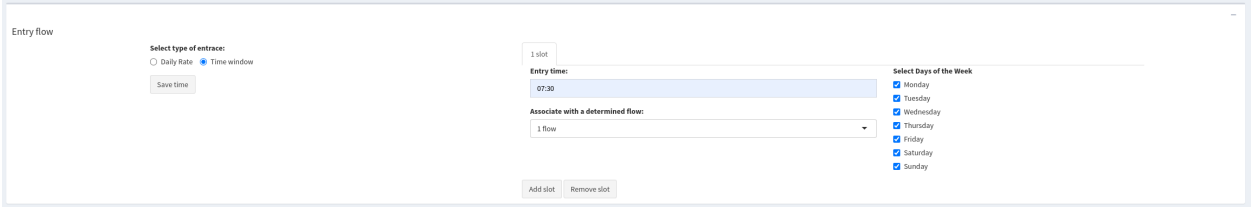


Figure B.15: An example of a time window entry flow. Agents will enter the environment at 7:30 A.M. every day of the week, following the first flow.

B.3.5 Resources

The **Resources** page allows the user to specify the maximum number of agents that can occupy a room simultaneously—see Figure B.16. It also provides options to set limits based on the type of agent.

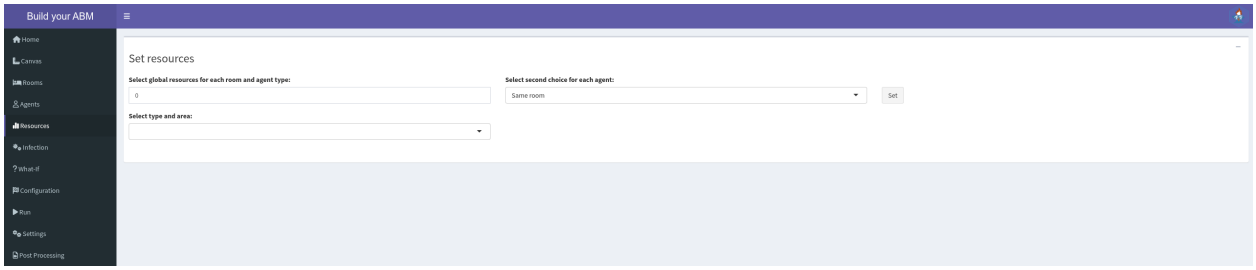


Figure B.16: The **Resources** page.

This page consists solely of the `Set Resources` panel, as illustrated in Figure B.17. Within this panel, two global values can be configured:

- **Maximum number of agents:** users can specify globally the maximum number of agents allowed to enter each room in the environment. This is particularly useful when the user is not interested in modeling resource limitations and prefers to set them to high values.
- **Global alternative for room entry:** if an agent is unable to enter a room in its flow due to capacity constraints, the user can define an alternative option based on the associated resources. The available alternatives include:
  - *Same room*: search for another room of the same type and area.
  - *Skip room*: proceed to the next room in the flow or skip the event (in the case an event occurred).
  - *Another room*: select a different room with a different type and area.

To set this value, simply enter the desired number in the input field and press the *Set* button.

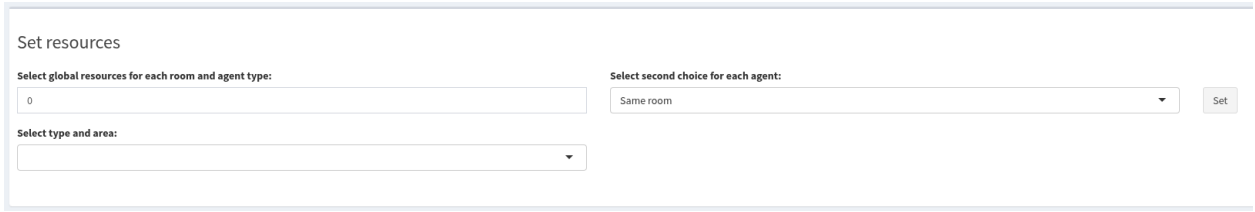


Figure B.17: The *Set resources* panel.

The third element of the section is a drop-down menu labeled *Select Type and Area*, which allows the user to choose a room to associate resources with, categorized by type and area. By selecting a room based on its type and area, the user can modify the associated resources. As shown in Figure B.18, a table displays all rooms of the selected type and area that are associated with at least one agent's flow.

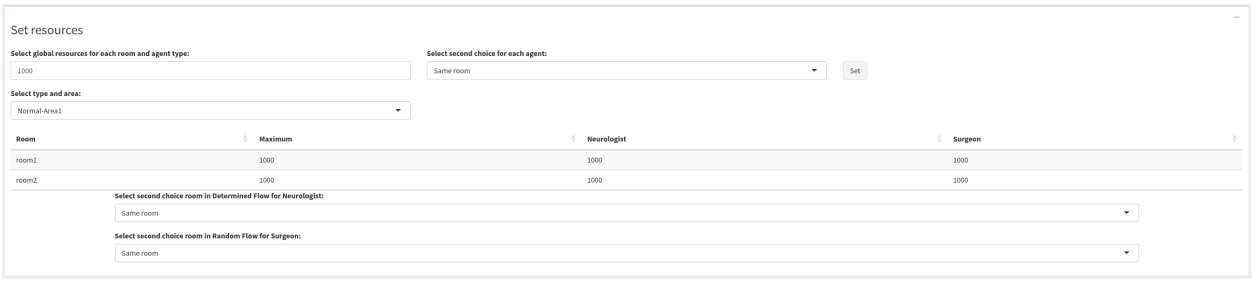


Figure B.18: The resources associated with rooms of type *Normal* and area *Area1* are displayed. This room appears in the predefined flows of the *Neurologist* agent and in the random flow of the *Surgeon* agent.

For each room, the user can view and modify the number of agents allowed to enter by double-clicking the respective field. The total number of agents permitted in the room can be adjusted in the *Maximum* column. Additionally, the user can specify the allowed number of agents for each type, with each type displayed in a separate column corresponding to the room. If an agent finds the room full when attempting to enter, the user can specify an alternative room using the *Select second choice room in Determined/Random Flow for* ***Agent*** drop-down menu for each agent. The user can assign different alternative rooms for an agent's determined and random flow, as shown in Figure B.18.

B.3.6 Infection

The **Infection** page—see Figure B.19—allows the user to choose from eight different disease compartmental models:

- SIR: Susceptible-Infected-Recovered.
- SIRD: Susceptible-Infected-Recovered-Deceased.
- SEIR: Susceptible-Exposed-Infected-Recovered.
- SEIRD: Susceptible-Exposed-Infected-Recovered-Deceased.
- SIRS: Susceptible-Infected-Recovered-Susceptible.
- SIRDS: Susceptible-Infected-Recovered-Deceased-Susceptible.
- SEIRS: Susceptible-Exposed-Infected-Recovered-Susceptible.
- SEIRDS: Susceptible-Exposed-Infected-Recovered-Deceased-Susceptible.

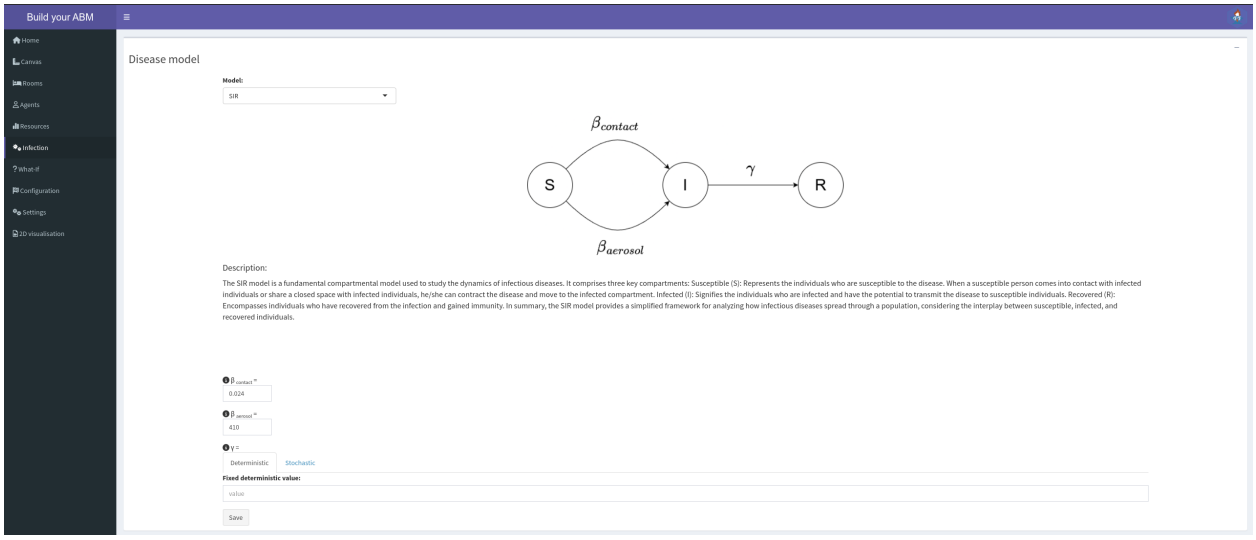


Figure B.19: The **Infection** page with the SIR compartmental model shown.

A brief description and a graphical schema are provided for each disease model. Based on the chosen model, the user must specify values for the various parameters of that model. In particular, all the disease models share three parameters: *i)* $\beta_{contact}$, which represents the contamination risk associated with infection from close-range contacts, as modeled in [133]—for example, for the COVID-19 disease the value to use is 0.024—, *ii)* $\beta_{aerosol}$, which represents the risk constant associated with infection due to aerosol, as modeled in [114]—for example, for the COVID-19 disease the value to use is 410—, and *iii)* γ, which represents the recovery rate. **Note that** $\beta_{contact}$ **and** $\beta_{aerosol}$ **are not the infection rates.** For each parameter, except $\beta_{contact}$ and $\beta_{aerosol}$, the user can choose a deterministic value or a stochastic distribution. After specifying the parameters, click on the *Save* button to save them.

In Figure B.20 the SEIRDS model is shown. This case is characterize by six parameters: *i)* $\beta_{contact}$, *ii)* $\beta_{aerosol}$, *iii)* γ, *iv)* α, which represents the incubation rate, *v)* λ, which represents the fatality rate, and *vi)* ν, which represents the end-of-immunization rate.

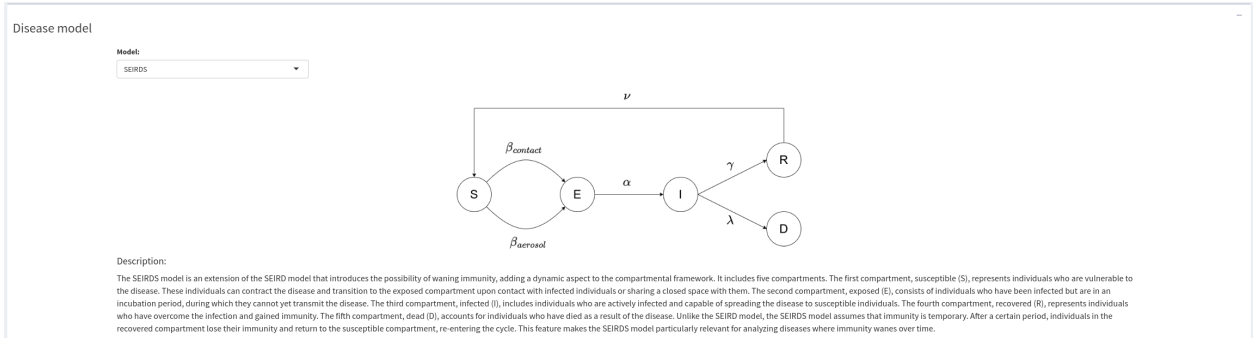


Figure B.20: The SEIRDS compartmental model.

The virus exposure is modeled through two different drivers: either close-range contacts [133] between susceptible and infected agents or through aerosol [114] within an indoor environment.

In the former case, the probability of successful contagious contact is associated with each susceptible agent ($\mathrm{S_i}$) that remains within an area ($\mathbf{A}$) of $3.46\,m^2$ surrounding an infected agent—in contrast to the model presented in Chapter 2, we are now considering a circle with a radius of $1.05\,m$ around each agent, instead of a square with area $2.1 \cdot 2.1\,m^2$; we chose $1.05\,m$ because it corresponds to the apothem of the square, but we are considering making this value a parameter that the user can specify through the application. This probability is determined by the following function [133]:

$$P_{S_i} = \beta_{contact} \frac{C_{S_i}}{\mathbf{A}} \tag{B.1}$$

where $\beta_{contact}$ represents the *contamination risk*, and C_{S_i} denotes the total time (in minutes) that the S_i agent remains within the vicinity of an infected agent.

In contrast, the spread of the pathogen via aerosol is calculated based on the *quanta* concentration in each room. A quanta is defined as the dose of airborne droplet nuclei necessary to infect 63% of susceptible individuals [62]. Following the approach in [114], we determine the number of quanta inhaled by each agent in a given room to compute the probability of infection without direct contact with an infected agent. This probability is given by the following function:

$$P_{S_i} = 1 - e^{-\frac{N_{\mathrm{virus}}}{\beta_{aerosol}}} \tag{B.2}$$

where $\beta_{aerosol}$ represents the reciprocal of the probability that a single pathogen particle will trigger an infection [114, 273], and N_{virus} denotes the number of quanta inhaled by each susceptible agent during a time interval Δt. This value is determined by solving the following differential equation:

$$dN_{virus} = (1 - \eta_{mask} f_{\mathrm{mask}})\; C(t) Q_{inh}\; dt \tag{B.3}$$

where η_{mask} is the efficacy of the masks, f_{mask} represents the proportion of agents correctly wearing a mask, $C(t)$ denotes the quanta concentration at time t, and Q_{inh} is the inhalation rate. Observe that quanta concentration at time t mainly depends on the number of people within a room, the number of infections, and the room's ventilation.

B.3.7 What-If

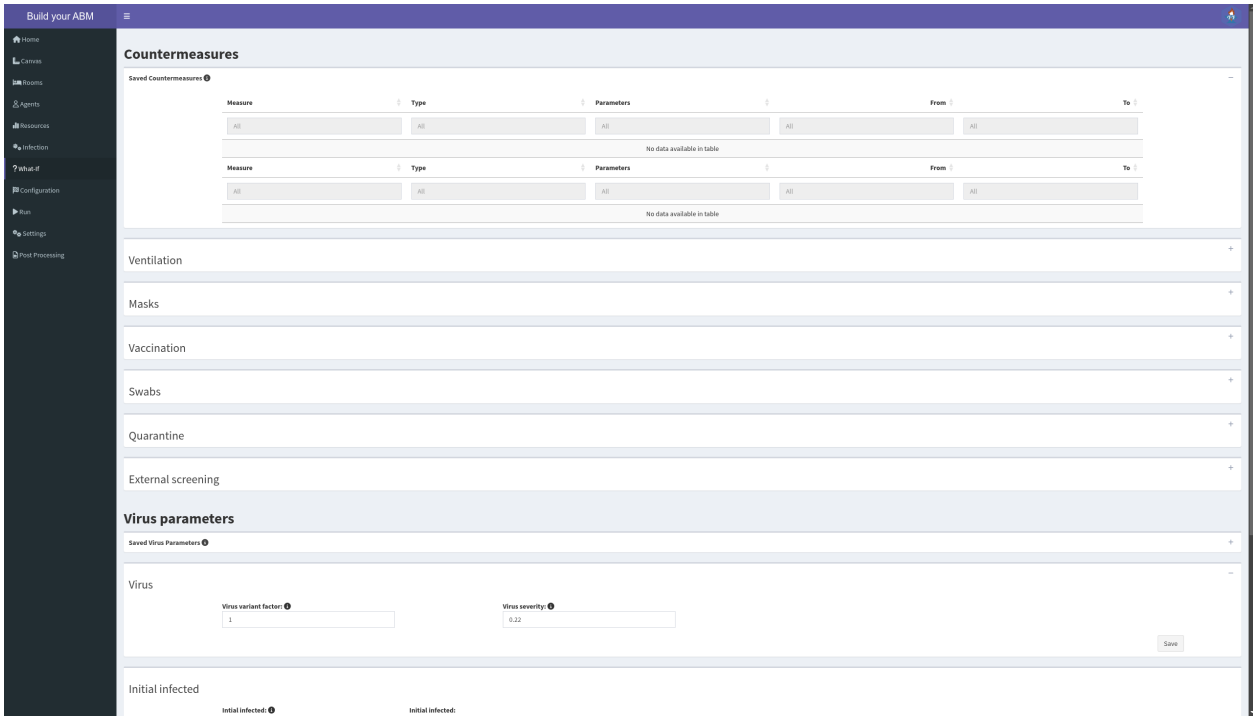


Figure B.21: The **What-If** page.

The **What-If** page—see Figure B.21—enables the user to conduct various what-if scenarios considering different *i)* levels of environmental ventilation, *ii)* masks policies, *iii)* levels of vaccination coverage and efficacy, *iv)* screening policies, *v)* quarantine policies, *vi)* external screening policies, *vii)* virus variants and severity of symptoms, *viii)* initial infected types and numbers, and *ix)* levels of outside contagion. The page is divided into two macro-sections: **Countermeasures** and **Virus Parameters**. If the user does not specify a particular countermeasure, it will not be applied. For the virus parameters, the default values will be used—see below.

For each countermeasure, the user must specify the period during which it will be in effect. For instance, the user can specify a particular type of mask with a defined fraction of agents wearing it from day d_1 to d_2 and apply a different setting from $d_2 + 1$ to d_3—with d_1 ¡ d_2 ¡ d_3. Clicking the *Save* button the specified settings will be saved—this applies to both the countermeasures and the virus parameters. In details the **Countermeasures** section contains:

- The `Ventilation` panel allows the user to define either a global ventilation level—Figure B.22—, which is the same for each room, or specific ventilation levels—Figure B.23—

which can differ for each room [114]. There are 7 possible ventilation levels in ACH[1], i.e. 0 (no ventilation), 0.3 (poorly ventilated), 1 (domestic), 3 (offices/schools), 5 (well-ventilated), 10 (typical maximum), and 20 (hospital setting). The ventilation level has an impact on the aerosol contagion process—see Section B.3.6.

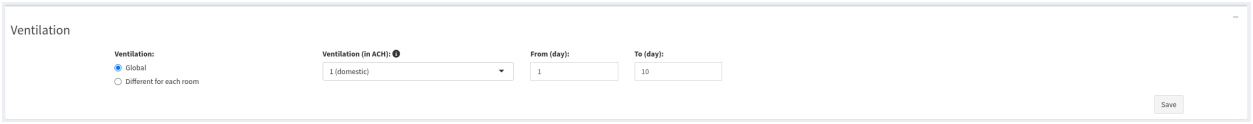


Figure B.22: Setting ventilation globally from day 1 to 10.

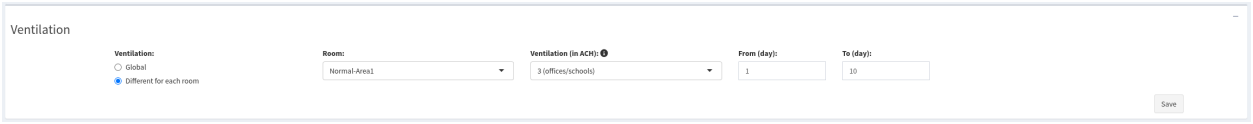


Figure B.23: Setting ventilation specifically to rooms with type *Normal* and area *Area1* from day 1 to 10.

- The `Masks` panel allows the user to define either a global mask policy—Figure B.24—, which is the same for each agent, or a specific mask policy—Figure B.25—which can differ for each agent. There are 3 different types of masks [200], i.e. no mask (0% of exhalation and inhalation efficacy), surgical mask (59% of exhalation and inhalation efficacy) and FFP2 mask (90% of exhalation and inhalation efficacy). The user can also specify the fraction of agents that will wear the selected mask, either globally or for a specific agent.

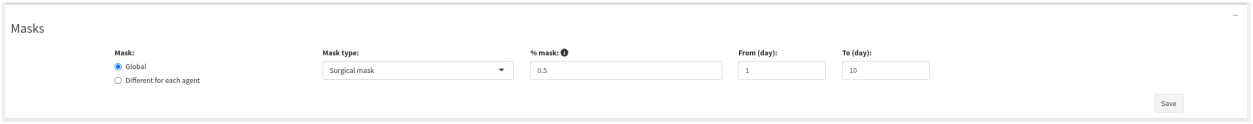


Figure B.24: Setting mask policy globally from day 1 to 10.

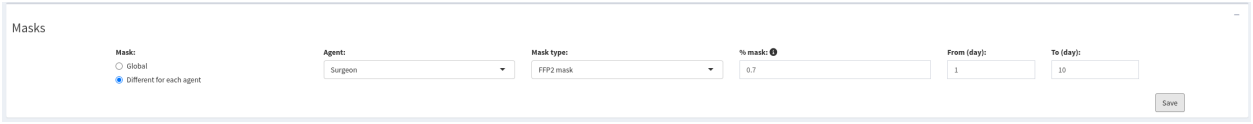


Figure B.25: Setting mask policy specifically to agents with type *Surgeon* from day 1 to 10.

- The `Vaccination` panel allows the user to define either a global vaccination policy—Figure B.26—, which is the same for each agent, or a specific vaccination policy—Figure B.27—which can differ for each agent. The user must specify the fraction of

[1]For instance, 3 Air Changes per Hour (ACH) means that in 1 hour 300.000 L (or analogous 300 squared meters) of external air are entered into the considered room.

agents that will be vaccinated and the efficacy of the vaccines. We define vaccination efficacy as the proportion of vaccinated agents that are effectively immune to infection, while the remainder remain susceptible despite being vaccinated. It is also necessary to set a vaccine coverage, which determines the number of days the vaccination will protect the agent. The user can configure this using either a deterministic value or a stochastic distribution, as described previously. In this case, the countermeasure will be applied on the specified day. For instance, if the user specifies a fraction f of vaccinated agents with an efficacy e and a coverage c on day d, those agents will be vaccinated on that day. In this case, there is no *From* and *To* period.

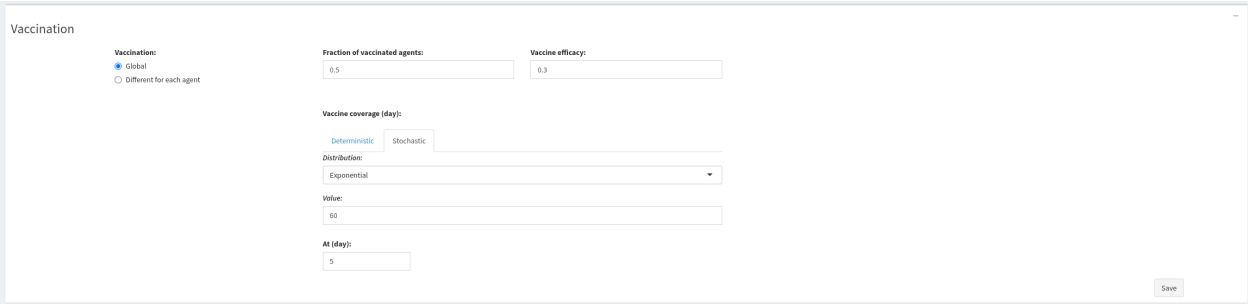


Figure B.26: Setting vaccination policy globally at day 5.

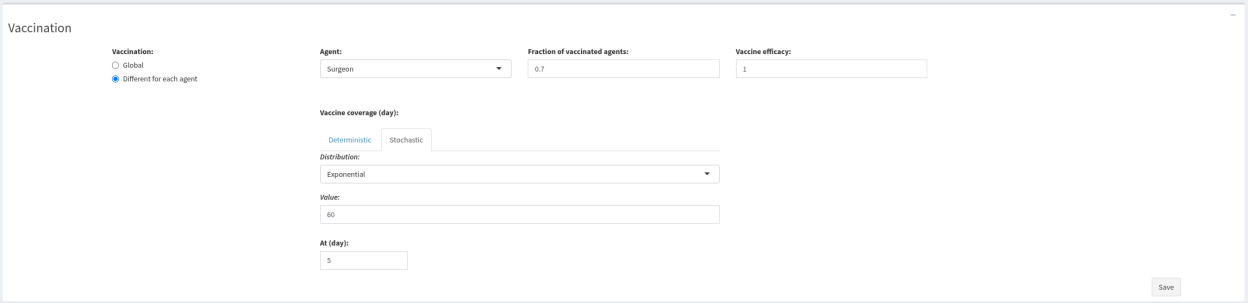


Figure B.27: Setting vaccination policy specifically to agents with type *Surgeon* at day 5.

- The `Swabs` panel allows the user to define either a global policy—Figure B.28—, which is the same for each agent, or a specific policy—Figure B.29—which can differ for each agent. The user must specify the sensitivity and specificity of the swabs, as well as the frequency (deterministic or stochastic) at which swabs are administered to each agent.

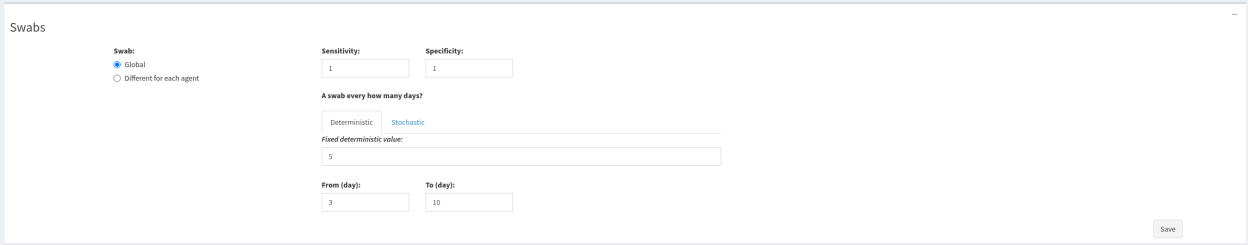


Figure B.28: Setting swab policy globally from day 3 to 10..

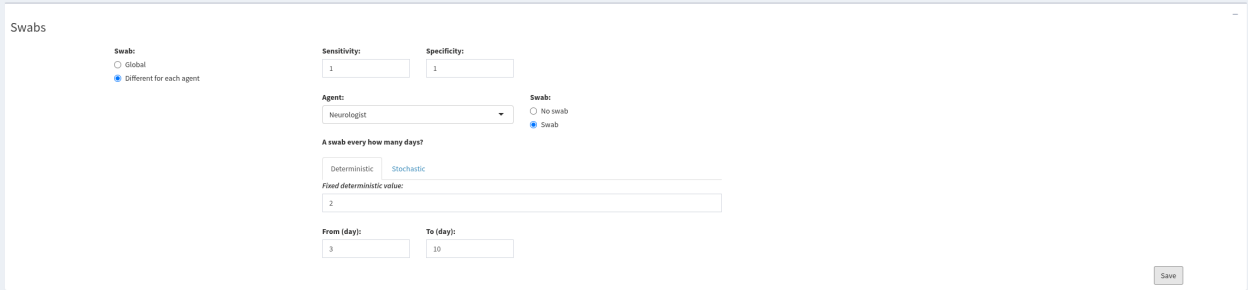


Figure B.29: Setting swab policy specifically to agents with type *Neurologist* from day 3 to 10.

