Quality & Quantity
https://doi.org/10.1007/s11135-024-01913-x

Pandemic data quality modelling: a Bayesian approach in the Italian case

Luisa Ferrari[1] · Giancarlo Manzi[2] · Alessandra Micheletti[3] · Federica Nicolussi[4] · Silvia Salini[2]

Accepted: 29 May 2024

Abstract
When pandemics like COVID-19 spread around the world, the rapidly evolving situation compels officials and executives to take prompt decisions and adapt policies depending on the current state of the disease. In this context, it is crucial for policymakers to always have a firm grasp on what is the current state of the pandemic, and envision how the number of infections and possible deaths is going to evolve shortly. However, as in many other situations involving compulsory registration of sensitive data from multiple collectors, cases might be reported with errors, often with delays deferring an up-to-date view of the state of things. Errors in collecting new cases affect the overall mortality, resulting in excess deaths reported by official statistics only months later. In this paper, we provide tools for evaluating the quality of pandemic mortality data. We accomplish this through a Bayesian approach accounting for the excess mortality pandemics might bring with respect to the normal level of mortality in the population.

Keywords Pandemics · Bayesian analysis · Variance models · Time-space models

✉ Giancarlo Manzi
giancarlo.manzi@unimi.it

Luisa Ferrari
luisa.ferrari5@unibo.it

Alessandra Micheletti
alessandra.micheletti@unimi.it

Federica Nicolussi
federica.nicolussi@polimi.it

Silvia Salini
silvia.salini@unimi.it

[1] Department of Statistical Sciences “Paolo Fortunati”, University of Bologna, Via Belle Arti, 41, 40126 Bologna, Italy

[2] Department of Economics, Management and Quantitative Methods, University of Milan, Via Conservatorio, 7, 20122 Milan, Italy

[3] Department of Environmental Science and Policy, University of Milan, Via Celoria, 2, 20133 Milan, Italy

[4] MOX-Department of Mathematics, Politecnico di Milano, Piazza Leonardo da Vinci, 32, 20133 Milan, Italy

Published online: 04 July 2024

1 Introduction

Historically, human populations have always dealt with major outbreaks of infectious diseases and effective reaction to these outbreaks has always been sought. The numerous black plague outbreaks in Europe in medieval and post-medieval times, the cholera pandemic in the late nineteenth- early twentieth century, and the Spanish flu in the early twentieth century are only a few examples in modern history.

However, the COVID-19 pandemic, which brought the world close to a halt in 2020 and 2021 and killed almost seven million people as of early 2023, has brought to light a new reality of a global pandemic never experienced before.

Worldwide governments initially underscored its urgency and their healthcare systems were quickly overwhelmed in a tsunami-like fashion. Nearly all of the affected countries progressively implemented measures to slow down the spread of the virus, ranging from recommending social distancing to introducing national lockdowns of social and economic activity.

These measures eventually proved to be effective, allowing numerous countries to relax restrictions, in an attempt to gradually return to normality. At the same time, with the threat posed by the virus still looming, decision-makers were forced to strike a balance between epidemiological risk and allowance of socioeconomic activity. In this type of context, surveillance of the number of new infections, mortality monitoring and quantification of the effects of social distancing became increasingly important (Colombo et al. 2020; Kantner and Koprucki 2020; Wu et al. 2020), particularly so at the regional level. Given the local nature of the phenomenon, such a regional view appears to be of crucial importance. One of the difficulties lies in the fact that exact numbers of new infections, reported deaths and recoveries from the disease are often available only with a certain probability to be later corrected or are reported with a delay of—sometimes—several days. Moreover, there are discrepancies among data reported at different levels (regional, provincial, etc.) sometimes crucially causing wrong responses to face the pandemic.

Another important aspect to keep in mind is that the deaths due to COVID-19 nationwide are underestimated by official data because the deaths of people who have not been tested for the disease are not counted. In addition, the pandemic had an indirect effect on mortality, preventing timely treatment for other diseases or limiting preventive examinations that could have anticipated critical situations. For all these reasons, official data on COVID-19 deaths in some countries were fairly inadequate to measure the effect of the pandemic (Covid-19 Excess Mortality Collaborators 2022). In aid of this, the World Health Organization (WHO) suggests studying excess mortality to assess the death burden (both direct and indirect) of COVID-19. Excess mortality refers to the number of deaths from all causes during the pandemic more than what we would have expected under "normal" conditions, which is a valid measure of the total effect of the COVID-19 pandemic (Beaney et al. 2020). Worldwide studies on excess mortality have been deepened as reported in Msemburi et al. (2023) and Shang et al. (2022).

Multiple model-based mortality estimates have been proposed in the literature, (Beaney et al. 2020; Blangiardo et al. 2020; Maruotti et al. 2022; Michelozzi et al. 2020; Achilleos et al. 2022) at national, regional and also county level. Several studies have shown a critical increase in excess mortality for higher age groups and for males, see, for instance, Gibertoni et al. (2021).

Excluding COVID-19, the ongoing global disease outbreaks reported by the WHO between 7/4/2022 and 7/4/2023 with at least one death are those listed in Table 1. Most

Table 1 Disease outbreaks with at least 1 death reported by WHO

Affected country	Last news update	Virus (presumed year of first reporting)	Case fatality ratio (%)	Deaths per day
Saudi Arabia	7/4/2022	MERS-Cov (2012)	66.67	0.02
Several countries	12/7/2022	Severe acute hepatitis of unknown aetiology in children (2022)	2.18	0.23
Malawi	9/2/2022	Cholera (19th century)	3.28	3.59
Dem. Rep. of Congo	28/4/2022	Ebola (1976)	100.00	0.02
Australia	28/4/2022	Japanese encephalitis (1871)	12.00	0.02
Qatar	12/5/2022	MERS-Cov (2012)	50.00	0.002
Cameroon	16/5/2022	Cholera (19th century)	2.01	0.73
Iraq	1/6/2022	Crimean-Congo Fever (1940s)	13.40	0.09
African region	10/6/2022	Monkeypox (1970s)	4.69	2.57
Pakistan	17/6/2022	Cholera (19th century)	0.70	0.07
Somalia	20/7/2022	Cholera (19th century)	0.47	0.19
Ghana	22/7/2022	Marburg virus disease (1967)	100.00	2.00
Bangladesh	28/11/2022	Dengue (18th century)	0.43	0.71
Tanzania	12/8/2022	Leptospirosis (18th century)	20.00	0.12
African region	3/1/2023	Yellow fever (17th century)	8.79	0.12
Argentina	5/9/2022	Legionellosis (1977)	36.35	0.25
Uganda	8/12/2022	Ebola disease (Sudan virus) (1976)	38.73	0.72
Nepal	10/10/2022	Dengue (18th century)	0.14	0.14
Haiti	13/12/2022	Cholera (19th century)	2.07	4.35
Pakistan	13/10/2022	Dengue (18th century)	0.24	0.23
Pakistan	17/10/2022	Malaria (Paleogene period)	–	–
Lebanon	19/10/2022	Cholera (19th century)	11.11	0.29
Mauritania	20/10/2022	Rift Valley fever (early 1900s)	48.94	0.48
Niger	8/2/2023	Meningitis (19th century)	16.22	0.21
Dem. Rep. of Congo	10/2/2023	Cholera (19th century)	0.36	0.31
Bangladesh	17/2/2023	Nipah virus infection (1999)	72.73	0.20
Mozambique	24/2/2023	Cholera (19th century)	0.71	0.23
Equatorial Guinea	25/2/2023	Marburg virus disease (1967)	96.55	0.38

7/4/2022—7/4/2023 (Source: Authors' calculations from WHO data)

of the diseases are endemic, especially in Africa and Asia. Diseases first reported in the last 25 years are, apart from Sars-Cov2, Mers-Cov (first reported in 2012), the Nipha virus (first reported in 1999) and a new type of hepatitis in children (first reported in 2022). Of the 16 diseases in an outbreak phase, 3 were first detected in the 18th century or before, 3 in the 19th century, 8 in the 20th century and 2 in the first 23 years of the 21st century.

Hence, it is essential to consider the likelihood of future pandemics. These pandemics may arise from various factors, such as global warming, as indicated in Christie (2021) where the author highlights the potential emergence of dangerous and unknown viruses from the thawing permafrost in Siberia. Additionally, advancements in intensive

farming and cultivation techniques pose a significant risk. Moreno-Madriñan and Kontowicz (2023) discuss how densely populated farming of single animal species can heighten the risk of mutation, re-assortment, and the generation of new pathogens. Furthermore, the escalating trends in migration, urbanization, and global conflicts, as mentioned in Hoiby (2020), contribute to the complex landscape.

