A Workflow for Estimating and Visualising Excess Mortality During the COVID-19 Pandemic

by Garyfallos Konstantinoudis, Virgilio Gómez-Rubio, Michela Cameletti, Monica Pirani, Gianluca Baio, and Marta Blangiardo

Abstract COVID-19 related deaths estimates underestimate the pandemic burden on mortality because they suffer from completeness and accuracy issues. Excess mortality is a popular alternative, as it compares the observed number of deaths versus the number that would be expected if the pandemic did not occur. The expected number of deaths depends on population trends, temperature, and spatio-temporal patterns. In addition to this, high geographical resolution is required to examine within country trends and the effectiveness of the different public health policies. In this tutorial, we propose a workflow using R for estimating and visualising excess mortality at high geographical resolution. We show a case study estimating excess deaths during 2020 in Italy. The proposed workflow is fast to implement and allows for combining different models and presenting aggregated results based on factors such as age, sex, and spatial location. This makes it a particularly powerful and appealing workflow for online monitoring of the pandemic burden and timely policy making.

1 Introduction

Estimating the burden of the effect of the COVID-19 pandemic on mortality is an important challenge (Weinberger et al., 2020). The estimates of COVID-19 related deaths are subject to testing capacity and changes in definition and reporting guidelines, raising accuracy and completeness considerations (Aburto et al., 2021; Konstantinoudis et al., 2022). In addition, the estimates of COVID-19-related deaths give an incomplete picture of the mortality burden of the COVID-19 pandemic, as they do not account for the indirect pandemic effects due to, for instance, disruption to health services (Kaczorowski and Del Grande, 2021). Alternatively, excess mortality has been extensively used to evaluate the impact of the COVID-19 pandemic on mortality (Rossen et al., 2020; Islam et al., 2021; Kontis et al., 2020; Konstantinoudis et al., 2022; Verbeeck et al., 2021).

Excess mortality is estimated by comparing the observed number of deaths in a particular time period with the expected number of deaths under the counterfactual scenario of the event (pandemic) not having had occurred. Typically, this counterfactual scenario is calculated using a comparison period, for instance, several previous years (https://www.euromomo.eu/). Calculating the expected number of deaths accurately requires accounting for factors such as population trends, seasonality, temperature, public holidays and spatio-temporal dependencies. While accounting for the above-mentioned factors, most studies to date have estimated excess mortality at the national level (see Rossen et al., 2020; Weinberger et al., 2020), and some have examined excess mortality across countries (Islam et al., 2021; Kontis et al., 2020, 2021).

While national-level estimates of excess mortality give valuable insights into the total number of excess deaths, they do not allow for evaluation of within-country geographical differences. However, such information is essential to understand the country's transmission patterns and effectiveness of local policies and measures to contain the pandemic (Kontopantelis et al., 2021). Temporal trends in excess mortality can substantially differ across regions of the same country (Blangiardo et al., 2020), which makes national-based comparisons even more challenging. Therefore, understanding the effect of the COVID-19 pandemic on mortality requires higher spatial resolution and models that account for spatial, temporal and spatio-temporal dependencies.