- The `Quarantine` panel allows the user to define either a global policy—Figure B.30—, the same for each agent, or a specific policy—Figure B.31—which can differ for each agent. The user must specify the duration of the quarantine in days (deterministic or stochastic) and the room used to put in quarantine agents with severe cases—the default is outside the environment. The user can also choose a swab policy, enabled only during quarantine, by specifying the frequency (deterministic or stochastic) at which swabs are administered to each agent, the sensitivity and the specificity.

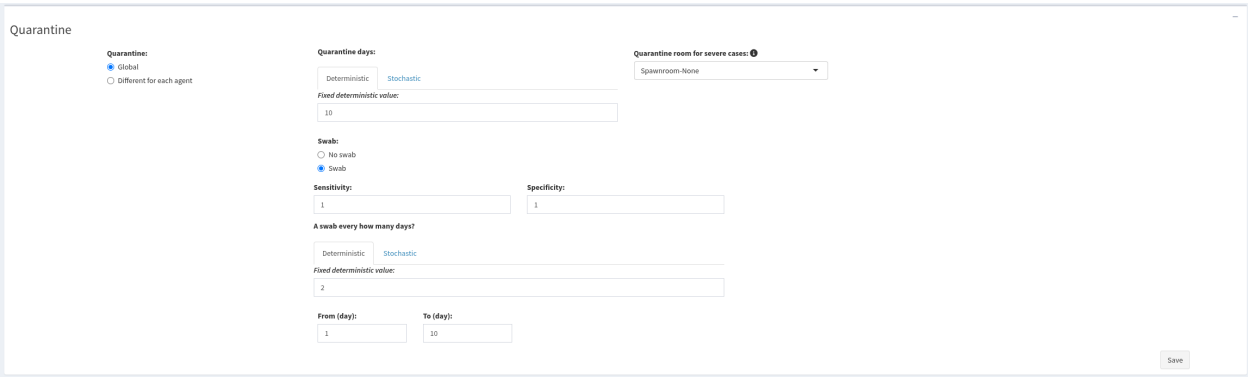


Figure B.30: Setting quarantine policy globally from day 1 to 10.

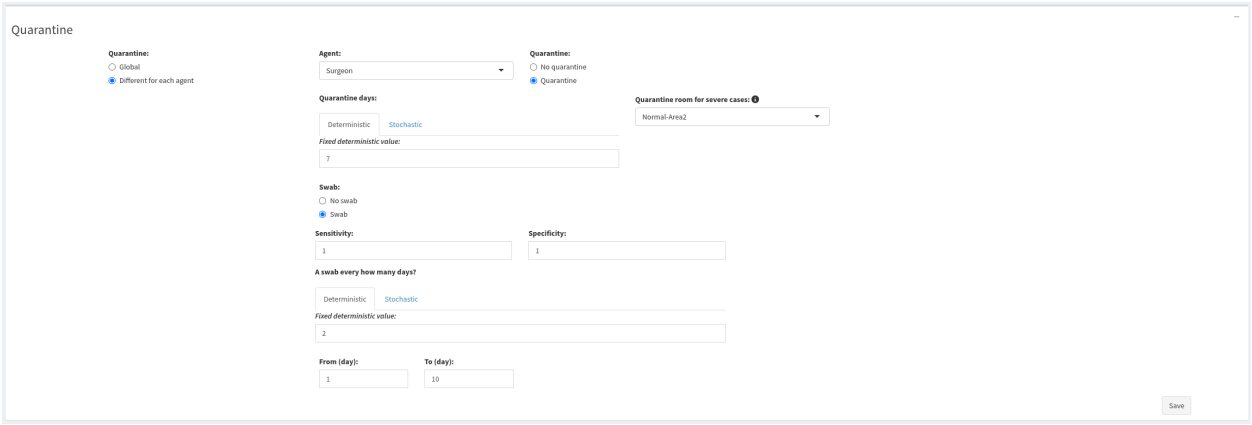


Figure B.31: Setting quarantine policy specifically for agents with type *Surgeon* from day 1 to 10.

- The `External screening` panel allows the user to define either a global policy—Figure B.32—,the same for each agent, or a specific policy—Figure B.33—which can differ for each agent. In particular, two different probabilities of external screening can be applied to agents each day: the probability of testing an agent due to activities that involve screening campaigns (e.g., sports practice) and the probability of testing an agent outside the environment due to the presence of symptoms.

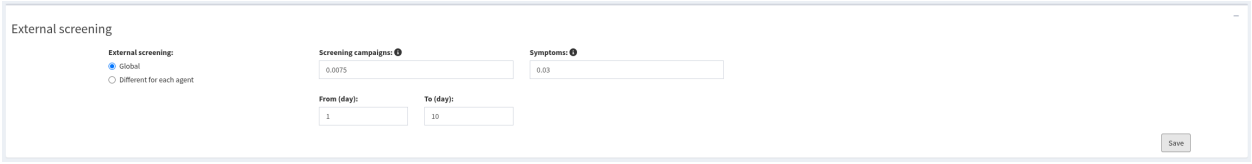


Figure B.32: Setting external screening policy globally from day 1 to 10.

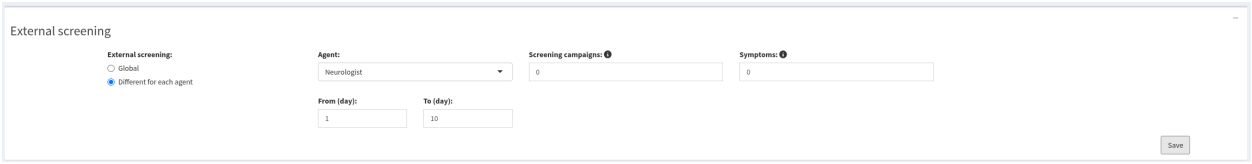


Figure B.33: Setting external screening policy specifically for agents with type *Neurologist* from day 1 to 10.

At the beginning of the **What-If** page, all countermeasures set by the user are displayed, as shown in Figure B.34, in the *Saved Countermeasures* panel. The first table contains countermeasures related to the rooms (i.e., ventilation), while the second table contains countermeasures related to agents. To remove one of the countermeasures, the user must click on the relevant entry and an alert will appear asking for confirmation to delete or cancel the action. It is also possible to display only specific countermeasures using the filter associated with each column. An example is shown in Figure B.35, where only the mask

policy is displayed. **Specific countermeasures will take precedence over global ones.** For instance, as shown in Figure B.34, 50% of agents will wear a surgical mask from day 1 to day 10, with the exception of surgeons, 70% of whom will wear an FFP2 mask.

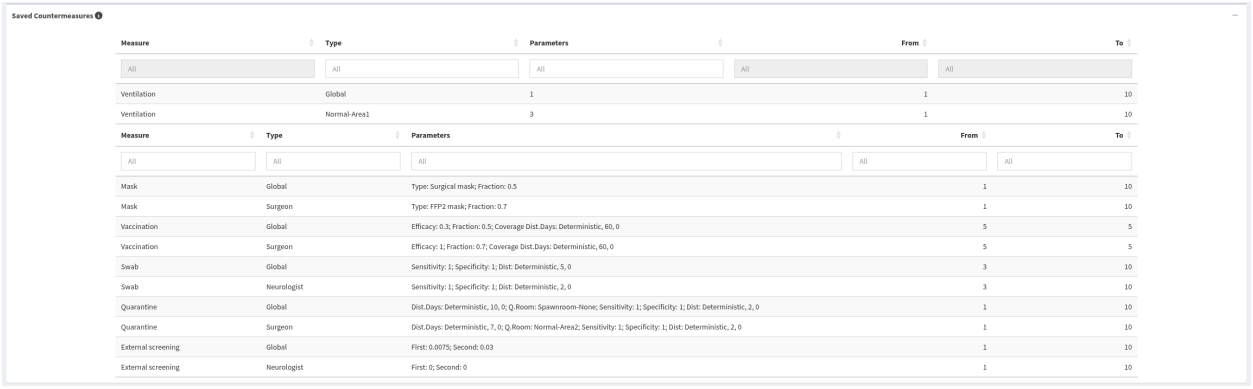


Figure B.34: Saved countermeasures listed in the panel *Countermeasures*.

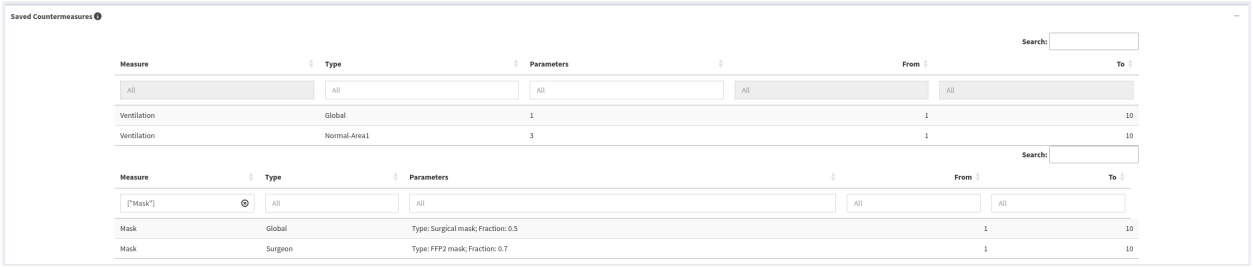


Figure B.35: Example of filter on the *Measure* column.

On the other hand, the **Virus parameters** section contains:

- The `Virus` panel allows the user to define the virus—or the variant of the virus—to model and the severity of symptoms—Figure B.36. Use the value 1 as *Virus variant factor* if there are no variants or if the user wants to model the base variant—in [200] the user can find some examples for COVID-19. The *Virus severity* represents the probability of showing severe symptoms—in [260] the user can find an example of COVID-19.

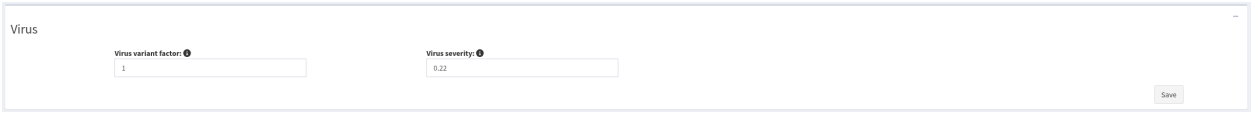


Figure B.36: Defining the virus parameters.

- The `Initial infected` panel allows the user to define the number of agents that will be initially infected. There are three options:

1. Random: if the number of initial infected is n, then these n agents will be randomly selected from all existing agents—Figure B.37.
2. Global: if the number of initial infected is n, then n agents of each agent type will be initially infected—Figure B.38.
3. Specific: for each agent type, the user can indicate the number of agents of that type that will be initially infected—Figure B.39.

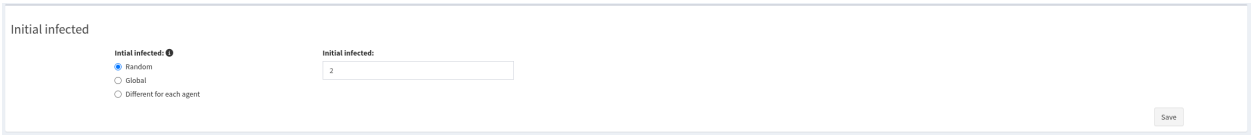


Figure B.37: Setting random initial infected.

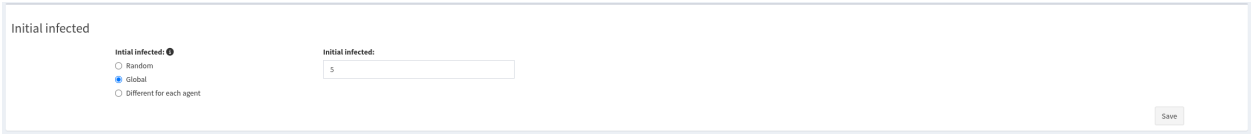


Figure B.38: Setting initially infected globally.

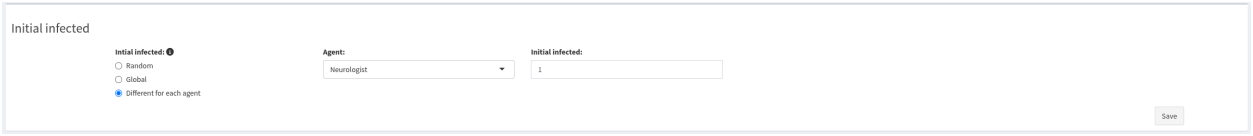


Figure B.39: Setting initially infected specifically.

- The `Outside contagion` panel—Figure B.40—allows users to model infections occurring outside the environment as a stochastic arrival process. This setup aligns the rate of external infections with the prevalence of the disease in the broader population. In particular, a probability is applied to each susceptible agent every day, representing the possibility of infection outside the environment. If the user wishes to model the infection outside the environment, they must upload a CSV file containing two columns: *day*, representing the simulation day, and *percentage_infected*, which indicates either the actual number of infected individuals in the considered population or the probability of acquiring the infection outside the environment on that day. If the former is provided, the user must specify the total population in the *Population* field; otherwise, they should set the population to 1. A plot is displayed after the file is uploaded—Figure B.41.

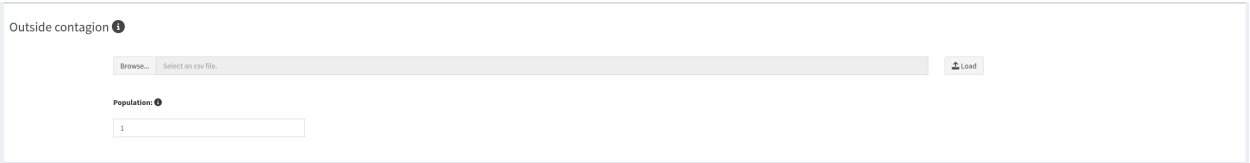


Figure B.40: Outside contagion panel.

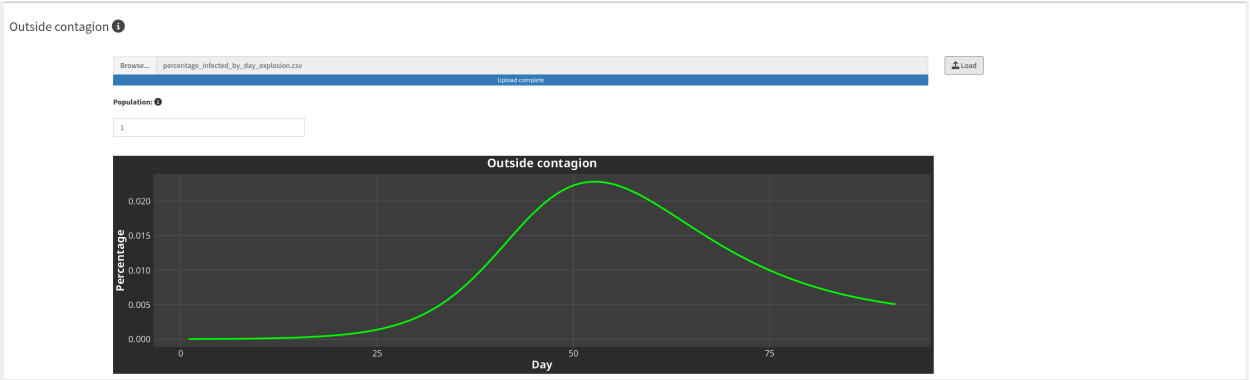


Figure B.41: Example of a plot generated after uploading an outside contagion file.

At the beginning of the *Virus Parameters* section, all configured virus parameters are displayed in the *Saved Virus Parameters* panel, as shown in Figure B.42. The first table shows the virus variant factor and virus severity, while the second one displays the initial infection set by the user. Entries can only be removed from the second table by clicking on them. An alert will appear, prompting the user to confirm the deletion or cancel the action. **Specific initial infections will take precedence over global ones, while randomly assigned infections are considered separately.**

Figure B.42: Saved countermeasures listed in the panel *Virus parameters*

B.3.8 Configuration

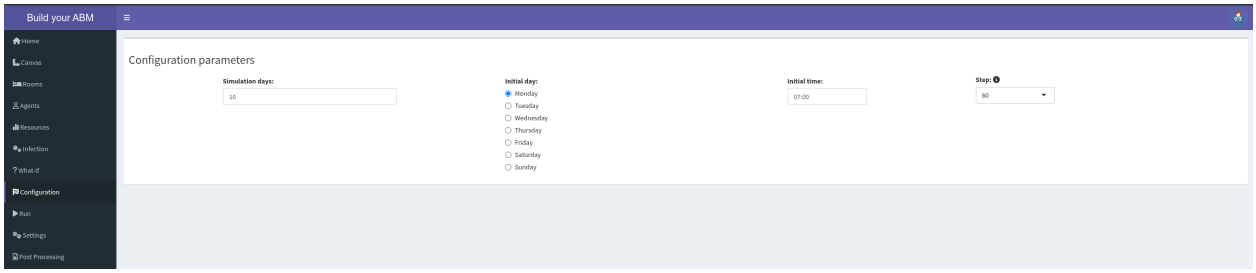


Figure B.43: The **Configuration** page.

On this page—see Figure B.43—, the user can configure various settings for the simulation:

- *Simulation days*: define the duration of the simulation in days.
- *Initial day*: set the start day of the week for the simulation.
- *Initial time*: determine the start time of the day for the simulation.
- *Step*: define the duration of a step in seconds; possible values are dividers of 60 seconds.

B.3.9 Run

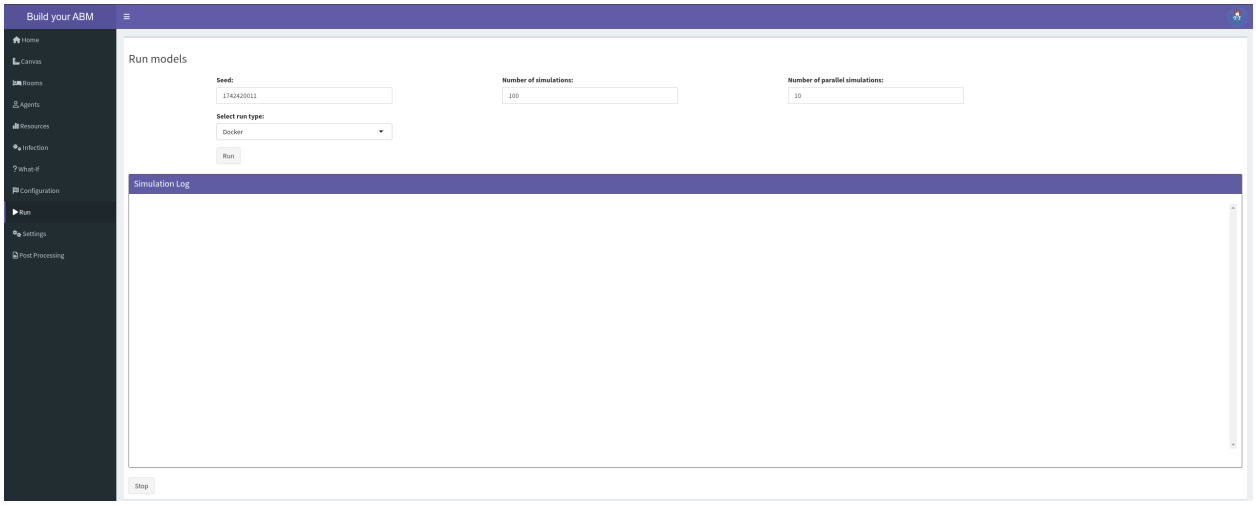


Figure B.44: The **Run** page.

On this page—see Figure B.44—, the user can configure various settings for the simulation:

- *Seed*: specify the starting seed for the simulation.

- *Number of simulations*: specify the number of simulations to run with a specific configuration of parameters.
- *Number of parallel simulations*: specify the number of simulations to run in parallel to speed up the computation.

Moreover, as shown in Figure B.45, the user can choose how to run the simulation they have just created using FLAME GPU 2. There are three options:

- Launch the simulation using Docker (see Section **??**); in this case, visualization will not be available.
- Launch the simulation locally with visualization.
- Launch the simulation locally without visualization.

These three options are available only when running *F4F* without Docker. If *F4F* is run using Docker Compose, the only available option will be *Docker*, while if *F4F* is run using Docker (without Compose), it will not be possible to run simulations using the *Run* button.

When the *Run* button is clicked, the system performs several checks, and the user must specify a location to save the results by clicking the *Select Folder* button, as shown in Figure B.46—when running the process with Docker Compose, the results are automatically saved in the directory containing the `docker_compose.yml` file. Additionally, the user must enter a name for the directory where the results will be stored. Finally, the user can terminate the running process by clicking the *Stop* button.

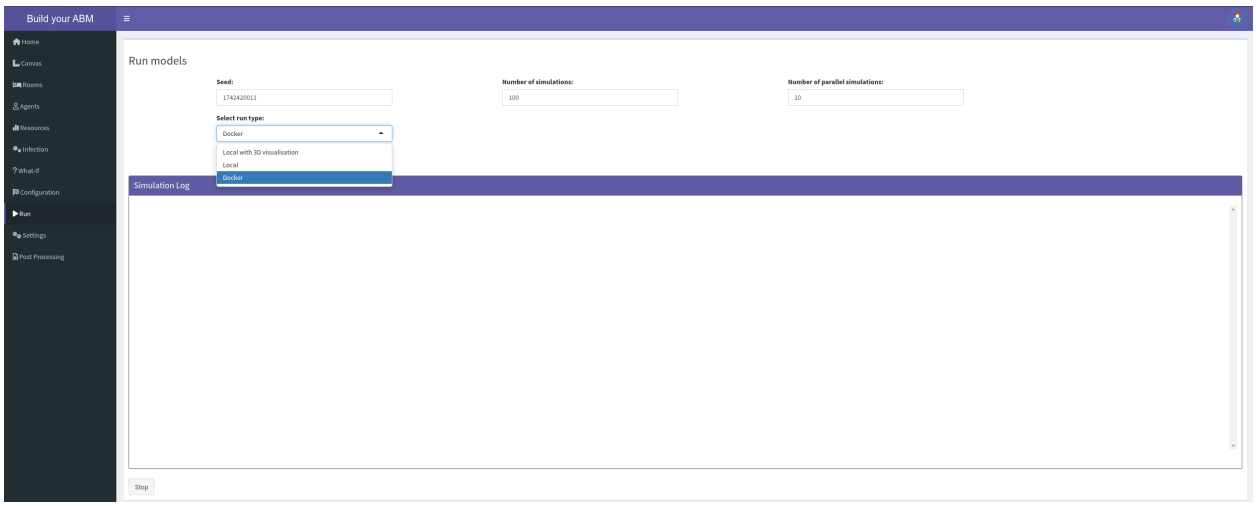


Figure B.45: Section where the user can select how to run their simulation.

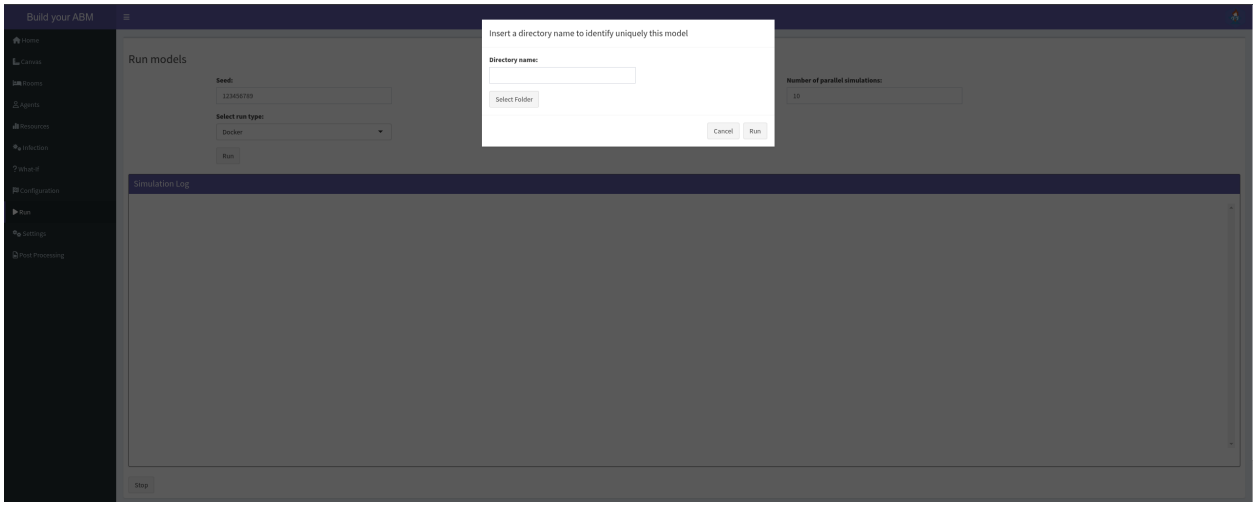


Figure B.46: Pop-up displayed when the *Run* button is clicked.

B.3.10 Settings

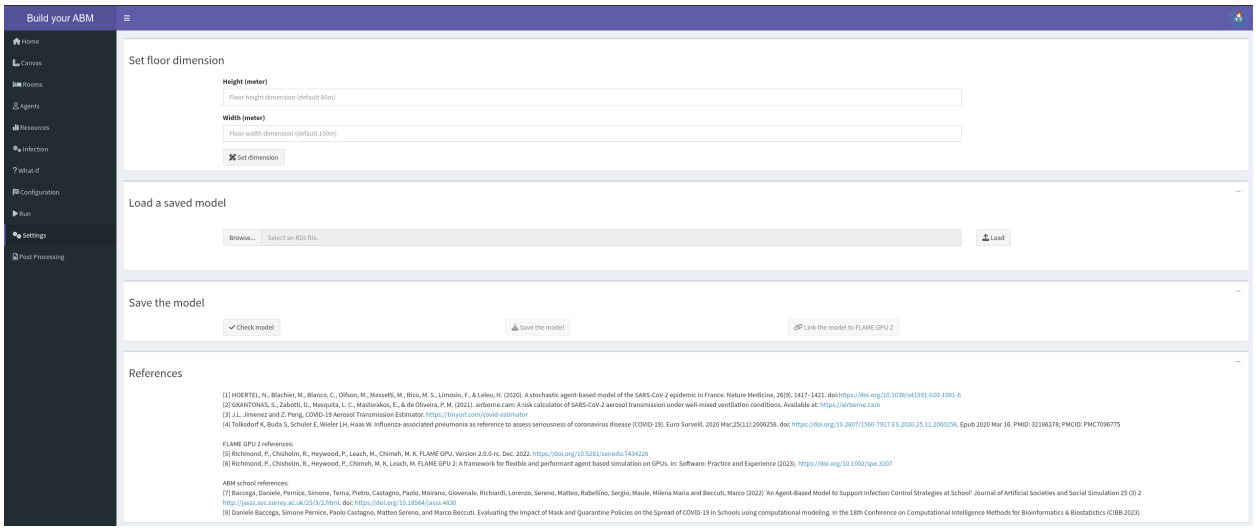


Figure B.47: The **Settings** page.

Inside the **Settings** page—see Figure B.47—the user can find three different panels:

- In the `Set floor dimension` panel, the user can adjust the dimensions of the canvas displayed on the Canvas page. The dimensions are specified in meters. The default dimensions are 80 meters in height and 100 meters in width. To set the desired dimensions, simply enter the values and press the *Set dimension* button.
- In the `Load a saved model` panel, the user can load a previously defined model. Click the *Browser* button to navigate through files and select the desired file. Once the

correct file is chosen, click the *Load* button to import the model into the application.