Given these factors, it becomes evident that addressing the "data challenge" associated with pandemics should be regarded as one of the foremost global priorities. This imperative is underscored by the expectation that future pandemics may share similarities with COVID-19 in terms of rapid spread, contagion rates, and societal impact. It is not a matter of chance that during the COVID-19 pandemic, among Western countries, those most hit had problems with their data collecting systems. Indeed, the viral characteristics of COVID-19 do not allow a clear country comparison for the efficiency of facing the pandemic. In fact, Martínez-Córdoba et al. (2021) found that European and American countries were less efficient than South Asian and African countries, but they also admitted that this was mainly due to demographic features of the populations, hidden mortality in African and South Asian countries and the virus hitting mostly the elderly and fragile people. However, other viruses with similar contagion speeds and rates but hitting youngsters rather than the elderly might be hard to face if data collection is not efficient.

In this work, our contribution to the literature on the monitoring of pandemics and in particular of COVID-19 can be summarized as follows. (i) The data collection concerned smaller areas where the virus spread than those provided by official statistics. This has allowed us to use space-time models in which the spatial component can be better represented and give a more appropriate forecasting contribution. (ii) The use of data quality metrics that better distinguish the different failures of how the pandemic was addressed allows us to provide more stringent information on how to address pandemics similar to COVID-19 in the future. Finally, (iii) the use of a model that reparametrizes the classic space-time model in terms of re-proportioning the variance (carried out through the choice of priors) allows us to better distinguish between the effects related to time, space and their interaction.

In this paper, we consider the bias between excess mortality and the official Italian COVID-19 data in the first 2020 outbreak for evaluating data quality in a space-time context. We model this bias following the Bayesian framework where two different quality measures ought to be evaluated: (i) the share in the population dying because of a particular infectious disease without being officially reported, so large values of this measure represent a worse scenario, and (ii) the coverage of the pandemic by health systems, which can be considered an adequate indicator of their quality and a proxy for the efficacy of the crisis response. Among the factors explaining the bias variability, the focus is on detecting the most important component among spatial, temporal and interaction components.

The paper is organized as follows. Section 2 is devoted to a description of the data used, and Sect. 3 focuses on proposed metrics for quantifying bias in official death data. In Sect. 4 we present the spatial, temporal, and spatio-temporal Bayesian model for the proposed metrics. The results are displayed in Sect. 5. Finally, Sect. 6 concludes the paper with an overview of possible future work.

2 Data

To evaluate the data quality of pandemic mortality in general and of COVID-19 mortality in particular, we considered two data sources: (i) official data on the pandemic evolution considering the essential variables for classic SE(I)RD models (daily or weekly new infections—i.e. cases—, susceptible, exposed, recovered and deceased people). Often a similar data collection starts *ad hoc* at the request of national and international organisms to face the pandemic as was the case for the COVID-19 pandemic. In the following, we refer to this data as *official data*. (ii) National or supranational statistical institute data on population mortality. Henceforth, we refer to this data as *ISTAT data* as our main application will be on the Italian case, being ISTAT the Italian national statistical agency.

2.1 Official data

In Italy, official data about deaths related to COVID-19 was reported daily from the 24th of February 2020 to the 30th of October 2022, at the European Union NUTS-2 level (i.e. regions), by the Italian Ministry of Health. While information about new cases was also published at the more refined NUTS-3 level (i.e. provinces), the number of daily COVID-19 new deaths was not officially available at this level. Nevertheless, it was possible to reconstruct the time series of COVID-19 at NUTS-3 level indirectly, using other official sources like regional authorities' daily bulletins on provincial new cases, hospitalization and deaths and other information sources (Ferrari et al. 2021). Bulletins were published by most of the region authorities in a *pdf* format so we were able to scrape data from these documents and retrieve the data of interest for the majority of the Italian provinces.

Table 2 presents an example to understand the difference between the ISS-ISTAT officially daily reported data on COVID-19 deaths in the Marche region and data on COVID-19 deaths reported on the regional bulletin which was scraped from the Marche regional government website. It can be noted from the example that ISS-ISTAT death data were always reported with a bit of delay compared to the bulletin data until a revision after two or three days.

In March 2022, ISTAT published data from Istituto Superiore di Sanità (ISS - the main public health institute in Italy) about the monthly evolution of COVID-19 deaths in each province. Using the regional weekly trends, it was possible to reconstruct an estimate for the weekly provincial COVID-19 officially reported deaths for the provinces which were missing from the scraped data (for example, the Lombardy region authority never published daily bulletins on provincial deaths). This technique has also been applied for some provinces for which the data had been scraped but the discrepancy with this monthly report was significant for some provinces, especially those experiencing low levels of mortality.

Official data about deaths caused by COVID-19 is assumed to have been subject to delays, errors, inconsistencies between reporting protocols, etc. Moreover, it is sensible to assume that the data is affected by a systematic underestimation and is therefore biased as an estimate for actual COVID-19-related deaths: especially in the first pandemic stages, testing and care were not available for all people infected with the disease, and the official data only reflected the share of people that went through at least a minimal contact with the health system (Castaldi et al. 2020, 2021; Rivieccio et al. 2021). Of course, this bias is supposed to change over time, as well as space, and, in general, it is assumed to be directly proportional to the severity of the pandemic.

Table 2 Example of differences in reporting daily deaths between ISTAT-ISS data and regional bulletin data

Date	Deaths reported by regional bulletins Provinces					Total from bulletins	Total from ISTAT-ISS	Diff.
	Ancona	Pesaro	Fermo	Ascoli	Macerata			
	(1)	(2)	(3)	(4)	(5)	(6)=sum[(1):(5)]	(7)	(8)=(7)-(6)
12/3/2020	2	3	0	0	0	5	4	− 1
13/3/2020	1	7	0	1	0	9	5	− 4
14/3/2020	6	3	0	0	1	10	9	− 1
15/3/2020	2	9	0	0	0	11	10	− 1
16/3/2020	0	9	1	0	2	12	11	− 1
17/3/2020	10	7	1	1	3	22	11	− 11
18/3/2020	2	18	0	0	3	23	23	0
19/3/2020	6	13	1	0	2	22	23	+ 1
20/3/2020	6	11	0	0	0	17	22	+ 5
21/3/2020	3	27	0	0	0	30	17	− 13
22/3/2020	5	14	0	0	0	19	30	+ 11
23/3/2020	11	13	1	0	3	28	19	− 9
24/3/2020	11	16	0	0	5	32	28	− 4
25/3/2020	7	11	0	1	4	23	56	+ 33
Total	72	161	4	3	23	263	269	+ 5

Additionally, COVID-19 did not only cause deaths directly because of infection, but the burden on the health system caused by the pandemic prevented to cure of other diseases and accidents, causing indirect deaths of other individuals as well. Although lockdown measures might have prevented some of these ”traditional” causes of death, this aspect should nevertheless be taken into account when assessing the impact on mortality caused by COVID-19.

2.2 ISTAT data

Each year, ISTAT provides a daily record of deaths reported in each municipality of Italy.[1] In this study, the ISTAT data is aggregated weekly by province: this is because the corresponding COVID-19 official data is affected by strong seasonality over the weekdays. This detailed time series data can help to estimate the actual deaths in 2020 caused by COVID-19, both directly and indirectly. A straightforward way to estimate this is through the concept of ”excess mortality”, i.e. the excess between the 2020 overall mortality and the average in the previous few years.

In this work, it is proposed to use a 5-year window for the period 2015-2019 to represent the stable mortality level. The excess mortality is then found by subtracting this stable level from the 2020 deaths data, in each province and week of the year. While the 2015-2019 average will be smoother, the 2020 time series is subject to a greater amount of noise,

[1] Source: https://www.istat.it/en/archivio/268504.

so we decided to apply minimal smoothing, consisting of a 3-week moving average, to make it more stable.

Since no other relevant information is available, this excess mortality can be imputed to COVID-19 and be an estimate for the actual number of deaths caused by COVID-19, both directly and indirectly. Indeed, in many works the indirect contribution of the excess mortality due to COVID-19 is highlighted, see, for example, Dorrucci et al. (2021); Achilleos et al. (2022); Modig et al. (2021) and Vanella et al. (2021). Mixing this estimate for the actual number of deaths with the official data allows us to create measures for the under-reporting bias present in official figures. Such metrics can be interpreted as a proxy for the quality of the health system response to the crisis and to assess how this evolved and changed over time and space. In the end, the spatial distribution of these metrics can be extremely useful for the policy-maker to identify hotspots from the health system network, with the best and worst response to the emergency. Finally, studying the local policies and procedures implemented in those highlighted areas can help in the definition of best practices for future emergencies and pandemics.

