MOX-Report No. 27/2021

A look at the spatio-temporal mortality patterns in Italy during the COVID-19 pandemic through the lens of mortality densities

Scimone, R.; Menafoglio, A.; Sangalli, L.M.; Secchi, P.

A look at the spatio-temporal mortality patterns in Italy during the COVID-19 pandemic through the lens of mortality densities

Riccardo Scimone[1,2], Alessandra Menafoglio[1], Laura M. Sangalli[1]
and Piercesare Secchi[1,2]

[1]MOX - Department of Mathematics, Politecnico di Milano, Milano, Italy
[2]Center for Analysis, Decisions and Society, Human Technopole, Milano, Italy
riccardo.scimone@polimi.it
alessandra.menafoglio@polimi.it
laura.sangalli@polimi.it
piercesare.secchi@polimi.it

April 29, 2021

Abstract

With the tools and perspective of Object Oriented Spatial Statistics, we analyze official daily data on mortality from all causes in the provinces and municipalities of Italy for the year 2020, the first of the COVID-19 pandemic. By comparison with mortality data from 2011 to 2019, we assess the local impact of the pandemic as perturbation factor of the natural spatio-temporal death process. For each Italian province, mortality data are represented by the densities of time of death in the year. Densities are regarded as functional data belonging to the Bayes space B^2 where we use functional-on-functional linear models to predict the expected mortality in 2020, based on mortality in previous years, and we compare predictions with actual observations to assess the impact of the pandemic. Through spatial downscaling of the provincial data down to the municipality level, we identify spatial clusters characterized by mortality densities anomalous with respect to those predicted based on mortality data of the nearby areas. This analysis pipeline could be extended to indexes different from death counts, measured at a granular spatio-temporal scale, and used as proxies for quantifying the local disruption generated by the pandemic.

1 Introduction

The year 2020 will be remembered as the first of the COVID-19 pandemic. The worldwide diffusion of the virus caused massive disruptive effects in our societies, both directly – the tragic toll of lives prematurely terminated by the disease – and indirectly, as a consequence of the measures that states and communities adopted to fight contagion and contain the spread of the virus. In 2021, we are still in a transient state and we cannot yet assess the medium and long term effects of the pandemic on our societies, on their health and welfare systems, on their economies, on their education networks, on their working dynamics. One key issue perceived by all analysts is however the heterogeneous impact, both in time and in space, of the shock generated by the pandemic. Exploring and quantifying this heterogeneity is a necessary step to build a quantitative platform on which decision makers can devise precision policies for recovery, and envision the new normal which will follow the pandemic. We believe that a key index which could be used as a proxy for measuring the immediate impact of the pandemic on Italian local communities, as well as those in other developed societies, is the number of deaths from all causes. Death counts from all causes are high-quality data, recorded at a fine granular scale over time and space, and not affected by varying definitions, as it happens, for instance, for deaths for which COVID-19 is recorded as the cause. Moreover death counts from all causes integrate the direct and indirect effects of the pandemic shock, registering also the undesirable consequences due to the containment policies designed to fight the virus and the disruption at the local level of the health and welfare systems overwhelmed by the struggle against COVID-19.

In this work we explore the official mortality data recorded and made publicly available by the Italian Statistical Institute (ISTAT)[1], considering deaths from all causes, in different age classes and at different spatial scales, for the last ten years. In particular, we focus on the densities of the distribution of time of death in the year (hereinafter named *mortality densities* or *death densities*), observed for each Italian province and also at the finer spatial resolution of municipalities. The analysis of yearly death densities, rather than of death counts, allows for a natural data standardization and enables us to bring into focus the temporal variability of the phenomenon, capturing its different expression over space, due to the different impact generated by the pandemic shock. We believe that the analysis offers an interesting perspective on the spatio-temporal evolution of the COVID-19 pandemic in 2020 in Italy.

Figure 1 shows daily death counts from all causes, for people aged 70 or more, for each Italian province and over the last four years. Figure 2 displays the corresponding empirical densities of the distribution of time of death, the main object of our study. We first conduct a preliminary exploratory data analysis based on the Moran Index Moran (1950); Anselin (1995), an indicator of local association,

---

[1] https://www.istat.it/it/archivio/240401

to evaluate the spatial correlation of mortality in different years, and to capture its perturbation in 2020. In each province and municipality, we also explore the differences between yearly mortality distributions in different years by computing their Wasserstein distances Villani (2008). This preliminary analysis already highlights the heterogeneous spatio-temporal impact of COVID-19 on Italian mortality from all causes in 2020. We then move to a functional data analysis

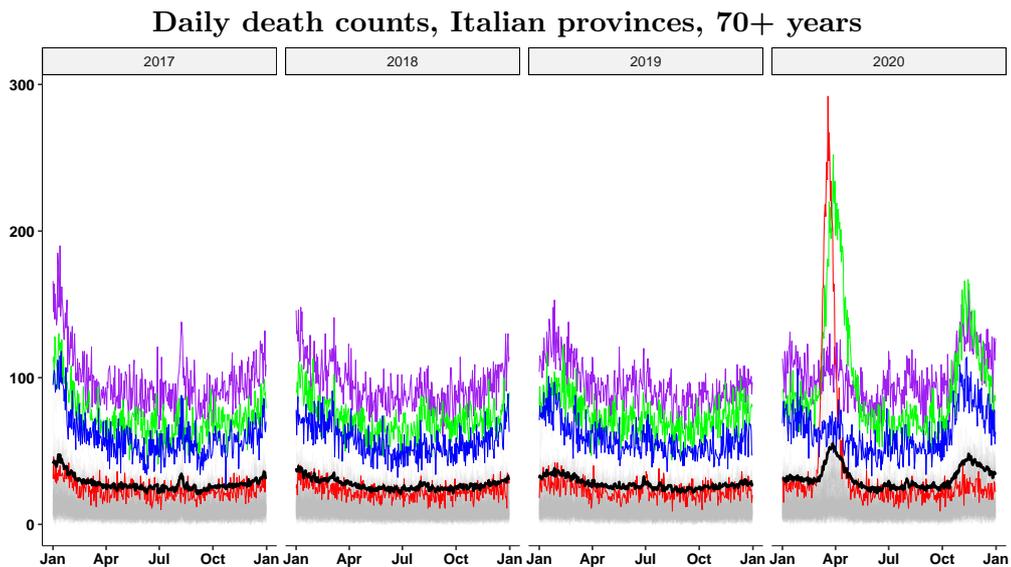**Daily death counts, Italian provinces, 70+ years**

Figure 1: Daily death counts during the last four year, for the Italian provinces. The plots refers to the elderly class (70+ years old). For each province, death counts along the year are plotted in light grey: curves are overlaid one on top of the other to visualize their variability. The black solid line is the weighted mean number of deaths, where each province has a weight proportional to its population. Death counts of four provinces are highlighted in color: Rome (purple), Milan (green), Naples (blue), and Bergamo (red).

framework, where we look at the mortality densities as functional compositional data Egozcue et al. (2006); van den Boogaart et al. (2010); Menafoglio et al. (2014) with spatial dependence. Indeed, densities represent the functional counterpart of compositional data Aitchison (1982, 1986); Filzmoser et al. (2018); Pawlowsky-Glahn and Egozcue (2001); Pawlowsky-Glahn et al. (2015), namely, vectors whose components represent parts (e.g., proportion or percentages) of a given total, thus conveying only relative information. To set the background, we recall that over the last decades considerable work has been done on extending classical statistical methods to the case of data embedded in functional Hilbert spaces (see, e.g., Ramsay and Silverman, 2005; Ferraty and Vieu, 2006; Horváth and Kokoszka, 2012; Wang et al., 2016, and references therein). Moreover, an im-

3

portant literature has focused on developing a consistent theoretical framework for geostatistical modeling and spatial prediction for functional data, developing the infinite-dimensional counterparts of techniques such as spatial variography and Kriging Giraldo et al. (2010); Nerini et al. (2010); Giraldo et al. (2011); Ruiz-Medina (2012); Menafoglio et al. (2013); Caballero et al. (2013); Ignaccolo et al. (2014); Menafoglio and Petris (2015); Menafoglio et al. (2016b,a). Recent works have specifically targeted the case of constrained functions and other object data, in a literature that we refer to as Object Oriented Spatial Statistics (O2S2) (Menafoglio and Secchi, 2017, and references therein). Functional compositional data are in fact a specific case of constrained functional data Mammen et al. (2001); Feng et al. (2014); Canale and Vantini (2016). Their analysis is not well formulated in the standard L^2 space typically used in functional data analysis. We hence embed the data in the so-called B^2 space (Egozcue et al., 2006; van den Boogaart et al., 2010; Pawlowsky-Glahn et al., 2014), which appropriately comply with the constrained nature of these data, extending to the functional setting the Aitchison's geometry established for compositional data Aitchison (1982); Pawlowsky-Glahn and Egozcue (2001). In this space, we formulate a linear model that allows us to decouple the mortality densities within provinces into a component that can be expected by looking at previous years, and a term that instead is unpredictable. We argue that, in 2020, this term precisely captures the impact of the pandemic shock. Through dimensionality reduction of this term, and a spatial analysis in B^2, we bring evidence and insights on the deep perturbation caused by the pandemic shock at different spatial scales. Spatial downscaling Kyriakidis (2004); Goovaerts (2008); Xiao et al. (2018) of the provincial data, down to the municipality level, finally allows us to identify hot-spots of local anomalies, associated with unexpected mortality densities in 2020, with respect to nearby provinces.

The paper is organized as follows. Section 2 presents the data and their preliminary exploration. In Section 3 we embed the data within the framework of functional compositional data analysis, and we analyse the spatio-temporal structure of the mortality densities over the last years. Section 4 focuses on 2020, and explores the perturbations in mortality patterns due to the pandemic shock, identifying local anomalies at the level of municipalities through spatial downscaling. Section 5 draws some conclusive remarks.

2 Data presentation and preliminary exploration

2.1 Official mortality data

ISTAT death data are collected with a very granular spatio-temporal scale. Moreover data are recorded per age class, with a 5 years segmentation, resulting in more than 20 classes. This subdivision is way too fine for our purposes, and in the following we consider only three age classes: the first class is composed by the youngest portion of the population, aged less than 49; the second class

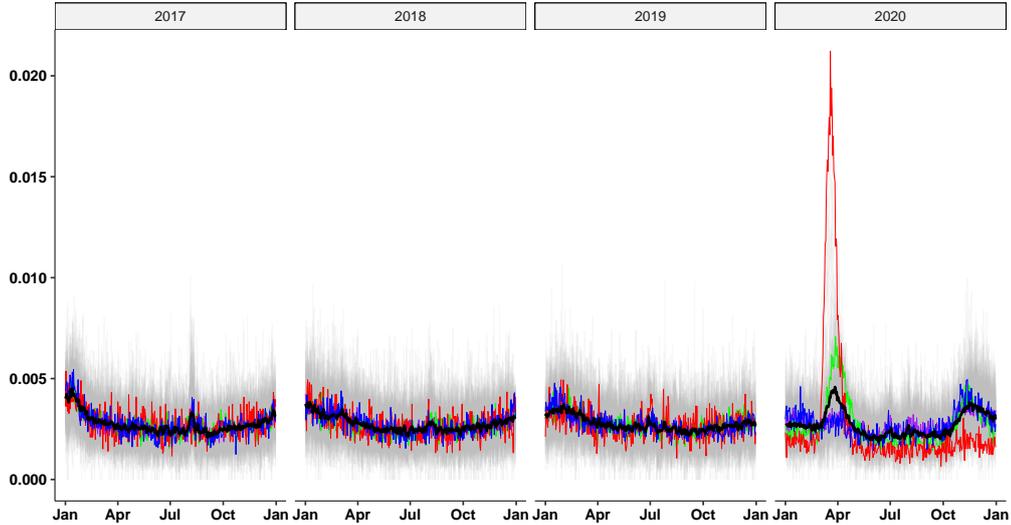**Empirical densities of daily mortality, provinces, 70+ years**

Figure 2: Empirical densities of daily mortality, for the elderly class, at the provincial scale. For each province, the empirical density of the daily mortality is plotted in light grey: densities are overlaid one on top of the other to visualize their variability. The black solid line is the weighted mean density, where the weight for each province has been set to be proportional to its population. Densities corresponding to four provinces are highlighted in color: Rome (purple), Milan (green), Naples (blue), and Bergamo (red).

groups middle aged people, from 50 to 69; while the last is for the elderly population, aged 70 or more. This partition is particularly convenient when exploring mortality during 2020, as the three population classes correspond to different risk classes. In particular, the first class (0-49 years) corresponds to people who are typically at low risk of death from COVID-19, the second (50-69 years) to people who have a medium risk, whilst the last (70+ years) to people who are at high risk. We carry out the analyses separately for the different age classes. The main text focuses on the elderly class, while insights for the other two age classes are given in the conclusive Discussion. In each age class, we denote death counts data by d_{iyt} where:

i identifies the province or the municipality, among the 107 provinces or the 7903 municipalities existing in Italy in 2020, depending on whether the analysis is carried out at the level of provinces or municipalities;

y refers to the year, from 2011 to 2020;

t refers to the day, within the year.