- In the `Save the model` panel, the user can save the defined model. Use the *Check model* button to perform basic validation checks before saving. Once the checks are complete, the user can save the entire model by clicking the *Save the model* button. Additionally, the user can link the newly created environment in the *F4F* application to FLAME GPU 2 by using the *Link the model to FLAME GPU 2* button, giving it a name. In this way, the model will be saved in every directory named `FORGE4FLAME/inst/FLAMEGPU-FORGE4FLAME/resources/f4f` within the user's home directory. **To run a model in FLAME GPU 2, it must be saved in this location.** The user can create a subdirectory within this path, placing the JSON file and the required RDs downloaded using the *Save the model* button. This process is automated when using the *Link the model to FLAME GPU 2* button.
- In the `References` panel, the user will find the sources and references that were used as a basis for the development of the tool.

B.3.11 Post Processing

On this page—see Figure B.48—users can analyze and visualize simulation results, filter specific simulations of interest, and observe virus evolution in each room, among other aspects. The section is divided into several panels, as described below. Additionally, if *F4F* is run using Docker without Compose, loading the simulation results will not be possible.

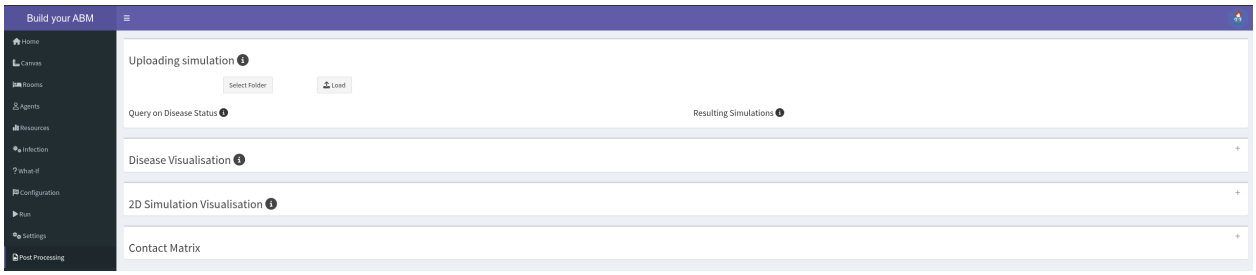


Figure B.48: The **Post Processing** page.

Uploading simulation

Here, the user must upload the folder containing all the files and directories generated by the simulation—see Section B.2 for more details on where results are saved. The required files, automatically generated by FLAME GPU 2, include `rooms_mapping.txt`, `seed.txt`, and all directories named `seed*`, each containing the following files: `evolution.csv`, `counters.csv`,

`AEROSOL.csv`, `CONTACTS_MATRIX.csv`, `CONTACT.csv`, `INFO.csv`, and `AGENT_POSITION_AND_STATUS.csv`. For more details on these files, see Section B.5.

Once the directory is chosen, a set of sliders will appear, as shown in Figure B.49. These sliders allow users to select a specific subset of simulations—listed on the right side of the figure—by querying them. Users can refine their selection by limiting the simulation days and filtering based on the number of Susceptible, Exposed, Infected, Recovered, and Deceased agents. Based on these filters, the *Resulting Simulations* section will display directories containing simulations that meet the user's specified ranges. For instance, in Figure B.49, only the simulations in which, between days 3 and 6, there were 20-30 susceptible agents, 0 exposed, 3-7 infected, 0-10 recovered, and 0 deaths were selected.

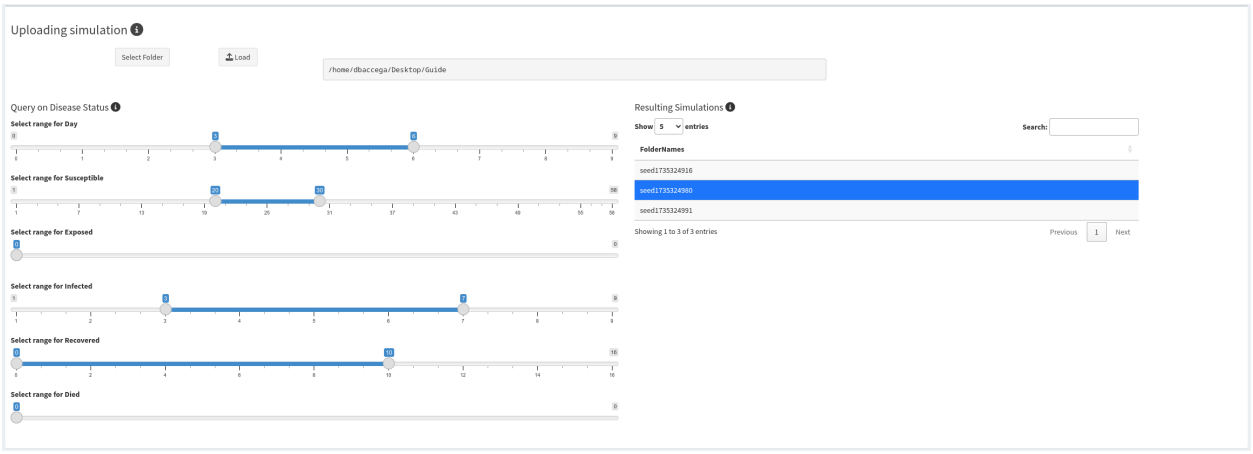


Figure B.49: Example of query.

By clicking on each listed directory, the corresponding plots will appear below. As shown in Figure B.50, these plots illustrate the daily evolution of the disease, along with additional data such as the number of agents birth or deaths (in the case of agents with a birth rate), the number of agents in quarantine, the number of agents infected outside the environment, and the number of swabs conducted. Additionally, the user can select a specific room and observe how contacts between susceptible and infected agents, as well as the virus concentration, have varied on an hourly basis—a contact between two agents is defined as the situation where they remain close to each other for a certain number of steps without ever separating. Finally, for each of these plots, the user can display the mean curves along with the areas representing all simulations.

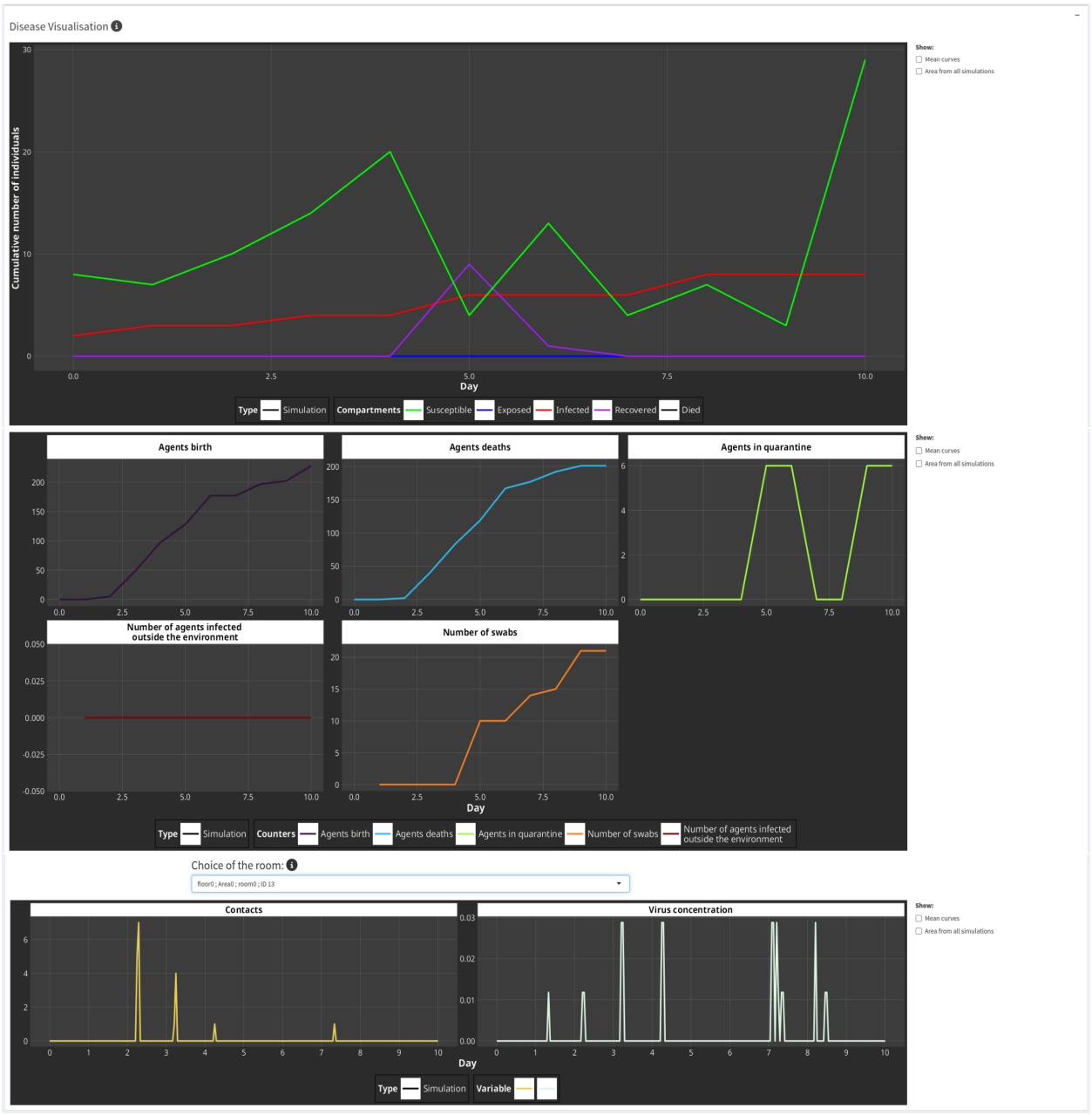


Figure B.50: Plots generated after selecting a listed directory.

2D Simulation Visualization

In this panel—see Figure B.51—the user can navigate through two main sections:

- In the *Features* panel, the user can configure various display options:
  - Choose which floors to display (default: all floors).
  - Choose which agent types to display (default: all agent types).

  - Choose room colors based on name, type, area, cumulative number of contacts, current or cumulative virus concentration (default: name).
  - Enable additional visual elements, such as room IDs, names, types or areas and agent IDs.

  Figure B.52 illustrates an example based on the model described in previous sections. Additionally, when an agent type is selected, the user can filter and display specific agents by their IDs—see Figure B.53.

- In the *2D Visualisation* panel, the dynamics of the simulation are visualized in a 2D plot. The user can observe agent movements and changes in disease states using a time slider. In particular:

  - Different shapes represent different agent types.
  - Colors indicate disease states.
  - If a shape or color is missing in the legend, it means no agent with those attributes is currently visible on the selected floor(s).

  The animation step can be adjusted via the input field next to the slider. The user can advance the simulation using the *Next* button, which increments the time step by the specified value and updates the 2D visualization automatically, eliminating the need for manual adjustments. Finally, Figure B.54 shows the cumulative virus concentration at the final time.

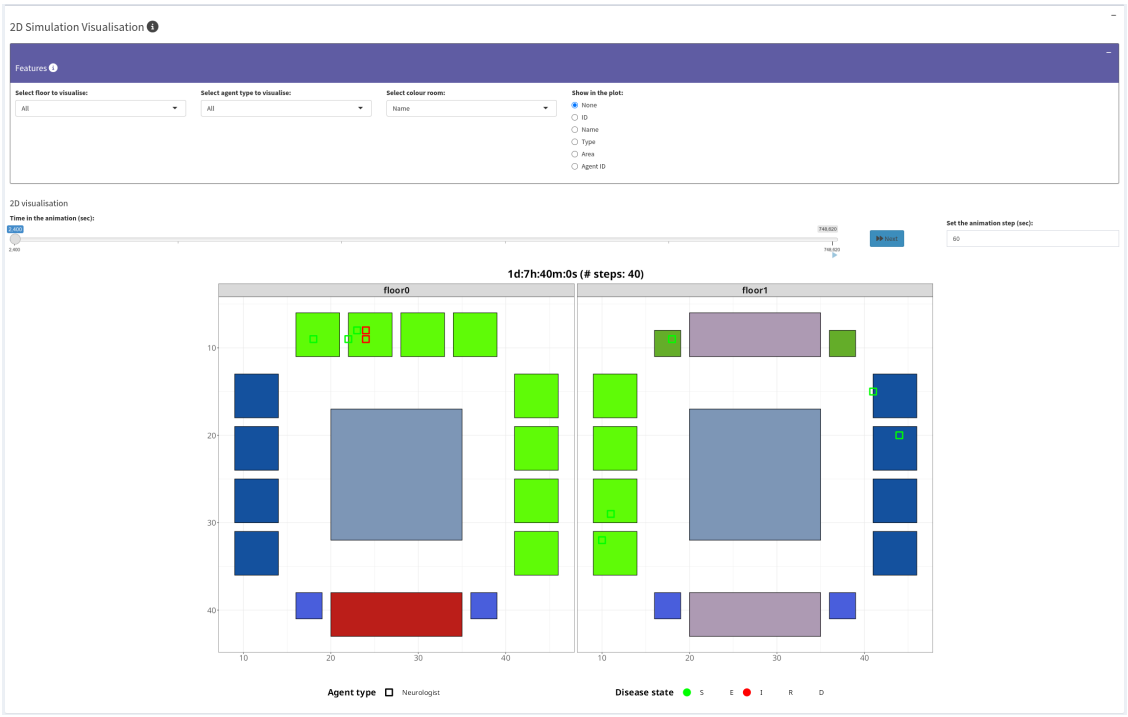


Figure B.51: The 2D visualization section.

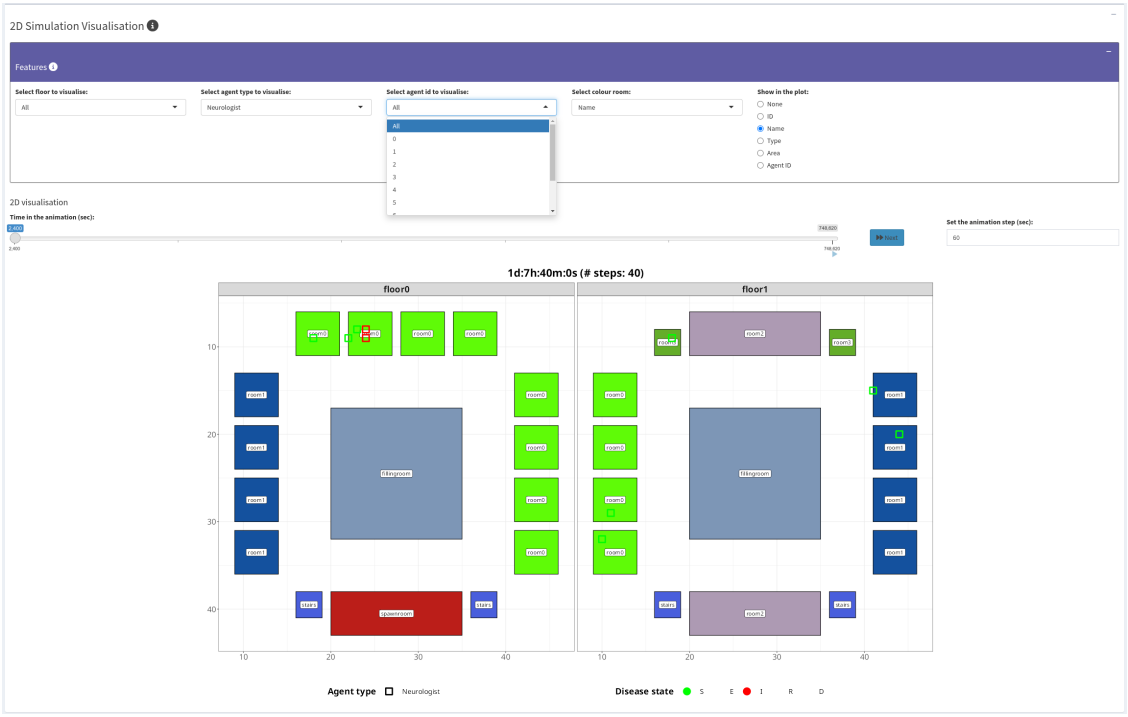


Figure B.52: The 2D visualization section in which rooms' names and only the agent type *Neurologist* are displayed.

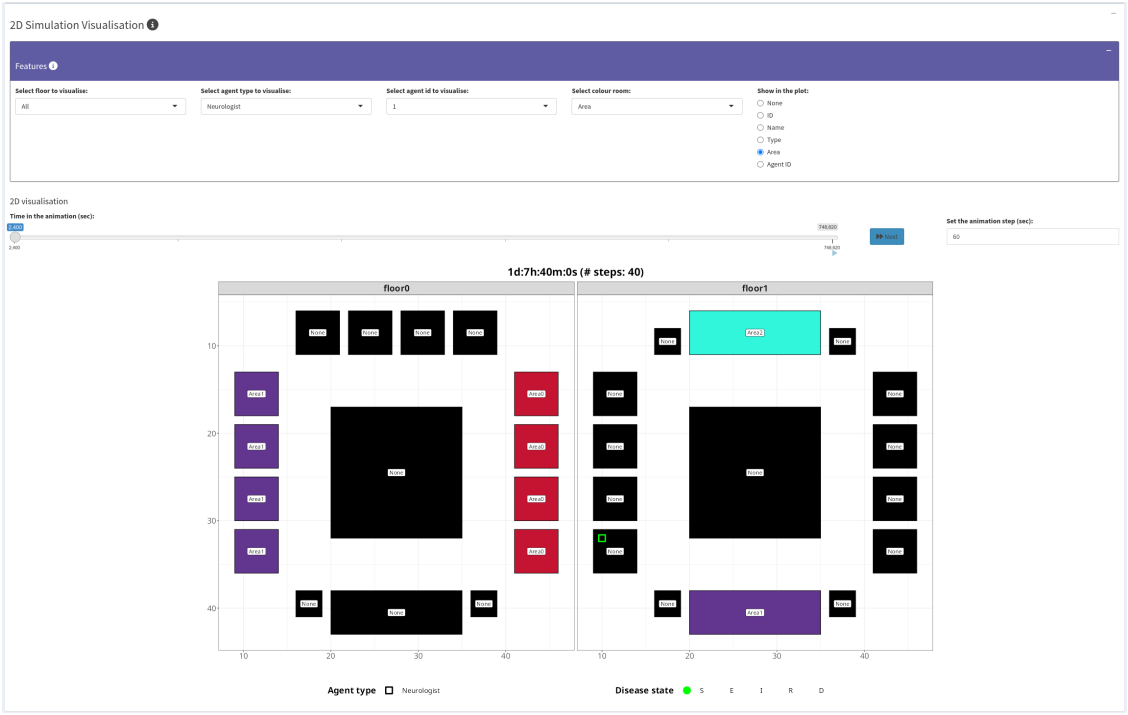


Figure B.53: The 2D visualization section in which rooms' areas and only the agent type *Neurologist* with id 1 are displayed.

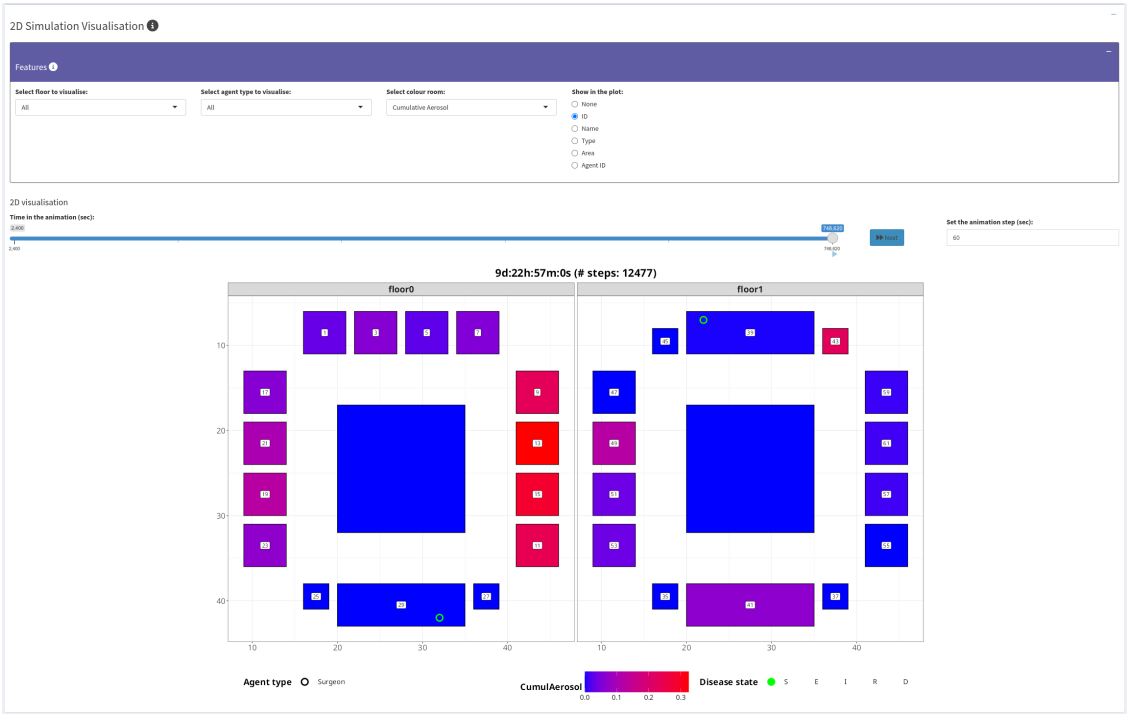


Figure B.54: The 2D visualization section in which the cumulative virus concentration at the final time is shown together with rooms' IDs.

Contact Matrix

In this panel—see Figure B.55—the lower triangular contact matrix shows the hourly average number of contacts among the different agent types.

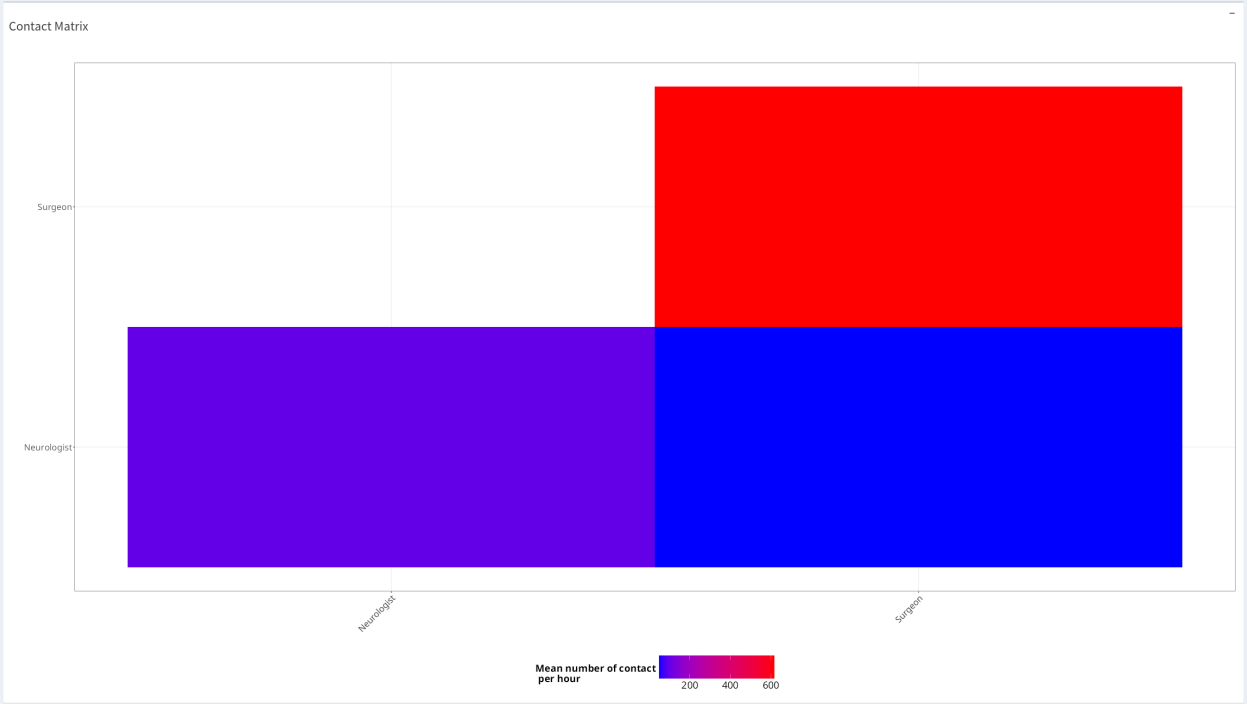


Figure B.55: The *Contact Matrix* panel.

B.4 School

In this section, we provide further details about the generic middle school environment described in Chapter 5. Figures B.56, B.57, and B.58 illustrate the three floors of the school defined using *F4F*. Specifically, *floor0* includes the spawn-room where agents enter the environment, along with four classrooms, two bathrooms—one for students and one for teachers—the gym, the teacher's room, the principal's office, the temperature measurement room, the stairs connecting the floors, a couple of filling rooms, and four graph points in the hallway. On *floor1* and *floor2*, there are four classrooms, a student bathroom, stairs, some filling rooms, and two graph points in the hallway.

Figures B.59, B.60, and B.61 illustrate the definition of the agent *student_1A*. Figure B.59 presents one of the two determined flows associated with the agent, where the agent spends 300 minutes in classroom 1A, with two 15-minute breaks—one between the second and third classes and another between the fourth and fifth classes. Figure B.60 highlights the random events—two random events are possible: going to the bathroom or visiting the principal's

office. Finally, Figure B.61 depicts one of the two entry flows linked to the first determined flow, which occurs every day except on Thursday, Saturday, and Sunday.

Figures B.62, B.63, and B.64 illustrate the definition of the agent *teacher_0*. Figure B.62 presents one of the three determined flows associated with the agent, where the agent will spend 115 minutes in classroom 1A. Figure B.63 highlights the random events—only one possible random event includes going to the bathroom. Finally, Figure B.64 depicts one of the three entry flows linked to the first determined flow, which occurs on Monday and Tuesday.

Figure B.65 illustrates the infection parameters of the SEIRS compartmental model used in the alert experiments in Chapter 5. Specifically, we set the average incubation and infection periods to 5 days and the average immunization period to one year and a half.

Figures B.66, B.67, B.68, and B.69 present screenshots from the *Post Processing* page considering the *No Countermeasures* scenario introduced in Chapter 5. The selection criteria for these simulations required that, between days 20 and 40, the number of agents fell within the following ranges: 100–200 susceptible, 30–60 exposed, 20–40 infected, 0–150 recovered, and 0 deaths. Figure B.67 illustrates the trajectories of one of the 56 simulations that meet the query conditions shown in Figure B.66. In the 2D visualization of the school—Figure B.68—, the snapshot captures the day 25^{th} of the simulation, depicting students in classrooms, the principal in his office, and a few teachers in the teacher's room. Additionally, the color of each room represents the hourly virus concentration. Figure B.69 shows the contact matrix among the different types of agents. Finally, we do not include screenshots of the *What-If*, *Configuration*, and *Settings* pages, as users have the flexibility to customize parameters as needed. Furthermore, no restrictions were imposed on room access based on resource availability, as each room was globally assigned 1,000 resources.