2.3 Time window

In this study, we consider the period starting from the first day of COVID-19 official data release in Italy (February 24, 2020) until May 11, 2020, after 11 weeks, when the first national lockdown was lifted in Italy. We opted to consider this first wave window, as underreporting and overall data quality are at their lowest at the beginning of a pandemic for obvious reasons, and they tend to improve over time.

3 Proposed metrics

In the context of what ? have defined as ”pandemetrics”, correctly reporting and monitoring the evolution of pandemics is indeed a hard task, especially when pandemics spread suddenly and little is known about the origin, mode of transmission, triggering factors, etc. as was the case with the COVID-19 pandemic. Regarding the years in which this pandemic developed (2020–2022, mainly), there was enormous scientific production on many aspects of it, especially in finding ways to slow down its spread. One of the aims of this production also concerned errors in the detection and timing of the main elements that characterize the evolution of a pandemic. For example, ? provide a stable tool for monitoring current infection levels in situations involving compulsory registration of sensitive data and when cases are reported with delay to a central register, with this delay deferring an up-to-date view of the state of things. ? explored the powerful interaction of demography and current age-specific mortality for COVID-19, which is also something we used in our analysis. The analyses on evaluating errors in monitoring and reporting cases, recoveries and deaths are important as a criterion to decide which containment measures are appropriate (see, for example, ? to understand how a cross-country comparison about errors in collecting and analysing data can help in improving pandemic forecasting).

In our work, the metrics to be defined aim to provide an estimate for the under-reporting mortality bias that the official data has been subject to. However, there is no unique way to define such bias. Here, two different metrics with different interpretations for the policy-makers are proposed.

3.1 Additive bias b_A

Let D_{ij} be the actual deaths in province i and week j due to COVID-19, either directly or indirectly. Let $\hat{D}_{ij}$ be the officially reported total number of deaths in province i and week j, which exceeds the average deaths of the previous 5 years in the same week and province. Let Y_{ij} be the officially reported number of COVID-19-related deaths in province i and week j. Finally, let POP_i be the average population in province i along the considered period. The additive bias is built starting from the difference between the *actual mortality* D_{ij}/POP_i and the *official mortality* Y_{ij}/POP_i. This bias is defined as "additive" because it must be added to the official mortality to get the unbiased value:

$$\tilde{b}_{A,ij} = \frac{\hat{D}_{ij} - Y_{ij}}{POP_i}$$

Although, at least theoretically, we have that $D_{ij} \geq Y_{ij}$, $\tilde{b}_{A,ij}$ can take in practice negative values because $\hat{D}_{ij}$ can take values lower than Y_{ij} and even be negative by design. Moreover, even assuming that $\hat{D}_{ij}$ correctly estimates D_{ij}, errors and delays potentially can occur in the reporting of Y_{ij}. Negative values of $\tilde{b}_{A,ij}$ can be removed and/or treated as 0.

In terms of interpretation, $\tilde{b}_A$ defines the share of the population that died because of COVID-19 without being reported in the official data, so large values represent a negative scenario. Its trend over time and space (but not its magnitude) is a rough proxy for the part of the pandemic that was concealed and undetected by the public administration. Also, it can be roughly interpreted for an individual in a population as the risk of dying because of COVID-19 before having access to the healthcare system due to capacity limits: this is, of course, dangerous for the population, since access to the healthcare system reduces the consequent risk of dying because of the disease. However, it is a personal risk because it does not consider the severity of the outbreak at each point in time and space, but it only assesses the remaining part of the pandemic that was unfortunately missed by the competent authorities.

For example, individuals from two populations with the same $\widetilde{b}_A$ have the same individual risk of dying of COVID-19 outside of the healthcare system, but the system itself may have to deal with two very different situations in terms of the level of spread of the disease: hence, the same value of $\widetilde{b}_A$ can be considered a satisfactory result in one setting but utterly unsatisfactory in others.

Overall, the bias can be interpreted as the damage or impact of under-reporting on the population's well-being, but it is not an appropriate indicator of the quality of the healthcare system. In order to make this metric comparable with the metric b_M which will be presented in the next Section, scaling is applied to the original formula:

$$b_{A,ij} = 1000 \cdot \tilde{b}_{A,ij} = 1000 \cdot \frac{\hat{D}_{ij} - Y_{ij}}{POP_i}$$

3.2 Multiplicative bias b_M

The ratio between Y_{ij} and D_{ij} assesses the probability of a COVID-19-related death being officially reported. To transform it into a biased metric, the complementary probability

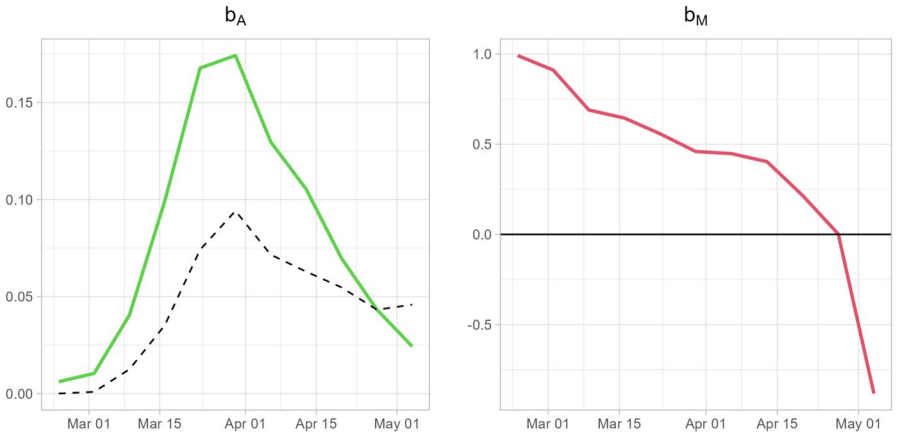

Fig. 1 Temporal trend of b_A and b_M for Italy. The dashed line in the left-hand plot represents the official COVID-19 reported mortality per 1000 people

of not being reported is considered instead and called $b_{M_{ij}}$. This is equivalent to the additive bias divided by the excess mortality rate:

$$b_{M,ij} = 1 - \frac{Y_{ij}}{\hat{D}_{ij}} = \frac{b_{A,ij}}{\hat{D}_{ij}/POP_{ij}}$$

This metric should also theoretically be bounded between 0 and 1, but for the same issues cited before, it can exceed these boundaries. Again, a larger bias indicates a bad situation. Regarding its interpretation, b_M measures the coverage of the pandemic by the healthcare system, thus it is an adequate indicator of its quality and a proxy for the efficacy of the crisis response.

The quality of the response is assumed to have been at least partially correlated to the evolving severity of the pandemic, e.g., a good emergency policy might nevertheless have performed poorly at the peak of the pandemic, while an ill-advised response policy may have exceeded expectations because of the negligibility of the crisis in the corresponding territory. As such, b_M may take into account this aspect by relating $b_{A,ij}$ to the severity of the epidemic, i.e. D_{ij}.

3.3 Comparison

To better understand the difference between the two proposed metrics, their trend can be compared over time for the whole Italian territory. Figure 1 shows the metrics obtained using the raw data sources from ISTAT and the official COVID-19 bulletins. It is clear to see that the two trends are starkly different from each other.

On the one hand, b_A has a bell-shaped trend, evidently correlated with the officially reported COVID-19 mortality (dashed line in the plot). As expected, this suggests that underreporting is a bigger issue during the worst period of a crisis, while the reported data becomes less biased as the situation is progressively better managed. However, b_A only evaluates the severity of the underreporting and does not capture the goodness of the reporting system despite the severity of the crisis.

Alternatively, b_M has a very different trend. A steady improvement is in fact what is expected from a metric evaluating the quality of the reported data, as it is sensible to assume that the data-collecting mechanism steadily improves over time because of the increase in resources and staff allocated to the task. Moreover, a metric capturing this aspect should not be influenced by the trend of the epidemic. Contrary to b_A, b_M appears to have both these desirable properties.

We can conclude that the analysis of both metrics can be useful for policymakers but that they must be carefully interpreted. b_A evaluates the magnitude of the underestimation in the data in absolute terms and thus it conveys information about how many people affected by the epidemic are not captured by the healthcare system: as such, a large value always suggests a more dangerous scenario. On the other hand, b_M assesses instead the share of the epidemic that is not captured and thus it highlights the quality of the reporting system itself rather than the gravity of the situation: therefore, a large value suggests that the system is failing to capture most of the epidemic casualties but does not directly says nothing about the size of the phenomenon, and thus its associated health risk.