5

These data are generally characterized by a high quality, since the collection of information about daily deaths is an efficient and consolidated process, which is carried out homogeneously among the different administrative units.

Figure 1 shows the death counts, for the elderly class, at the province spatial scale. For each province, death counts along the year are plotted in light grey: curves are overlaid one on top of the other to visualize their variability. The black solid line is the weighted mean number of deaths, where each province has a weight proportional to its population. Death counts of four provinces are highlighted in color: Rome (purple), Milan (green), Naples (blue), and Bergamo (red). A visual inspection of the years from 2017 to 2019 highlights that the mortality in this age group has a typical seasonal behavior, with highest mortality in colder months and lower mortality in warmer months, except for heat waves during the summer, that may claim a big toll on elderly people, as clearly visible for instance in 2017. The year 2020, on the other hand, clearly presents a totally abnormal behaviour of the death process, in many provinces, due to the dramatic effect of the COVID-19 pandemic. The province of Milan (green) displays two sharp mortality peaks in correspondence of the two main waves of the pandemic in 2020, in March-April and November-December. The province of Naples (blue) appears instead relatively unaffected by the first wave, being spared by the pandemic during the spring, as many other parts of Italy, mainly thanks to the severe national lockdown that prevented the spread of the virus. The most tragic toll during the first wave is paid by the province of Bergamo (red), where the death counts surpass those of provinces with a population size several times larger; the pandemic hits Bergamo so strongly during the first wave, that the second wave is here almost imperceptible, presumably because the population at high risk has already significantly shrunk, and survivors have developed a natural immunity, due to the large circulation of the virus during the first wave.

As anticipated in the Introduction, we do not directly study the mortality counts d_{iyt}, but we are rather interested in the corresponding mortality densities. In particular, we start from

$$p_{iy}(t) = \frac{d_{iyt}}{\sum_t d_{iyt}}, \qquad t = 1, ..., 365.$$

The p_{iy} can be seen as the empirical discrete probability density associated with the absolutely continuous random variables T_{iy}, modeling the instant of death for a person in the considered age class, living in area i, and passing away during year y. Figure 2 shows these data for the elderly population at the provincial scale, with the same color choices as in Figure 1. Similar considerations can be drawn, with the difference that data are now normalized and no longer affected by the variability due to the different population sizes of the provinces. This natural alignment of amplitudes provides a better insight on the temporal structure of the phenomenon, if compared to the death counts in Figure 1. For example, the sharp increase in deaths during the heat wave of summer 2017 (the second

hottest summer over the last hundred years, closely following 2003), is much more visible. The same can be commented of the pandemic peaks, with the empirical mortality density in the province of Bergamo appearing particularly striking.

2.2 Spatial association analysis of the incidence of mortality

We enrich death data with the total population data (referring to January 1, 2020). Specifically, we denote by r_i the total population in the considered age class, in area i. These data enable us to compute the index $m_{iy} = \frac{\sum_t d_{iyt}}{r_i}$, which can be interpreted as a proxy of the death incidence per area and year, for the considered age class. Figure 3 shows the map of Italy, for the years from 2017 to 2020, where each Italian province is colored according to the index m_{iy}, computed for the elderly population. The same figure also shows the results of a spatial association analysis carried out on the m_{iy}, at both province and municipality scale, based on the Moran Index Moran (1950); Anselin (1995). This statistic is commonly used in the evaluation of spatial correlation (Getis and Ord, 1992; Anselin et al., 2005; Anselin, 1995). In particular, we use the Anselin's Local Moran I statistic Anselin (1995): for an area (say province) i and a given feature $X \in \mathbb{R}$, the corresponding Local Moran statistic is defined as

$$I_i = \frac{x_i - \overline{x}}{S_i^2} \sum_{r \neq i} w_{ir}(x_r - \overline{x}),$$

where $S_i^2 = \frac{\sum_{r \neq i}(x_r - \overline{x})^2}{n-1}$, n is the total number of areas in the Country, $\overline{x}$ is the sample mean for the Country of the feature X, and w_{ir} are weights codifying proximity between areas. We compute this statistics with $x_i := m_{iy}$, for the considered year y, and setting the proximity weights to $w_{ir} = 1$ if r is adjacent to i, and $w_{ir} = 0$ otherwise. If the index I_i is significantly greater than 0, then the feature X in area i has a value which deviates (positively or negatively) from the overall mean in the same direction as the average behavior of the neighbors of i: in this case, i is part of a *spatial cluster*. The case in which I_i is significantly lower than 0 corresponds to the situation of *spatial outliers*, since in this case the feature X in area i has a value which deviates from the overall mean in opposite direction as the average behavior of the neighbors of i. Significance can be assessed through tests and pseudo p-values as described in Anselin (1995), and is here evaluated via the R package SPDEP Bivand et al. (2013), using a significance level of 5%. This analysis highlights various spatial clusters, at both province and municipality scales, as shown respectively in the central row panels and bottom row panel of Figure 3; no spatial outlier is instead detected. Clusters characterized by a significantly high mortality incidence are highlighted in red, whilst clusters characterized by a significantly low mortality incidence are highlighted in blue. The analysis identifies spatial association patterns which are rather stable along all the years from 2011 up to 2019: this includes significantly
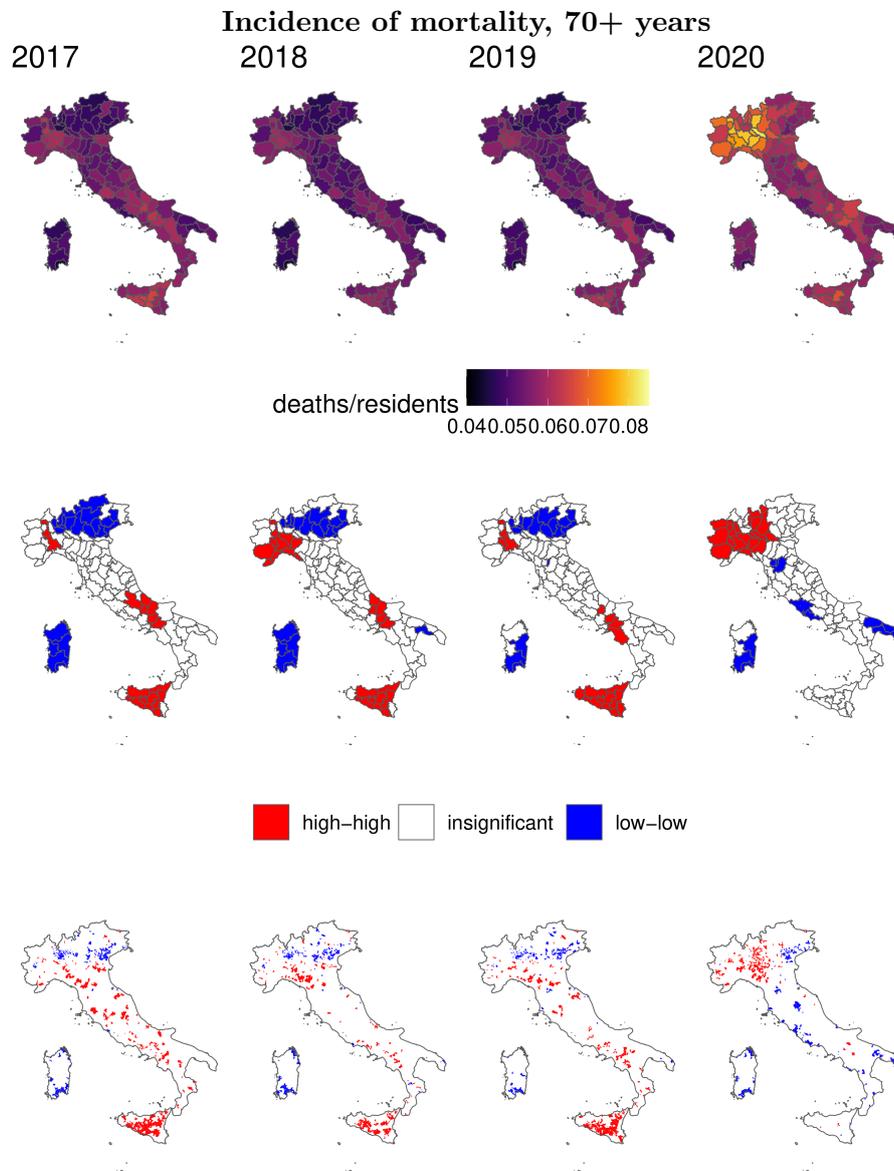
Figure 3: First row: heat-maps of the incidence of mortality m_{iy}, for the elderly class. Central and bottom rows: corresponding spatial cluster, identified using Anselin's Local Moran Index, at the province and municipality scale, respectively. The plots highlight a stationary spatial correlation structure of the mortality up to 2019, whilst 2020 displays completely different mortality patterns.

high mortality incidences in Sicily, and significantly low mortality incidence in the north-east of Italy and in Sardinia. On the other hand, the events of 2020 totally overturn such patterns: the large red cluster in the north-west of Italy corresponds to the part of the Country with higher incidence of mortality during 2020, as highlighted by the heat map in the first row of the same figure; these are the areas most interested by the first pandemic wave.

2.3 Exploration of the yearly variability of the empirical mortality distributions using the Wasserstein distance

Wasserstein distances of mortality with respect to previous years
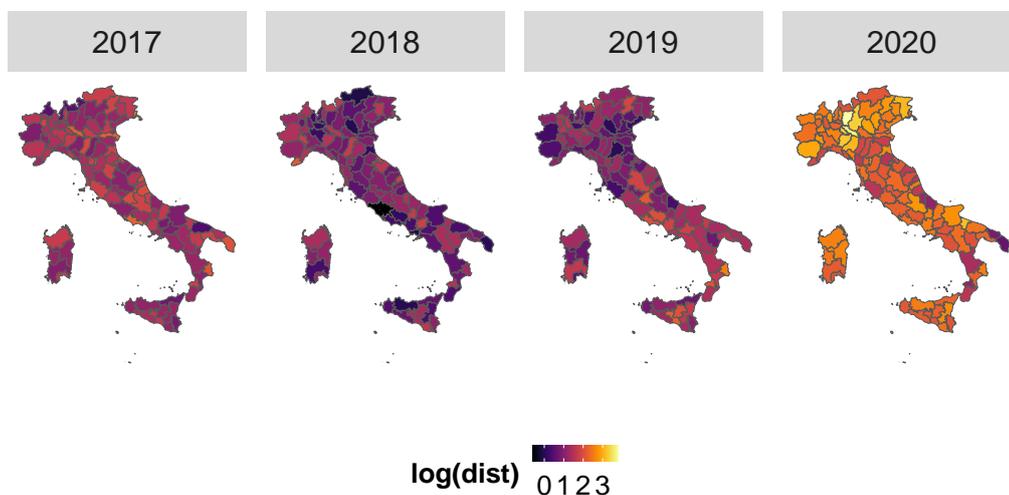

Figure 4: Heatmaps of quadratic Wasserstein distance between the mortality density of the considered year and the average mortality density over the 4 preceding years, in logarithmic scale, for the elderly class, at the provincial spatial scale.