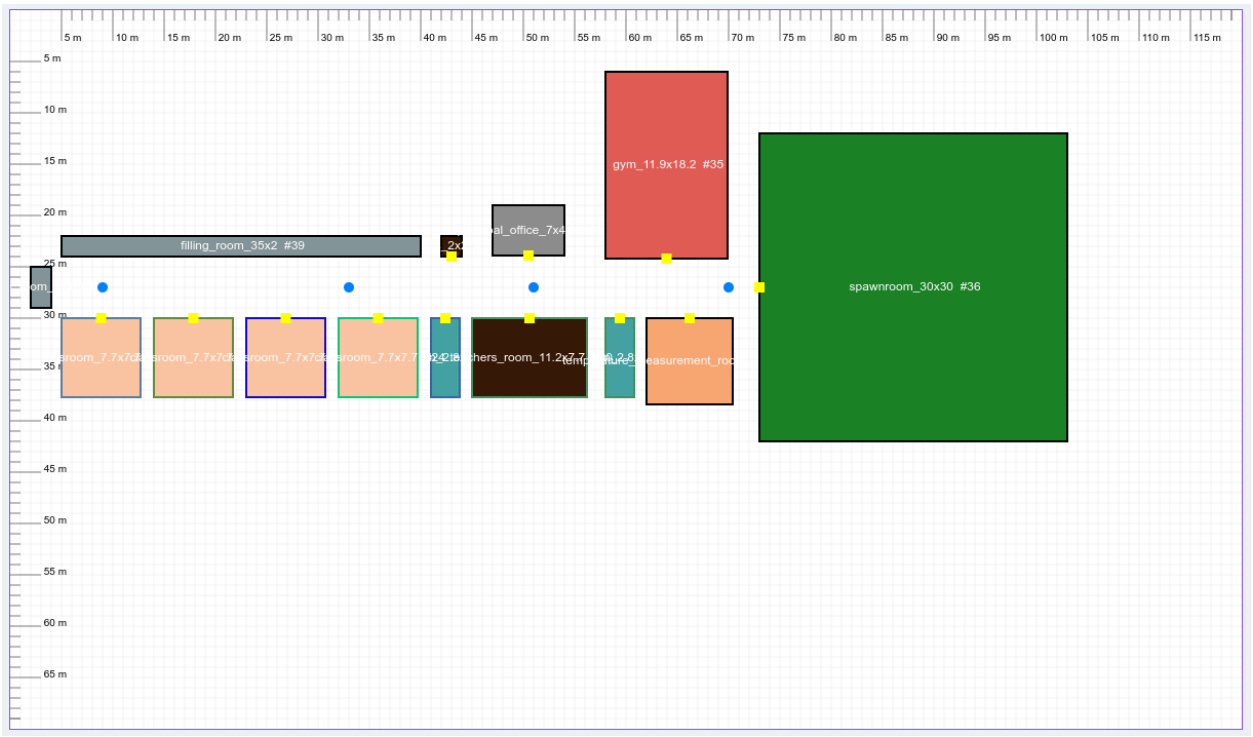


Figure B.56: *floor0* of the school environment visualized in the canvas of *F4F*.

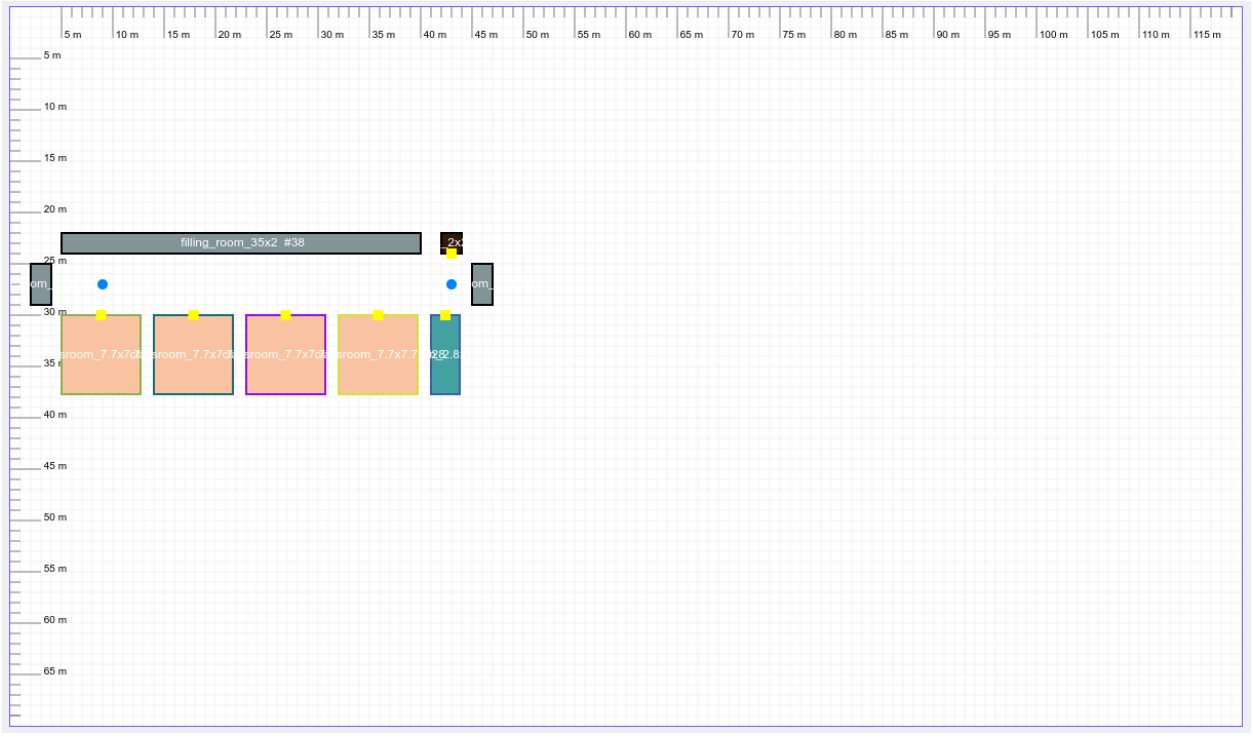


Figure B.57: *floor1* of the school environment visualized in the canvas of *F4F*.

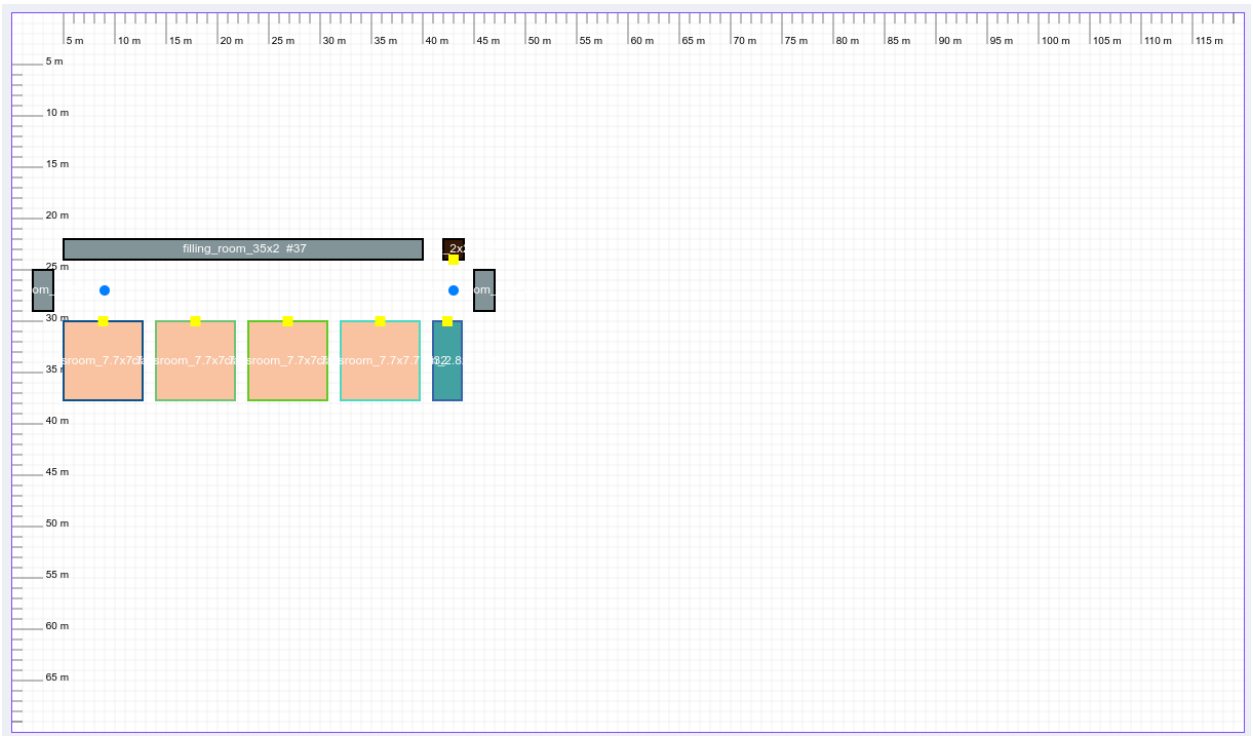


Figure B.58: *floor2* of the school environment visualized in the canvas of *F4F*.

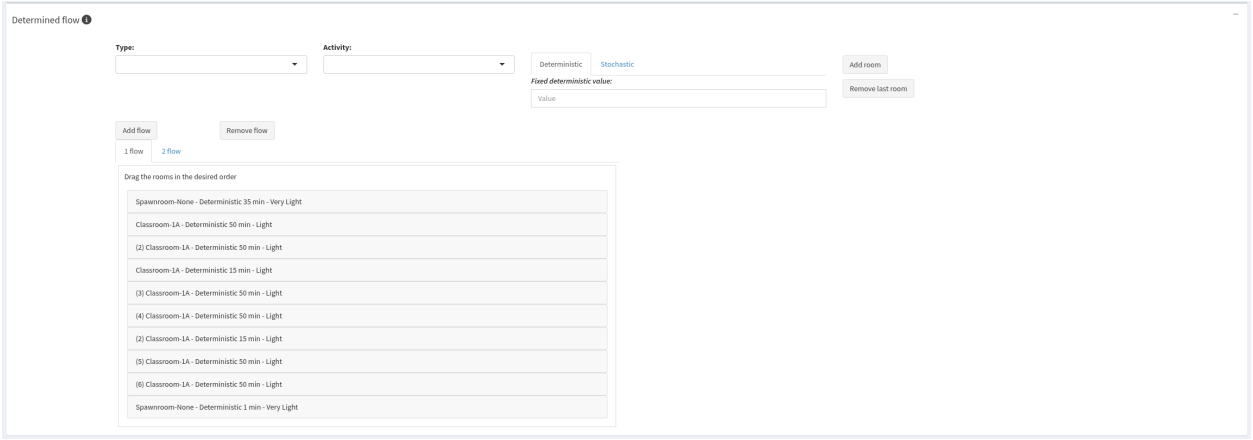


Figure B.59: Determined flow of agent *student_1A*.

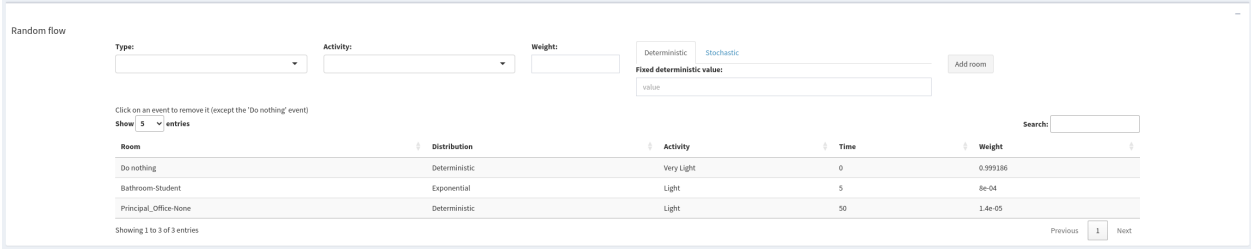


Figure B.60: Random flow of agent *student_1A*.

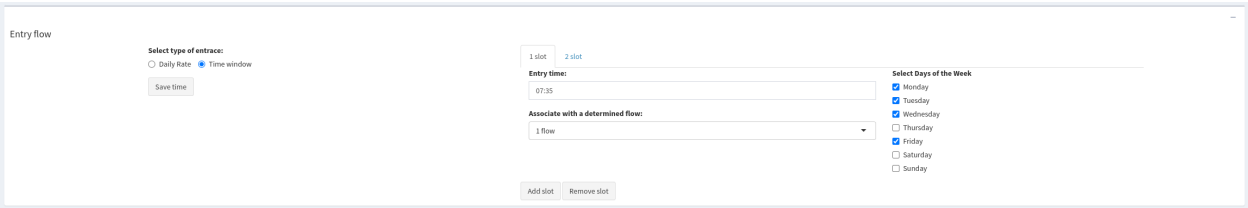


Figure B.61: Entry flow of agent *student_1A*.

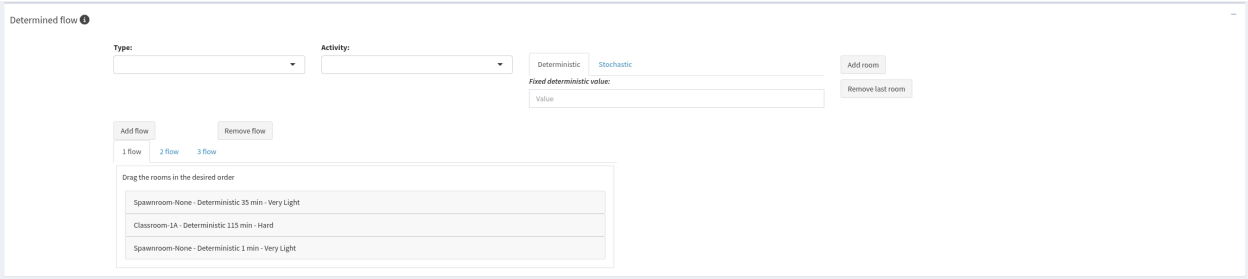


Figure B.62: Determined flow of agent *teacher_0*.

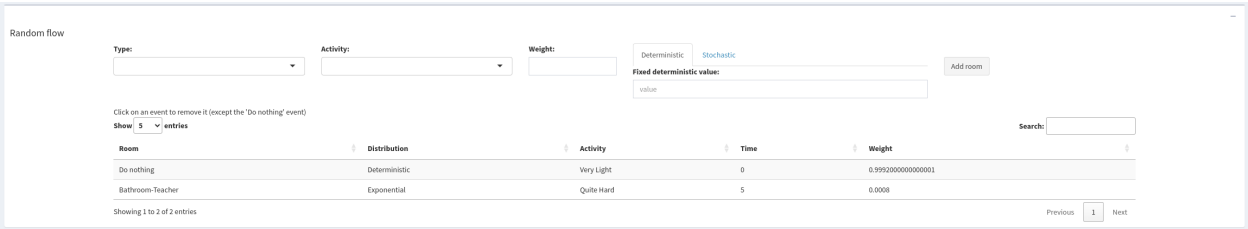


Figure B.63: Random flow of agent *teacher_0*.

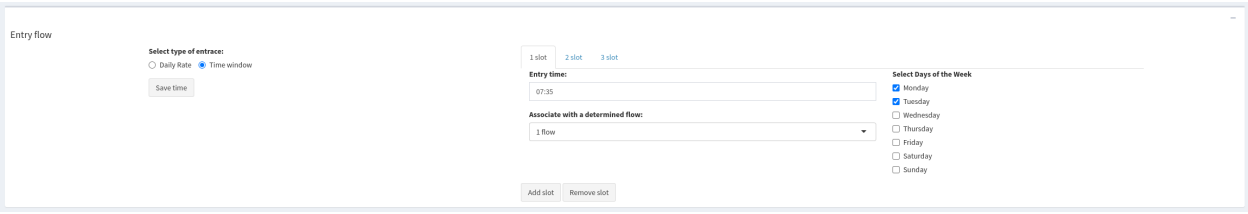


Figure B.64: Entry flow of agent *teacher_0*.

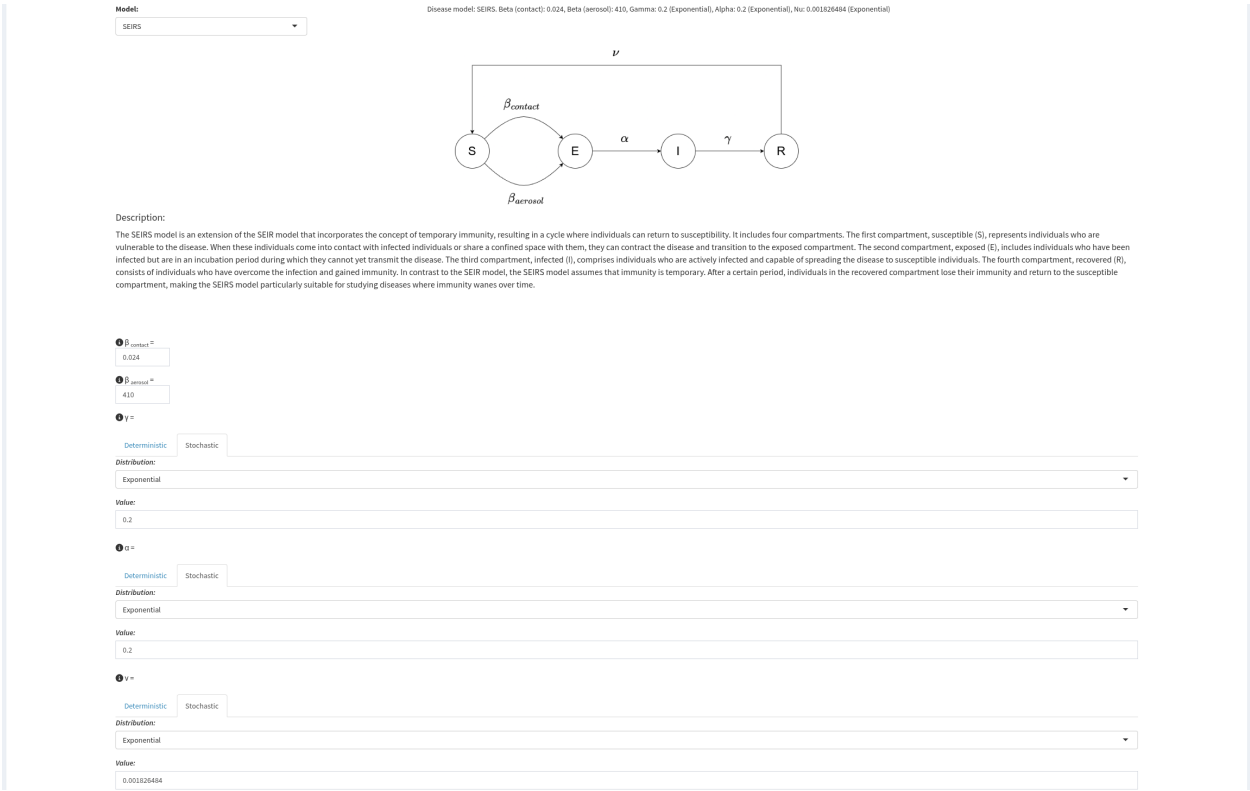


Figure B.65: Infection parameters used to validate the school model.

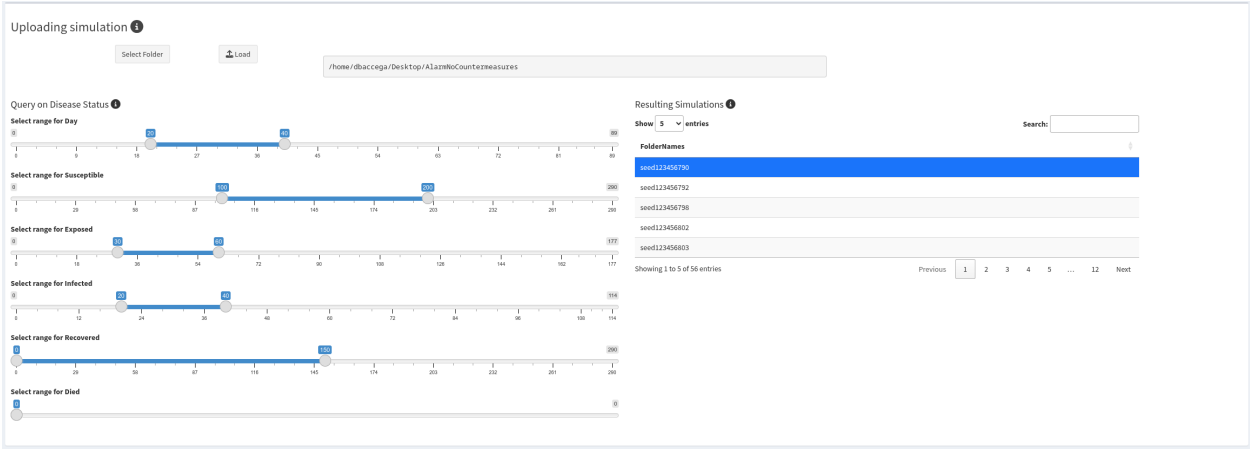


Figure B.66: Example of query on the results obtained using the school model on the *No Countermeasures* scenario.

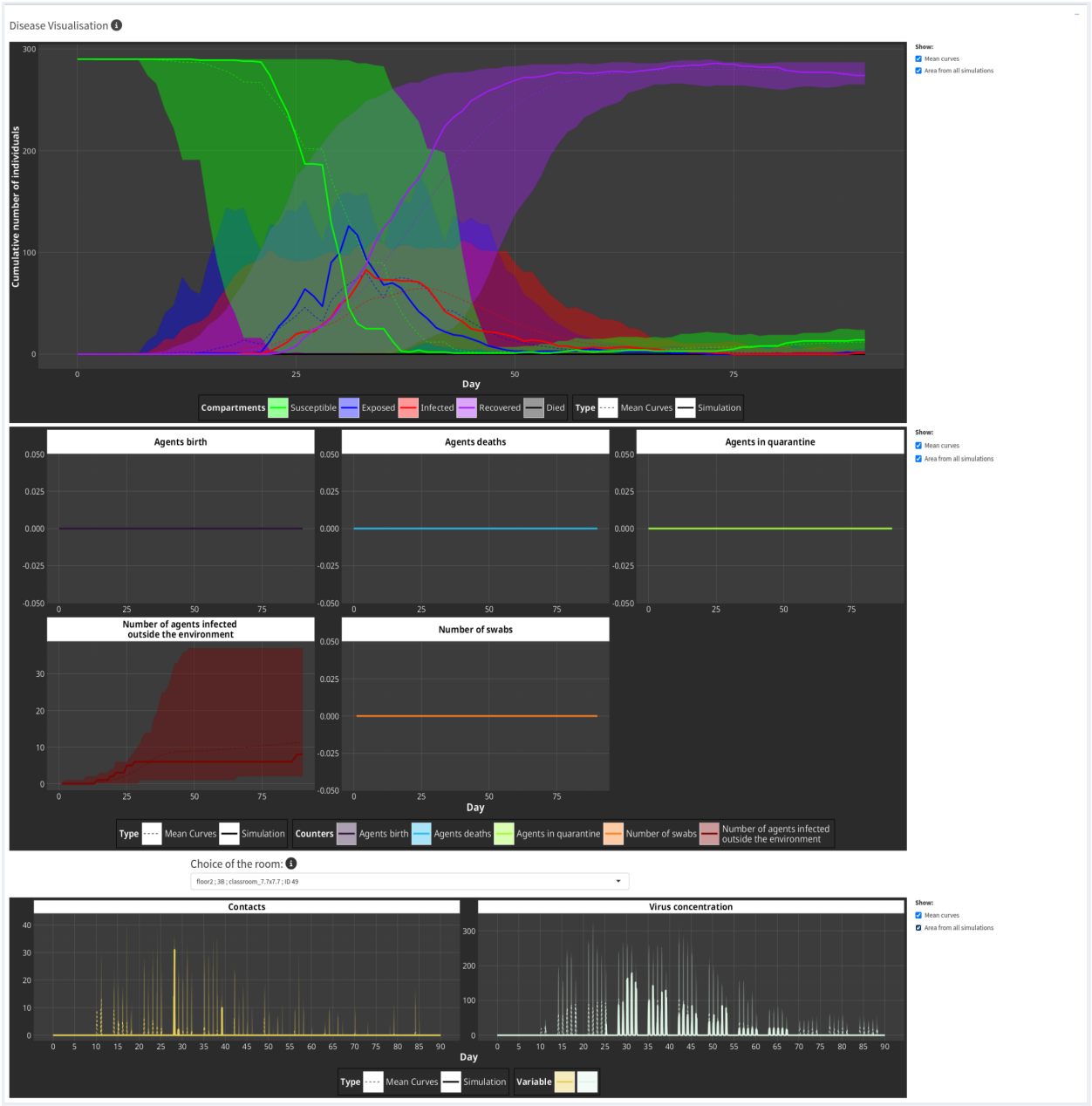


Figure B.67: Plots generated after selecting a listed directory using the school model on the *No Countermeasures* scenario (with means and areas).

Figure B.68: 2D visualization of the school where the color of each room represents the hourly virus concentration on the *No Countermeasures* scenario.

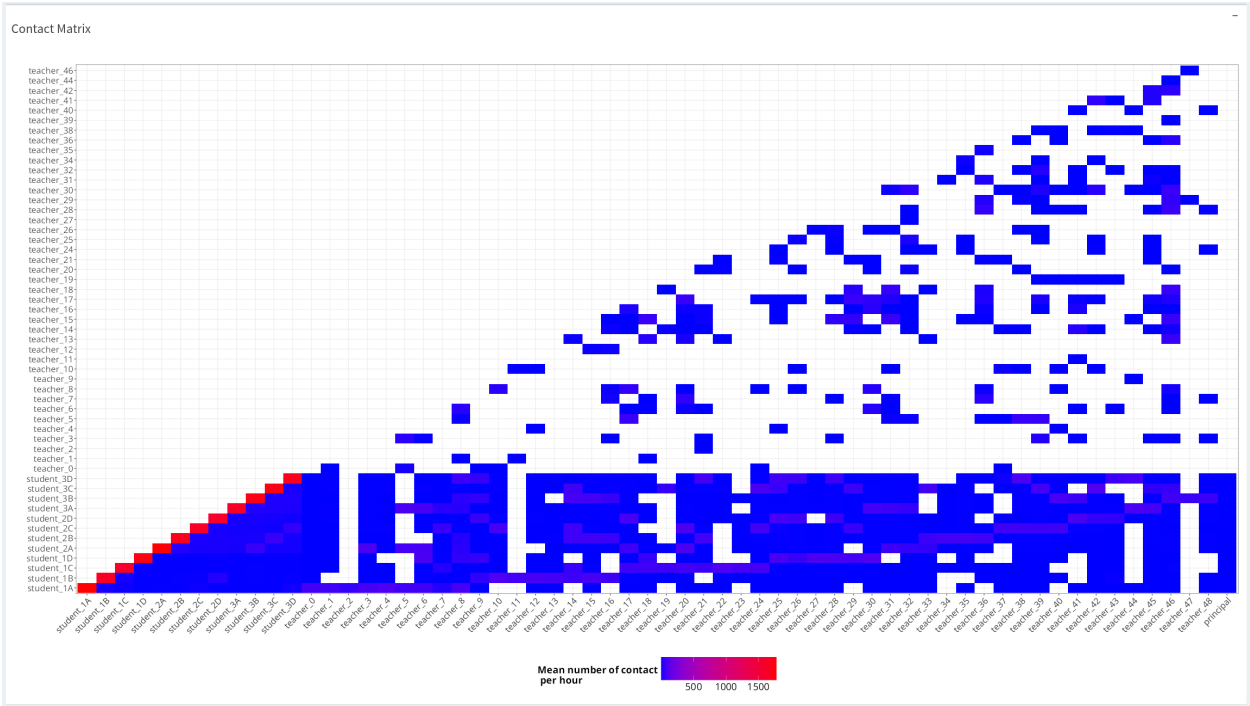


Figure B.69: Contact matrix of the school model on the *No Countermeasures* scenario.