4 Model

Our model aims to consider the spatial, temporal and spatio-temporal structures (as well as potential fixed effects) for $b_{A,ij}$ and $b_{M,ij}$ (Franco-Villoria et al. 2022). The objectives of such models are: (i) assessing the overall spatial distribution of these quality indicators on the Italian territory, as well as the temporal trend in the first wave of COVID-19; (ii) assessing the significance of the interaction structure in the model; (iii) estimating the importance of each of these components, including also the contribution of the potential fixed effects.

The chosen specification is a Latent Gaussian Model (LGM) on the logit transformation of the response with random effects. LGMs are additive regression models. In the Bayesian setting, LGMs have three components: a data model $p(\mathbf{y}|\mathbf{x}, \boldsymbol{\theta})$ (i.e. the likelihood), a process model $p(\mathbf{x}|\boldsymbol{\theta})$ (i.e. the density of a random Gaussian field $\mathbf{x}$, given the parameter $\boldsymbol{\theta}$, with $\mathbf{x}$ having Markov properties, i.e., for some $i \neq j$, x_i and x_j are independent on $\mathbf{x}_{-ij}$), and a parameter model $p(\theta)$ (Rue et al. 2009). In our case we have:

$$\mathbf{b} = [b_{1,1}, ..., b_{1,j}, ..., b_{1,J}, ..., b_{i,1}, ..., b_{i,J}, ..., b_{I,1}, ..., b_{I,J}]^T \tag{1}$$

$$\text{logit}(\mathbf{b})|\boldsymbol{\eta}, \sigma_\epsilon^2 \sim N_{I.J}(\boldsymbol{\eta}, \sigma_\epsilon^2 \mathbf{I}) \tag{2}$$

$$\boldsymbol{\eta} = \boldsymbol{\mu} + \sigma_u(\mathbf{I}_I \otimes \mathbf{1}_J)\mathbf{u} + \sigma_v(\mathbf{1}_I \otimes \mathbf{I}_J)\mathbf{v} + \sigma_w \mathbf{w} \tag{3}$$

$$\mathbf{u} = [u_1, ..., u_I]^T \sim N_I(\mathbf{0}, \boldsymbol{\Sigma}_u) \qquad \mathbf{v} = [v_1, ..., v_J]^T \sim N_J(\mathbf{0}, \boldsymbol{\Sigma}_v) \tag{4}$$

$$\mathbf{w} = [w_{1,1}, ..., w_{1,J}, ..., w_{I,1}, ..., w_{I,J}]^T \sim N_{I.J}(\mathbf{0}, \boldsymbol{\Sigma}_w = \boldsymbol{\Sigma}_u \otimes \boldsymbol{\Sigma}_v) \tag{5}$$

where $\mathbf{b}$ is one of the two metrics presented in Sect. 3, $\boldsymbol{\mu}$ represents the mean value at all times and spaces, u_i and v_j respectively represent the spatial and temporal main effects, and w_{ij} is the interaction random effect. This structure allows us to clearly distinguish between the three main sources of variation under investigation, i.e. a spatial association, a changing pattern over time, and an interaction between these two elements.

The parameters of this model are $\boldsymbol{\mu}, \boldsymbol{\beta}, \boldsymbol{\sigma}_u, \boldsymbol{\sigma}_v, \boldsymbol{\sigma}_w, \sigma_\epsilon$, while the covariance matrices $\boldsymbol{\Sigma}_u$ and $\boldsymbol{\Sigma}_v$ respectively define appropriate spatial and temporal models and are treated as fixed and known. The covariance matrix of the interaction term is defined as the Kronecker product of the covariance matrix of the two main effects, following the work of Knorr-Held (Knorr-Held 2000).

4.1 Temporal component specification

The existing literature on model components is wide and articulated, especially in pandemic monitoring and prediction. For example, ? considered data on confirmed and recovered cases and deaths, the growth rate and the trend of COVID-19 infections in Australia, Italy and the UK. ? employed machine learning models to understand the correlation between population movements and virus spread and to predict possible new outbreaks. ? relaxed the Normality assumption for modelling COVID-19 data, but focused exclusively on temporal effects.

Our model temporal component specification aims to retrieve the main temporal pattern behind the response and assess the impact of this factor. Hence, a Random Walk model of order 1 could be a simple but flexible choice, appropriate for the goal of the project:

$$\mathbf{v} \sim N_J(\mathbf{0}, \boldsymbol{\Sigma}_v = \boldsymbol{\Sigma}_{RW1}).$$

Note that this is not assumed to be the actual temporal model behind the data but it is exclusively used because of its convenience (see Franco-Villoria et al. (2022) for details on this).

4.2 Spatial component specification

One of the most popular models for lattice data is the ICAR model (Besag 1974; Besag et al. 1991), which consists of an improper Normal distribution for the spatial component, defined using a square symmetric matrix $\mathbf{M}$ with only non-negative entries and a null diagonal, and a corresponding diagonal matrix $\mathbf{D}$ where $d_{p,q} = \sum_q m_{p,q}$.

$$\mathbf{u} \sim N_I\left(\mathbf{0}; \boldsymbol{\Sigma}_u = (\mathbf{I_J} - \mathbf{D^{-1}M})^{-1}\mathbf{D}^{-1}\right).$$

The weights $m_{p,q}$ represent the neighbourhood structure between the spatial areas, and are usually based on the adjacency matrix, according to which regions that share a border are assigned $m_{p,q} = 1$ and 0 otherwise. This approach is appropriate when adjacency is the most reliable factor providing information about the potential association between regions. However, it can be argued that other variables should be considered instead, whenever they are more reasonable estimators of the actual connection between different areas. Adjacency matrices often consider geographical factors inappropriately: e.g. islands are isolated from the mainland, while geographical barriers such as mountains are ignored.

This is particularly relevant in the context of epidemiology, where the main factor causing spatial association is human mobility. This is considered the prevalent factor behind spatial association so traditional definitions of the weight matrix M, such as adjacency or distance-based matrices are unable to capture factors such as commuting routes, physical boundaries, air and highways traffic, metropolitan areas, etc.

Hence, it is proposed to replace the traditional adjacency matrix with smartphone location data, which estimates the average commuting of individuals between two

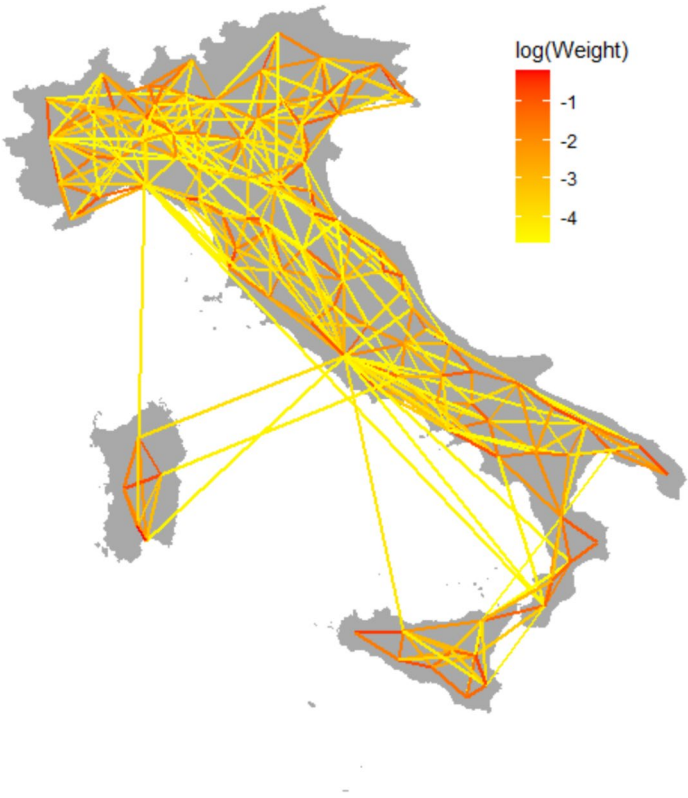

Fig. 2 Mobility-based weight matrix in log scale

provinces, no matter their actual geographical location. The spatial weight matrix used in this application was derived starting from data provided by Pepe et al. (2020), who built a daily time series of an origin–destination flow matrix processing smartphones' location data across Italy. The daily matrices from 18th January to 21st February 2020 were averaged to create a mean matrix representing the average mobility across provinces before the beginning of the COVID-19 outbreak. The matrices were built to have rows summing to 1, including the within-province mobility. To respect the conditions necessary for the CAR model, the diagonal entries were set to 0 and the remaining entries were normalised through a division by row sum. The origin–destination matrix was not symmetric as the flows were directional: hence, the inward and outward flow between two provinces were averaged to obtain an appropriate **M**.