We now explore the change over the years of the empirical mortality distributions, p_{iy}, using the quadratic Wasserstein distance w_2, a popular metric for density functions Ambrosio et al. (2003); Villani (2003, 2008). This metric, that is more generally defined for probability measures, has the interesting physical interpretation of minimal cost necessary to re-arrange a mass distribution into another Villani (2008). Distances are computed using the R package `transport` Schuhmacher et al. (2020).

Figure 4 shows the distances computed between the density of mortality in the considered year and the average density of mortality over the 4 preceding years, for the elderly population, at the province level. Specifically, the figure shows the distances $w_2(p_{iy}, \overline{p}_{iy})$, for years $y = 2017, ..., 2020$, where $\overline{p}_{iy} = \frac{1}{4} \sum_{v=y-4}^{y-1} p_{iv}$ is the mean empirical density taken over the 4 years preced-

ing y. Up to 2019, these distances are rather low, in most provinces, indicating no major variability along time for the mortality process of the elderly population. Among these years, 2017 shows slightly higher values, presumably due to the unusually high mortality during the winter months, and the summer peak of mortality caused by a very strong heat wave registered in that year. The picture for 2020 is instead completely different: high values of w_2 are observed across all Italy, with few exceptions; particularly high are the distances in the provinces of Bergamo, Lodi and Cremona, dramatically hit by the first pandemic wave.
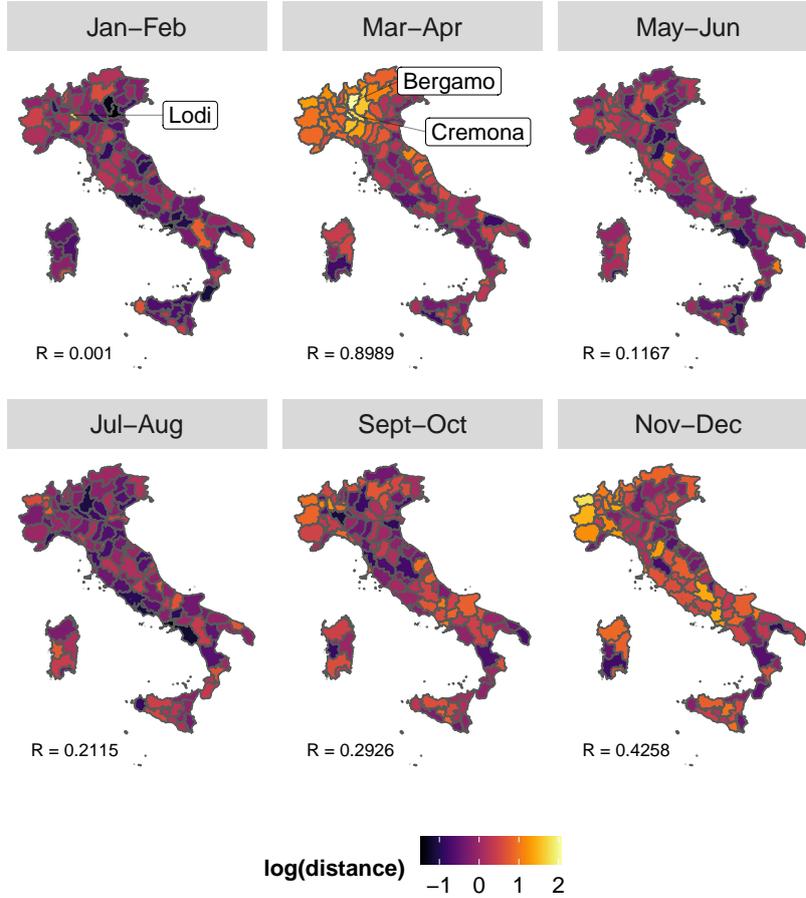

Figure 5: Heatmaps of $w_2(p_{iyT}, \bar{p}_{iT})$, with $y = 2020$, in logarithmic scale, for the elderly class, at province level. In each plot, the mortality density distributions are conditioned to be supported in the corresponding time window T, and the corresponding Wasserstein distance is computed. The heatmaps faithfully represents the spatio-temporal evolution of the pandemic. For each time window we also show the value of the correlation coefficient between the distances $w_2(p_{iyT}, \bar{p}_{iT})$ and the mortality indexes m_{iyT}.

We then use the Wasserstein distance, at a finer temporal resolution, to better explore the spatio-temporal pattern of the death anomalies in 2020, with respect to all previous years. We achieve this by conditioning the empirical distributions on sequential two-months time windows. Formally, given a time window T, we compute $w_2(p_{iyT}, \overline{p}_{iT})$, where $\overline{p}_i = \frac{1}{9}\sum_{y=2011}^{2019} p_{iy}$ is the mean empirical density taken on all previous years, while $p_{iyT}(t) = \frac{p_{iy}(t)}{\sum_{\tau \in T} p_{iy}(\tau)}$, $t \in T$ and $\overline{p}_{iT}(t) = \frac{\overline{p}_i(t)}{\sum_{\tau \in T} \overline{p}_i(\tau)}$, $t \in T$, are the corresponding conditional distributions on T. This distance, shown in Figure 5, neatly highlights the spatio-temporal diffusion of the epidemic in Italy. In January-February, only one province has an abnormal value of this distance: this is the province of Lodi, where the municipality of Codogno is located, the first hot-spot in Italy, where the first local case of COVID-19 is diagnosed. In March-April the epidemic spreads across most of Northern Italy, with a particular tragic situation in the provinces of Lodi and Bergamo. The central region Marche is also hit hard in this period. Nevertheless, the epidemic still has a local character, with many regions of Italy left untouched. In May-June the epidemic has regressed almost everywhere, as an effect of the severe lockdown. The months of July and August do not present noteworthy anomalies, whilst September and October mark a first upswing of the epidemic. The second wave arrives in the months of November and December, and it involves most of Italy, though some provinces are less affected. Among the less affected areas, there are many provinces, especially in the North-East of Italy and the region of Marche and Emilia Romagna, that were instead hardly hit by the first epidemic wave. To further corroborate this interpretation, Figure 5 also displays, for each window T, the sample correlation coefficient between the distances $w_2(p_{iyT}, \overline{p}_{iT})$ and the corresponding mortality index in the window, that is $m_{iyT} := \frac{\sum_{t \in T} d_{iyt}}{r_i}$, showing them in the same figure. This correlation coefficient is low during periods where the pandemic has low intensity; during the same periods the mortality densities do not significantly differ from previous years, as indicated by low Wasserstein distances. Instead, the correlation is high during the pandemic waves, especially during first wave, highlighting that the different temporal mortality during these periods is strongly associated with a higher incidence of mortality.

3 Modeling death data as functional compositions

3.1 The appropriate functional space for functional compositions

We now embed the mortality densities in a functional data analysis framework. It is in fact natural to think of the p_{iy}'s presented in Section 2 as a discretizations, on the grid of days, of the family $f_{iy} : \Theta \to \mathbb{R}^+$ of the continuous density functions of the random variables T_{iy}, where Θ stands indeed for the continuous time domain of the year. The crucial point here is the identification of

an appropriate Hilbert space where these functional data should be embedded, smoothly estimated and analyzed. In standard spatial data analysis this space is usually chosen to be $L^2(\Theta)$, the space of square-integrable functions (see, e.g., Ramsay and Silverman, 2005). However, it has been shown that L^2 is not a proper Hilbert embedding for the analysis of density functions, since its geometry is not coherent with the constrained nature of densities Egozcue et al. (2006); van den Boogaart et al. (2010); Delicado (2011); Pawlowsky-Glahn et al. (2014); Hron et al. (2016). On the other hand, almost all functional data analysis techniques, formerly developed for the L^2 case, can actually be extended, with some effort, to any separable Hilbert space Horváth and Kokoszka (2012). In Section 2, we implicitly considered the real Wasserstein space W^2, defined as the set $W^2 := \{\mu \text{ prob. meas. on } \mathbb{R} \text{ s.t. } \int_{-\infty}^{+\infty} x^2 \, d\mu(x) < \infty\}$, equipped with the metric w_2. Although W^2 has been proved to be a very regular metric space (it is for instance a complete geodesic metric space, see e.g., Villani (2008)), it is not a vector space and it lacks an inner product. This makes it hard to perform typical functional data analyses which rely on these elements, as for instance functional principal component analysis and regression by linear models.

To understand what space is more appropriate, it is convenient to think at the more classical setting of compositional data analysis Aitchison (1982), the area of statistics that deals with data that are proportions, whose natural embedding has been proved to be the Aitchison simplex, rather than the usual Euclidean space Pearson (1897); Aitchison (1986); Barcelo-Vidal et al. (2001); Filzmoser et al. (2018). The infinite dimensional analogue of the Aitchison simplex is the Bayes Space $B^2(\Theta)$. This is the set (of equivalence classes) of functions

$$B^2(\Theta) = \{f : \Theta \to \mathbb{R}^+ \, s.t. \, f > 0, \, log(f) \in L^2(\Theta)\}$$

where the equivalence relation in $B^2(\Theta)$ is defined among *proportional* functions, i.e., $f =_{B^2} g$ if $f = \alpha g$ for a constant α in $\mathbb{R}$. In $B^2(\Theta)$, the operations of sum and external product are defined as

$$(f + g)(t) =_{B^2} f(t)g(t) \quad \text{and} \quad (\alpha \cdot f)(t) =_{B^2} f(t)^\alpha \tag{1}$$

for $t \in \Theta$, f, g in $B^2(\Theta)$, and α in $\mathbb{R}$. The space $B^2(\Theta)$ defines a Hilbert space when equipped with the generalization of the Aitchison inner product, defined as

$$\langle f, g \rangle_{B^2} = \frac{1}{2|\Theta|} \int_\Theta \int_\Theta \log \frac{f(t)}{f(s)} \log \frac{g(t)}{g(s)} dt ds \tag{2}$$

for f, g in $B^2(\Theta)$, $|\Theta|$ being the length of the interval Θ. We refer the reader to Egozcue et al. (2006); van den Boogaart et al. (2010); Pawlowsky-Glahn et al. (2014) for a complete description of $B^2(\Theta)$ and its Hilbert structure. From now on, we always look at the f_{iy}'s as realizations of a spatio-temporal functional random process taking values in $B^2(\Theta)$, and we develop our analyses consistently.
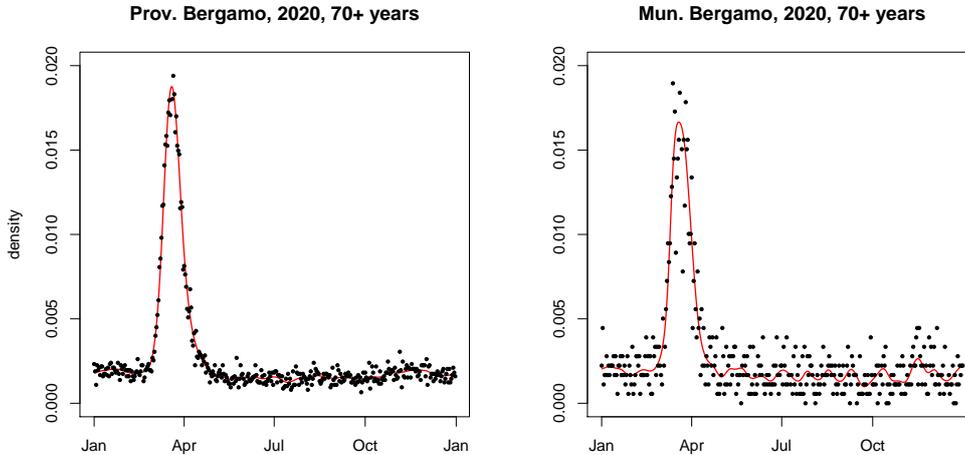
Figure 6: Empirical mortality distributions p_{iyt} (black dots) and corresponding smooth estimate of f_{iy} (red lines), for the elderly class, in 2020, for the province (left) and the municipality (right) of Bergamo.