B.5 Link between F4F and FLAME GPU 2

The output of *F4F*—a JSON file including the whole model—is linked to FLAME GPU 2 through a Python script which automatically generates the necessary code to run the simulation. In particular, the information required to simulate the model is stored in FLAME GPU 2 using environmental properties and macro-properties, agents' variables, and various C++ *defines*.

One of the key components of FLAME GPU 2 simulations is the movement graph, briefly introduced in Section B.3. This graph is automatically generated using room centers, room doors, user-defined graph points, and stairs—the latter connects floors to each other. It is stored in FLAME GPU 2 environment and macro-environment variables to handle agent movements. Using the A* algorithm, the graph enables agents to determine the shortest path between a source and a destination whenever they need to navigate the modeled environment. For example, Figure B.70 illustrates all possible edges of the graph for *floor0* in the model defined in Section B.3.

When running the FLAME GPU 2 simulation with the visualization enabled—without Docker, see Section B.2 for more details—, the environment created using *F4F*—as defined in Section B.3.2—will be displayed, as shown in Figure B.71. Moreover, in Figure B.72 a 3D visualization of the school environment is shown.

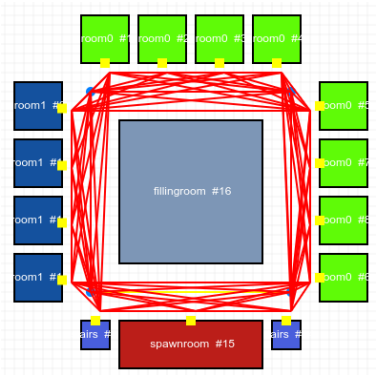


Figure B.70: All possible edges of the graph for *floor0* in the model defined in Section B.3

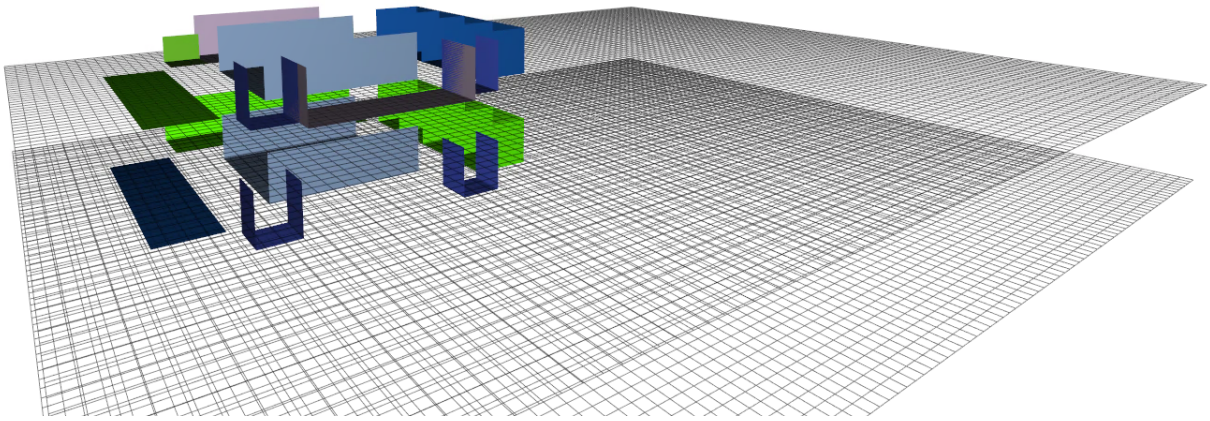


Figure B.71: 3D visualization in FLAME GPU 2 of the example depicted in Section B.3.

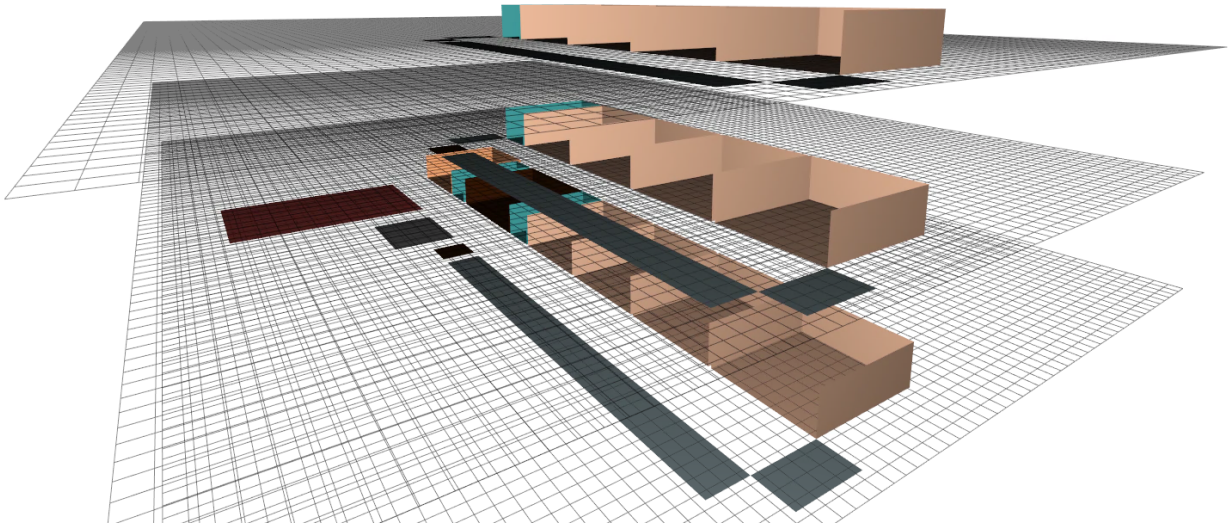


Figure B.72: 3D visualization in FLAME GPU 2 of the school depicted in Section B.4 and used in Chapter 5.

In Chapter 5 and in Section B.2, we outline the ways users can run *F4F* and FLAME GPU 2 explaining how users can effectively run simulations. When FLAME GPU 2 completes the simulation runs, it generates CSV files containing various information required to post-analyze the simulated scenario (e.g., disease evolution, contact matrix, virus diffusion within the environment, and agents' movements). This information is stored in separate files—one for disease evolution, one for the contact matrix, and so on. These files can be passed to *F4F* for post-analysis and visualization of disease dynamics in the modeled system—see Section B.3.11.

Specifically, each folder containing the model's results includes:

- `rooms_mapping.txt` stores the mapping between each room's ID in the graph and the corresponding position of its centre;
- `seed.txt` stores the simulation seed specified using *F4F*;
- various directories named `seed*`—there are as many directories as there are simulation runs—, each containing the following files:
  - `evolution.csv` stores the daily evolution of the disease; the file contains six columns: *Day*, *Susceptible*, *Exposed*, *Infected*, *Recovered*, and *Died*;
  - `counters.csv` stores the daily evolution of some interesting pre-defined counters—already introduced in Section B.3.11; the file contains five columns: *Day*, *COUNTERS_CREATED_AGENTS_WITH_RATE*, *COUNTERS_KILLED_AGENTS_WITH_RATE*, *AGENTS_IN_QUARANTINE*, *SWABS*, and *NUM_INFECTED_OUTSIDE*;
  - `AEROSOL.csv` stores the hourly evolution of virus concentration in each room—entries are only logged when the virus concentration is greater than zero; the file contains three columns: the first represents the FLAME GPU 2 step—i.e., simulation time—, the second indicates the virus concentration, and the third corresponds to the room's ID in the graph;
  - `CONTACTS_MATRIX.csv` stores the hourly lower triangular contact matrix—entries are logged only if the number of contacts in the last hour is greater than zero; the file contains four columns: the first represents the FLAME GPU 2 step, the second and third indicate the agent types, and the fourth records the number of contacts in the last hour between those agent types;
  - `CONTACT.csv` stores interactions between susceptible and infected agents; it contains four columns: the first indicates the FLAME GPU 2 step, the second and third specify the types of the susceptible and infected agents, respectively, and the fourth provides the ID of the room where the contact occurred; if agents remain in the contact area from step t_1 to t_2—where t_1 ¡ t_2—, the user will find t_2 - t_1 rows in the file, corresponding to the time steps t_1, $t_1 + 1$, ..., t_2;
  - `AGENT_POSITION_AND_STATUS.csv` stores the position of each agent whenever it changes from the previous FLAME GPU 2 step; the file contains seven columns: the first indicates the FLAME GPU 2 step, the second represents the agent's ID, the third specifies the agent type, the fourth, fifth, and sixth provide the agent's position on the x, y, and z axes, respectively, and the seventh denotes the agent's disease state.

- `INFO.csv` contains information about the simulation. Its columns are formatted as follows: the first column indicates the FLAME GPU 2 step, and the second column is a string.
- `host_rng_state.txt` stores the state of the host's random number generator at the end of the simulation.

Appendix C

Other Works

C.1 Convolutional neural network (CNN)-enabled electrocardiogram (ECG) analysis: a comparison between standard twelve-lead and single-lead setups

Artificial intelligence (AI) has shown promise in the early detection of various cardiac conditions from a standard 12-lead electrocardiogram (ECG). However, the ability of AI to identify abnormalities from single-lead recordings across a range of pathological conditions remains to be systematically investigated. This study aims to assess the performance of a convolutional neural network (CNN) using a single-lead (D1) rather than a standard 12-lead setup for accurate identification of ECG abnormalities.

We designed and trained a lightweight CNN—Figure C.1—to identify 20 different cardiac abnormalities on ECGs, using data from the PTB-XL dataset [232]—Figure C.2. With a relatively simple architecture, the network was designed to accommodate different combinations of leads as input (¡100,000 learnable parameters). We compared various lead setups such as the standard 12-lead, D1 alone, and D1 paired with an additional lead.

The CNN based on single-lead ECG (D1) achieved satisfactory performance compared to the standard 12-lead framework (average percentage AUC difference: -8.7%)—Figure C.3. Notably, for certain diagnostic classes, there was no difference in the diagnostic AUC between the single-lead and the standard 12-lead setups. When a second lead was detected in the CNN in addition to D1, the AUC gap was further reduced to an average percentage difference of -2.8% compared with that of the standard 12-lead setup—Figure C.4.

A relatively lightweight CNN can predict different classes of cardiac abnormalities from D1 alone and the standard 12-lead ECG. Considering the growing availability of wearable

devices capable of recording a D1-like single-lead ECG, we discuss how our findings contribute to the foundation of a large-scale screening of cardiac abnormalities.

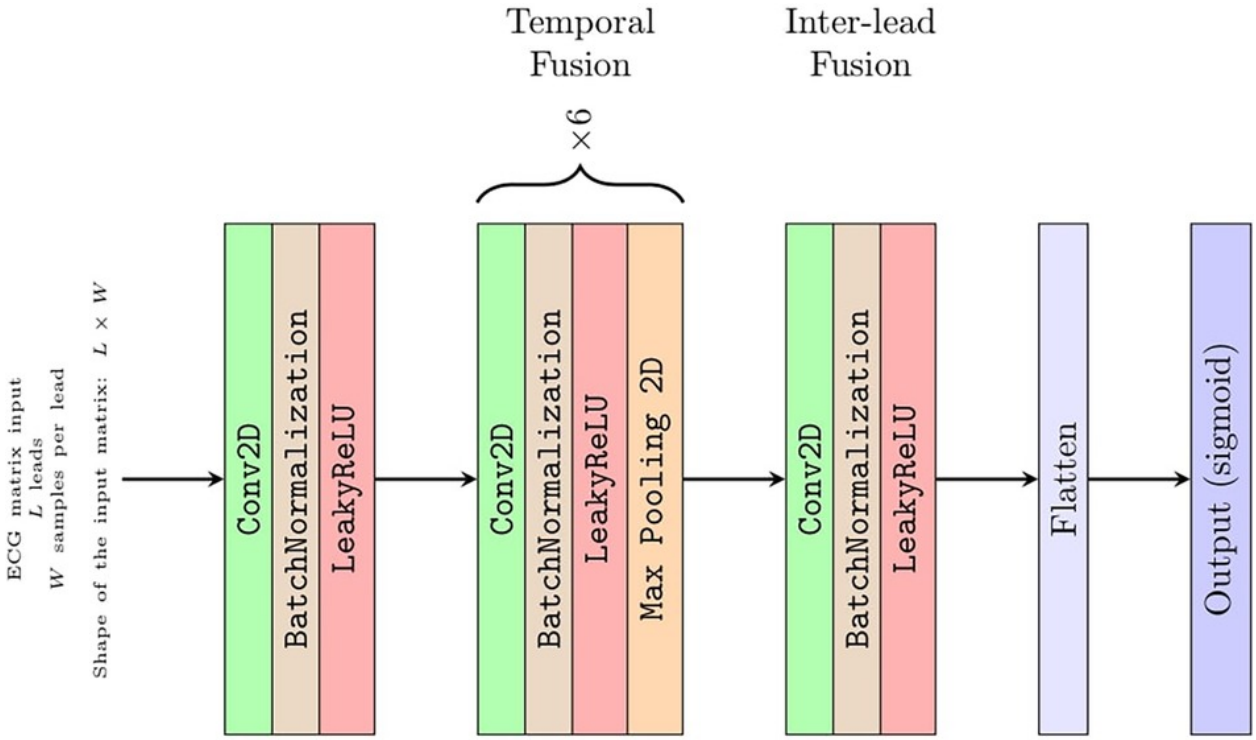


Figure C.1: Detailed CNN architecture.

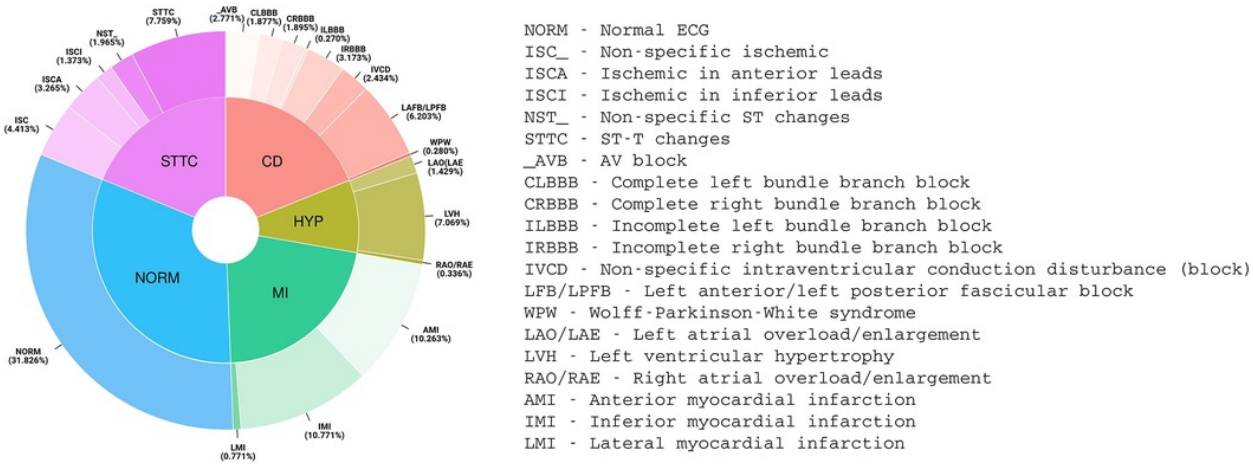


Figure C.2: Distribution of PTB-XL ECG diagnoses (aggregated into 20 diagnostic classes and 5 diagnostic superclasses).

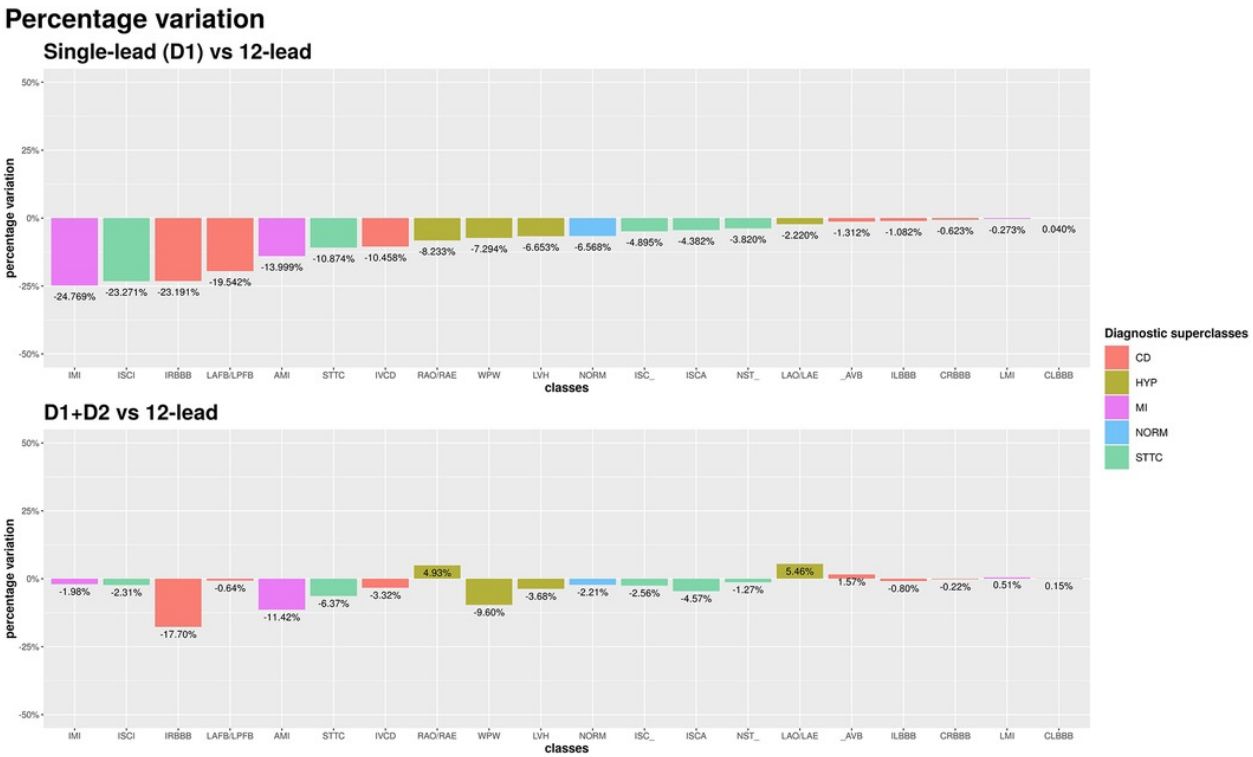


Figure C.3: AUC percentage variation among single-lead (D1) and 12-lead scenarios (top chart) and among D1 + D2 and 12-lead scenarios (bottom chart). Percentage differences are reported relative to the 12-lead setup.

Classes	D1	D1 + D2	D1 + V1	D1 + V2	D1 + V3	D1 + V4	D1 + V5	D1 + V6	8 leads	12 leads	12 w/o aug	Ref
NORM	89.76 [89.69, 89.83]	93.95 [93.91, 93.99]	92.00 [91.94, 92.06]	91.91 [91.84, 91.98]	92.23 [92.17, 92.29]	92.90 [92.86, 92.94]	93.07 [93.01, 93.13]	92.79 [92.75, 92.83]	96.06 [96.02, 96.10]	**96.07** [96.03, 96.11]	96.06 [95.96, 96.16]	95.2
ISC_	90.73 [90.64, 90.82]	92.96 [92.88, 93.04]	91.81 [91.72, 91.90]	91.68 [91.58, 91.78]	91.64 [91.54, 91.74]	93.46 [93.38, 93.54]	95.18 [95.10, 95.26]	95.57 [95.50, 95.64]	95.28 [95.19, 95.37]	95.40 [95.31, 95.49]	95.10 [94.96, 95.24]	**96.5**
ISCA	87.94 [87.81, 88.07]	87.77 [87.65, 87.89]	89.11 [88.98, 89.24]	88.76 [88.62, 88.90]	89.68 [89.53, 89.83]	89.74 [89.60, 89.88]	89.44 [89.29, 89.59]	88.47 [88.34, 88.60]	91.85 [91.70, 92.00]	91.97 [91.80, 92.14]	90.38 [90.00, 90.76]	**93.4**
ISCI	72.01 [71.37, 72.65]	91.68 [91.40, 91.96]	71.75 [71.17, 72.33]	77.32 [76.75, 77.89]	76.59 [76.04, 77.14]	79.16 [78.64, 79.68]	81.60 [81.16, 82.04]	81.90 [81.42, 82.38]	92.68 [92.40, 92.96]	**93.85** [93.48, 94.22]	90.72 [90.16, 91.28]	91.5
NST_	84.86 [84.59, 85.13]	87.11 [86.86, 87.36]	85.09 [84.82, 85.36]	84.60 [84.29, 84.91]	84.86 [84.64, 85.08]	84.97 [84.71, 85.23]	87.32 [87.07, 87.57]	86.57 [86.30, 86.84]	**88.24** [88.03, 88.45]	88.23 [87.96, 88.50]	86.92 [86.45, 87.39]	86.7
STTC	80.73 [80.58, 80.88]	84.81 [84.69, 84.93]	81.97 [81.84, 82.10]	84.52 [84.38, 84.66]	85.87 [85.76, 85.98]	87.53 [87.42, 87.64]	88.29 [88.19, 88.39]	86.53 [86.43, 86.63]	90.11 [90.00, 90.22]	90.58 [90.48, 90.68]	89.60 [89.45, 89.75]	**91.1**
_AVB	92.50 [92.21, 92.79]	95.20 [95.09, 95.31]	93.27 [93.09, 93.45]	92.97 [92.79, 93.15]	93.21 [93.06, 93.36]	93.18 [92.97, 93.39]	93.59 [93.48, 93.70]	93.60 [93.38, 93.82]	94.24 [94.09, 94.39]	93.73 [93.58, 93.88]	93.75 [93.55, 93.95]	**96.9**
CLBBB	99.56 [99.52, 99.60]	99.67 [99.65, 99.69]	99.75 [99.72, 99.78]	99.76 [99.74, 99.78]	99.68 [99.66, 99.70]	99.65 [99.62, 99.68]	99.72 [99.71, 99.73]	99.71 [99.69, 99.73]	99.56 [99.51, 99.61]	99.52 [99.46, 99.58]	99.20 [99.08, 99.32]	**99.9**
CRBBB	98.94 [98.84, 99.04]	99.34 [99.31, 99.37]	99.69 [99.68, 99.70]	99.58 [99.56, 99.60]	99.22 [99.14, 99.30]	99.16 [99.10, 99.22]	99.24 [99.20, 99.28]	99.16 [99.11, 99.21]	99.58 [99.56, 99.60]	99.56 [99.54, 99.58]	99.60 [99.58, 99.62]	**99.8**
ILBBB	90.48 [89.92, 91.04]	90.74 [90.26, 91.22]	92.07 [91.70, 92.44]	93.07 [92.71, 93.43]	91.18 [90.55, 91.81]	93.39 [92.89, 93.89]	**93.72** [93.22, 94.22]	92.87 [92.13, 93.61]	92.32 [91.73, 92.91]	91.47 [90.87, 92.07]	87.43 [86.34, 88.52]	91.9
IRBBB	73.86 [73.48, 74.24]	79.14 [78.82, 79.46]	96.17 [96.09, 96.25]	91.41 [91.25, 91.57]	76.17 [75.91, 76.43]	75.33 [75.03, 75.63]	76.07 [75.77, 76.37]	79.75 [79.46, 80.04]	95.90 [95.78, 96.02]	96.16 [96.03, 96.29]	95.02 [94.79, 95.25]	**98.0**
IVCD	68.41 [68.10, 68.72]	73.86 [73.57, 74.15]	71.46 [71.17, 71.75]	72.52 [72.24, 72.80]	71.41 [71.14, 71.68]	69.89 [69.62, 70.16]	68.94 [68.60, 69.28]	70.82 [70.51, 71.13]	75.47 [75.23, 75.71]	76.40 [76.03, 76.77]	**78.52** [78.14, 78.90]	74.4
LAFB/ LPFB	79.01 [78.78, 79.24]	97.57 [97.51, 97.63]	83.56 [83.40, 83.72]	80.94 [80.72, 81.16]	81.72 [81.55, 81.89]	85.87 [85.71, 86.03]	89.58 [89.48, 89.68]	90.86 [90.74, 90.98]	97.83 [97.76, 97.90]	**98.20** [98.15, 98.25]	97.67 [97.59, 97.75]	97.5
WPW	86.55 [85.44, 87.66]	84.40 [83.16, 85.64]	90.70 [89.56, 91.84]	84.33 [82.85, 85.81]	85.12 [83.96, 86.28]	84.68 [83.69, 85.67]	83.07 [81.55, 84.59]	83.02 [81.31, 84.73]	92.38 [91.40, 93.36]	**93.36** [92.22, 94.50]	90.26 [88.81, 91.71]	85.5
LAO/LAE	80.60 [80.19, 81.01]	**86.93** [86.57, 87.29]	83.52 [83.08, 83.96]	80.27 [79.74, 80.80]	79.92 [79.55, 80.29]	81.30 [80.83, 81.77]	82.16 [81.70, 82.62]	81.90 [81.44, 82.36]	81.22 [80.59, 81.85]	82.43 [81.75, 83.11]	80.37 [79.64, 81.10]	78.2
LVH	87.97 [87.85, 88.09]	90.77 [90.62, 90.92]	91.42 [91.27, 91.57]	90.92 [90.76, 91.08]	90.98 [90.84, 91.12]	92.06 [91.92, 92.20]	93.23 [93.10, 93.36]	94.08 [93.98, 94.18]	94.44 [94.33, 94.55]	94.24 [94.11, 94.37]	95.10 [94.95, 95.25]	**95.3**
RAO/ RAE	84.93 [84.14, 85.72]	**97.11** [96.83, 97.39]	87.60 [86.83, 88.37]	84.21 [83.34, 85.08]	80.86 [79.63, 82.09]	88.83 [87.83, 89.83]	89.35 [88.56, 90.14]	89.41 [88.62, 90.20]	93.81 [93.15, 94.47]	92.55 [91.83, 93.27]	92.02 [91.01, 93.03]	95.9
AMI	83.18 [83.03, 83.33]	85.67 [85.57, 85.77]	90.29 [90.18, 90.40]	94.94 [94.87, 95.01]	94.16 [94.09, 94.23]	90.17 [90.04, 90.30]	86.26 [86.16, 86.36]	86.35 [86.21, 86.49]	96.70 [96.64, 96.76]	96.72 [96.64, 96.80]	96.21 [96.06, 96.36]	**96.9**
IMI	71.59 [71.23, 71.95]	93.28 [93.16, 93.40]	75.18 [74.96, 75.40]	74.34 [74.06, 74.62]	75.40 [75.18, 75.62]	77.05 [76.78, 77.32]	80.26 [80.04, 80.48]	83.05 [82.84, 83.26]	93.05 [92.93, 93.17]	**95.16** [95.08, 95.24]	95.11 [94.97, 95.25]	94.6
LMI	98.81 [98.58, 99.04]	**99.59** [99.51, 99.67]	98.61 [98.40, 98.82]	99.17 [99.01, 99.33]	99.29 [99.10, 99.48]	99.36 [99.26, 99.46]	99.36 [99.27, 99.45]	99.58 [99.51, 99.65]	98.25 [97.62, 98.88]	99.08 [98.86, 99.30]	95.88 [94.72, 97.04]	91.4
Avg	85.12 [85.00, 85.24]	90.58 [90.50, 90.66]	88.25 [88.16, 88.34]	87.86 [87.76, 87.96]	86.96 [86.86, 87.06]	87.88 [87.80, 87.96]	88.47 [88.36, 88.58]	88.80 [88.68, 88.92]	92.95 [92.85, 93.05]	**93.23** [93.15, 93.31]	92.25 [92.08, 92.42]	93.10

The last column represents the results of the best architecture proposed by Strodthoff et al. (15). The AUCs are reported as percentages. The bold numbers correspond to the best performer for each class. The underlined numbers correspond to the competitive results.