The only disadvantage of this approach consists of the loss of the sparsity of the matrix. To obtain a matrix sparse enough for computation efficiency, it was chosen to consider only the last quintile, i.e. the highest 20% of the weights, assuming the impact of smaller associations irrelevant. The resulting weights can be visualized in Fig. 2: it is clear how this weight matrix is an improvement concerning the naive adjacency one, as it can recognize important industrial hubs and big cities, as well as air and sea traffic, giving an overall accurate representation of the spatial structure.

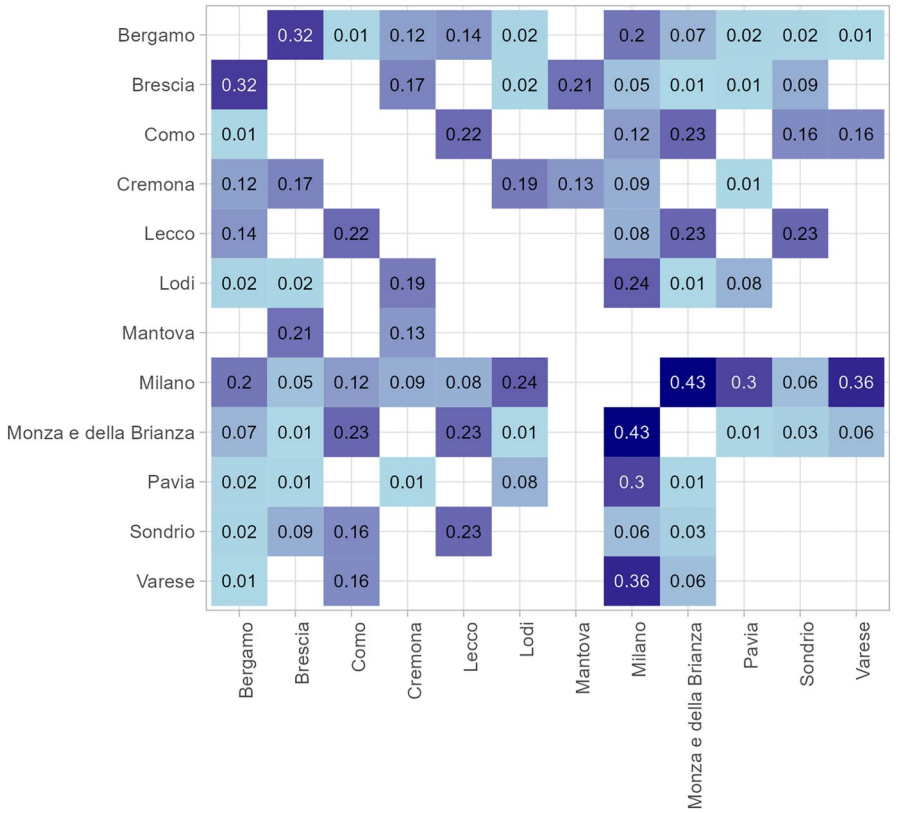

Fig. 3 Entries of M for the subset of rows and columns representing the provinces of Lombardy

A snippet of $\mathbf{M}$ for the provinces of Lombardy is represented in Fig. 3. The Lombardy region has been one of those most affected by the COVID-19 pandemic. Furthermore, however, there has been a very heterogeneous situation if we consider its provinces: there have been very affected provinces and others where the situation has been not so dramatic. Finally, the first hotspot occurred in a municipality in the South of the region. This snippet shows how the smartphone location data precisely reconstructs human mobility patterns not only at the national level but also locally.

4.3 Prior specification

In terms of prior specification, we take an approach similar to that of (Franco-Villoria et al. 2022), built on the innovative hierarchical variance decomposition method introduced by (Fuglstad et al. 2020), employed in the context of spatio-temporal epidemiological data. The method is based on a reparametrization of the variance parameters in terms of a single total variance and a set of proportions, defined through a decomposition tree. In this setting, it becomes easier to translate prior assumptions on the parameters into prior distributions and hyperparameters. Moreover, it gives more intuitive results in the posterior analysis, as the parameters are already set as adimensional proportional contributions, rather than variances or standard deviations.

As in (Franco-Villoria et al. 2022), here we choose to define the first split to separate the main from the interaction term, and secondly, the spatial and temporal effects are divided. This results in a new reparametrization of the three original variances $\sigma_u^2, \sigma_v^2, \sigma_w^2$ into a total residual variance V, the proportion ψ of this V given by the interaction term, and the proportion ϕ of main effects variance imputable to the spatial effect. The prior specification is then chosen on this new set of parameters.Specifically, the Integrated Nested Laplace Approximation (INLA) default prior on variance parameters is assumed on σ_ϵ^2, which is gamma with shape parameter 1 and rate parameter $5e^{-5}$. The INLA approach implements Laplace's method approximation to solve nasty integrals by Taylor expansion around the mode with a nested version of it to get the posterior of interest in a computationally feasible way. Details are, for example, in (?). Then, we chose a Uniform on ϕ, and a Penalized Complexity (PC) prior (Simpson et al. 2017) on ψ with base model $\phi_0 = 0$, and a PC prior on V with base model $V_0 = 0$:

$$
\begin{aligned}
&\sigma_u^2, \sigma_v^2, \sigma_w^2 \longrightarrow V, \phi, \psi \\
&\boldsymbol{\eta} = \sqrt{V}\left\{ \sqrt{1-\psi}\left[\sqrt{1-\phi}(\mathbf{I}_I \otimes \mathbf{1}_J)\mathbf{u} + \sqrt{\phi}(\mathbf{1}_I \otimes \mathbf{I}_J)\mathbf{v}\right] + \sqrt{\psi}\mathbf{w}\right\} \\
&V \sim PC_0(U, \alpha = 0.05) \\
&\phi \sim Unif(0,1) \\
&\psi \sim PC_0(\lambda_\psi = 1) \\
&\sigma_\epsilon^2 \sim \text{Inv-Gamma}(1, 5e^{-5}).
\end{aligned}
$$

5 Results

5.1 Fitting the models for b_A and b_M

The models were fitted in R using the INLA package (Rue et al. 2009). The same model was fitted for b_A and b_M, specifically using $U_A = U_M = 0.1$. The INLA R implementation is presented in the Appendix.

First of all, the different contributions of the random components to the total variability can be assessed considering the marginal posterior distribution of proportions of total variance V, see Fig. 4. For both b_A and b_M, it appears that the spatial component is the most relevant, followed by the interaction effect, while the temporal trend explains a much smaller share of the total variability. This shows how the multiplicative bias b_M might be more correlated with the specific policies put in place during the crisis period (and differently throughout Italy), than with the intrinsic characteristics of local areas, such as the population density, which may instead be more correlated with b_A.

Secondly, the posterior means of the random effects offer a summary of the different contributions to the response. Figure 5 shows the posterior mean of the spatial random effects over the provinces of Italy. Regarding b_A, provinces in Northern Italy experienced a larger share of underreported deaths concerning the overall population. However, the spatial distribution completely changes for b_M, as most of the Northern provinces show small values, while the highest effects are found in Southern and North-Eastern provinces. These figures display how the two metrics measure very different quantities, with b_M being much more consistent with the literature on the spatial distribution of health system quality indicators in Italy.

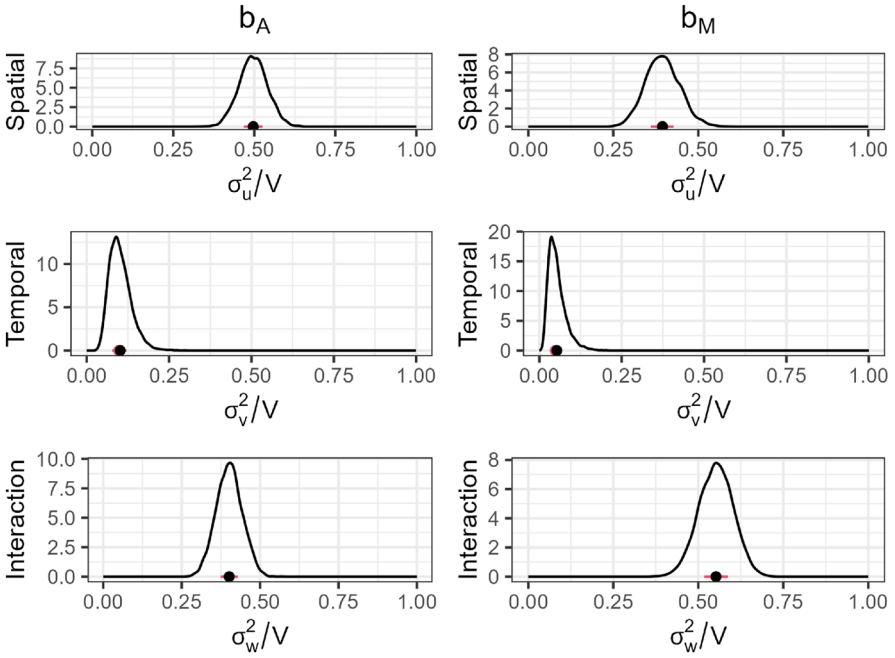

Fig. 4 Posterior densities of the variance parameters as proportions of the total residual variance V in the b_A and b_M models