3.2 Estimating smooth mortality densities

The estimation of the continuous densities f_{iy} from their discretization p_{iy} cannot be tackled with classical smoothing techniques, since the constrained nature of density data must be taken into account during the smoothing phase (Machalová et al., 2015). Several splines expansions have been proposed for the B^2 case, like the constrained $B-$splines described in Machalová et al. (2015), or the compositional spline basis introduced in Machalová et al. (2020). Here we consider the procedure described in Machalová et al. (2015), and implemented in the R package `robCompositions` Templ et al. (2011), which is chosen for its simplicity and solid implementation. We use cubic splines, with a penalization on the second derivative of the estimate, and evenly spaced knots, one per week (52 internal knots in total). Figure 6 shows an example of this smoothing, for the province and for the municipality of Bergamo, respectively in the left and in the right panel. Death processes at the municipality scale are much noisier than those at the province scale. This leads to rougher estimates at the municipality level, as exemplified by Figure 6. In particular, we favor here the ability to well capture the sharp peak in the mortality, at the cost of obtaining a somewhat wiggly estimate. However, this less accurate reconstruction of the functional composition at the municipality level is something that we take into account in subsequent analyses. The smooth estimates obtained at the province level appear instead very good, and capable to well describe the data and their sharp peaks, without producing any artifacts. Figure 7 shows the estimates of f_{iy} for the 107 different

provinces, for the elderly population. The usual pattern of mortality is visible till 2019, with higher mortality during the coldest months and during summer heat waves, that may hit only some regions. The functional process is very different in 2020, and dominated by the two pandemic waves, clearly captured by the estimated densities: a first wave in March-April, that hits some regions to a dramatic extent, and a second less dramatic wave in November-December, which however involves all provinces.

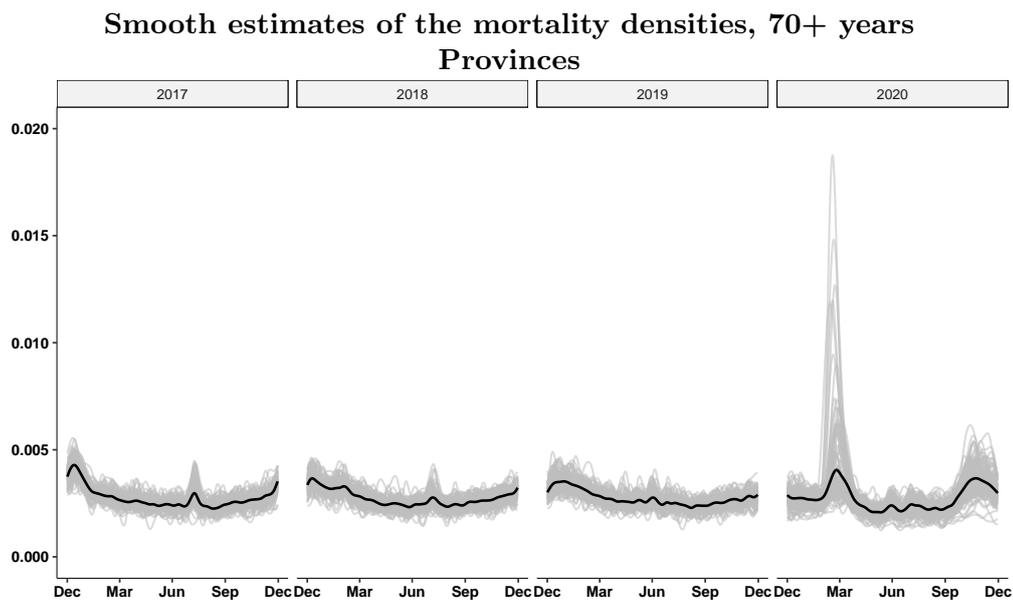**Smooth estimates of the mortality densities, 70+ years**
Provinces

Figure 7: Smooth estimates of the mortality densities over the 107 Italian provinces, for the elderly population. The usual pattern of mortality is visible till 2019, with higher mortality during the coldest months and during the summer heat waves, that may hit only some regions. The functional process is completely different in 2020, with the two pandemic waves clearly captured by the estimated densities. The black thick lines represent the mean density, computed in B^2, with weights proportional to the population in each area.

3.3 Modeling the temporal variability of mortality densities

Starting from the smooth functional samples f_{iy}, we then construct a functional linear model in the B^2 space Talská et al. (2018), in order to predict the mortality density in each province in one year, using as predictors the corresponding mortality densities in previous years in the same province, in an attempt to account for the variability associated with recurrent seasonal effects. For each province i, we compute $\overline{f}_{iy} = \frac{1}{4} \sum_{r=y-4}^{y-1} f_{ir}$, the B^2 mean of the observed densities in the four years preceding year y, where summation and multiplication by

14

a scalar should be intended in the B^2 sense of equation (1). We then formulate the following functional linear model in $B^2(\Theta)$

$$f_{iy}(t) = \beta_{0y}(t) + \langle \beta_y(\cdot, t), \overline{f}_{iy} \rangle_{B^2} + \epsilon_{iy}(t), \quad i = 1, ...107, \quad t \in \Theta \quad (3)$$

where the functional parameters $\beta_{0y}(t), \beta_y(s, t)$ and residual terms $\epsilon_{iy}(t)$ are defined in the B^2 sense. In model (3), both β_{0y} and ϵ_{iy} are probability densities in $B^2(\Theta)$, while β_y is the kernel of a $B^2(\Theta)$ linear operator, so that also $\langle \beta_y(\cdot, t), \overline{f}_{iy} \rangle_{B^2}$ is a probability density in $B^2(\Theta)$. Model (3) is an extension to the $B^2(\Theta)$ space of the classical function-on-function regression models formulated for the L^2 case (e.g., Ramsay and Silverman, 2005). The functional parameters $\beta_{0y}(t)$ and $\beta_y(s, t)$ can be estimated via ordinary or penalized least squares, using similar techniques, such as basis approximations. In the present work, we use ordinary least squares, using the same constrained splines basis adopted in the smoothing phase; see Section 3.2. Model fitting is assessed via the functional regression routines implemented in the R package `fda.usc` Febrero-Bande and Oviedo de la Fuente (2012). We obtain a family of estimated linear models, each one defined by the estimates $\hat{\beta}_{0y}, \hat{\beta}_y, y = 2015, ...2020$, where y starts from 2015 since we need four previous years to fit the corresponding model. It is interesting to look at the densities δ_{iy} defined by (sum and differences are again taken in the B^2 sense)

$$\delta_{iy}(t) = f_{iy}(t) - \hat{f}_{iy}(t) \quad (4)$$

where

$$\hat{f}_{iy}(t) := \beta_{0y-1}(t) + \langle \beta_{y-1}(\cdot, t), \overline{f}_{iy} \rangle_{B^2}. \quad (5)$$

The δ_{iy} represent the prediction errors of model (3) calibrated at year $y-1$, when this is used to produce forecasts for the following year y. As such, unlike the residuals of model (3), the prediction errors $\delta_{iy}(t)$ may have a non-null sample mean. We then look at the δ_{iy} as a proxy of what happens in year y, that cannot be predicted from the previous years $y - 1, ..., y - 4$.

Figure 8 shows the δ_{iy} for different years, at the provincial scale; these prediction errors should be compared with the corresponding smooth densities in Figure 7. We remark that the zero of the B^2 space is the uniform distribution, so that good predictions coincide with almost uniform errors. As already noted, the linear model (3) can be seen as a way to filter out seasonal effects, which can be predicted on the basis of previous years, so that the prediction errors show to what extent a given year has unusual mortality patterns. If we look at the prediction errors in 2017, it is clear that what happens during summer is to a large extent unpredictable. This is reasonable, *a posteriori*, because 2017 is characterized by the second hottest summer in the last century, closely following 2003, with clear effects on the elderly population. Regarding 2020, the characteristic of the death process in March-April and in November-December appears completely unpredictable from past years, leading us to conclude that the effect of the COVID-19 pandemic is by far the most important factor in determining the mortality during this year.

15

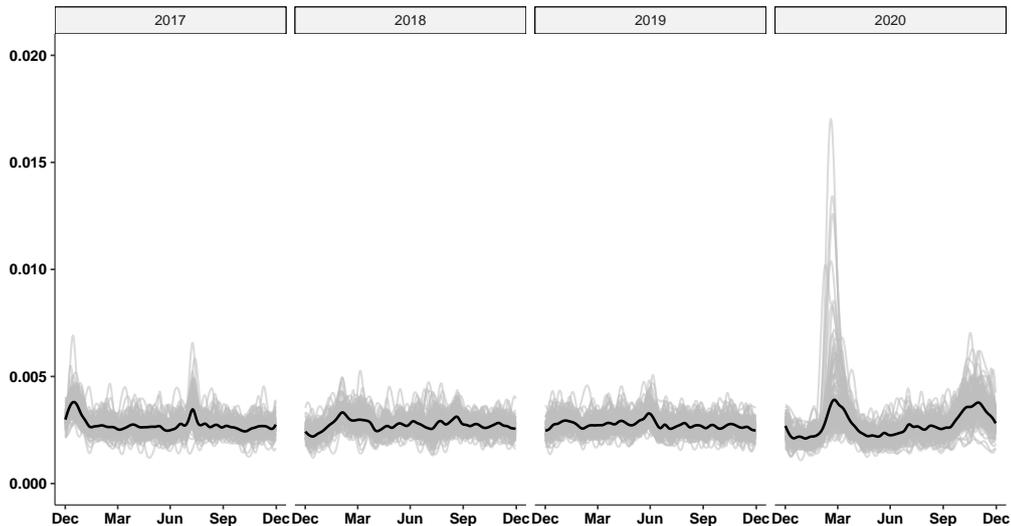**Prediction errors, provinces, 70+ years**

Figure 8: Prediction errors δ_{iy} for the mortality densities of the elderly population generated, in each province i and for a given year y, by the linear model (3) fitted for the year $y-1$ and applied to the average of the mortality densities of the four years preceding y, in the same province. Visual inspection of 2017 identifies the sharp increases in mortality during winter and summer which could not be predicted based on the mortality densities of the four previous years. Note that the peak in mortality observed also during the summer of 2018 is correctly predicted by the linear model (3) fitted for 2017, while mortality in winter is slightly overestimated. It is visually evident that the dramatic patterns of the 2020 mortality densities cannot be predicted by the linear model (3) fitted on the mortality densities of the previous years.

Figure 9 shows the maps of the B^2 norm of the prediction errors for the same years. The map for 2020 highlights the spatial diffusion of the pandemic, clearly identifying those provinces where the death density process is highly perturbed with respect to previous years. Figure 10 shows the results of a functional K-means clustering Abraham et al. (2003) on the prediction errors, in the B^2 space, with $K = 3$ centers, for the year 2020. It is interesting to note that the norms of the prediction errors capture the effect of the pandemics in the different areas, and that the clustering almost perfectly divides the provinces among: provinces very hardly hit by the first wave but left mostly untouched by the second wave (red cluster), provinces hit to a less dramatic extent by the first wave but significantly affected also by the second wave (green cluster), and provinces which only experience the second wave (blue cluster).

Prediction error norms, provinces, 70+ years

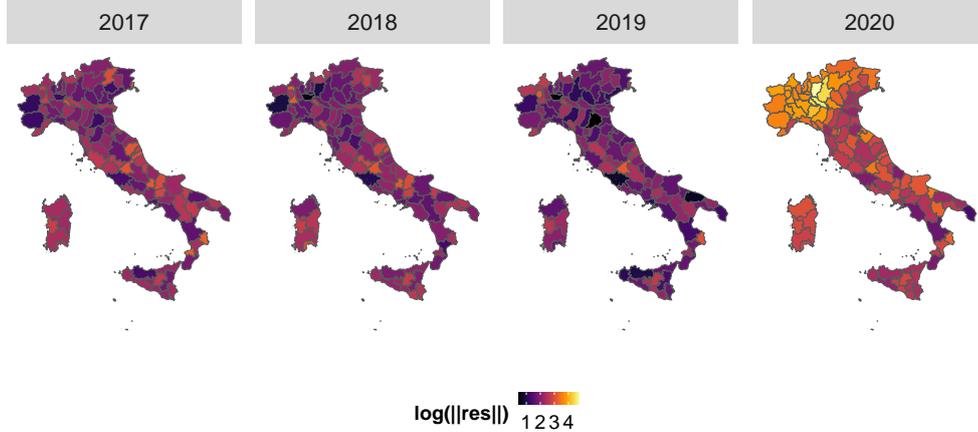| 2017 | 2018 | 2019 | 2020 |

log(||res||) 1 2 3 4

Figure 9: Heatmaps of the B^2 norm of the prediction errors in Figure 8, in logarithmic scale, for the elderly class. The map for 2020 highlights the spatial characteristic of the pandemic diffusion, clearly identifying those provinces where the death density process is highly perturbed with respect to previous years.