Figure C.4: Average AUC with 95% CI (over 50 runs) of the considered scenarios.

C.2 PHeP: TrustAlert Open-Source Platform for Enhancing Predictive Healthcare with Deep Learning

Predictive healthcare, driven by the availability of large medical datasets and computational advancements, plays a crucial role in enhancing patient care and optimizing healthcare system performance. However, training and deploying predictive models in healthcare presents substantial challenges, such as specialized expertise, and the need for state-of-the-art and costly hardware. To address these barriers, we introduce the Predictive Healthcare Platform (PHeP) [86]—Figure C.5—, an open-source platform designed to simplify the use of pre-trained models. Through a real-world case study, we showcase PHeP functionality in making advanced predictive healthcare accessible to users without extensive computational skills, or expensive computational resources. Specifically, demonstrate its effectiveness by creating an online service which uses three pre-trained BERT models to predict hospitalization at three months from the last access to the national health service in a pediatric (under 14 years old) population—Tables C.1 and C.2.

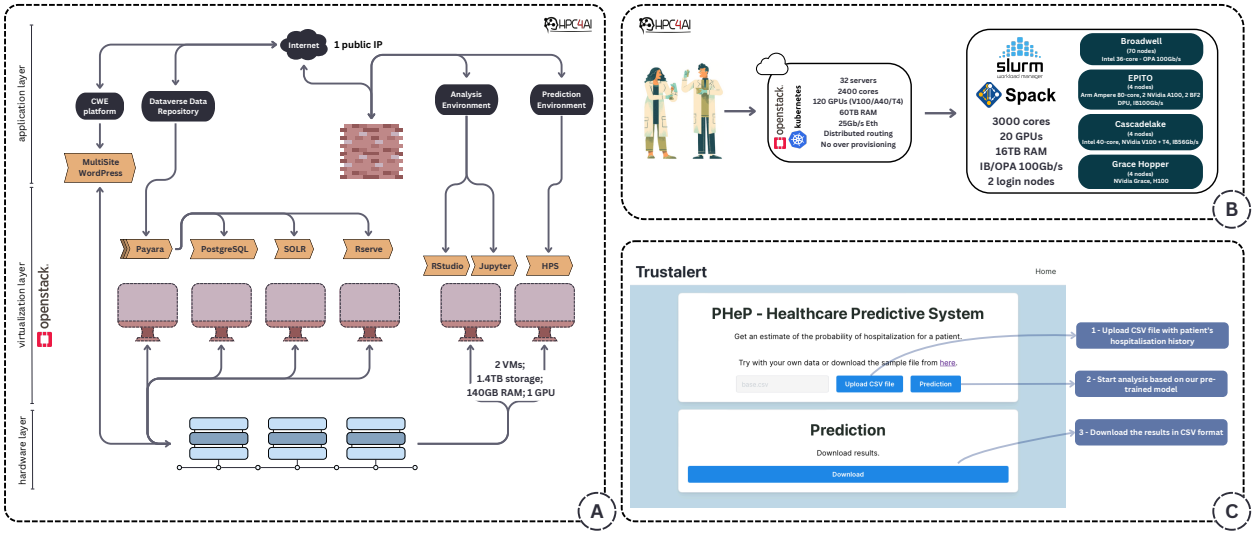


Figure C.5: (A) Overview of the hardware and software stack supporting PHeP. Black boxes represent PHeP subcomponents, orange boxes represent microservices for each subcomponent, and arrows illustrate their interdependencies. (B) The open-access High-Performance Computing for Artificial Intelligence (HPC4AI) architecture [21]. (C) The PHeP user interface for our case study. The figure shows how to (1) load (2) process the data, and (3) download the prediction results.

		Overall	Training set	Testing set	P-value
N		1,647	1,317	330	-
Age group	< 1	160 (9.7)	135 (10.3)	25 (7.6)	0.50
	1-4	567 (34.4)	454 (34.5)	113 (34.2)	
	5-9	588 (35.7)	465 (35.3)	123 (37.3)	
	10-14	332 (20.2)	263 (20.0)	69 (20.9)	
N of unique diagnoses/procedures		4 [1, 9]	4 [1, 9]	4 [2, 9]	0.37
N of unique medications[a]		0 [0, 0]	0 [0, 0]	0 [0, 0]	0.62
N of hospitalizations		6 [3, 10]	6 [3, 10]	6 [4, 11]	0.55

[a] Zero values for this distribution are due to the fact that > 90% of the children and adolescents in this dataset do not assume any medication.

Table C.1: Sample characteristics. Categorical variables are presented as numbers (percentage), and the comparison between training and testing set was performed *via* χ^2 test. Count variables are presented as median (interquartile range) and compared using the Wilcoxon test.

		MLM	NSP	MLM+NSP
Scenario 1	5-fold CV	0.848 ± 0.015	0.825 ± 0.005	0.824 ± 0.053
	testing set	0.578	0.579	0.673
Scenario 2	5-fold CV	0.765 ± 0.024	0.546 ± 0.027	0.729 ± 0.048
	testing set	0.724	0.579	0.766

Table C.2: Accuracy in predicting short-term hospitalization in a cohort of 1,647 pediatrics patients. The table shows, for each task used for pre-training, the results of the stability test (5-fold cross-validation) as well as the accuracies obtained on the testing set by the models trained on the complete training set. Accuracies obtained during the 5-fold cross-validation are presented as mean ± standard deviation. Scenarios 1 and 2 perform fine-tuning using data with or without augmentation, respectively.

Acronyms

ABMs Agent-Based Models. 31–33, 40, 41, 43, 56, 58, 61, 62, 183–185, 188, 189, 201, 202

ACH Air Changes per Hour. 68, 73, 193, 194

ACTS Attention Crossing Time Series. 49

AI Artificial Intelligence. 28, 35–37, 51, 285

AIDS Acquired Immune Deficiency Syndrome. 26

ALOTT A Little Of The Time. 152

AO Always On. 69, 79–82, 86–89

ARIMA Auto-Regressive Integrated Moving Average. 48, 49, 51, 96, 100, 103, 110–112, 116, 117, 142

ATT All The Time. 152

AUC Areo Under the Curve. 285, 287, 288

BERT Bidirectional Encoder Representations from Transformers. 289

Bi-LSTM Bidirectional LSTM. 48, 49

CDC Centers for Disease Control and Prevention. 101, 131, 144

CDF Cumulative Distribution Function. 81–85

CFR Case Fatality Rate. 34

CI Confidence Interval. 288

CLI COVID-like illness. 149

CMS Continuous Monitoring System. 27, 36, 37, 44

CMU Carnegie Mellon University. 178, 179

CN-ABM Contact Network Agent-Based Model. 45

CNN Convolutional Neural Network. 19, 37, 47, 49, 285, 286

Covasim COVID-19 Agent-based Simulator. 56

CQ Class Quarantine. 68, 79–82, 86–89

CTIS Center COVID-19 Trends and Impact Survey. 179

CTMCs Continuous-Time Markov Chains. 31

CV Cross Validation. 290

DEFSI Deep Learning Based Epidemic Forecasting with Synthetic Information. 49

DFDNM Densely Connected Flexible Dendritic Neuron Model. 48

DL Deep Learning. 47, 48, 51

DNN Deep Neural Network. 51

DQN Deep Q-Networks. 37

ECDC European Center for Disease Control. 101, 120, 144

ECG Electrocardiogram. 285, 286

EINNs Epidemiologically-Informed Neural Networks. 51

EU European Union. 62

F4F Forge4Flame. 19, 43, 59, 61, 183, 184, 188–194, 197, 198, 201, 202, 206, 207, 233–237, 239, 241, 265, 267, 273, 275, 276, 281–283

fDB Functional Davies–Bouldin. 16, 167, 169

FFP2 Filtering Face Piece 2. 69, 73, 80, 81, 193, 198–200, 257, 261

FOI Force Of Infection. 198

FWR FFP2 When Required. 69, 79–81

GDELT Global Database of Events, Language, and Tone. 29

GIS Geographic Information Systems. 27

GLMNet Generalized Linear Model Elastic Net. 48

GRU Gated Recurrent Unit. 47–49, 96, 103, 110–112, 116, 117, 142

HHS Health and Human Services. 101, 131

HIV Human Immunodeficiency Virus. 26

HPC High-Performance Computing. 43, 59, 184, 187, 188, 201, 202, 206, 233, 236, 239

HPC4AI High-Performance Computing for Artificial Intelligence. 236, 241, 289

ICU Intensive Care Unit. 45, 46

IFR Infection Fatality Rate. 34

ILI Influenza-Like Illness. 49

ITP Interval Testing Procedure. 192, 195, 196, 241

LR Linear Regression. 48

LSTM Long Short-Term Memory. 47–49, 51, 96, 103, 110–112, 116, 117, 142

MaS-LaS Macro simulation-Large scale. 182

MaS-SmS Macro simulation-Small scale. 182, 183

MCMC Markov Chain monte Carlo. 100

MCNN Multivariate CNN. 49

MERS Middle East Respiratory Syndrome. 26

MiS-SmS Micro simulation-Small scale. 182, 183

ML Machine Learning. 36, 47, 48, 96, 141, 153

MLM Masked Language Modeling. 290

MLP Multi Layer Perceptron. 49

MLR Multiple Linear Regression. 48

MOTT Most Of The Time. 152

Neural ODEs Neural Ordinary Differential Equations. 207

NLP Natural Language Processing. 36, 37

NM No Mask. 69, 74, 75, 79–82, 86–88

NMAE Normalized Mean Absolute Error. 173, 174

NN None. 68, 74–76, 79–82, 86–88

NOTT None Of The Time. 152

NPIs Non-Pharmaceutical Interventions. 28, 40–43, 45, 46, 52–55, 57, 60–64, 68, 72, 73, 75, 76, 79, 81, 89, 90, 92, 146–148, 150, 155, 161, 177, 183, 189, 192, 193, 195, 196, 204–206, 218–230

NQ No Quarantine. 68, 79–82, 86–88

NSP Next Sentence Prediction. 290

ODEs Ordinary Differential Equations. 30, 31, 96, 97

OxCGRT Oxford Covid-19 Government Response Tracker. 149, 150, 154–156, 160

PHeP Predictive Healthcare Platform. 289

PINNs Physics-Informed Neural Networks. 207

PIs Pharmaceutical Interventions. 28

PPE Personal Protective Equipment. 198, 199

$\mathbf{R_0}$ Basic Reproduction Number. 34, 53, 70–72, 76, 182

$\mathbf{R_t}$ Effective Reproduction Number. 34, 40, 42, 43, 50, 54, 56, 58, 60, 61, 96, 146–148, 151, 153, 157, 173–177, 182, 206, 218

RF Random Forest. 48

RMSE Root Mean Squared Error. 54, 103, 105, 107, 108, 110–112, 114–117, 135, 173, 174

RNN Recurrent Neural Network. 47–49, 51

RR Round Robin. 68, 74–77, 79–82

RT-PCR Reverse Transcription-Polymerase Chain Reaction. 149

SAARC South Asian Association for Regional Cooperation. 48

SARIMA Seasonal ARIMA. 48, 49, 96, 103, 110–112, 116, 117, 142

SARS Severe Acute Respiratory Syndrome. 26

SATT Surgical All The Time. 69, 74–76

SEIR Susceptible-Exposed-Infected-Recovered. 50, 51, 57, 64–66, 193, 196, 253

SEIRD Susceptible-Exposed-Infected-Removed-Deceased. 253

SEIRDS Susceptible-Exposed-Infected-Removed-Deceased-Susceptible. 18, 197, 253, 254

SEIRS Susceptible-Exposed-Infected-Removed-Susceptible. 197, 253, 274

SI Susceptible-Infected. 33

SI$^{\mathrm{V}}$RDS Susceptible-Infected$^{\mathrm{V}}$-Recovered-Deceased-Susceptible. 95, 98, 100

SIR Susceptible-Infected-Recovered. 31, 57, 253, 254

SIRD Susceptible-Infected-Removed-Deceased. 253

SIRDS Susceptible-Infected-Recovered-Deceased-Susceptible. 95–97, 153, 253

SIRS Susceptible-Infected-Removed-Susceptible. 253

Slurm Simple Linux Utility for Resource Management. 43, 59, 184, 187, 188, 191, 201, 202, 206, 233, 236, 239, 241

SOTT Some Of The Time. 152

SQ Single Quarantine. 68, 75, 79–81

SRR Sparse Round Robin. 68, 74–77

SSA Stochastic Simulation Algorithm. 31

SVM Support Vector Machine. 48

SVR Support Vector Regression. 48

TA Test All. 68, 74–77, 79

UAE United Arab Emirates. 49

UMD University of Maryland. 149, 178, 179

UMD Global CTIS University of Maryland Social Data Science Center Global COVID-19 Trends and Impact Survey. 149, 150, 152, 178, 179

VoC Variants of Concern. 50

WHO World Health Organization. 57, 70, 71, 101

YAML YAML Ain't Markup Language. 186, 239

Bibliography

[1] Data on SARS-CoV-2 variants in the EU/EEA. European Centre for Disease Prevention and Control (ECDC).

[2] Docker Engine. `https://docs.docker.com/engine/`. [Accessed: 2024-12-02].

[3] Docker Hub. `https://hub.docker.com/`. [Accessed: 2024-12-02].

[4] Association between mobility patterns and COVID-19 transmission in the USA: a mathematical modelling study, author=Badr, Hamada s and du, hongru and marshall, maximilian and dong, ensheng and squire, marietta m and gardner, lauren m. *The Lancet Infectious Diseases*, 20(11):1247–1254, 2020.

[5] Estimating the impact of mobility patterns on COVID-19 infection rates in 11 European countries, author=Bryant, Patrick and elofsson, arne. *PeerJ*, 8:e9879, 2020.

[6] Facebook Data for Good. COVID-19 symptom survey – request for data access., 2020.

[7] Early detection of covID-19 outbreaks using human mobility data, author=Guan, Grace and dery, yotam and yechezkel, matan and ben-gal, irad and yamin, dan and brandeau, margaret l. *PloS one*, 16(7):e0253865, 2021.

[8] Improving prediction of COVID-19 evolution by fusing epidemiological and mobility data, author=García-cremades, santi and morales-garcía, juan and hernández-sanjaime, rocío and martínez-españa, raquel and bueno-crespo, andrés and hernández-orallo, enrique and lópez-espín, josé j and cecilia, josé m. *Scientific Reports*, 11(1):15173, 2021.

[9] Public mobility data enables COVID-19 forecasting and management at local and global scales, author=Ilin, Cornelia and annan-phan, sébastien and tai, xiao hui and mehra, shikhar and hsiang, solomon and blumenstock, joshua e. *Scientific reports*, 11(1):13531, 2021.

[10] Spatial association of mobility and COVID-19 infection rate in the USA: A county-level study using mobile phone location data, author=Tokey, Ahmad ilderim. *Journal of Transport & Health*, 22:101135, 2021.

[11] A causal inference approach for estimating effects of non-pharmaceutical interventions during Covid-19 pandemic, author=Barros, Vesna and manes, itay and akinwande, victor and cintas, celia and bar-shira, osnat and ozery-flato, michal and shimoni, yishai and rosen-zvi, michal. *Plos one*, 17(9):e0265289, 2022.

[12] Sameera Abar, Georgios K Theodoropoulos, Pierre Lemarinier, and Gregory MP O'Hare. Agent Based Modelling and Simulation tools: A review of the state-of-art software. *Computer Science Review*, 24:13–33, 2017.

[13] Elaheh Abdollahi, Margaret Haworth-Brockman, Yoav Keynan, Joanne M Langley, and Seyed M Moghadas. Simulating the effect of school closure during COVID-19 outbreaks in Ontario, Canada. *BMC medicine*, 18:1–8, 2020.

[14] Paulette Acheson, Cihan Dagli, and Nil Kilicay-Ergin. Model based systems engineering for system of systems using agent-based modeling. *Procedia Computer Science*, 16:11–19, 2013.

[15] Michael O Adeniyi, Segun I Oke, Matthew I Ekum, Temitope Benson, and Matthew O Adewole. Assessing the impact of public compliance on the use of non-pharmaceutical intervention with cost-effectiveness analysis on the transmission dynamics of COVID-19: Insight from mathematical modeling. *Modeling, control and drug development for COVID-19 outbreak prevention*, pages 579–618, 2022.

[16] Haneen Alabdulrazzaq, Mohammed N. Alenezi, Yasmeen Rawajfih, Bareeq A. Alghannam, Abeer A. Al-Hassan, and Fawaz S. Al-Anzi. On the accuracy of ARIMA based prediction of COVID-19 spread. *Results in Physics*, 27:104509, 2021.

[17] Teodoro Alamo, Pablo Millán, Daniel G Reina, Victor M Preciado, and Giulia Giordano. Challenges and future directions in pandemic control. *IEEE Control Systems Letters*, 6:722–727, 2021.

[18] Teodoro Alamo, Daniel G Reina, Pablo Millán Gata, Victor M Preciado, and Giulia Giordano. Data-driven methods for present and future pandemics: Monitoring, modelling and managing. *Annual Reviews in Control*, 52:448–464, 2021.

[19] Hakan Alcik, Oguz Ozel, Nurdan Apaydin, and Mustafa Erdik. A study on warning algorithms for Istanbul earthquake early warning system. *Geophysical Research Letters*, 36(5), 2009.

[20] M Aldinucci, S Bagnasco, S Lusso, P Pasteris, S Rabellino, and S Vallero. OCCAM: a flexible, multi-purpose and extendable HPC cluster. *Journal of Physics: Conference Series*, 898(8):082039, 2017.

[21] Marco Aldinucci, Sergio Rabellino, Marco Pironti, Filippo Spiga, Paolo Viviani, Maurizio Drocco, Marco Guerzoni, Guido Boella, Marco Mellia, Paolo Margara, et al. HPC4AI: an AI-on-demand federated platform endeavour. In *Proceedings of the 15th ACM International Conference on Computing Frontiers*, pages 279–286, 2018.

[22] Laura Alessandretti. What human mobility data tell us about COVID-19 spread. *Nature Reviews Physics*, 4(1):12–13, 2022.

[23] Richard M Allen, Paolo Gasparini, Osamu Kamigaichi, and Maren Bose. The status of earthquake early warning around the world: An introductory overview. *Seismological Research Letters*, 80(5):682–693, 2009.

[24] Javier Álvarez, Carlos Baquero, Elisa Cabana, Jaya Prakash Champati, Antonio Fernández Anta, Davide Frey, Augusto Garcia-Agundez, Chryssis Georgiou, Mathieu Goessens, Harold Hernández, et al. Estimating active cases of COVID-19. *MedRxiv*, pages 2021–12, 2021.

[25] Gary An, Qi Mi, Joyeeta Dutta-Moscato, and Yoram Vodovotz. Agent-based models in translational systems biology. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*, 1(2):159–171, 2009.

[26] Sercan Arik, Chun-Liang Li, Jinsung Yoon, Rajarishi Sinha, Arkady Epshteyn, Long Le, Vikas Menon, Shashank Singh, Leyou Zhang, Martin Nikoltchev, et al. Interpretable sequence learning for covid-19 forecasting. *Advances in Neural Information Processing Systems*, 33:18807–18818, 2020.

[27] Sercan Ö Arık, Joel Shor, Rajarishi Sinha, Jinsung Yoon, Joseph R Ledsam, Long T Le, Michael W Dusenberry, Nathanael C Yoder, Kris Popendorf, Arkady Epshteyn, et al. A prospective evaluation of ai-augmented epidemiology to forecast covid-19 in the usa and japan. *NPJ digital medicine*, 4(1):146, 2021.

[28] Parul Arora, Himanshu Kumar, and Bijaya Ketan Panigrahi. Prediction and analysis of COVID-19 positive cases using deep learning models: A descriptive case study of India. *Chaos, solitons & fractals*, 139:110017, 2020.

[29] K.E. ArunKumar, Dinesh V. Kalaga, Ch. Mohan Sai Kumar, Masahiro Kawaji, and Timothy M. Brenza. Comparative analysis of Gated Recurrent Units (GRU), long Short-Term memory (LSTM) cells, autoregressive Integrated moving average (ARIMA), seasonal autoregressive Integrated moving average (SARIMA) for forecasting COVID-19 trends. *Alexandria Engineering Journal*, 61(10):7585–7603, 2022.

[30] Badar Nadeem Ashraf. Socioeconomic conditions, government interventions and health outcomes during COVID-19. *Covid Economics*, 37(14):141–162, 2020.

[31] Nikolaos Askitas, Konstantinos Tatsiramos, and Bertrand Verheyden. Estimating worldwide effects of non-pharmaceutical interventions on COVID-19 incidence and population mobility patterns using a multiple-event study. *Scientific reports*, 11(1):1972, 2021.

[32] Christina M Astley, Gaurav Tuli, Kimberly A Mc Cord, Emily L Cohn, Benjamin Rader, Tanner J Varrelman, Samantha L Chiu, Xiaoyi Deng, Kathleen Stewart, Tamer H Farag, et al. Global monitoring of the impact of the COVID-19 pandemic through online surveys sampled from the Facebook user base. *Proceedings of the National Academy of Sciences*, 118(51):e2111455118, 2021.

[33] H Ceren Ates, Ali K Yetisen, Firat Güder, and Can Dincer. Wearable devices for the detection of COVID-19. *Nature Electronics*, 4(1):13–14, 2021.

[34] D Attali. Colourpicker: A colour picker tool for shiny and for selecting colours in plots. *R package version*, 1, 2017.

[35] D Attali, T Edwards, and Z Wang. shinyalert: Easily create pretty popup messages (modals) in 'shiny'. *R package version*, 2(0), 2018.

[36] Dean Attali. Shinyjs: Easily improve the user experience of your shiny apps in seconds. *R package version*, 2(0), 2020.

[37] Seyed Mohammad Ayyoubzadeh, Seyed Mehdi Ayyoubzadeh, Hoda Zahedi, Mahnaz Ahmadi, and Sharareh R Niakan Kalhori. Predicting covid-19 incidence through analysis of google trends data in iran: data mining and deep learning pilot study. *JMIR public health and surveillance*, 6(2):e18828, 2020.

[38] Daniele Baccega, Jose Aguilar, Carlos Baquero, Antonio Fernandez Anta, and Juan Marcos Ramirez. Social compliance with npis, mobility patterns, and reproduction number: Lessons from covid-19 in europe. *medRxiv*, pages 2025–01, 2025.

[39] Daniele Baccega, Paolo Castagno, Antonio Fernández Anta, and Matteo Sereno. Enhancing COVID-19 forecasting precision through the integration of compartmental models, machine learning and variants. *Scientific Reports*, 14(1):19220, 2024.

[40] Daniele Baccega, Simone Pernice, Paolo Castagno[1], Matteo Sereno, and Marco Beccuti. Living along covid-19: Assessing contention policies through agent-based. In *Computational Intelligence Methods for Bioinformatics and Biostatistics: 18th International Meeting, CIBB 2023, Padova, Italy, September 6–8, 2023, Revised Selected Papers*, volume 14513, page 143. Springer Nature, 2025.

[41] Daniele Baccega, Simone Pernice, Pietro Terna, Paolo Castagno, Giovenale Moirano, Lorenzo Richiardi, Matteo Sereno, Sergio Rabellino, Milena Maria Maule, Marco Beccuti, et al. An agent-based model to support infection control strategies at school. *JASSS*, 25(3):1–15, 2022.

[42] Daniele Baccega, Irene Terrone, Peter Heywood, Robert Chisholm, P Richmond, Sandro Gepiro Contaldo, Lorenzo Bosio, Simone Pernice, and Marco Beccuti. Forge4flame: An intuitive dashboard for designing gpu agent-based models to simulate infectious disease spread. *Available at SSRN 5194584*.

[43] Michela Baccini and Giulia Cereda. Screening plans for SARS-CoV-2 based on sampling and rotation: An example in a European school setting. *PLOS ONE*, 16(9):1–13, 09 2021.

[44] Zhihua Bai, Yue Gong, Xiaodong Tian, Ying Cao, Wenjun Liu, and Jing Li. The rapid assessment and early warning models for COVID-19. *Virologica Sinica*, 35:272–279, 2020.

[45] Eric Bailey. shinybs: Twitter bootstrap components for shiny. *R package version 0.61*, 2015.

[46] D Banerjee, A Campbell, J Gruger, et al. What has the pandemic revealed about the US health care system—and what needs to change. *MIT News*, 2021.

[47] Nicolas Banholzer, Eva van Weenen, Adrian Lison, Alberto Cenedese, Arne Seeliger, Bernhard Kratzwald, Daniel Tschernutter, Joan Puig Salles, Pierluigi Bottrighi,

Sonja Lehtinen, Stefan Feuerriegel, and Werner Vach. Estimating the effects of non-pharmaceutical interventions on the number of new infections with COVID-19 during the first epidemic wave. *PLOS ONE*, 16(6):1–16, 06 2021.

[48] Reid Basher. Global early warning systems for natural hazards: systematic and people-centred. *Philosophical transactions of the royal society a: mathematical, physical and engineering sciences*, 364(1845):2167–2182, 2006.

[49] Gopi Battineni, Nalini Chintalapudi, and Francesco Amenta. Forecasting of COVID-19 epidemic size in four high hitting nations (USA, Brazil, India and Russia) by Fb-Prophet machine learning model. *Applied Computing and Informatics*, 2020.

[50] Marco Beccuti, Paolo Castagno, Giuliana Franceschinis, Marzio Pennisi, and Simone Pernice. A Petri net formalism to study systems at different scales exploiting agent-based and stochastic simulations. In *European Workshop on Performance Engineering*, pages 22–43. Springer, 2021.

[51] Hans Walter Behrens, K Selçuk Candan, Xilun Chen, Yash Garg, Mao-Lin Li, Xinsheng Li, Sicong Liu, Maria Luisa Sapino, Md Shadab, Dalton Turner, et al. DataStorm: Coupled, Continuous Simulations for Complex Urban Environments. *ACM/IMS Transactions on Data Science*, 2(3):1–37, 2021.

[52] Francisco Benita. Human mobility behavior in COVID-19: A systematic literature review and bibliometric analysis. *Sustainable Cities and Society*, 70:102916, 2021.

[53] Ethan M Berke, Lori M Newman, Suzana Jemsby, Natasha Bhalla, Natalie E Sheils, Nandini Oomman, John Reppas, Bethany Hyde, Pratek Verma, and Gerard A Cangelosi. Pooling-in-a-pod: A Strategy for COVID-19 Testing to Facilitate Safe Return to School. *medRxiv*, 2021.

[54] Lisa F Berkman, Ichirō Kawachi, and M Maria Glymour. *Social epidemiology*. Oxford University Press, 2014.

[55] Shilpa N Bhupathiraju and Frank B Hu. Epidemiology of obesity and diabetes and their cardiovascular complications. *Circulation research*, 118(11):1723–1735, 2016.