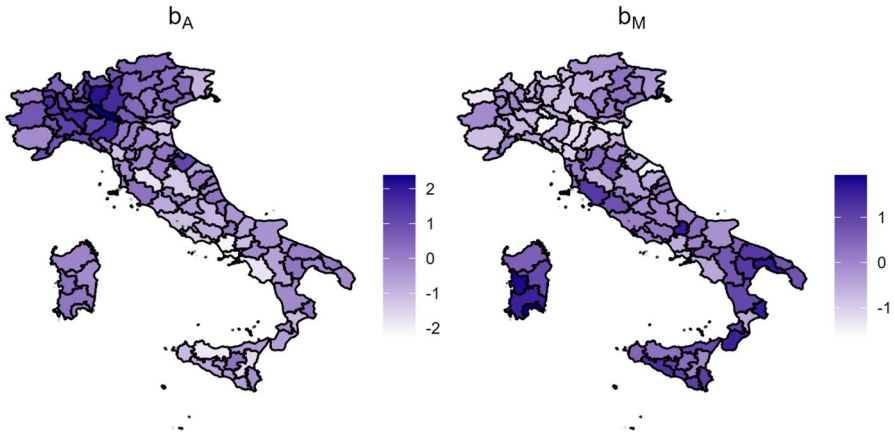

Fig. 5 Posterior mean of the spatial random effects **u** on b_A and b_M

Concerning the temporal pattern, the two metrics also show differences. Figure 6 shows that the indicator b_A performed the worst at the peak of the "official" pandemic evolution, plus a delay due to the fact that deaths are considered instead of cases. Hence, this confirms the assumption that b_A is related to the level of stress of the health system, rather than to the quality of its response to a certain amount of stress.

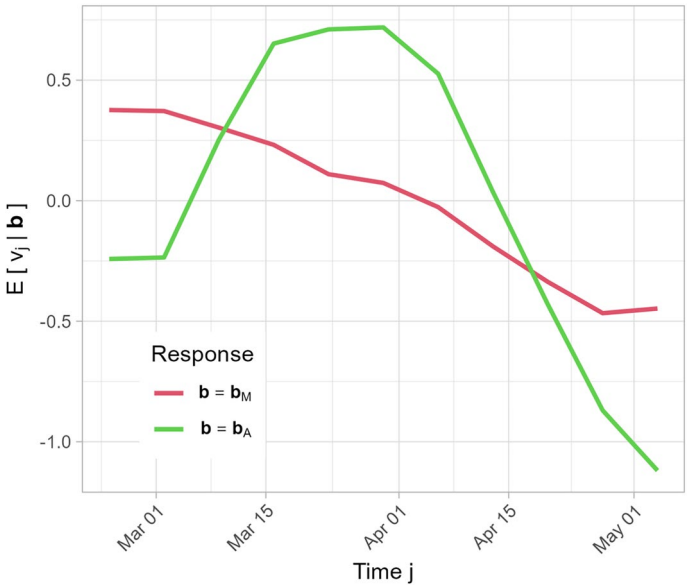

Fig. 6 Posterior mean of the temporal random effect for b_A and b_M

The results for b_M are again completely different. Figure 6, red curve, shows that its level is at its maximum at the beginning of the data reporting period and then, the quality level of the emergency response increases as the COVID-19 situation is taken more and more seriously and more effective policies and practices are put in place. Of course, this shows how the official data are unreliable, not only in magnitude but also in their trend, and should not be used rawly to evaluate the evolution of a pandemic, particularly at the beginning of the reporting period, as the quality of the official data tends to improve significantly, i.e. a decrease in the b_M indicator.

While the temporal and spatial effects can offer a general evaluation of the overall evolution, the interaction factor seems to play a vital role in the total variability, as shown in Fig. 4. Therefore, any conclusion on the behaviour of individual provinces must be drawn on the basis of the entire linear predictor.

In Table 3 and 4 we considered the posterior distribution of the fitted values reparametrized in the original scale of the b metrics, through a logistic transformation. In Table 3 the 10 provinces having reached the highest b_A value during the considered period are listed, whereas in Table 4 the 10 provinces experiencing the biggest change for b_M between the start and the end of the period are displayed.

5.2 Clustering for the b_M indicator

To classify provinces based on the b_M metric, i.e. a data quality indicator, clustering was applied to the posterior means of the fitted values from our model presented in equations 1 and following. Specifically, partitional clustering using a dynamic time warping distance was chosen as an appropriate partitioning method for time series (Sardá-Espinosa, 2017). The posterior mean of the fitted values for each province was computed to form the time

Table 3 Top 10 provinces with highest peak of b_A according to the model

Province	Peak Week	Maximum
Cremona	4	0.62
Bergamo	4	0.59
Lodi	3	0.36
Parma	4	0.33
Brescia	5	0.33
Piacenza	4	0.33
Lecco	5	0.25
Vercelli	5	0.24
Biella	5	0.21
Alessandria	5	0.21

Table 4 Top 10 provinces with respect to b_M biggest change between Week 1 and Week 11

Province	Difference
Trento	− 0.78
Imperia	− 0.75
Lucca	− 0.75
Chieti	− 0.74
Forli-Cesena	− 0.74
Torino	− 0.73
Massa Carrara	− 0.72
Bolzano	− 0.72
Pescara	− 0.70
Ravenna	− 0.70

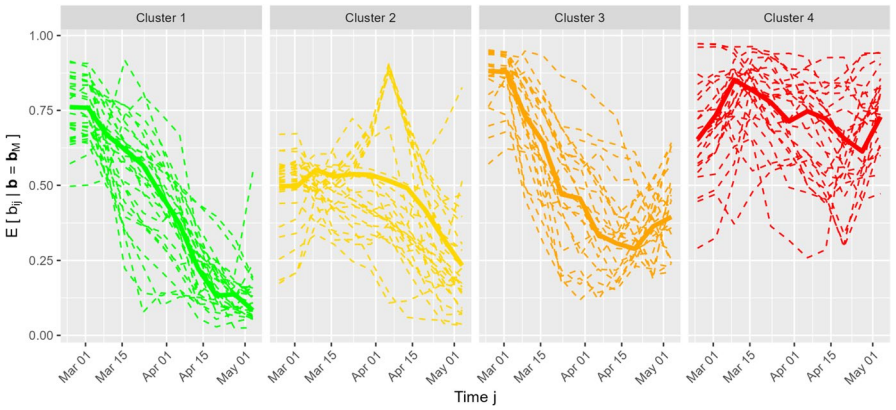

Fig. 7 Posterior mean of fitted values divided into 4 clusters with corresponding centroids in the provinces of Aosta, Rimini, Catanzaro, and Cosenza

series, to remove the residual noise present in the original raw data. Evaluating different performance metrics, the optimal cluster number was found to be 4. Figure 7 shows the 4 different classes and their centroid, ordered by best to worst overall performance.

The first cluster, coloured in green, includes all the provinces with a steady and relevant decrease over time. The red cluster identifies the worst-behaved class, as the index b_M, being already high at the beginning of the period, is mostly stable, and does not display significant signs of improvement during this wave.

The other two clusters represent intermediate behaviours and they are more ambiguous and heterogeneous. First, the orange cluster identifies provinces that performed better than the ones in the red cluster, but worse than the green provinces, because of a much lower rate of decrease. There are two main subtypes of trends in this cluster. Most of the provinces display an initial decrease, followed by a mild rise in the last few weeks, which could be a sign of a decline in alertness and a possible relaxation of some emergency measures initially put in place. Few provinces in the same cluster followed an inverse pattern, with an initial stable level of b_M followed by a quite steep improvement (e.g. Udine, Pordenone, Verona, Belluno, Vicenza, Padova, Venezia, Treviso), as if the severity of the pandemic was downplayed after the initial shock. The two groups have in common an insufficient degree of vigilance and promptness in dealing with a pandemic emergency, which became manifest as either a delay in putting in place effective response policies or a relaxation of the same way too early.