3.4 Analyzing the spatial correlation of the prediction errors

We now proceed to analyze the spatial dependence structure of the prediction errors. We can model the δ_{iy}'s as realizations of a spatial random process $\{X_s\}_{s \in D}$, where D is the spatial domain identified by the Country, and X_s, for $s \in D$, takes values in $B^2(\Theta)$. We assume square integrability, that is $\mathbb{E}[\|X_s\|_{B^2}^2] < \infty$ for all $s \in D$. We can thus define the expectation m_s of the process as $m_s := \mathbb{E}[X_s]$, and the trace-covariogram Menafoglio et al. (2013); Menafoglio and Petris (2015); Menafoglio et al. (2016a) of the process as

$$C : D \times D \to \mathbb{R}$$

where, for all $(s_k, s_j) \in D \times D$,

$$C(s_k, s_j) := \mathbb{E}[\langle X_{s_k} - m_{s_k}, X_{s_j} - m_{s_j} \rangle_{B^2}]$$

that is the infinite-dimensional analogue of the covariogram of a real-valued process Cressie (1993). If we assume second-order stationarity and isotropy, the trace-covariogram reduces to a function of a real non-negative variable, i.e.

$$C(h) = \mathbb{E}[\langle X_{s_k} - m_{s_k}, X_{s_j} - m_{s_j} \rangle_{B^2}], \text{ for } (s_k, s_j) \in D(h) \tag{6}$$

where $D(h) := \{(s_k, s_j) \in D \times D \,|\, d(s_k, s_j) = h\}$ and d is the Euclidean distance. The definition of the trace-covariogram in equation (6) allows for the definition

17
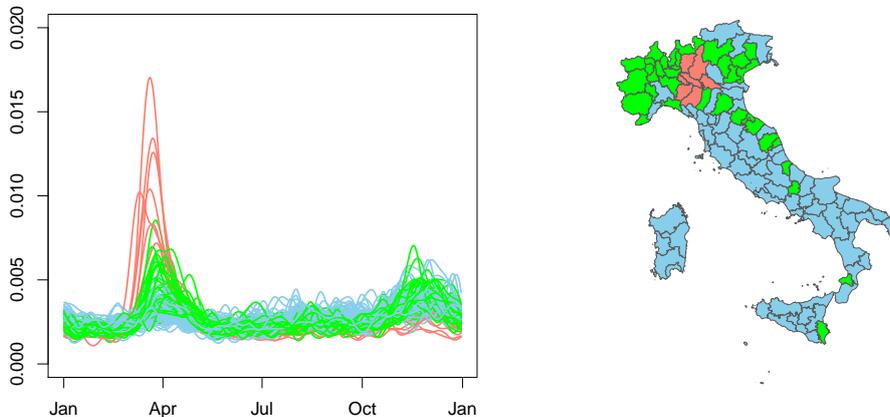
Clustering of the 2020 prediction errors, 70+ years

Figure 10: Functional K-means clustering ($K = 3$) of the prediction errors in Figure 8, for 2020. The prediction errors appear rather good descriptors of the pandemic development. The red cluster identifies the provinces that are very hardly hit by the first wave but are left mostly untouched by the second wave; the green cluster groups the provinces hit to a less dramatic extent by the first wave but significantly affected also by the second wave; finally, the blue cluster mostly collects provinces which only experience the second wave.

of the trace-semivariogram, which in this setting is defined as

$$\gamma(h) = \frac{1}{2}\mathbb{E}[\|X_{s_k} - X_{s_j}\|^2_{B^2}], \text{ for } (s_k, s_j) \in D(h) \tag{7}$$

and linked to (6) by the relation $\gamma(h) = C(0) - C(h)$, in analogy with the scalar case Cressie (1993). In this simplified setting, classical geostatistical tasks such as empirical variography, variogram estimation and Kriging prediction with scalar weigths can be performed in a similar way to the finite-dimensional case Menafoglio and Petris (2015). In particular, the definition of the B^2 empirical trace-semivariogram is extended from the Euclidean case Cressie (1993) as

$$\hat{\gamma}(h) = \frac{1}{2N(h)} \sum_{(s_k, s_j) \in D(h)} \|X_{s_k} - X_{s_j}\|^2_{B^2} \tag{8}$$

where $N(h)$ is the cardinality of $D(h)$. Classical fitting of theoretical (semi)variogram models (e.g., exponential, spherical, Matérn models) Cressie (1993) can then be applied to $\hat{\gamma}$ without difficulties.

Figure 11 shows the empirical trace-semivariograms of the prediction errors δ_{iy}, from 2017 to 2020, for the elderly population, at the provincial scale.

Functional trace-semivariograms, provinces

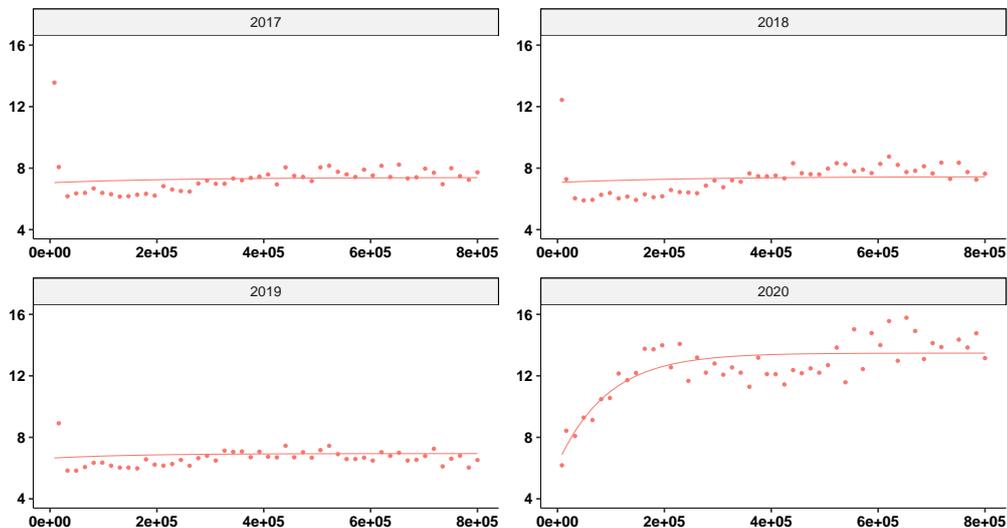Figure 11: Empirical trace-semivariograms for the prediction errors δ_{iy}, in the elderly population. The red lines are the corresponding fitted exponential models. Distances on the x-axes are expressed in meters. The last panel shows the 2020 severe perturbation of the spatial dependence structure of the process generating the prediction errors.

The empirical trace-semivariograms are computed with the R package `fdagstat` Menafoglio and Grujic (2017) and all the fitted semivariograms are exponential models with nugget Cressie (1993). It is interesting to note that up to 2019 the stochastic processes generating the δ_{iy}'s are almost spatially uncorrelated. In 2020, the spatial dependence structure of the process generating the prediction errors δ_{iy} is deeply perturbed; spatial correlation is now evident, with a sharp increase of the partial sill with respect to previous years and, consequently, of the total variance of the process. We interpret this as the effect of the pandemic, which introduces spatial correlation, being a natural process with an intrinsic spatial diffusive nature. In the next section, we further explore this conjecture, by studying the main sources of variability in the δ_{iy}'s, and their associated spatial structure. The subsequent dimensionality reduction will then allow us to identify local anomalies in the 2020 mortality densities, likely associated with COVID-19 pandemic, through a downscaling analysis of the provincial densities at a municipality spatial scale.

19

4 Studying the perturbation of mortality densities in 2020

4.1 Simplicial Functional Principal Component Analysis

From now on, we focus our analysis on 2020; we thus drop the index y, since all quantities and functional objects refer to the same year. We study the variability structure of the prediction errors δ_i's, by decomposing them in their main modes of variability. To this end, we use Simplicial Functional Principal Component Analysis (SFPCA) Hron et al. (2016), which is an extension of functional principal component analysis Ramsay and Silverman (2005) to functional data in the B^2 space. In this setting, the prediction error δ_i is represented by its (truncated) expansion on the principal components, namely

$$\widetilde{\delta}_i = \overline{\delta} + \sum_{j=1}^{K} w_{ij} \cdot \psi_j \tag{9}$$

where operations are computed in $B^2(\Theta)$, as defined in Section 3.1, K denotes the chosen truncation order, $\overline{\delta}$ is the B^2 overall sample mean of the δ_i's with respect to all the considered areas, and $\psi_1, \ldots, \psi_K$ are $B^2(\Theta)$ functions representing the first K elements of the orthonormal principal component basis. For $j = 1, \ldots, K$, the scalar w_{ij} is the score of the projection of δ_i on ψ_j. Note that the total (empirical) variability of the prediction errors δ_i is precisely obtained as the sum of the variability expressed by the scores along the principal components, i.e.,

$$\frac{1}{n} \sum_{i=1}^{n} \|\delta_i - \overline{\delta}\|^2 = \frac{1}{n} \sum_{j=1}^{\infty} \sum_{i=1}^{n} w_{ij}^2 = \sum_{j=1}^{\infty} \lambda_j,$$

λ_j being the (empirical) variance along the j-th Simplicial Functional Principal Component (SFPC). An analogous result is obtained when looking at the spatial variability, as represented by the trace-semivariogram (7) and its empirical version (8). In this sense, studying the SFPCs allows us to interpret the sources of variability within the δ_i's and within their spatial dependence.

Figure 12 shows the results of the SFPCA applied to the δ_i's for 2020, at the province level. The top panels of this figure display in black the mean of the prediction errors $\overline{\delta}$, and in color two additional curves obtained by perturbing the mean in the direction of each principal component, for the first three SFPCs. Specifically, the blue curve is $\overline{\delta} - 3\sqrt{\lambda_j}\psi_j$ and the red curve is $\overline{\delta} + 3\sqrt{\lambda_j}\psi_j$ for $j = 1, 2, 3$, respectively in the top left, top center and top right panels of Figure 12. By inspecting these plots, we notice that the first SFPC represents a contrast between provinces severely hit by the first pandemic wave against those provinces basically untouched by the first wave. Indeed, provinces with a high positive score on the first component are characterized by a peak in death
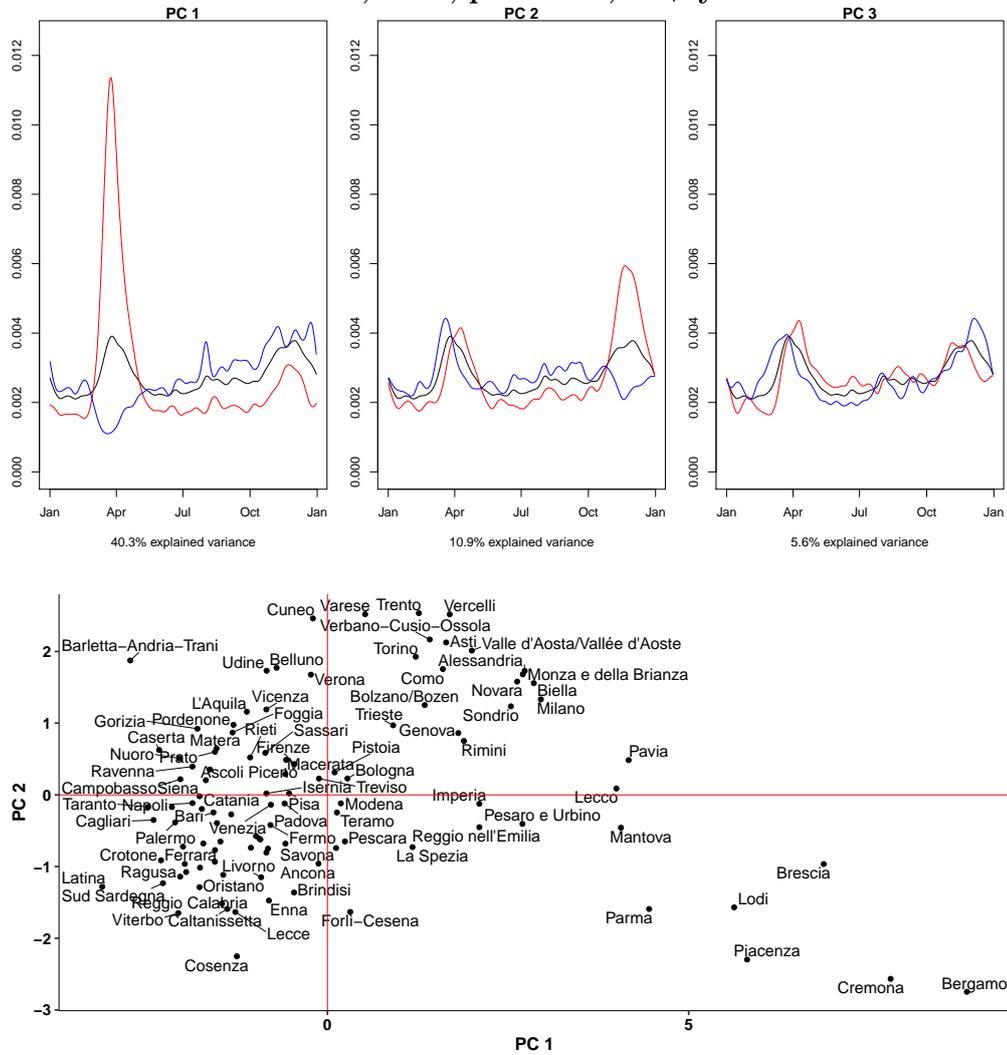
Figure 12: Top: for the first three principal components (left, center and right), the plot shows the mean $\overline{\delta}$ of the prediction errors (black) and its perturbations along the direction of the corresponding principal component; the perturbation is equal to plus (in red) or minus (in blue) three standard deviations of the corresponding score. Bottom: scatterplot of the scores of the first two principal components.