[56] Qifang Bi, Katherine E Goodman, Joshua Kaminsky, and Justin Lessler. What is machine learning? A primer for the epidemiologist. *American journal of epidemiology*, 188(12):2222–2239, 2019.

[57] Toheeb A Biala and AQM Khaliq. A fractional-order compartmental model for the spread of the COVID-19 pandemic. *Communications in Nonlinear Science and Numerical Simulation*, 98:105764, 2021.

[58] C Borau, R Chisholm, P Richmond, and D Walker. An agent-based model for cell microenvironment simulation using FLAMEGPU2. *Computers in Biology and Medicine*, 179:108831, 2024.

[59] Philip S. Brachman. Epidemiology. In Samuel Baron, editor, *Medical Microbiology*, chapter 9. University of Texas Medical Branch at Galveston, Galveston, TX, 4th edition, 1996.

[60] Fred Brauer. Compartmental models in epidemiology. *Mathematical epidemiology*, pages 19–79, 2008.

[61] Peter J. Brockwell and Richard A. Davis, editors. *State-Space Models*, pages 259–316. Springer New York, New York, NY, 2002.

[62] Giorgio Buonanno, Luca Stabile, and Lidia Morawska. Estimation of airborne viral emission: Quanta emission rate of SARS-CoV-2 for infection risk assessment. *Environment International*, 141:105794, 2020.

[63] Frada Burstein, Clyde W Holsapple, Bartel van de Walle, and Murray Turoff. Decision support for emergency situations. *Handbook on Decision Support Systems 2: Variations*, pages 39–63, 2008.

[64] Matthieu Bussiere and Marcel Fratzscher. Towards a new early warning system of financial crises. *journal of International Money and Finance*, 25(6):953–973, 2006.

[65] Andrew William Byrne, David McEvoy, Aine B Collins, Kevin Hunt, Miriam Casey, Ann Barber, Francis Butler, John Griffin, Elizabeth A Lane, Conor McAloon, et al. Inferred duration of infectious period of SARS-CoV-2: rapid scoping review and analysis of available evidence for asymptomatic and symptomatic COVID-19 cases. *BMJ open*, 10(8):e039856, 2020.

[66] Andrew William Byrne, David McEvoy, Aine B Collins, Kevin Hunt, Miriam Casey, Ann Barber, Francis Butler, John Griffin, Elizabeth A Lane, Conor McAloon, Kirsty O'Brien, Patrick Wall, Kieran A Walsh, and Simon J More. Inferred duration of infectious period of SARS-CoV-2: rapid scoping review and analysis of available evidence for asymptomatic and symptomatic COVID-19 cases. *BMJ Open*, 10(8), 2020.

[67] Muffy Calder, Claire Craig, Dave Culley, Richard De Cani, Christl A Donnelly, Rowan Douglas, Bruce Edmonds, Jonathon Gascoigne, Nigel Gilbert, Caroline Hargrove, et al. Computational modelling for decision-making: where, why, what, who and how. *Royal Society open science*, 5(6):172096, 2018.

[68] Jamie M Caldwell, Xuan Le, Lorin McIntosh, Michael T Meehan, Samson Ogunlade, Romain Ragonnet, Genevieve K O'Neill, James M Trauer, and Emma S McBryde. Vaccines and variants: Modelling insights into emerging issues in COVID-19 epidemiology. *Paediatric Respiratory Reviews*, 39:32–39, 2021.

[69] Rocio Cardenas, Laith Hussain-Alkhateeb, David Benitez-Valladares, Gustavo Sánchez-Tejeda, and Axel Kroeger. The Early Warning and Response System (EWARS-TDR) for dengue outbreaks: can it also be applied to chikungunya and Zika outbreak warning? *BMC Infectious Diseases*, 22(1):235, 2022.

[70] Armando Cartenì, Luigi Di Francesco, and Maria Martino. How mobility habits influenced the spread of the COVID-19 pandemic: Results from the Italian case study. *Science of the Total Environment*, 741:140489, 2020.

[71] Centers for Disease Control and Prevention. COVID Data Tracker. Atlanta, GA: U.S. Department of Health and Human Services, CDC, June 13 2024. `https://covid.cdc.gov/covid-data-tracker`. Accessed: 2024-06-13.

[72] Ho Fai Chan, Martin Brumpton, Alison Macintyre, Jefferson Arapoc, David A Savage, Ahmed Skali, David Stadelmann, and Benno Torgler. How confidence in health care systems affects mobility and compliance during the COVID-19 pandemic. *PloS one*, 15(10):e0240644, 2020.

[73] Winston Chang, Joe Cheng, J Allaire, Yihui Xie, and Jonathan McPherson. Package 'shiny'. *See http://citeseerx. ist. psu. edu/viewdoc/download*, 2015.

[74] Winston Chang, Joe Cheng, JJ Allaire, Carson Sievert, Barret Schloerke, Yihui Xie, Jeff Allen, Jonathan McPherson, Alan Dipert, and Barbara Borges. *shiny: Web Application Framework for R*, 2024. R package version 1.10.0, https://github.com/rstudio/shiny.

[75] Winston Chang, T Park, L Dziedzic, N Willis, and M McInerney. shinythemes: Themes for Shiny. *R package version*, 1(2), 2018.

[76] Winston Chang, Barbara Borges Ribeiro, Almasaeed Studio, and Maintainer Winston Chang. Package 'shinydashboard'. 2022.

[77] Tianqi Chen and Carlos Guestrin. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, pages 785–794, 2016.

[78] Xuwei Chen, John W Meaker, and F Benjamin Zhan. Agent-based modeling and analysis of hurricane evacuation procedures for the Florida Keys. *Natural Hazards*, 38:321–338, 2006.

[79] Kyunghyun Cho, Bart Van Merriënboer, Caglar Gulcehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, and Yoshua Bengio. Learning phrase representations using RNN encoder-decoder for statistical machine translation. *arXiv preprint arXiv:1406.1078*, 2014.

[80] Derek K Chu, Elie A Akl, Stephanie Duda, Karla Solo, Sally Yaacoub, Holger J Schünemann, Amena El-harakeh, Antonio Bognanni, Tamara Lotfi, Mark Loeb, et al. Physical distancing, face masks, and eye protection to prevent person-to-person transmission of SARS-CoV-2 and COVID-19: a systematic review and meta-analysis. *The lancet*, 395(10242):1973–1987, 2020.

[81] Jin Young Chung, Choong-Ki Lee, and Yae-Na Park. Trust in social non-pharmaceutical interventions and travel intention during a pandemic. *Journal of Vacation Marketing*, 27(4):437–448, 2021.

[82] Philip Coatsworth, Laura Gonzalez-Macia, Alexander Silva Pinto Collins, Tolga Bozkurt, and Firat Güder. Continuous monitoring of chemical signals in plants under stress. *Nature Reviews Chemistry*, 7(1):7–25, 2023.

[83] Vittoria Colizza, Marc Barthélemy, Alain Barrat, and Alessandro Vespignani. Epidemic modeling in complex realities. *Comptes rendus. Biologies*, 330(4):364–374, 2007.

[84] Nicholson Collier and Michael North. Repast hpc: A platform for large-scale agent-based modeling. *Large-Scale Computing*, pages 81–109, 2011.

[85] Manola Comar, Simone Benvenuto, Marzia Lazzerini, Giorgio Fedele, Egidio Barbi, Alessandro Amaddeo, Francesco Maria Risso, Tamara Strajn, Paola Di Rocco, Paola Stefanelli, et al. Prevalence of SARS-CoV-2 infection in Italian pediatric population: a regional seroepidemiological study. *Italian journal of pediatrics*, 47(1):131, 2021.

[86] Sandro Gepiro Contaldo, Emanuele Pietropaolo, Lorenzo Bosio, Simone Pernice, Irene Terrone, Daniele Baccega, Yuting Wang, Rahul Kumar Sahoo, Giuseppe Rizzo, Alessia

Visconti, et al. PHeP: TrustAlert Open-Source Platform for Enhancing Predictive Healthcare with Deep Learning. *medRxiv*, pages 2024–08, 2024.

[87] B. Copeland. (2024, September 11). artificial intelligence. Encyclopedia Britannica. `https://www.britannica.com/technology/artificial-intelligence`.

[88] Richard M Crowder, Mark A Robinson, Helen PN Hughes, and Yee-Wai Sim. The development of an agent-based modeling framework for simulating engineering team work. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, 42(6):1425–1439, 2012.

[89] Gábor Csárdi, Kuba Podgórski, and Rich Geldreich. zip: Cross-platform "zip" compression, 2020.

[90] Gábor Csárdi, Jim Hester, Hadley Wickham, Winston Chang, Martin Morgan, and Dan Tenenbaum. *remotes: R Package Installation from Remote Repositories, Including 'GitHub'*, 2024. R package version 2.5.0, https://github.com/r-lib/remotes#readme.

[91] Erik Cuevas. An agent-based model to evaluate the COVID-19 transmission risks in facilities. *Computers in biology and medicine*, 121:103827, 2020.

[92] David L Davies and Donald W Bouldin. A cluster separation measure. *IEEE transactions on pattern analysis and machine intelligence*, pages 224–227, 1979.

[93] C De Boor. A practical guide to splines. *Springer-Verlag google schola*, 2:4135–4195, 1978.

[94] Scott De Marchi and Scott E Page. Agent-based models. *Annual Review of political science*, 17(1):1–20, 2014.

[95] Michael John De Smith, Michael F Goodchild, and Paul Longley. *Geospatial analysis: a comprehensive guide to principles, techniques and software tools*. Troubador publishing ltd, 2007.

[96] A de Vries, B Schloerke, and K Russell. sortable: Drag-and-drop in 'shiny'apps with 'sortablejs'. r package version 0.5. 0, 2024.

[97] Byomakesh Debata, Pooja Patnaik, and Abhisek Mishra. COVID-19 pandemic! It's impact on people, economy, and environment. *Journal of public affairs*, 20(4):e2372, 2020.

[98] Qi Deng. Dynamics and development of the COVID-19 epidemic in the United States: a compartmental model enhanced with deep learning techniques. *Journal of Medical Internet Research*, 22(8):e21173, 2020.

[99] Huijue Kelly Duan and Hanxin Hu. Continuous intelligent pandemic monitoring (CIPM). *Journal of Emerging Technologies in Accounting*, 18(1):185–194, 2021.

[100] Ken TD Eames and Matt J Keeling. Contact tracing and disease control. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 270(1533):2565–2571, 2003.

[101] ECDC. COVID-19 in children and the role of school settings in COVID-19 transmission. Technical report, European Center for Disease Prevention and Control, 08 2020.

[102] ECDC. COVID-19 in children and the role of school settings in transmission - second update. Technical report, European Center for Disease Prevention and Control, 07 2021.

[103] Joshua M Epstein. Agent-based computational models and generative social science. *Complexity*, 4(5):41–60, 1999.

[104] J Fan, Y Li, K Stewart, AR Kommareddy, A Garcia, J O'Brien, A Bradford, X Deng, S Chiu, F Kreuter, et al. The University of Maryland Social Data Science Center Global COVID-19 Trends and Impact Survey. *Partnership with Facebook*, 2020.

[105] Yingzheng Fan, Xingyu Wang, Thomas Funk, Ishrat Rashid, Brianna Herman, Nefeli Bompoti, MD Shaad Mahmud, Maria Chrysochoou, Meijian Yang, Timothy M Vadas, et al. A critical review for real-time continuous soil monitoring: Advantages, challenges, and perspectives. *Environmental Science & Technology*, 56(19):13546–13564, 2022.

[106] Elena Farina, Ilenia Eboli, Teresa Spadea, Carlo Saugo, Lorenzo Richiardi, Milena Maule, Pietro Presti, and Antonella Bena. 'Scuola sicura': a school screening testing programme to prevent the spread of COVID-19 in students in Piedmont. *Epidemiologia e Prevenzione*, 45:In–press, 2021.

[107] J Doyne Farmer and Duncan Foley. The economy needs agent-based modelling. *Nature*, 460(7256):685–686, 2009.

[108] Harvey V. Fineberg and Mary Elizabeth Wilson. Epidemic Science in Real Time. *Science*, 324(5930):987–987, 2009.

[109] Miha Fošnarič, Tina Kamenšek, Jerneja Žganec Gros, and Janez Žibert. Extended compartmental model for modeling COVID-19 epidemic in Slovenia. *Scientific reports*, 12(1):16916, 2022.

[110] Arlindo R. Galvão Filho, Lauro C. Martins de Paula, Clarimar José Coelho, Telma Woerle de Lima, and Anderson da Silva Soares. Cuda parallel programming for simulation of epidemiological models based on individuals. *Mathematical Methods in the Applied Sciences*, 39(3):405–411, 2016.

[111] Nigel Gilbert and Pietro Terna. How to build and use agent-based models in social science. *Mind & Society*, 1:57–72, 2000.

[112] Daniel T Gillespie. Exact stochastic simulation of coupled chemical reactions. *The journal of physical chemistry*, 81(25):2340–2361, 1977.

[113] Daniel T Gillespie. Approximate accelerated stochastic simulation of chemically reacting systems. *The Journal of chemical physics*, 115(4):1716–1733, 2001.

[114] Savvas Gkantonas, Daniel Zabotti, Leo CC Mesquita, Epaminondas Mastorakos, and Pedro M de Oliveira. airborne. cam: A risk calculator of SARS-CoV-2 aerosol transmission under well-mixed ventilation conditions. 2021.

[115] Janyce Eunice Gnanvi, Kolawolé Valère Salako, Gaëtan Brezesky Kotanmi, and Romain Glèlè Kakaï. On the reliability of predictions on Covid-19 dynamics: A systematic and critical review of modelling techniques. *Infectious Disease Modelling*, 6:258–272, 2021.

[116] Gabriela B Gomez, Cedric Mahé, and Sandra S Chaves. Uncertain effects of the pandemic on respiratory viruses. *Science*, 372(6546):1043–1044, 2021.

[117] Jonatan Gomez, Jeisson Prieto, Elizabeth Leon, and Arles Rodríguez. Infekta—an agent-based model for transmission of infectious diseases: The covid-19 case in bogotá, colombia. *PloS one*, 16(2):e0245787, 2021.

[118] Katelyn M Gostic, Lauren McGough, Edward B Baskerville, Sam Abbott, Keya Joshi, Christine Tedijanto, Rebecca Kahn, Rene Niehus, James A Hay, Pablo M De Salazar, et al. Practical considerations for measuring the effective reproductive number, Rt. *PLoS computational biology*, 16(12):e1008409, 2020.

[119] IM Gould. The epidemiology of antibiotic resistance. *International journal of antimicrobial agents*, 32:S2–S9, 2008.

[120] Emanuele Guidotti. A worldwide epidemiological database for COVID-19 at fine-grained spatial resolution. *Scientific Data*, 9(1):112, 2022.

[121] Emanuele Guidotti and David Ardia. COVID-19 Data Hub. *Journal of Open Source Software*, 5(51):2376, 2020.

[122] Amit Kumar Gupta, Vijander Singh, Priya Mathur, and Carlos M Travieso-Gonzalez. Prediction of COVID-19 pandemic measuring criteria using support vector machine, prophet and linear regression models in Indian scenario. *Journal of Interdisciplinary Mathematics*, 24(1):89–108, 2021.

[123] Thomas Hale, Noam Angrist, Rafael Goldszmidt, Beatriz Kira, Anna Petherick, Toby Phillips, Samuel Webster, Emily Cameron-Blake, Laura Hallas, Saptarshi Majumdar, et al. A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker). *Nature human behaviour*, 5(4):529–538, 2021.

[124] Lynne Hamill and Nigel Gilbert. *Agent-based modelling in economics*. John Wiley & Sons, 2015.

[125] John A Hartigan, Manchek A Wong, et al. A k-means clustering algorithm. *Applied statistics*, 28(1):100–108, 1979.

[126] Sha He, Sanyi Tang, Libin Rong, et al. A discrete stochastic model of the COVID-19 outbreak: Forecast and control. *Math. Biosci. Eng*, 17(4):2792–2804, 2020.

[127] Randy Heiland, Daniel Bergman, Blair Lyons, Grant Waldow, Julie Cass, Heber Lima da Rocha, Marco Ruscone, Vincent Noël, and Paul Macklin. PhysiCell Studio: a graphical tool to make agent-based modeling more accessible. *Gigabyte*, 2024:gigabyte128, 2024.

[128] Ana Maria Hernandez-Hernandez and Rodrigo Huerta-Quintanilla. Managing school interaction networks during the COVID-19 pandemic: Agent-based modeling for evaluating possible scenarios when students go back to classrooms. *Plos one*, 16(8):e0256363, 2021.

[129] Jim Hester and Jennifer Bryan. Glue: Interpreted string literals. *R package version*, 1(2), 2020.

[130] Robert Hinch, William JM Probert, Anel Nurtay, Michelle Kendall, Chris Wymant, Matthew Hall, Katrina Lythgoe, Ana Bulas Cruz, Lele Zhao, Andrea Stewart, et al.

Openabm-covid19—an agent-based model for non-pharmaceutical interventions against covid-19 including contact tracing. *PLoS computational biology*, 17(7):e1009146, 2021.

[131] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, 9(8):1735–1780, 1997.

[132] Emma B. Hodcroft. CoVariants: SARS-CoV-2 Mutations and Variants of Interest., 2021. `https://covariants.org/`.

[133] Nicolas Hoertel, Martin Blachier, Carlos Blanco, Mark Olfson, Marc Massetti, Marina Sánchez Rico, Frédéric Limosin, and Henri Leleu. A stochastic agent-based model of the SARS-CoV-2 epidemic in France. *Nature medicine*, 26(9):1417–1421, 2020.

[134] Kaori Honjo. Social epidemiology: Definition, history, and research examples. *Environmental health and preventive medicine*, 9:193–199, 2004.

[135] Nai-Chi Hsiao, Yih-Min Wu, Tzay-Chyn Shin, Li Zhao, and Ta-Liang Teng. Development of earthquake early warning system in Taiwan. *Geophysical research letters*, 36(5), 2009.

[136] Tao Hu, Siqin Wang, Bing She, Mengxi Zhang, Xiao Huang, Yunhe Cui, Jacob Khuri, Yaxin Hu, Xiaokang Fu, Xiaoyue Wang, et al. Human mobility data in the COVID-19 pandemic: characteristics, applications, and challenges. *International Journal of Digital Earth*, 14(9):1126–1147, 2021.

[137] Laith Hussain-Alkhateeb, Tatiana Rivera Ramirez, Axel Kroeger, Ernesto Gozzer, and Silvia Runge-Ranzinger. Early warning systems (EWSs) for chikungunya, dengue, malaria, yellow fever, and Zika outbreaks: What is the evidence? A scoping review. *PLoS neglected tropical diseases*, 15(9):e0009686, 2021.

[138] Eveline Otte Im Kampe, Ann-Sophie Lehfeld, Silke Buda, Udo Buchholz, and Walter Haas. Surveillance of COVID-19 school outbreaks, Germany, March to August 2020. *Eurosurveillance*, 25(38):2001645, 2020.

[139] GM James, TJ Hastie, and CA Sugar. Principal component models for sparse functional data. *Biometrika*, 87(3):587–602, 09 2000.

[140] Kinnary Jangla and Kinnary Jangla. Docker compose. *Accelerating Development Velocity Using Docker: Docker Across Microservices*, pages 77–98, 2018.

[141] Xiaoyong Jin, Yu-Xiang Wang, and Xifeng Yan. Inter-series attention model for covid-19 forecasting. In *Proceedings of the 2021 SIAM International Conference on Data Mining (SDM)*, pages 495–503. SIAM, 2021.

[142] A.V. Joshi. *Machine Learning and Artificial Intelligence*. Springer International Publishing, 2019.

[143] Firuz Kamalov, Khairan Rajab, Aswani Kumar Cherukuri, Ashraf Elnagar, and Murodbek Safaraliev. Deep learning for Covid-19 forecasting: State-of-the-art review. *Neurocomputing*, 511:142–154, 2022.

[144] Naim Kapucu and Alpaslan Özerdem. *Managing emergencies and crises*. Jones & Bartlett Publishers, 2011.

[145] Jackie Kazil, David Masad, and Andrew Crooks. Utilizing python for agent-based modeling: The mesa framework. In *Social, Cultural, and Behavioral Modeling: 13th International Conference, SBP-BRiMS 2020, Washington, DC, USA, October 18–21, 2020, Proceedings 13*, pages 308–317. Springer, 2020.

[146] Cliff C Kerr, Robyn M Stuart, Dina Mistry, Romesh G Abeysuriya, Katherine Rosenfeld, Gregory R Hart, Rafael C Núñez, Jamie A Cohen, Prashanth Selvaraj, Brittany Hagedorn, et al. Covasim: an agent-based model of COVID-19 dynamics and interventions. *PLOS Computational Biology*, 17(7):e1009149, 2021.

[147] Khaled M Khalil, M Abdel-Aziz, Taymour T Nazmy, and Abdel-Badeeh M Salem. An agent-based modeling for pandemic influenza in Egypt. *Handbook on Decision Making: Vol 2: Risk Management in Decision Making*, pages 205–218, 2012.

[148] Shruti Khare, Céline Gurry, Lucas Freitas, Mark B Schultz, Gunter Bach, Amadou Diallo, Nancy Akite, Joses Ho, Raphael TC Lee, Winston Yeo, et al. GISAID's role in pandemic response. *China CDC weekly*, 3(49):1049, 2021.

[149] Joyce Kiarie, Samuel Mwalili, and Rachel Mbogo. Forecasting the spread of the COVID-19 pandemic in Kenya using SEIR and ARIMA models. *Infectious Disease Modelling*, 7(2):179–188, 2022.

[150] Ali Serhan Koyuncugil and Nermin Ozgulbas. Financial early warning system model and data mining application for risk detection. *Expert systems with Applications*, 39(6):6238–6253, 2012.

[151] Frauke Kreuter, Neta Barkay, Alyssa Bilinski, Adrianne Bradford, Samantha Chiu, Roee Eliat, Junchuan Fan, Tal Galili, Daniel Haimovich, Brian Kim, et al. Partnering with a global platform to inform research and public policy making: What needs to be in place to make a global COVID-19 survey work? In *Survey Research Methods*, volume 14, pages 159–163, 2020.

[152] Rajalakshmi Krishnamurthi, Dhanalekshmi Gopinathan, and Adarsh Kumar. Wearable devices and COVID-19: state of the art, framework, and challenges. *Emerging Technologies for Battling Covid-19: Applications and Innovations*, pages 157–180, 2021.

[153] Valeria V Krzhizhanovskaya, GS Shirshov, Natalia B Melnikova, Robert G Belleman, FI Rusadi, BJ Broekhuijsen, Ben P Gouldby, Julien Lhomme, Bartosz Balis, Marian Bubak, et al. Flood early warning system: design, implementation and computational modules. *Procedia Computer Science*, 4:106–115, 2011.

[154] Adam J Kucharski, Timothy W Russell, Charlie Diamond, Yang Liu, John Edmunds, Sebastian Funk, Rosalind M Eggo, Fiona Sun, Mark Jit, James D Munday, et al. Early dynamics of transmission and control of COVID-19: a mathematical modelling study. *The lancet infectious diseases*, 20(5):553–558, 2020.

[155] Chin-Chi Kuo, Katherine A Moon, Shu-Li Wang, Ellen Silbergeld, and Ana Navas-Acien. The association of arsenic metabolism with cancer, cardiovascular disease, and diabetes: a systematic review of the epidemiological evidence. *Environmental health perspectives*, 125(8):087001, 2017.

[156] Chao-Chih Lai, Chen-Yang Hsu, Hsiao-Hsuan Jen, Amy Ming-Fang Yen, Chang-Chuan Chan, and Hsiu-Hsi Chen. The Bayesian Susceptible-Exposed-Infected-Recovered model for the outbreak of COVID-19 on the Diamond Princess Cruise Ship. *Stochastic Environmental Research and Risk Assessment*, pages 1–15, 2021.

[157] Andrew Lakoff. *Unprepared: global health in a time of emergency*. Univ of California Press, 2017.

[158] Harry T Lawless, Hildegarde Heymann, Harry T Lawless, and Hildegarde Heymann. Descriptive analysis. *Sensory evaluation of food: Principles and practices*, pages 227–257, 2010.

[159] Kalev Leetaru and Philip A Schrodt. Gdelt: Global data on events, location, and tone, 1979–2012. In *ISA annual convention*, volume 2, pages 1–49. Citeseer, 2013.

[160] Rui Li, Bowen Ying, Yingxiao Qian, Danni Chen, Xiaoxiao Li, Hui Zhu, and Hu Liu. Prevalence of self-reported symptoms of computer vision syndrome and associated risk factors among school students in China during the COVID-19 pandemic. *Ophthalmic Epidemiology*, 29(4):363–373, 2022.

[161] Chuyao Liao, Xiang Chen, Li Zhuo, Yuan Liu, Haiyan Tao, and Christopher G Burton. Reopen schools safely: Simulating COVID-19 transmission on campus with a contact network agent-based model. *International Journal of Digital Earth*, 15(1):381–396, 2022.

[162] Changjun Liu, Liang Guo, Lei Ye, Shunfu Zhang, Yanzeng Zhao, and Tianyu Song. A review of advances in China's flash flood early-warning system. *Natural hazards*, 92:619–634, 2018.

[163] Longjian Liu. *Heart failure: epidemiology and research methods*. Elsevier Health Sciences, 2017.

[164] Yu Liu, Richard B Saltman, and Ming-Jui Yeh. From bureaucratic administration to effective intervention: Comparing early governmental responses to the COVID-19 virus across East Asian and western health systems. *Health Services Management Research*, 36(3):193–204, 2023.

[165] Fabian Lorig, Emil Johansson, and Paul Davidsson. Agent-based social simulation of the COVID-19 pandemic: A systematic review. *Journal of Artificial Societies and Social Simulation*, 24(3), 2021.

[166] Fabian Lorig, Emil Johansson, and Paul Davidsson. Agent-Based Social Simulation of the Covid-19 Pandemic: A Systematic Review. *Journal of Artificial Societies and Social Simulation*, 24(3):5, 2021.

[167] Jie Luan, Jianbo Ba, Bin Liu, Xiongli Xu, and Dong Shu. 2021–2022 monitoring, early warning, and forecasting of global infectious diseases. *Journal of Biosafety and Biosecurity*, 4(2):98–104, 2022.

[168] Jonas F Ludvigsson. Children are unlikely to be the main drivers of the COVID-19 pandemic–a systematic review. *Acta Paediatrica*, 109(8):1525–1530, 2020.

[169] Sean Luke, Claudio Cioffi, Liviu Panait, Keith Sullivan, and Gabriel Balan. Mason: A multiagent simulation environment. *Simulation*, 81:517–527, 01 2005.

[170] Jianxi Luo. Forecasting COVID-19 pandemic: Unknown unknowns and predictive monitoring. *Technological forecasting and social change*, 166:120602, 2021.

[171] Nason Maani. The True Costs of the COVID-19 Pandemic. *Scientific American Blog Network*. URL: `https://blogs.scientificamerican.com/observations/the-true-costs-of-the-covid-19-pandemic/`, April 2020.

[172] Charles M Macal and Michael J North. Agent-based modeling and simulation. In *Proceedings of the 2009 winter simulation conference (WSC)*, pages 86–98. IEEE, 2009.