For the yellow group, the general trend, followed by the majority of the units in this cluster, is decreasing but more slowly than for the green provinces, hence indicating a worse performance in data reporting. However, looking more closely, we can identify a small subgroup of provinces in this cluster that displays an unusual concave U-shape trend: Sondrio, Verbano-Cusio-Ossola, Monza e della Brianza, Varese, Lecco, Como. The areas with this distinct behaviour are all located in North-West Italy and all fall into this class because they do not tend to reach extremes, either positive or negative, but are more or less stable around the 0.4 threshold. One hypothesis to explain this behaviour could be that these provinces might have experienced an actual delay in the spread of the pandemic concerning the general evolution, which might have led the peak to be shifted later than the other provinces, for about 5 or 6 weeks. This delay was probably responsible for keeping b_M low even in the worst moments for these provinces, as they had a temporal advantage consisting of all the experience and knowledge accumulated and shared by the other regional health system administrations over the first few weeks.

The map presented in Fig. 8 summarizes the cluster assignments over the Italian provinces, along with the location of the first COVID-19 hotspot in Italy in Codogno, Lodi. At first glance, it may seem counter-intuitive that areas closer to the first outbreak performed better than the ones far away. However, it may be that provinces in the North considered the emergency more seriously and sooner than the rest of Italy, as they perceived the risks earlier and may have put in place effective response protocols before the Southern areas.

6 Conclusions and future work

In this study, we considered two metrics to evaluate the quality of pandemic official data. This has been achieved using a measure of excess mortality registered during a pandemic period, concerning the average level in the previous years. While the metric b_A is a proxy for the underestimation of the pandemic severity, it has been shown how b_M can be used as a quality indicator of the health system in terms of monitoring the pandemic situation and reporting accurately the data. Both metrics were smoothed through a spatio-temporal model, which highlighted the importance of the spatial components in the variability

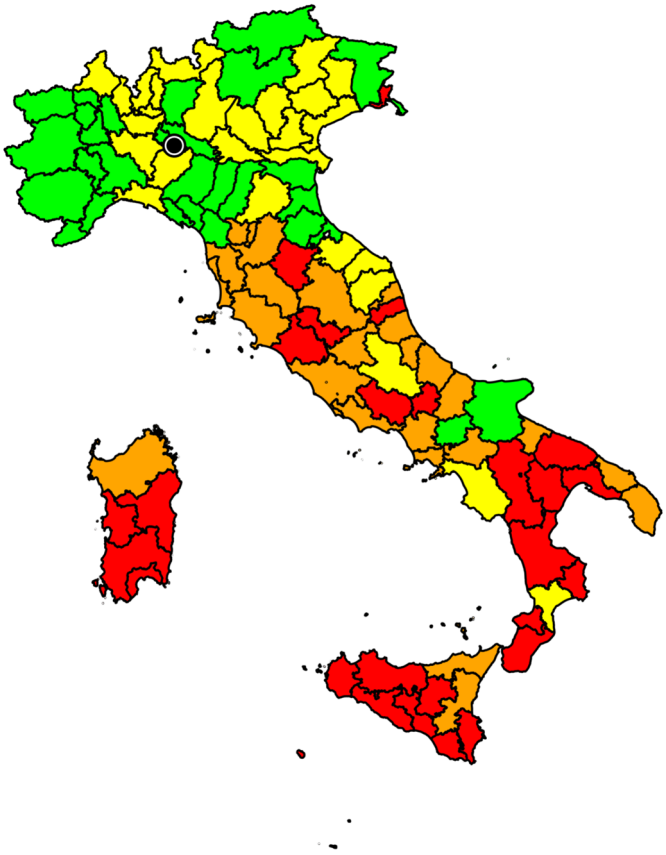

Fig. 8 Provinces by cluster (represented with four colours) and the city of Codogno (in the province of Lodi) marked by a black dot

of both responses. Results for b_M are consistent with the assumptions about data quality: in particular, the general temporal trend shows a steady decrease and it is in line with an improvement in the data quality in the first period of reporting.

Finally, the Italian provinces were grouped based on the b_M indicator fitted values, to classify them into clusters of different quality levels. This could be extremely useful for policymakers, since it offers an overview of the reporting performance of the different health systems and may suggest strategies employed locally, which may become best practices for future pandemics.

Future work will entail an implementation of the quality metrics to other pandemics and nations. We may be interested in comparing the overall data quality between nations facing the same pandemic, rather than within a single nation. For the Italian case, the time series could be extended to the following waves if official data about COVID-19 were available at the provincial level. Covariates related to the health system could be included in the models as fixed effects, to check whether they reduce the spatial variability and may explain the b_M metric.

In conclusion, the multiplicative bias is the main proposal of this work and it consists of a simple metric that could in theory be computed with a short delay, i.e. as soon

as registered deaths are aggregated by the governmental institutes for official statistics. Monitoring this indicator could help identify which areas display better results and performance in terms of the response to the pandemic, as well as areas where the severity of the situation is underestimated. This type of information could significantly accelerate the process of identification of effective and ineffective protocols and prevention measures, which consequently may save many lives down the line.

Appendix

INLA implementation

The model presented in Eqs. 1 and following has been fitted using INLA (Rue et al. 2009) and the code has been implemented following the one presented in the Supplementary Material of Franco-Villoria et al. (2022).

```
rm(list=ls())
library(INLA)
inla.setOption(num.threads = "1")
## load the data
data(mortality_italy)
## prepare the data stack for the INLA call
data_stack <- inla.stack(
  data=list(y=logit_b_M),
  A=list(1,S_matrix,T_matrix,I_matrix
  ),
  effects=list(data.frame(intercept=rep(1,N_week*N_prov)),
               list(spatial=1:N_prov),
               list(temporal=1:N_week),
               list(inter=1:(N_week*N_prov))
  ),
  remove.unused = F)
## prepare the joint prior
VP_prior_specification <- function(theta) {
#set the Penalized Complexity priors parameters
lambda_psi <- 1  #rate for the interaction proportion PC prior
upper_bound <-1 # upper bound for sqrt(V)
tail_prob <-0.05 # upper bound probability for sqrt(V)
# define each of the original parameters separately
t_0 <- theta[1] # log precision of residual effect
t_1 <- theta[2] # log precision of spatial effect
t_2 <- theta[3] # log precision of temporal effect
t_3 <- theta[4] # log precision of interaction effect
#Step A: write the mapping between theta and the new parameters
#residual precision
tau_epsilon <- expression(exp(t_0))
#total random effects precision
random_tau <- expression(exp(t_1+t_2+t_3)/(exp(t_1+t_2)+exp(t_2+t_3)+exp(t_1+t_3)))
#proportion of interaction term:
psi <- expression(exp(t_1+t_2)/(exp(t_1+t_2)+exp(t_2+t_3)+exp(t_1+t_3)))
#proportion of spatial term in main effect variance:
phi <- expression(exp(t_1)/(exp(t_1)+exp(t_2)))
#Step B: derive the Jacobian of the transformation
D_matrix <- matrix(data =
                     c(
```

```
                                eval(D(tau_epsilon,"t_0")),
                                eval(D(random_tau,"t_0")),
                                eval(D(psi,"t_0")),
                                eval(D(phi,"t_0")),
                                eval(D(tau_epsilon,"t_1")),
                                eval(D(random_tau,"t_1")),
                                eval(D(psi,"t_1")),
                                eval(D(phi,"t_1")),
                                eval(D(tau_epsilon,"t_2")),
                                eval(D(random_tau,"t_2")),
                                eval(D(psi,"t_2")),
                                eval(D(phi,"t_2")),
                                eval(D(tau_epsilon,"t_3")),
                                eval(D(random_tau,"t_3")),
                                eval(D(psi,"t_3")),
                                eval(D(phi,"t_3"))
                              ),nrow = 4)
#Step C: function for the density of a PC prior with base model 0 on a proportion parameter
log_PC_psi <- function(psi, lambda) {
    return(
      log(lambda*exp(-lambda*sqrt(psi))/(2*sqrt(psi)*(1-exp(-lambda))))
    )
  }
  #Step D: compute the log prior for theta through change of variable formula
log_prior <-
#Gamma default for tau_epsilon
 dgamma(eval(tau_epsilon),shape = 1,rate = 0.00005,log = T)+
        #PC prior for random_tau with base model 0
        inla.pc.dprec(eval(random_tau),u=upper_bound,alpha=tail_prob,log=T)+
        #Uniform distribution for phi
        dunif(eval(w_time),log=T)+
        #PC prior for psi with base model 0
        log_PC_psi(eval(psi),lambda_psi)+
        # absolute value of the determinant of the Jacobian
        log(abs(det(D_matrix)))
  return(log_prior)
}
Model_M <- inla(
  # Model formula
  y ~ -1+intercept +
    f(spatial, model="generic0", Cmatrix = S_precision,
      constr = T)+ # constrained to 0 sum
    f(temporal, model="generic0",Cmatrix = T_precision,
      constr = T)+   # constrained to 0 sum
    f(inter, model="generic0",Cmatrix = I_precision,

      constr = F,
      extraconstr = list(A=I_constraint,
                         e=matrix(0,N_week+N_prov))),
# Call to the data stack created
data=inla.stack.data(data_stack),
# Call to the joint prior just defined:
control.expert = list(jp=inla.jp.define(VP_prior_specification)),
control.compute=list(return.marginals.predictor=TRUE),
control.predictor=list(A=inla.stack.A(data_stack), compute=TRUE))
```

Acknowledgements The authors acknowledge financial support from the Italian Ministry of University and Research (MUR) under the Department of Excellence 2023-2027 grant agreement Centre of Excellence in Economics and Data Science (CEEDS). FN acknowledges the initiative "Dipartimento di Eccellenza 2023–2027", MUR, Italy, Dipartimento di Matematica, Politecnico di Milano.