When working at high spatio-temporal resolution, data are generally sparse, and this leads to highly variable excess mortality estimates. This is aggravated by the fact that excess deaths tend to be subject to spatial and temporal correlation. This makes it essential to use statistical methods that can account for these dependencies and provide robust and accurate estimates. The disease mapping framework, which is commonly employed in epidemiology to study the spatio-temporal variation of diseases (Waller et al., 1997; Moraga, 2018), can be exploited to estimate excess mortality at subnational and weekly level. The Bayesian approach is naturally suited in this context, as it is able to model complex dependency structures, as well as to incorporate uncertainty in the data generating and modelling process. In addition, while fully propagating the uncertainty, it allows for summaries of results at any desired level of spatio-temporal aggregation (using for instance coarser geographical units suitable for policy implementation). This in combination with fast approximate methods to inference, such as the Integrated Laplace Approximation (INLA; Rue et al., 2009), make this framework
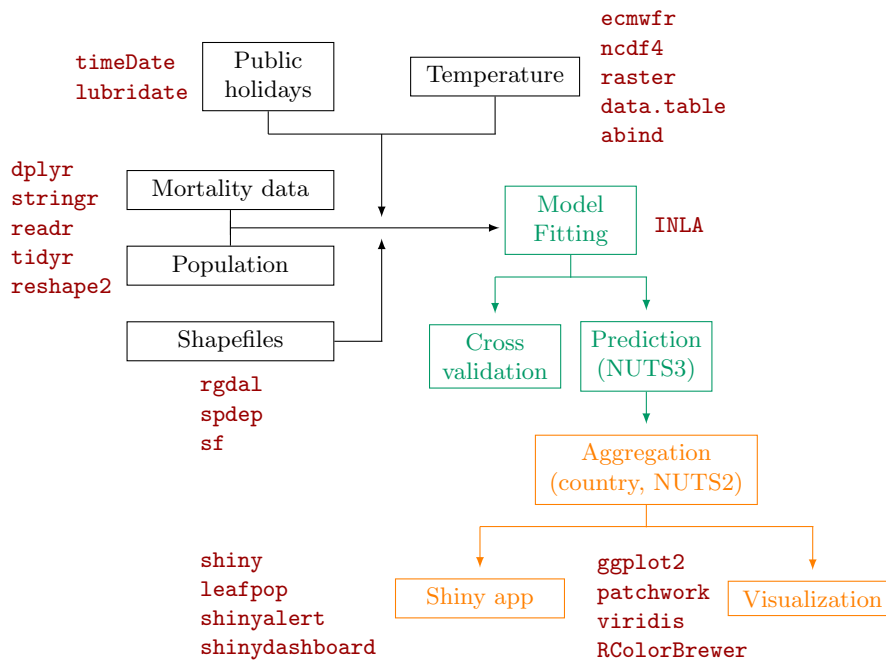
Figure 1: Diagram of the workflow: data wrangling in black, analysis in green, post-processing in orange and packages in red (timeDate; Wuertz et al., 2023, lubridate; Grolemund and Wickham, 2011, dplyr; Wickham et al., 2023a, stringr; Wickham, 2022, readr; Wickham et al., 2022, tidyr; Wickham et al., 2023b, reshape2; Wickham, 2007, rgdal; Bivand et al., 2023, spdep; Bivand, 2022, sf; Pebesma, 2018, shiny; Chang et al., 2022, leafpop; Appelhans and Detsch, 2021, shinyalert; Attali and Edwards, 2021, shinydashboard; Chang and Borges Ribeiro, 2021, ecmwfr; Hufkens et al., 2019, ncdf4; Pierce, 2023, raster; Hijmans (2023), data.table; Dowle and Srinivasan, 2022, abind; Plate and Heiberger, 2016, ggplot2; Wickham, 2016, patchwork; Pedersen, 2022, viridis; Garnier et al., 2021, RColorBrewer; Neuwirth, 2022). NUTS stands for Nomenclature of Territorial Units for Statistics with NUTS3 being the highest spatial resolution available and NUTS2 coarser but appropriate for policy making.

particularly powerful and appealing for the monitoring of the pandemic burden and timely policy making.

Here, we describe how to run a study on excess mortality at high spatial and temporal resolution using a Bayesian approach and R. This tutorial provides a detailed explanation of the modelling approach used previously to quantify excess mortality in 5 European regions (Konstantinoudis et al., 2022). We have further modified the way we model long-term trends in Konstantinoudis et al. (2022), as we showed that it provides more accurate predictions (Riou et al., 2023). Figure 1 illustrates the workflow followed in this paper together with the main R packages used. Source code for replicating the data wrangling (R scripts prefixed with 01, 02 or 03), analysis (R scripts prefixed with 04) and post-processing steps (R scripts prefixed with 05 or 06 and the App folder) and data files are available from GitHub at https://github.com/gkonstantinoudis/TutorialExcess.

This tutorial is structured as follows: we first describe the modelling framework and present the case study in Italy. We then show how to run and evaluate the model, and extract and visualise the results. Finally, we present an R-shiny app which makes the results effectively and easily disseminated.

2 Bayesian hierarchical spatio-temporal model to estimate excess mortality

We propose a Bayesian hierarchical model to quantify the effect of spatio-temporal location on excess mortality under extreme events such as the COVID-19 pandemic, stratified by specific age-sex population groups. To do so, we first describe the statistical model for predicting the number of deaths from all-causes based on historical data, in the counterfactual scenario in which the pandemic did not take place. Then, we show how to estimate the magnitude of excess deaths over a specific period of time, with associated uncertainty, by comparing the predicted versus the actual number of deaths.