[173] Manotosh Mandal, Soovoojeet Jana, Swapan Kumar Nandi, Anupam Khatua, Sayani Adak, and TK Kar. A model based study on the dynamics of COVID-19: Prediction and control. *Chaos, Solitons & Fractals*, 136:109889, 2020.

[174] Madhav Marathe and Anil Kumar S Vullikanti. Computational epidemiology. *Communications of the ACM*, 56(7):88–96, 2013.

[175] Madhav V Marathe and Naren Ramakrishnan. Recent advances in computational epidemiology. *IEEE intelligent systems*, 28(4):96–101, 2013.

[176] Edouard Mathieu, Hannah Ritchie, Lucas Rodés-Guirao, Cameron Appel, Charlie Giattino, Joe Hasell, Bobbie Macdonald, Saloni Dattani, Diana Beltekian, Esteban Ortiz-Ospina, and Max Roser. Coronavirus Pandemic (COVID-19). *Our World in Data*, 2020. https://ourworldindata.org/coronavirus.

[177] Margaret McCartney. Margaret McCartney: Christmas covid-19 testing for students needs independent oversight. URL: `https://blogs.bmj.com/bmj/2020/12/01/margaret-mccartney-christmas-covid-19-testing-for-students-needs-independent-oversight/`. Archived at `https://web.archive.org/web/20220104180927/https://blogs.bmj.com/bmj/2020/12/01/margaret-mccartney-christmas-covid-19-testing-for-students-needs-independent-oversight/`, 2020.

[178] Darcy Jones McMaughan, Oluyomi Oloruntoba, and Matthew Lee Smith. Socioeconomic status and access to healthcare: interrelated drivers for healthy aging. *Frontiers in public health*, 8:231, 2020.

[179] Rehab Meckawy, David Stuckler, Adityavarman Mehta, Tareq Al-Ahdal, and Bradley N Doebbeling. Effectiveness of early warning systems in the detection of infectious diseases outbreaks: a systematic review. *BMC public health*, 22(1):2216, 2022.

[180] WG Meikle and Niels Holst. Application of continuous monitoring of honeybee colonies. *Apidologie*, 46:10–22, 2015.

[181] Dirk Merkel. Docker: lightweight linux containers for consistent development and deployment. *Linux journal*, 2014(239):2, 2014.

[182] Fanny Meyer and Victor Perrier. shinybusy: Busy indicator for "shiny"applications, 2020.

[183] James Kyle Miller, Kimberly Elenberg, and Artur Dubrawski. Forecasting emergence of COVID-19 variants of concern. *Plos one*, 17(2):e0264198, 2022.

[184] Karel Mls, Milan Kořínek, Kamila Štekerová, Petr Tučník, Vladimír Bureš, Pavel Čech, Martina Husáková, Peter Mikuleckỳ, Tomáš Nacházel, Daniela Ponce, et al. Agent-based models of human response to natural hazards: systematic review of tsunami evacuation. *Natural hazards*, 115(3):1887–1908, 2023.

[185] Zahra Mohammadi, Monica Gabriela Cojocaru, and Edward Wolfgang Thommes. Human behaviour, NPI and mobility reduction effects on COVID-19 transmission in different countries of the world. *BMC Public Health*, 22(1):1594, 2022.

[186] Stephen S Morse. Public health surveillance and infectious disease detection. *Biosecurity and bioterrorism: biodefense strategy, practice, and science*, 10(1):6–16, 2012.

[187] Khondoker Nazmoon Nabi, Md Toki Tahmid, Abdur Rafi, Muhammad Ehsanul Kader, and Md. Asif Haider. Forecasting COVID-19 cases: A comparative analysis between recurrent and convolutional neural networks. *Results in Physics*, 24:104137, 2021.

[188] Sk Shahid Nadim, Indrajit Ghosh, and Joydev Chattopadhyay. Short-term predictions and prevention strategies for COVID-19: a model-based study. *Applied mathematics and computation*, 404:126251, 2021.

[189] Michael J. North, Nicholson T. Collier, Jonathan Ozik, Eric R. Tatara, Charles M. Macal, Mark Bragen, and Pam Sydelko. Complex adaptive systems modeling with repast simphony. *Complex Adaptive Systems Modeling*, 1(1):3, Mar 2013.

[190] Barbara Nussbaumer-Streit, Verena Mayr, Andreea Iulia Dobrescu, Andrea Chapman, Emma Persad, Irma Klerings, Gernot Wagner, Uwe Siebert, Dominic Ledinger, Casey Zachariah, et al. Quarantine alone or in combination with other public health measures to control COVID-19: a rapid review. *Cochrane Database of Systematic Reviews*, 2020.

[191] Abu Quwsar Ohi, MF Mridha, Muhammad Mostafa Monowar, and Md Abdul Hamid. Exploring optimal control of epidemic spread using reinforcement learning. *Scientific reports*, 10(1):22106, 2020.

[192] William A Oleckno. *Epidemiology: concepts and methods*. Waveland Press, 2008.

[193] Jeroen Ooms, Duncan T Lang, and Lloyd Hilaiel. jsonlite: A robust, high performance JSON parser and generator for R. *R package version*, 1, 2018.

[194] World Health Organization. Coronavirus disease 2019 (COVID-19): situation report, 73. Technical documents, 2020.

[195] David Ornelles. Ebimage extra: Helper functions for ebimage, 2019. License: GPL-3.

[196] Marcus Painter and Tian Qiu. Political beliefs affect compliance with covid-19 social distancing orders. *Covid Economics*, 4(April):103–23, 2020.

[197] Jasmina Panovska-Griffiths, Cliff C Kerr, Robyn M Stuart, Dina Mistry, Daniel J Klein, Russell M Viner, and Chris Bonell. Determining the optimal strategy for reopening schools, the impact of test and trace interventions, and the risk of occurrence of a second COVID-19 epidemic wave in the UK: a modelling study. *The Lancet Child & Adolescent Health*, 4(11):817–827, 2020.

[198] Grégoire Pau, Florian Fuchs, Oleg Sklyar, Michael Boutros, and Wolfgang Huber. Ebimage—an r package for image processing with applications to cellular phenotypes. *Bioinformatics*, 26(7):979–981, 2010.

[199] Ruoling Peng, Kang Liu, and Po Yang. Flares: A framework for large-scale agent-based rapid epidemic simulation. In *2024 IEEE 22nd International Conference on Industrial Informatics (INDIN)*, pages 1–6, 2024.

[200] Zhe Peng, AL Pineda Rojas, Emilio Kropff, William Bahnfleth, Giorgio Buonanno, Stephanie J Dancer, Jarek Kurnitski, Yuguo Li, Marcel GLC Loomans, Linsey C Marr, et al. Practical indicators for risk of airborne transmission in shared indoor environments and their application to COVID-19 outbreaks. *Environmental science & technology*, 56(2):1125–1137, 2022.

[201] Marzio Pennisi, Giuliana Franceschinis, Daniele Baccega, Simone Pernice, and Irene Terrone. Exploiting gpu computing for effective agent-based simulation: initial experiments. In *2025 33rd Euromicro International Conference on Parallel, Distributed, and Network-Based Processing (PDP)*, pages 303–308. IEEE, 2025.

[202] Simone Pernice, Paolo Castagno, Linda Marcotulli, Milena Maria Maule, Lorenzo Richiardi, Giovenale Moirano, Matteo Sereno, Francesca Cordero, and Marco Beccuti. Impacts of reopening strategies for COVID-19 epidemic: a modeling study in Piedmont region. *BMC Infectious Diseases*, 20:1–9, 2020.

[203] Simone Pernice, Roberta Sirovich, Elena Grassi, Marco Viviani, Martina Ferri, Francesco Sassi, Luca Alessandrì, Dora Tortarolo, Raffaele A Calogero, Livio Trusolino, et al. CONNECTOR, fitting and clustering of longitudinal data to reveal a new risk stratification system. *Bioinformatics*, 39(5):btad201, 2023.

[204] Nicola Perra. Non-pharmaceutical interventions during the COVID-19 pandemic: A review. *Physics Reports*, 913:1–52, 2021.

[205] Victor Perrier, Fanny Meyer, David Granjon, et al. shinywidgets: Custom inputs widgets for shiny. *R package version 0.4*, 7, 2019.

[206] James L. Peterson. Petri Nets. *ACM Comput. Surv.*, 9(3):223–252, September 1977.

[207] Brendon Phillips, Dillon T Browne, Madhur Anand, and Chris T Bauch. Model-based projections for COVID-19 outbreak size and student-days lost to closure in Ontario childcare centres and primary schools. *Scientific reports*, 11(1):1–14, 2021.

[208] Regione Piemonte. Scuola: operative entro il 17 novembre le nuove regole per le quarantene. `https://www.regione.piemonte.it/web/pinforma/notizie/scuola-operative-entro-17-novembre-nuove-regole-per-quarantene`.

[209] Alessia Pini and Simone Vantini. The interval testing procedure: A general framework for inference in functional data analysis. *Biometrics*, 72:835–845, 2016.

[210] Jocelyne Piret and Guy Boivin. Pandemics throughout history. *Frontiers in microbiology*, 11:631736, 2021.

[211] Kiesha Prem, Alex R Cook, and Mark Jit. Projecting social contact matrices in 152 countries using contact surveys and demographic data. *PLoS Comput Biol*, 13(9):e1005697, September 2017.

[212] Trustalert Project. Trustalert. `https://www.trustalert.it`, 2023. [Accessed: 2024-10-18].

[213] Zhaozhi Qian, Ahmed M Alaa, and Mihaela van der Schaar. When and how to lift the lockdown? global covid-19 scenario analysis and policy assessment using compartmental

gaussian processes. *Advances in neural information processing systems*, 33:10729–10740, 2020.

[214] Joseph E Quansah, Bernard Engel, and Gilbert L Rochon. Early warning systems: a review. *Journal of Terrestrial Observation*, 2(2):5, 2010.

[215] Iman Rahimi, Fang Chen, and Amir H Gandomi. A review on COVID-19 forecasting models. *Neural Computing and Applications*, pages 1–11, 2021.

[216] Iman Rahimi, Fang Chen, and Amir H Gandomi. A review on covid-19 forecasting models. *Neural computing and applications*, 35(33):23671–23681, 2023.

[217] Somayeh Bakhtiari Ramezani, Amin Amirlatifi, and Shahram Rahimi. A novel compartmental model to capture the nonlinear trend of COVID-19. *Computers in Biology and Medicine*, 134:104421, 2021.

[218] Regione Piemonte. Coronavirus e scuola, le indicazioni. URL: `https://www.regione.piemonte.it/web/temi/coronavirus-piemonte/coronavirus-scuola-indicazioni`, 2020.

[219] Muhammad Rendana and Wan Mohd Razi Idris. New COVID-19 variant (B. 1.1. 7): forecasting the occasion of virus and the related meteorological factors. *Journal of infection and public health*, 14(10):1320–1327, 2021.

[220] David K Rensin. Kubernetes: Scheduling the Future at Cloud Scale. 2015. [Accessed: 2024-12-02].

[221] Carl Orge Retzlaff, Martina Ziefle, and André Calero Valdez. The history of agent-based modeling in the social sciences. In *International Conference on Human-Computer Interaction*, pages 304–319. Springer, 2021.

[222] Paul Richmond, Robert Chisholm, Peter Heywood, Mozhgan Kabiri Chimeh, and Matthew Leach. FLAME GPU 2: A framework for flexible and performant agent based simulation on GPUs. *Software: Practice and Experience*, 2023.

[223] Paul Richmond, Robert Chisholm, Peter Heywood, Matthew Leach, and Mozhgan Kabiri Chimeh. FLAME GPU, December 2022.

[224] Steven Riley, Haowei Wang, Oliver Eales, David Haw, Caroline E Walters, Kylie EC Ainslie, Christina Atchison, Claudio Fronterre, Peter J Diggle, Andrew J Page, et al. REACT-1 round 12 report: resurgence of SARS-CoV-2 infections in England associated with increased frequency of the Delta variant. *MedRxiv*, pages 2021–06, 2021.

[225] Ian D Roberts and Cendri A Hutcherson. Affect and decision making: Insights and predictions from computational models. *Trends in cognitive sciences*, 23(7):602–614, 2019.

[226] Joacim Rocklöv, Henrik Sjödin, and Annelies Wilder-Smith. COVID-19 outbreak on the Diamond Princess cruise ship: estimating the epidemic potential and effectiveness of public health countermeasures. *Journal of travel medicine*, 27(3):taaa030, 2020.

[227] David Rodbard. Continuous glucose monitoring: a review of successes, challenges, and opportunities. *Diabetes technology & therapeutics*, 18(S2):S2–3, 2016.

[228] Alexander Rodríguez, Jiaming Cui, Naren Ramakrishnan, Bijaya Adhikari, and B Aditya Prakash. Einns: epidemiologically-informed neural networks. In *Proceedings of the AAAI conference on artificial intelligence*, volume 37, pages 14453–14460, 2023.

[229] Alexander Rodríguez, Harshavardhan Kamarthi, Pulak Agarwal, Javen Ho, Mira Patel, Suchet Sapre, and B Aditya Prakash. Machine learning for data-centric epidemic forecasting. *Nature Machine Intelligence*, 6(10):1122–1131, 2024.

[230] David Rogers and Vladimir Tsirkunov. Costs and benefits of early warning systems. *Global assessment rep*, 2011.

[231] Jesús Rufino, Carlos Baquero, Davide Frey, Christin A Glorioso, Antonio Ortega, Nina Reščič, Julian Charles Roberts, Rosa E Lillo, Raquel Menezes, Jaya Prakash Champati, et al. Using survey data to estimate the impact of the omicron variant on vaccine efficacy against COVID-19 infection. *Scientific Reports*, 13(1):900, 2023.

[232] Andrea Saglietto, Daniele Baccega, Roberto Esposito, Matteo Anselmino, Veronica Dusi, Attilio Fiandrotti, and Gaetano Maria De Ferrari. Convolutional neural network (CNN)-enabled electrocardiogram (ECG) analysis: a comparison between standard twelve-lead and single-lead setups. *Frontiers in Cardiovascular Medicine*, 11:1327179, 2024.

[233] Sweeti Sah, B Surendiran, R Dhanalakshmi, Sachi Nandan Mohanty, Fayadh Alenezi, Kemal Polat, et al. Forecasting COVID-19 pandemic using Prophet, ARIMA, and hybrid stacked LSTM-GRU models in India. *Computational and Mathematical Methods in Medicine*, 2022, 2022.

[234] Joshua A Salomon, Alex Reinhart, Alyssa Bilinski, Eu Jing Chua, Wichada La Motte-Kerr, Minttu M Rönn, Marissa B Reitsma, Katherine A Morris, Sarah LaRocca,

Tamer H Farag, et al. The US COVID-19 Trends and Impact Survey: Continuous real-time measurement of COVID-19 symptoms, risks, protective behaviors, testing, and vaccination. *Proceedings of the National Academy of Sciences*, 118(51):e2111454118, 2021.

[235] Sam Abbott, Joel Hellewell, Katharine Sherratt, Katelyn Gostic, Joe Hickson, Hamada S. Badr, Michael DeWitt, Robin Thompson, EpiForecasts, and Sebastian Funk. *EpiNow2: Estimate Real-Time Case Counts and Time-Varying Epidemiological Parameters*, 2020.

[236] Egle Samanidou, Elmar Zschischang, Dietrich Stauffer, and Thomas Lux. Agent-based models of financial markets. *Reports on Progress in Physics*, 70(3):409, 2007.

[237] Iqra Sardar, Muhammad Azeem Akbar, Víctor Leiva, Ahmed Alsanad, and Pradeep Mishra. Machine learning and automatic ARIMA/Prophet models-based forecasting of COVID-19: Methodology, evaluation, and case study in SAARC countries. *Stochastic Environmental Research and Risk Assessment*, 37(1):345–359, 2023.

[238] Md Shohel Sayeed, Siti Najihah Hishamuddin, and Ong Thian Song. Covid-19 forecasting model based on machine learning approaches: a review. *Bulletin of Electrical Engineering and Informatics*, 13(6):4335–4345, 2024.

[239] Ketan Rajshekhar Shahapure and Charles Nicholas. Cluster quality analysis using silhouette score. In *2020 IEEE 7th international conference on data science and advanced analytics (DSAA)*, pages 747–748. IEEE, 2020.

[240] Farah Shahid, Aneela Zameer, and Muhammad Muneeb. Predictions for COVID-19 with deep learning models of LSTM, GRU and Bi-LSTM. *Chaos, Solitons & Fractals*, 140:110212, 2020.

[241] Md Salman Shamil, Farhanaz Farheen, Nabil Ibtehaz, Irtesam Mahmud Khan, and M Sohel Rahman. An agent-based modeling of covid-19: validation, analysis, and recommendations. *Cognitive computation*, pages 1–12, 2021.

[242] Adil E Shamoo and David B Resnik. *Responsible conduct of research*. Oxford University Press, 2009.

[243] Wenjing Shao, Jingui Xie, and Yongjian Zhu. Mediation by human mobility of the association between temperature and COVID-19 transmission rate. *Environmental research*, 194:110608, 2021.

[244] Mohammad Shirvani, Georges Kesserwani, and Paul Richmond. Agent-based simulator of dynamic flood-people interactions. *Journal of flood risk management*, 14(2):e12695, 2021.

[245] Petrônio CL Silva, Paulo VC Batista, Hélder S Lima, Marcos A Alves, Frederico G Guimarães, and Rodrigo CP Silva. COVID-ABS: An agent-based model of COVID-19 epidemic to simulate health and economic effects of social distancing interventions. *Chaos, Solitons & Fractals*, 139:110088, 2020.

[246] Berber T Snoeijer, Mariska Burger, Shaoxiong Sun, Richard JB Dobson, and Amos A Folarin. Measuring the effect of Non-Pharmaceutical Interventions (NPIs) on mobility during the COVID-19 pandemic using global mobility data. *NPJ digital medicine*, 4(1):81, 2021.

[247] John Snow. *On the mode of communication of cholera*. John Churchill, 1855.

[248] Mohammad Soheilypour and Mohammad RK Mofrad. Agent-based modeling in molecular systems biology. *BioEssays*, 40(7):1800020, 2018.

[249] Charles Spearman. The proof and measurement of association between two things. 1961.

[250] Chen Stein-Zamir, Nitza Abramson, Hanna Shoob, Erez Libal, Menachem Bitan, Tanya Cardash, Refael Cayam, and Ian Miskin. A large COVID-19 outbreak in a high school 10 days after schools' reopening, Israel, May 2020. *Eurosurveillance*, 25(29):2001352, 2020.

[251] Sybert H Stroeve and Mariken HC Everdij. Agent-based modelling and mental simulation for resilience engineering in air transport. *Safety science*, 93:29–49, 2017.

[252] Charles S Taber and Richard J Timpone. *Computational modeling*, volume 113. Sage, 1996.

[253] Carrick Talmadge and R Waxler. Infrasound from tornados: Theory, measurement, and prospects for their use in early warning systems. *Acoust. Today*, 12(1):43–51, 2016.

[254] Cheng Tang, Yuki Todo, Sachiko Kodera, Rong Sun, Atsushi Shimada, and Akimasa Hirata. A novel multivariate time series forecasting dendritic neuron model for covid-19 pandemic transmission tendency. *Neural Networks*, 179:106527, 2024.

[255] Muhammad Usman Tariq and Shuhaida Binti Ismail. Deep learning in public health: Comparative predictive models for covid-19 case forecasting. *Plos one*, 19(3):e0294289, 2024.

[256] Hanisha Tatapudi and Tapas K Das. Impact of school reopening on pandemic spread: A case study using an agent-based model for COVID-19. *Infectious Disease Modelling*, 6:839–847, 2021.

[257] Sean J Taylor and Benjamin Letham. Forecasting at scale. *The American Statistician*, 72(1):37–45, 2018.

[258] The University of Maryland Social Data Science Center. COVID19_symptom_survey_intl_V11_noneu, 2021. Accessed 2 July 2024.

[259] Lindsay Thomas. Pandemics of the future: Disease surveillance in real time. *Surveillance & Society*, 12(2):287–300, 2014.

[260] Kristin Tolksdorf, Silke Buda, Ekkehard Schuler, Lothar H Wieler, and Walter Haas. Influenza-associated pneumonia as reference to assess seriousness of coronavirus disease (COVID-19). *Eurosurveillance*, 25(11):2000258, 2020.

[261] Juliana Tolles and ThaiBinh Luong. Modeling epidemics with compartmental models. *Jama*, 323(24):2515–2516, 2020.

[262] John L Tonry. An early warning system for asteroid impact. *Publications of the Astronomical Society of the Pacific*, 123(899):58, 2010.

[263] Oskar Triebe, Hansika Hewamalage, Polina Pilyugina, Nikolay Laptev, Christoph Bergmeir, and Ram Rajagopal. Neuralprophet: Explainable forecasting at scale. *arXiv preprint arXiv:2111.15397*, 2021.

[264] Jörg Unbehauen, Sebastian Hellmann, Sören Auer, and Claus Stadler. Knowledge extraction from structured sources. *Search computing: Broadening web search*, pages 34–52, 2012.

[265] Maarten K Van Aalst. The impacts of climate change on the risk of natural disasters. *Disasters*, 30(1):5–18, 2006.

[266] H Rogier van Doorn. The epidemiology of emerging infectious diseases and pandemics. *Medicine*, 49(10):659–662, 2021.

[267] Benjamin Van Rooij, Anne Leonore de Bruijn, Chris Reinders Folmer, Emmeke Barbara Kooistra, Malouke Esra Kuiper, Megan Brownlee, Elke Olthuis, and Adam Fine. Compliance with COVID-19 mitigation measures in the United States. *Amsterdam law school research paper*, 2020.

[268] Alexey Voinov and Carl Cerco. Model integration and the role of data. *Environmental Modelling & Software*, 25(8):965–969, 2010.

[269] Kevin G Volpp, Bruce H Kraut, Smita Ghosh, and John Neatherlin. Minimal SARS-CoV-2 transmission after implementation of a comprehensive mitigation strategy at a school—New Jersey, August 20–November 27, 2020. *Morbidity and Mortality Weekly Report*, 70(11):377, 2021.

[270] Jian Wang, Ke Wang, Jing Li, Jianmin Jiang, Yanfei Wang, Jing Mei, and Shaochun Li. Accelerating epidemiological investigation analysis by using NLP and knowledge reasoning: a case study on COVID-19. In *AMIA Annual Symposium Proceedings*, volume 2020, page 1258, 2021.

[271] Lijing Wang, Jiangzhuo Chen, and Madhav Marathe. Defsi: Deep learning based epidemic forecasting with synthetic information. In *Proceedings of the AAAI conference on artificial intelligence*, volume 33, pages 9607–9612, 2019.

[272] Peipei Wang, Xinqi Zheng, Jiayang Li, and Bangren Zhu. Prediction of epidemic trends in COVID-19 with logistic model and machine learning technics. *Chaos, Solitons & Fractals*, 139:110058, 2020.

[273] Toru Watanabe, Timothy A Bartrand, Mark H Weir, Tatsuo Omura, and Charles N Haas. Development of a dose-response model for SARS coronavirus. *Risk Analysis: An International Journal*, 30(7):1129–1138, 2010.

[274] Gregory L Watson, Di Xiong, Lu Zhang, Joseph A Zoller, John Shamshoian, Phillip Sundin, Teresa Bufford, Anne W Rimoin, Marc A Suchard, and Christina M Ramirez. Pandemic velocity: Forecasting COVID-19 in the US with a machine learning & Bayesian time series compartmental model. *PLoS computational biology*, 17(3):e1008837, 2021.

[275] Oliver J Watson, Gregory Barnsley, Jaspreet Toor, Alexandra B Hogan, Peter Winskill, and Azra C Ghani. Global impact of the first year of COVID-19 vaccination: a mathematical modelling study. *The Lancet Infectious Diseases*, 22(9):1293–1302, 2022.

[276] Duncan J Watts and Steven H Strogatz. Collective dynamics of small-world networks. *nature*, 393(6684):440–442, 1998.

[277] Jack West, Serenydd Everden, and Nikitas Nikitas. A case of COVID-19 reinfection in the UK. *Clinical medicine*, 21(1):e52, 2021.

[278] Hadley Wickham. stringr: modern, consistent string processing. *The R Journal*, 2(2):38–40, 2010.

[279] Hadley Wickham. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York, 2016.

[280] Hadley Wickham. Package 'tidyr'. *Easily Tidy Data with'spread'and'gather ()'Functions*, 2017.

[281] Hadley Wickham, Romain François, Lionel Henry, Kirill Müller, and Davis Vaughan. *dplyr: A Grammar of Data Manipulation*, 2023. R package version 1.1.4.

[282] Hadley Wickham, Jim Hester, Winston Chang, and Jennifer Bryan. *devtools: Tools to Make Developing R Packages Easier*, 2022. https://devtools.r-lib.org/, https://github.com/r-lib/devtools.

[283] Hadley Wickham, Jim Hester, Romain Francois, Jennifer Bryan, Shelby Bearrows, J Jylänki, and M Jørgensen. Package 'readr'. *Read Rectangular Text Data. Available online: https://cran. r-project. org/web/packages/readr/readr. pdf (accessed on 23 August 2023)*, 2024.

[284] Jenna Wiens and Erica S Shenoy. Machine learning for healthcare: on the verge of a major shift in healthcare epidemiology. *Clinical infectious diseases*, 66(1):149–153, 2018.

[285] Uri Wilensky. NetLogo: Center for connected learning and computer-based modeling, Northwestern University. Evanston, IL, 1999, 2014.

[286] Yihui Xie, Joe Cheng, Xianying Tan, et al. Dt: a wrapper of the javascript library 'datatables'. *R package version 0.4*, 2018.

[287] Wei W Xun, Aneire E Khan, Edwin Michael, and Paolo Vineis. Climate change epidemiology: methodological challenges. *International Journal of Public Health*, 55:85–96, 2010.

[288] Andy B Yoo, Morris A Jette, and Mark Grondona. Slurm: Simple linux utility for resource management. In *Workshop on job scheduling strategies for parallel processing*, pages 44–60. Springer, 2003.

[289] Nazar Zaki and Elfadil A. Mohamed. The estimations of the COVID-19 incubation period: A scoping reviews of the literature. *Journal of Infection and Public Health*, 14(5):638–646, 2021.

[290] Mengxi Zhang, Siqin Wang, Tao Hu, Xiaokang Fu, Xiaoyue Wang, Yaxin Hu, Briana Halloran, Zhenlong Li, Yunhe Cui, Haokun Liu, et al. Human mobility and COVID-19 transmission: a systematic review and future directions. *Annals of GIS*, 28(4):501–514, 2022.

[291] Peijue Zhang, Kairui Feng, Yuqing Gong, Jieon Lee, Sara Lomonaco, and Liang Zhao. Usage of Compartmental Models in Predicting COVID-19 Outbreaks. *The AAPS Journal*, 24(5):98, 2022.

[292] Daren Zhao, Ruihua Zhang, Huiwu Zhang, and Sizhang He. Prediction of global omicron pandemic using ARIMA, MLR, and Prophet models. *Scientific reports*, 12(1):18138, 2022.

[293] Yuan Zhou, Alexander Nikolaev, Ling Bian, Li Lin, and Lin Li. Investigating transmission dynamics of influenza in a public indoor venue: An agent-based modeling approach. *Computers & Industrial Engineering*, 157:107327, 2021.

[294] Jochen Zschau and Andreas N Küppers. *Early warning systems for natural disaster reduction*. Springer Science & Business Media, 2013.