Author contributions All authors contributed equally to the study conception and design. Material preparation, data collection and analysis were performed by Luisa Ferrari and Giancarlo Manzi. The first draft of the manuscript was written by Luisa Ferrari and Giancarlo Manzi and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding Open access funding provided by Università degli Studi di Milano within the CRUI-CARE Agreement. This work was supported by a financial support from the Italian Ministry of University and Research

(MUR) under the Department of Excellence 2023-2027 grant agreement Centre of Excellence in Economics and Data Science (CEEDS).

Declarations

Conflict of interest All the author declare they have no financial interests

References

Achilleos, S., Quattrocchi, A., Gabel, J., Heraclides, A., Kolokotroni, O., Constantinou, C., et al.: Excess all-cause mortality and Covid-19-related mortality: a temporal analysis in 22 countries, from January until August 2020. Int. J. Epidemiol. **51**(1), 35–53 (2022)

Beaney, T., Clarke, J.M., Jain, V., Golestaneh, A.K., Lyons, G., Salman, D., et al.: Excess mortality: the gold standard in measuring the impact of Covid-19 worldwide? J. Roy. Soc. Med. **113**(9), 329–334 (2020)

Besag, J.: Spatial interaction and the statistical analysis of lattice systems (with discussion). J R Stat Soc B. **36**, 192–225 (1974)

Besag, J., York, J., Mollié, A.: Bayesian image restoration, with two applications in spatial statistics. Ann. Inst. Stat. Math. **43**(1), 1–20 (1991)

Blangiardo, M., Cameletti, M., Pirani, M., Corsetti, G., Battaglini, M., Baio, G.: Estimating weekly excess mortality at sub-national level in Italy during the Covid-19 pandemic. PLoS ONE **15**(10), e0240286 (2020)

Castaldi, S., Maffeo, M., Rivieccio, B.A., Zignani, M., Manzi, G., Nicolussi, F., et al.: Monitoring emergency calls and social networks for Covid-19 surveillance to learn for the future: the outbreak experience of the Lombardia region in Italy. Acta Biomed. **91**(9S), 29–33 (2020)

Castaldi, S., Luconi, E., Rivieccio, B.A., Boracchi, P., Marano, G., Pariani, E., et al.: Are epidemiological estimates able to describe the ability of health systems to cope with Covid-19 epidemic? Risk Manag. **14**, 2221–2229 (2021)

Christie, A.: Blast from the Past: Pathogen Release from Thawing Permafrost could lead to Future Pandemics. Cambridge J. Sci. Pol. **2**(2), 1–8 (2021)

Colombo, R.M., Garavello, M., Marcellini, F., Rossi, E.: An age and space structured SIR model describing the Covid-19 pandemic. J. Math. Ind. **10**, 22 (2020)

Covid-19 Excess Mortality Collaborators. Estimating excess mortality due to the Covid-19 pandemic: a systematic analysis of Covid-19-related mortality, 2020-21. The Lancet. **399**, 1513–1536 (2022)

Dorrucci, M., Minelli, G., Boros, S., Manno, V., Prati, S., Battaglini, M., et al.: Excess mortality in Italy during the COVID-19 pandemic: assessing the differences between the first and the second wave, year 2020. Front. Public Health **16**, 927 (2021)

Ferrari, L., Gerardi, G., Manzi, G., Micheletti, A., Nicolussi, F., Biganzoli, E., et al.: Modeling Provincial Covid-19 Epidemic Data Using an Adjusted Time-Dependent SIRD Model. Int. J. Environ. Res. **18**, 6563 (2021)

Franco-Villoria, M., Ventrucci, M., Rue, H.: Variance partitioning in spatio-temporal disease mapping models. Stat. Methods Med. Res. **31**(8), 1566–1578 (2022)

Fuglstad, G.A., Hem, I.G., Knight, A., Rue, H., Riebler, A.: Intuitive joint priors for variance parameters. Bayesian Anal. **15**(4), 1109–1137 (2020)

Gibertoni, D., Adja, K.Y.C., Golinelli, D., Reno, C., Regazzi, L., Lenzi, J., et al.: Patterns of Covid-19 related excess mortality in the municipalities of Northern Italy during the first wave of the pandemic. Health Place **67**, 102508 (2021)

Hoiby, N.: Pandemics: past, present, future that is like choosing between cholera and plague. APMIS. **129**(7), 352–371 (2020)

Kantner, M., Koprucki, T.: Beyond just "flattening the curve'': optimal control of epidemics with purely non-pharmaceutical interventions. J. Math. Ind. **10**, 23 (2020)
Knorr-Held, L.: Bayesian modelling of inseparable space-time variation in disease risk. Stat. Med. **19**(17–18), 2555–2567 (2000)
Martínez-Córdoba, P., Benito, B., García-Sánchez, I.: Efficiency in the governance of the COVID-19 pandemic: political and territorial factors. Globalization Health. **17**, 113 (2021)
Maruotti, A., Jona Lasinio, G., Divino, F., Lovison, G., Ciccozzi, M., Farcomeni, A.: Estimating Covid-19-induced excess mortality in Lombardy, Italy. Aging clinical and experimental research. Aging Clin. Exp. Res. **34**(2), 475–479 (2022)
Michelozzi P, de'Donato F, Scortichini M, Pezzotti P, Stafoggia M, De Sario M, et al.: Temporal dynamics in total excess mortality and Covid-19 deaths in Italian cities. BMC Public Health. **20**, 1–8 (2020)
Modig, K., Ahlbom, A., Ebeling, M.: Excess mortality from Covid-19: weekly excess death rates by age and sex for Sweden and its most affected region. Eur. J. Public Health **31**, 17–22 (2021)
Moreno-Madriñan, M.J., Kontowicz, E.: Stocking density and homogeneity. Consider. Pandemic Potential. Zoonotic Dis. **3**, 85–92 (2023)
Msemburi, W., Karlinsky, A., Knutson, V., Aleshin-Guendel, S., Chatterji, S., Wakefield, J.: The WHO estimates of excess mortality associated with the Covid-19 pandemic. Nature **613**(7942), 130–137 (2023)
Pepe, E., Bajardi, P., Gauvin, L., Privitera, F., Lake, B., Cattuto, C., et al.: Covid-19 outbreak response, a dataset to assess mobility changes in Italy following national lockdown. Sci. Data. **7**(1), 1–7 (2020)
Rivieccio, B.A., Micheletti, A., Maffeo, M., Zignani, M., Comunian, A., Nicolussi, F., et al.: Covid-19, learning from the past: a wavelet and cross-correlation analysis of the epidemic dynamics looking to emergency calls and Twitter trends in Italian Lombardy region. PLoS ONE **16**(2), e0247854 (2021)
Rue, H., Martino, S., Chopin, N.: Approximate Bayesian inference for latent Gaussian models using integrated nested Laplace approximations (with discussion). J. R. Stat. Soc. B. **71**(2), 319–392 (2009)
Sardá-Espinosa, A.: Comparing time-series clustering algorithms in R using the dtwclust package. R package vignette, **12**, 41 (2017)
Shang, W., Wang, Y., Yuan, J., Guo, Z., Liu, J., Liu, M.: Global excess mortality during Covid-19 pandemic: a systematic review and meta-analysis. Vaccines **10**(10), 1702 (2022)
Simpson, D., Rue, H., Riebler, A., Martins, T.G., Sørbye, S.H.: Penalising model component complexity: a principled, practical approach to constructing priors. Stat. Sci. **32**, 1–28 (2017)
Vanella, P., Basellini, U., Lange, B.: Assessing excess mortality in times of pandemics based on principal component analysis of weekly mortality data. The case of Covid-19. Genus **77**, 16 (2021)
Wu, J., Tang, B., Bragazzi, N.L., Nah, K., McCarthy, Z.: Quantifying the role of social distancing, personal protection and case detection in mitigating Covid-19 outbreak in Ontario. J. Math. Ind. **10**, 15 (2020)