Let y_{jstk} and P_{jstk} be the number of all-cause deaths and the population at risk for the j-th week, $j = 1, \ldots, J_t$, where J_t is the total number of weeks of the year t ($t = 2015, \ldots 2019$), the s-th spatial

unit ($s = 1, \ldots S$, where S is the number of provinces in Italy), and k-th age-sex group ($k = 1, \ldots 10$). Also let x_{1jt}, x_{2t} and z_{jst} denote the adjustment covariates, respectively: public holidays ($x_{1jt} = 1$ if week j of the t-th year contains a public holiday and 0 otherwise), the year that the j-the week belongs to, and temperature. We assume that the historical observed number of deaths, conditional on the risk r_{jstk}, follows a Poisson distribution, with a log-linear model for the risk. To simplify notation, we omit k, although the following model was fitted to all k age-sex groups (with the elements of the linear predictor also depending on k):

$$
\begin{aligned}
y_{jst}|r_{jst} &\sim \mathrm{Poisson}(\mu_{jst} = r_{jst}P_{jst}), \\
\log\left(r_{jst}\right) &= \beta_{0jt} + \beta_1 x_{1jt} + \beta_2 x_{2t} + f(z_{jst}) + b_s + w_{jt}.
\end{aligned}
\tag{1}
$$

Here, β_{0jt} is the week-specific intercept in year t given by $\beta_{0jt} = \beta_0 + \epsilon_{jt}$ for the k-th age-sex group, where β_0 is the global intercept and ϵ_{jt} is an unstructured random effect representing the deviation of each week from the global intercept, which is modelled using independent and identically (iid) distributed Gaussian prior distribution with zero-mean and variance equal to τ_ϵ^{-1}. The parameters β_1 and β_2 are unknown regression coefficients associated to the public holidays covariate x_{1jt} and a long term linear trend. The effect of temperature, $f(\cdot)$, is allowed to be non-linear by specifying a second-order random walk (RW2) model:

$$
z_{jst} \mid z_{(j-1)st}, z_{(j-2)st}, \tau_z \sim \mathrm{Normal}\left(2z_{(j-1)st} + z_{(j-2)st}, \tau_z^{-1}\right),
$$

where τ_z^{-1} is the variance.

Terms b_s and w_j are spatial and temporal random effects, respectively. We specify the spatial random effect term, b_s, with a convolution prior (Besag et al., 1991), and the temporal random effect term, w_j with a non-stationary in time prior. In detail, we model $\boldsymbol{b}$ using a reparameterisation of the popular Besag-York-Mollié prior, which is a convolution of an intrinsic conditional autoregressive (CAR) model and an iid Gaussian model. Let u_s be the spatially structured component defined by an intrinsic CAR (Besag, 1974) prior $u_s|u_i, i \in \partial_s \sim (\bar{u}, \tau_u^{-1}/m_s)$, where $\bar{u}$ is the mean of the neighbours and ∂_s and m_s are respectively the set and the number of neighbours of area s, τ_u^{-1} the conditional variance, and v_s the unstructured component with prior $v_s \overset{iid}{\sim} \mathrm{Normal}(0, \tau_v^{-1})$. We model b_s as follows (Besag et al., 1991; Riebler et al., 2016; Konstantinoudis et al., 2020):

$$
b_s = \frac{1}{\sqrt{\tau_b}}\left(\sqrt{1-\phi}v_s^\star + \sqrt{\phi}u_s^\star\right)
$$

where $u_s^\star$ and $v_s^\star$ are standardised versions of u_s and v_s such that their variance is equal to 1 (Simpson et al., 2017). The term $0 \leq \phi \leq 1$ is a mixing parameter, which measures the proportion of the marginal variance explained by the structured effect. Finally, we assign to the temporal random effect term, w_{jt}, a Gaussian random walk model of order 1 (RW1). This component captures seasonality and is specified as:

$$
w_{jt} \mid w_{(j-1)t}, \tau_w \sim \mathrm{Normal}(w_{(j-1)t}, \tau_w^{-1}).
$$

The Bayesian representation of the above model is completed once we select priors for the fixed effects β_0 and $\boldsymbol{\beta}$ and the hyperparameters: $\tau_\epsilon, \tau_z, \tau_b, \tau_w$, and ϕ. For the fixed effects we selected minimally informative Normal distributions, whereas for the hyperparameters we specified "penalising complexity" (PC) priors (Simpson et al., 2017). PC priors are defined by penalising deviations from a "base" model (e.g., specified in terms of a specific value of the hyperparameters) and have the effect of regularising inference, while not implying too strong prior information. Technically, PC priors imply an exponential distribution on a function of the Kullback–Leibler divergence between the base model and an alternative model in which the relevant parameter is unrestricted. This translates to a suitable "minimally informative", regularising prior on the natural scale of the parameter.