density corresponding to the first wave of the pandemic, followed by a below-average death density during the second wave. The second principal component captures a time shift in the pandemic development; positive scores characterize provinces where the pandemic starts later during the first wave and which are strongly hit by the second wave. The scatterplot of the scores of the first two principal components, in the bottom part of the same figure, confirms this interpretation: provinces like Bergamo, Cremona, Piacenza and Lodi, characterized by high positive scores along the first component and high negative scores along the second, are in fact among those dramatically hit by the first pandemic wave, and where the most at-risk segment of population has shrunk consistently by the time of the second wave, because of death or acquired immunity. Similar consid-

Maps of SFPCA scores
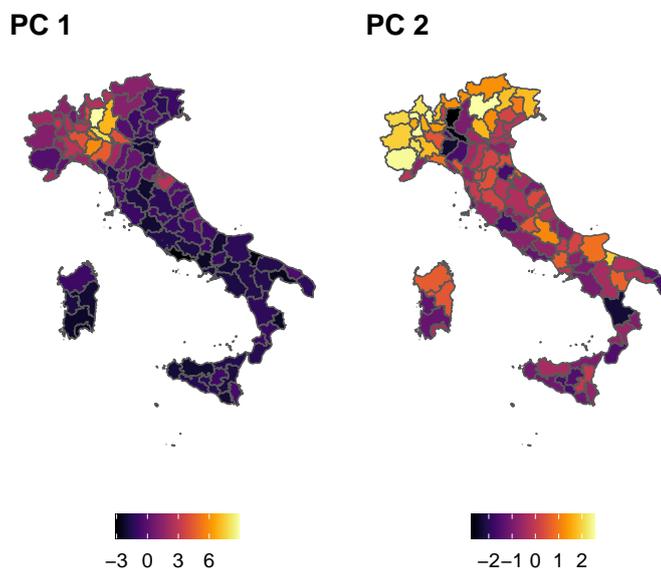
PC 1   **PC 2**

Figure 13: Heatmap of the scores of the errors δ_i on the first (left) and second (right) SFPC.

erations can be elicited for other provinces: for instance, we can see that Milano, and other northern provinces like Sondrio and Torino, are represented among those less severely hit by the first wave, but appear to also suffer heavy losses during the second one. Figure 13 offers a global picture of the geographical behaviour of the first two SFPCs, representing the heatmaps of the corresponding scores. Such maps, together with the interpretation of the principal components highlighted by Figure 12, efficiently capture the spatio-temporal behavior of the main pandemic waves, allowing to immediately visualize their characteristics in

22

different areas.

4.2 The spatial structure of the SFPCA scores

The dimensionality reduction induced by the SFPCA procedure leads to a representation of each δ_i through the corresponding $K-$ dimensional vector $\mathbf{w}_i \in \mathbb{R}^K$ of scores. This representation allows us to express the spatial dependence structure of the spatial field generating the prediction errors through the multivariate cross-variogram structure of the $\mathbf{w}_i$'s, as defined by their K (empirical) semivariograms $\hat{\gamma}_{rr}$, for $r = 1, \ldots, K$, and $K(K-1)/2$ (empirical) cross-semivariograms $\hat{\gamma}_{rs}$, where $r, s = 1, \ldots, K$, with $r < s$. For a formal account, see, e.g., Menafoglio et al. (2013); Menafoglio and Petris (2015). In particular, the spatial dependence of the field generating the δ_i's captured by the trace-semivariogram (7), and its empirical version (8), can be expressed as the composition of the (empirical) semivariograms of the $\mathbf{w}_i$'s, i.e., the $\hat{\gamma}_{rr}$. In this sense, studying the empirical semivariograms $\hat{\gamma}_{rr}$ enables us to further support the conjecture according to which the perturbation observed in 2020 in the δ_i's can indeed be interpreted in terms of the COVID-19 pandemic, consistently with the observation that the first two SFPCs can be precisely ascribed to the first and second pandemic waves. Figure 14 shows the estimated semivariograms and cross-semivariograms for the first three principal components. Here, blue symbols represent the empirical estimates, while the blue lines represent the fitted model of coregionalization (exponential model with nugget). It is clear that the first component, associated with the first pandemic wave (as highlighted in Figure 12), is characterized by a very strong spatial correlation. The spatial dependence in the other components appears instead much weaker. This suggests that the first wave is the event which mainly affects the change in spatial correlation within the δ_i in 2020, with respect to previous years. The second component, associated with the second pandemic wave, and to a time shift in the first one, still presents an appreciable spatial correlation structure, although much less evident than the first SFPC. The correlation structure becomes instead negligible when looking at the third component (and to subsequent components, not shown here), corroborating the hypothesis that the spatial structure within the δ_i's is indeed associated with the effects of the COVID-19 pandemic.

4.3 Analysing the mortality densities at a finer spatial scale through spatial downscaling

We finally use the representation provided by SFPCA to obtain predictions of the mortality densities at a municipality scale, through geostatistical downscaling of the provincial mortality densities. The relevance of the analysis is twofold. First, it allows us to gain robustness with respect to the direct consideration of death counts at a municipality level, as these are affected by heavier noise than province data, being related to areas with smaller population, as discussed in
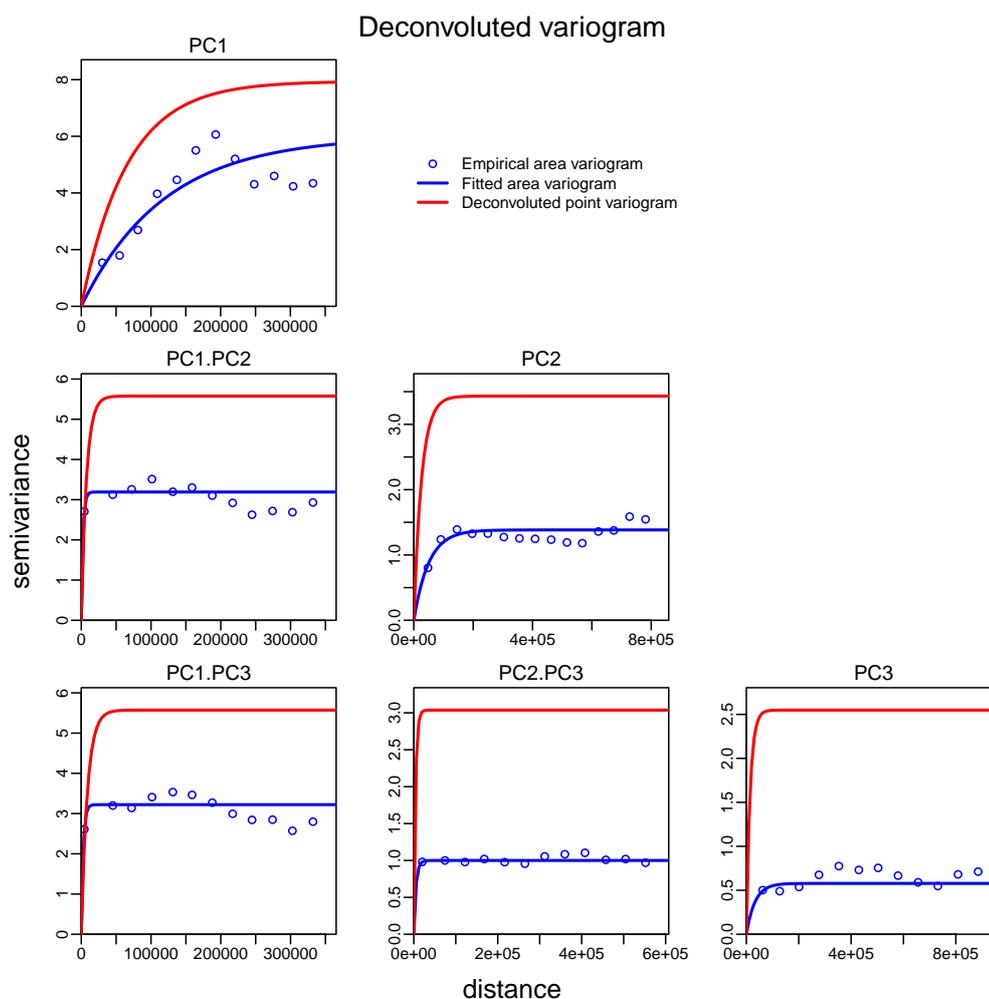
Figure 14: Empirical and fitted semivariograms and cross-semivariograms at province level (blue lines and dots) and corresponding deconvoluted semivariograms for prediction at municipality scale (red lines). The spatial correlation of the functional process is almost completely explained by the action of the first two components. Distances on the x-axes are expressed in meters.

Section 3.2 and highlighted by Figure 7. Second, this analysis enables us to identify municipalities which present *anomalies* with respect to the corresponding provinces, in the sense that their death densities are dissimilar to those that would be expected by 'projecting' the provincial mortality densities down to the municipality spatial scale.

To avoid confusion, in this section we use the index i to refer to provinces and the index ℓ to refer to municipalities. To provide a prediction $\hat{f}_\ell$ of the mortality density for the ℓ-th municipality, we use the spatio-temporal information embedded within the provincial sample f_i, this being decomposed in its time-repeatable ($\hat{f}_i$) and time-non-repeatable (δ_i) components. In fact, the crucial point is to provide a (spatial) prediction $\hat{\delta}_\ell$ for the error term, so that the corresponding death density is estimated as

$$\hat{f}_\ell = \hat{f}_i + \hat{\delta}_\ell \tag{10}$$

where $\hat{f}_i$ represents the component of the mortality density in the province i containing municipality ℓ, predictable from the previous years, as from equation (5). The problem of obtaining predictions $\hat{\delta}_\ell$ from the $\hat{\delta}_i$'s is a problem of (functional) *downscaling*, requiring to change the spatial resolution of the data from a provincial scale down to a municipality scale. We tackle this problem by relying on the finite-dimensional representation (9) induced by SFPCA, thus reducing the functional problem to a multivariate one. In this context, obtaining $\hat{\delta}_\ell$ requires to produce predictions $\hat{\mathbf{w}}_\ell$ by downscaling the province realizations $\mathbf{w}_i$ Kyriakidis (2004); Goovaerts (2008); Xiao et al. (2018), which is here done via Area-to-Point Co-Kriging (ATPCoK), as implemented in the R packages `gstat` Pebesma (2004) and `atakrig` Hu (2020). The main issue in ATPCoK is the estimation of a family of semivariograms and cross-semivariogram models $\eta_{r,s}$ for the finer geographical scale (i.e., municipality) from the models $\gamma_{r,s}$ estimated at the coarser scale (i.e., province). This problem is classically solved via the so-called *variogram deconvolution* technique Goovaerts (2008). Figure 14 shows the deconvoluted semivariograms and cross-semivariograms $\eta_{r,s}$ (red lines) for the first three principal components, as estimated from the fitted provincial semivariograms $\gamma_{r,s}$ (blue lines). The $\eta_{r,s}$'s indeed have analogous interpretations as those at a provincial scale. Note that all the *cross*-semivariograms are characterized by a negligible spatial structure, suggesting that different SFPCs identify spatially uncorrelated components in the δ_i's. Their cross-correlation is thus neglected for prediction purposes.

The fitted deconvoluted semivariograms are then used to obtain ATPCoK predictions $\hat{\mathbf{w}}_\ell$ Goovaerts (2008), which in turn lead to predict the death density for the ℓ-th municipality as

$$\hat{f}_\ell = \hat{f}_i + \sum_{j=1}^{K} \hat{w}_{\ell j} \psi_j \tag{11}$$

25

where the ψ_j are the SFPCs appearing in equation (9), and $K = 3$. The predicted densities are shown in the first row of Figure 15, colored according to the province-level clustering proposed in Figure 10. Note that the cluster information is only proposed to ease interpretation, as it plays no role in producing the predictions $\hat{f}_\ell$. These results suggest that the combination of dimensionality reduction via SFPCA and downscaling allows us to obtain much smoother estimates of the municipality densities than those produced directly from the municipality death counts, reducing their wiggliness while preserving their key features (e.g., the sharp peaks due to the pandemic). For example, the second row of Figure 15 compares the predicted densities $\hat{f}_\ell$ with the histograms of the municipality death counts p_ℓ, for the municipalities of Bergamo (≈ 120000 inhabitants), Lodi (≈ 45000 inhab.), and Clusone (≈ 8000 inhab., in the province of Bergamo). Note that the predictions remain fairly reasonable even when the interested area is scarcely populated (e.g., Clusone).

To identify municipalities characterized by an anomalous behaviour with respect to nearby provinces, for each municipality ℓ we compute the quadratic Wasserstein distance $w_2(\hat{f}_{\ell y}, p_{\ell y})$ between the downscaled predicted density and the respective empirical distribution binned on a weekly basis. A high distance is indicative of a municipality whose death distribution is anomalous with respect what is expected considering the mortality densities of the nearby provinces. Using the Local Moran Index, described in Section 2.2, we hence look for spatial clusters of *anomalous* municipalities, which can be associated with areas where local and unusual events cause a perturbation: indeed, these unusual events, during 2020, could be related to local pandemic hot-spots. Figure 16 shows the result of a spatial association analysis carried out on the logarithms of the Wasserstein distances, using the Local Moran Index. We only draw the high-high spatial clusters, since they are associated with areas where downscaled predictions are farther away from observed data. We also show a zoomed map of northern Italy, which is characterized by the presence of several interesting clusters: some significative provinces are indicated in purple, and some representative municipalities in the spatial clusters are indicated black. The bottom part of the same figure displays the comparison between downscaled predicted densities and observations, for the highlighted municipalities. Several anomalous areas are signaled. In province of Cuneo, for instance, the municipality of Mondovì (≈ 23000 inhab.) is the center of a cluster in which the impact of the first wave is highly underestimated, while the second wave is delayed with respect to predictions. At the border between the provinces of Alessandria and Pavia, the red area containing Tortona (≈ 27000 inhab.) is hit harder and earlier than expected by the first wave, while the second wave has a smaller impact than predicted. Another interesting case is Galbiate (≈ 9000 inhab., in the province of Lecco), whose cluster is characterized by an evident delay of the first wave, that has a much stronger effect with respect to the surrounding areas. On the other hand, the pandemic hit sooner than expected, and harder, in some area of the province of Bergamo, as for example the surroundings of San Pellegrino Terme
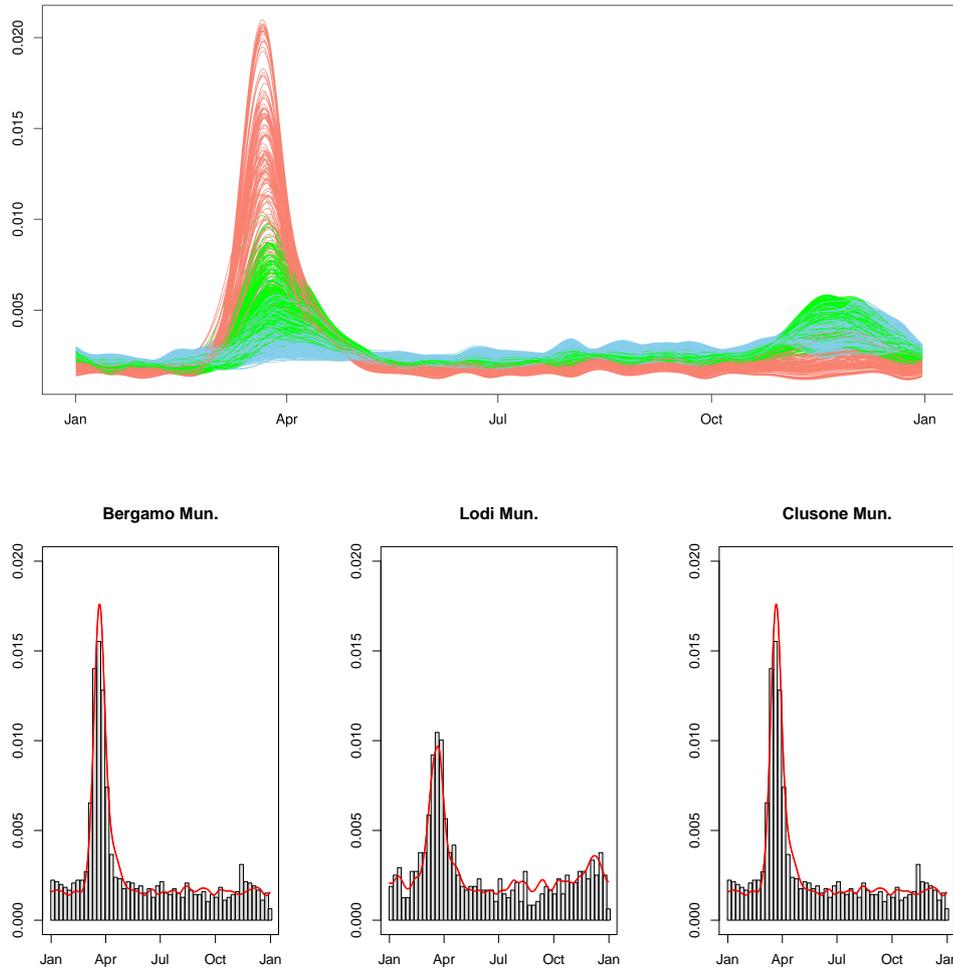
Figure 15: Top: densities $\hat{f}_\ell$ predicted through spatial downscaling, and colored according to the provincial clustering of Figure 10. Bottom: comparison of downscaling predictions (red lines) with the actual observations (histograms obtained aggregating the p_ℓ on a weekly basis) for the municipalities of Bergamo (≈ 120000 inhab.), Lodi (≈ 45000 inhab.), and Clusone (≈ 8000 inhab., in the province of Bergamo).
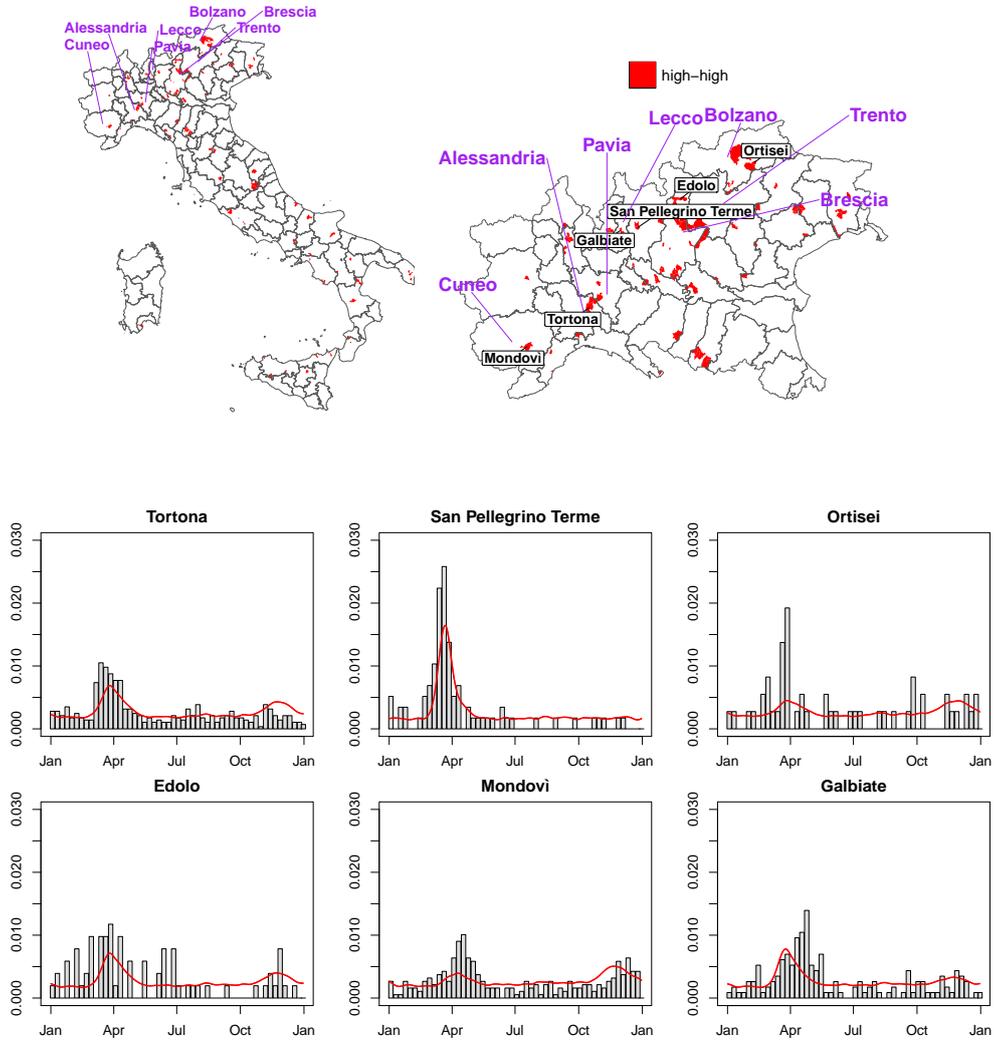
27

Figure 16: Top: spatial clusters, identified using Anselin's Local Moran Index on the logarithm of the Wasserstein distance between the predicted downscaled death density in the municipality, and the corresponding empirical density binned on a weakly basis. Only clusters with high distances are shown. Bottom: comparison between downscaled mortality densities (red lines) and actual observations (histograms obtained aggregating the p_ℓ on a weekly basis), for some selected anomalous municipalities.

($\approx$ 5000 inhab.). Similar anomalies are present also in the province of Brescia, close to the border with the province of Trento, as in the case of municipality of Edolo ($\approx$ 5000 inhab.) and its neighbors. Finally, several anomalies are present in the province of Bolzano, where several areas are hit by the first pandemic wave with much greater magnitude than expected, as in the municipality of Ortisei ($\approx$ 5000 inhab.).

5 Discussion

Up to 2019, the Italian yearly mortality within the elderly class shows consolidated and almost stationary spatio-temporal patterns. These are completely overturned in 2020, with a deep perturbation in the spatio-temportal structure of the mortality process. While the link between this and the pandemic shock can be conjectured when looking at the trace-variography of our model (Section 3), the interpretative tools of SFPCA allow us to consistently attribute such upheavals in the mortality to the COVID-19 pandemic, with particular regard to its first wave of contagion (Section 4). Analyses performed on the younger age classes, not reported here for brevity, lead to consistent results, with some important differences. As reasonable, spatial patterns and pandemic effects are reduced in the age class 50-69 with respect to the 70+ class, and they become almost negligible in the class 0-49. The results of the analyses on the younger classes are available at `https://github.com/RiccardoScimone/Mortality-densities-italy-analysis.git`, together with the R codes to reproduce the analyses presented in this work.

Although the analyses here presented focus on mortality data, the pipeline we propose is entirely general, and applicable to a broad spectrum of indicators. Quarantined people over time, occupancy of intensive care units, emergency calls or contacts with family doctors are only a few examples of indicators which may offer complementary views of the disruptive effects of the pandemic in our societies. In this sense, our work offers a methodological viewpoint rather than a single analysis, to assess the spatio-temporal impacts of an exogenous shock on health, economy and society, supporting the identification of local anomalies in space and time. In fact, while our analyses on mortality data allow us to quantify the impact *a posteriori* – being delayed with respect to the contagion itself – the same analyses based on other types of indicators could act as early warning signals of anomalies, to be possibly used to support decision makers in designing timely prevention policies. In this context, the same analyses could be carried out at larger spatial scales (e.g., identifying anomalous provinces within regions) or at shorter time scales (i.e., conditioning the densities on suitable time windows), to detect, in near real-time, abnormal behaviors at the desired spatio-temporal granularity.

References

Abraham, C., Cornillon, P., Matzner-Løber, E., Molinari, N., 2003. Unsupervised curve clustering using b-splines. Scandinavian Journal of Statistics 30, 581 – 595.

Aitchison, J., 1982. The statistical analysis of compositional data. Journal of the Royal Statistical Society. Series B (Methodological) 44, 139–177.

Aitchison, J., 1986. The Statistical Analysis of Compositional Data. Chapman & Hall, Ltd. London.

Ambrosio, L., Caffarelli, L.A., Brenier, Y., Buttazzo, G., Villani, C., Salsa, S., 2003. Optimal Transportation and Applications. Springer Berlin Heidelberg.

Anselin, L., 1995. Local indicators of spatial association—lisa. Geographical Analysis 27, 93 – 115.

Anselin, L., Syabri, I., Kho, Y., 2005. Geoda : An introduction to spatial data analysis. Geographical Analysis 38, 5–22.

Barcelo-Vidal, C., Martín-Fernandez, J.A., Pawlowsky-Glahn, V., 2001. Mathematical foundations of compositional data analysis, in: Proceedings o fIAMG'01 - The 6th annual conference of the Int. Ass. for Mathematical Geology, p. 20.

Bivand, R.S., Pebesma, E., Gomez-Rubio, V., 2013. Applied spatial data analysis with R, Second edition. Springer, NY.

van den Boogaart, K.G., Egozcue, J.J., Pawlowsky-Glahn, V., 2010. Bayes linear spaces. SORT 34, 201– 222.

Caballero, W., Giraldo, R., Mateu, J., 2013. A universal kriging approach for spatial functional data. Stochastic Environmental Research and Risk Assessment 27, 1553–1563.

Canale, A., Vantini, S., 2016. Constrained functional time series: Applications to the italian gas market. International Journal of Forecasting 32, 1340–1351.

Cressie, N.A.C., 1993. Statistics for spatial data. Wiley series in probability and mathematical statistics Applied probability and statistics. revised edition ed., J. Wiley & Sons, New York.

Delicado, P., 2011. Dimensionality reduction when data are density functions. Computational Statistics & Data Analysis 55, 401–420.

Egozcue, J.J., L. Díaz–Barrero, J., Pawlowsky-Glahn, V., 2006. Hilbert space of probability density functions based on aitchison geometry. Acta Mathematica Sinica, English Series 22, 1175–1182.

Febrero-Bande, M., Oviedo de la Fuente, M., 2012. Statistical computing in functional data analysis: The R package fda.usc. Journal of Statistical Software 51, 1–28.

Feng, X., Sedransk, N., Xia, J.Q., 2014. Calibration using constrained smoothing with applications to mass spectrometry data. Biometrics 70, 398–408.

Ferraty, F., Vieu, P., 2006. Nonparametric functional data analysis. Springer Series in Statistics, Springer, New York. Theory and practice.

Filzmoser, P., Hron, K., Templ, M., 2018. Applied Compositional Data Analysis. Springer.

Getis, A., Ord, J.K., 1992. The analysis of spatial association by use of distance statistics. Geographical Analysis 24, 189–206.

Giraldo, R., Delicado, P., Mateu, J., 2010. Continuous time-varying kriging for spatial prediction of functional data: An environmental application. Journal of Agricultural, Biological, and Environmental Statistics 15, 66–82.

Giraldo, R., Delicado, P., Mateu, J., 2011. Ordinary kriging for function-valued spatial data. Environmental and Ecological Statistics 18, 411–426.

Goovaerts, P., 2008. Kriging and semivariogram deconvolution in the presence of irregular geographical units. Mathematical Geosciences 40, 101–128.

Horváth, L., Kokoszka, P., 2012. Inference for Functional Data with Applications. Springer Series in Statistics, Springer New York.

Hron, K., Menafoglio, A., Templ, M., Hrůzová, K., Filzmoser, P., 2016. Simplicial principal component analysis for density functions in bayes spaces. Computational Statistics & Data Analysis 94, 330–350.

Hu, M., 2020. atakrig: Area-to-Area Kriging. R package version 0.9.7.

Ignaccolo, R., Mateu, J., Giraldo, R., 2014. Kriging with external drift for functional data for air quality monitoring. Stochastic Environmental Research and Risk Assessment 28, 1171–1186. doi:10.1007/s00477-013-0806-y.

Kyriakidis, P., 2004. A geostatistical framework for area-to-point spatial interpolation. Geographical Analysis 36, 259–289.

Machalová, J., Hron, K., Monti, G., 2015. Preprocessing of centred logratio transformed density functions using smoothing splines. Journal of Applied Statistics 43.

Machalová, J., Talská, R., Hron, K., Gába, A., 2020. Compositional splines for representation of density functions. Computational Statistics , 1–34.

Mammen, E., Marron, J.S., Turlach, B.A., Wand, M.P., 2001. A general projection framework for constrained smoothing. Statistical Science 16, 232–248.

Menafoglio, A., Grujic, O., 2017. fdagstat: A functional geostatistics package. R package version 1.0.

Menafoglio, A., Grujic, O., Caers, J., 2016a. Universal Kriging of functional data: Trace-variography vs cross-variography? Application to gas forecasting in unconventional shales. Spatial Statistics 15, 39–55.

Menafoglio, A., Guadagnini, A., Secchi, P., 2014. A Kriging approach based on Aitchison geometry for the characterization of particle-size curves in heterogeneous aquifers. Stochastic Environmental Research and Risk Assessment 28, 1835–1851.

Menafoglio, A., Petris, G., 2015. Kriging for hilbert-space valued random fields: The operatorial point of view. Journal of Multivariate Analysis 146.

Menafoglio, A., Secchi, P., 2017. Statistical analysis of complex and spatially dependent data: A review of object oriented spatial statistics. European Journal of Operational Research 258, 401–410.

Menafoglio, A., Secchi, P., Dalla Rosa, M., 2013. A Universal Kriging predictor for spatially dependent functional data of a Hilbert Space. Electronic Journal of Statistics 7, 2209–2240.

Menafoglio, A., Secchi, P., Guadagnini, A., 2016b. A Class-Kriging Predictor for Functional Compositions with Application to Particle-Size Curves in Heterogeneous Aquifers. Mathematical Geosciences 48, 463–485.

Moran, P.A.P., 1950. Notes on continuous stochastic phenomena. Biometrika 37, 17–23.

Nerini, D., Monestiez, P., Manté, C., 2010. Cokriging for spatial functional data. Journal of Multivariate Analysis 101, 409–418.

Pawlowsky-Glahn, V., Egozcue, J., 2001. Geometric approach to statistical analysis in the symplex. Stochastic Environmental Research and Risk Assessment 15, 384–398.

Pawlowsky-Glahn, V., Egozcue, J., Tolosana-Delgado, R., 2015. Modelling and analysis of compositional data. John Wiley & Sons, Ltd.

Pawlowsky-Glahn, V., Egozcue, J.J., Van den Boogaart, K., 2014. Bayes hilbert spaces. Australia and New Zealand Journal of Statistics 56, 171–194.

Pearson, K., 1897. Mathematical contributions to the theory of evolution. on a form of spurious correlation which may arise when indices are used in the

measurement of organs, in: Proceedings of the Royal Society of London LX, pp. 489–502.

Pebesma, E.J., 2004. Multivariable geostatistics in S: the gstat package. Computers & Geosciences 30, 683–691.

Ramsay, J., Silverman, B.W., 2005. Functional Data Analysis. Springer.

Ruiz-Medina, M.D., 2012. Spatial functional prediction from spatial autoregressive Hilbertian processes. Environmetrics 23, 119–128.

Schuhmacher, D., Bähre, B., Gottschlich, C., Hartmann, V., Heinemann, F., Schmitzer, B., 2020. transport: Computation of Optimal Transport Plans and Wasserstein Distances. R package version 0.12-2.

Talská, R., Menafoglio, A., Machalová, J., Hron, K., Fišerová, E., 2018. Compositional regression with functional response. Computational Statistics and Data Analysis 123, 66–85.

Templ, M., Hron, K., Filzmoser, P., 2011. robCompositions: an R-package for robust statistical analysis of compositional data. John Wiley and Sons.

Villani, C., 2003. Topics in Optimal Transportation. Graduate studies in mathematics, American Mathematical Society.

Villani, C., 2008. Optimal transport – Old and new. Grundlehren der mathematischen Wissenschaften, Springer-Verlag Berlin Heidelberg.

Wang, J.L., Chiou, J.M., Müller, H.G., 2016. Functional Data Analysis.

Xiao, M., Zhang, G., Breitkopf, P., Villon, P., Zhang, W., 2018. Extended co-kriging interpolation method based on multi-fidelity data. Applied Mathematics and Computation 323, 120–131.

MOX Technical Reports, last issues

Dipartimento di Matematica
Politecnico di Milano, Via Bonardi 9 - 20133 Milano (Italy)

26/2021   Vigano, L.; Sollini, M.; Ieva, F.; Fiz, F.; Torzilli, G.
Chemotherapy-Associated Liver Injuries: Unmet Needs and New Insights for Surgical Oncologists

25/2021   Tenderini, R.; Pagani, S.; Quarteroni, A.; Deparis S.
PDE-aware deep learning for inverse problems in cardiac electrophysiology

24/2021   Regazzoni, F.; Chapelle, D.; Moireau, P.
Combining Data Assimilation and Machine Learning to build data-driven models for unknown long time dynamics - Applications in cardiovascular modeling

22/2021   Domanin, M.; Bennati, L.; Vergara, C.; Bissacco, D.; Malloggi, C.; Silani, V.; Parati, G.; Trim
Fluid structure interaction analysis to stratify the behavior of different atheromatous carotid plaques

23/2021   Scimone, R.; Taormina, T.; Colosimo, B. M.; Grasso, M.; Menafoglio, A.; Secchi, P.
Statistical modeling and monitoring of geometrical deviations in complex shapes with application to Additive Manufacturing

20/2021   Pasquale, A.; Ammar, A.; Falcó, A.; Perotto, S.; Cueto, E.; Duval, J.-L.; Chinesta, F.
A separated representation involving multiple time scales within the Proper Generalized Decomposition framework

21/2021   Torti, A.; Galvani, M.; Menafoglio, A.; Secchi, P.; Vantini S.
A General Bi-clustering Algorithm for Hilbert Data: Analysis of the Lombardy Railway Service

16/2021   Salvador, M.; Dede', L.; Manzoni, A.
Non intrusive reduced order modeling of parametrized PDEs by kernel POD and neural networks

17/2021   Chew, R.; Benacchio, T.; Hastermann, G.; Klein, R.
Balanced data assimilation with a blended numerical model

18/2021   Gigante, G.; Vergara, C.
On the choice of interface parameters in Robin-Robin loosely coupled schemes for fluid-structure interaction