In order to quantify the weekly excess mortality at sub-national level for specific age-sex population groups, we need to predict the number of deaths that would be expected if the COVID-19 pandemic had not occurred. In Bayesian analysis, this can be performed by drawing random samples from the posterior predictive distribution (that is, the distribution of unobserved values conditional on the observed values from previous years). Specifically, letting $\boldsymbol{\theta}$ be the model parameters, $\mathcal{D}$ be the observed data, and y_{jst^*} be the count of deaths that we want to predict, we have:

$$
p(y_{jst^*} \mid \mathcal{D}) = \int p(y_{jst^*} \mid \boldsymbol{\theta})p(\boldsymbol{\theta} \mid \mathcal{D})d\boldsymbol{\theta}.
\tag{2}
$$

Operationally, we first generate random samples from the joint posterior marginal of the parame-

ters specified in Equation (1) at the highest spatial resolution available (NUTS3 regions; Nomenclature of Territorial Units for Statistics 3 regions, https://ec.europa.eu/eurostat/web/nuts/background/). Following that, we use the samples to compute the linear predictor, compute the mean parameter of the Poisson distribution via inverse link function, and obtain the predicted number of deaths, which represents the baseline number of deaths assuming the pandemic did not take place.

Finally, to estimate the excess deaths, the predicted number of deaths is compared against the actual observed number of deaths. Further, this allows us to compute the relative change in mortality (relative to what we would expect if the pandemic did not occur). This is obtained by (i) subtracting the predicted number of deaths from the observed number of deaths in each time point j in the t^*-th year and spatial unit s (number of excess deaths or NED), and (ii) dividing NED by the predicted number of deaths for each sample and multiplying by 100 (yielding % relative excess mortality or REM).

The model estimates are computed using Integrated Nested Laplace Approximation (INLA; Rue et al., 2009, which performs approximate Bayesian inference on the class of latent Gaussian models (Rue and Held, 2005). Unlike simulation based Markov chain Monte Carlo method, INLA is a deterministic algorithm, which employs analytical approximations and efficient numerical integration schemes to provide accurate approximations of the posterior distributions in short computing times. The INLA software is provided through the R package INLA, which can be downloaded from https://www.r-inla.org/.

3 Motivating example: Italy

3.1 Outcome data

We retrieved all-cause mortality data during 2015-2020 in Italy from the Italian National Institute of Statistics (https://www.istat.it/). Data were available weekly (ISO week), by age (5-year age groups), sex and NUTS3 regions. As the COVID-19 mortality rates increase with age, we aggregated mortality counts based on the following age groups: <40, 40-59, 60-69, 70-79 and 80 years and older (Davies et al., 2021).

3.2 Population data

Population data in Italy during 2015-2020 were retrieved from the Italian National Institute of Statistics. The data represent the population in Italy on first of January of every year stratified by age (5-year age groups), sex and NUTS3 regions. To retrieve weekly population, we performed linear interpolation by the selected age groups (<40, 40-59, 60-69, 70-79 and 80+), sex, and NUTS3 regions using populations on the first of January of the current and next year. Population counts on the first of January 2021, which takes COVID-19 deaths in 2020 into consideration, were available at the time of analysis. Our goal was, however, to predict mortality for 2020, as if the pandemic had not occurred. Thus we performed an additional linear interpolation by age, sex and NUTS3 regions to predict the population at January 1st 2021, using the years 2015-2020 (Figure 2, panel A). Object pop is a tibble containing the NUTS3 region ID (NUTS318CD), age group (ageg), sex (sex), year (year) and population counts (population):

```
pop
# A tibble: 6,420 x 5
# Groups:   ageg, sex [10]
  NUTS318CD ageg    sex       year population
  <chr>     <fct>   <chr>    <dbl>      <dbl>
 1 TO       less40  female    2015     435758
 2 TO       less40  female    2016     427702
 3 TO       less40  female    2017     420498
 4 TO       less40  female    2018     413141
 5 TO       less40  female    2019     406937
 6 TO       less40  female    2020     402768
 7 TO       less40  male      2015     449605
 8 TO       less40  male      2016     443941
 9 TO       less40  male      2017     439522
10 TO       less40  male      2018     433365
# ... with 6,410 more rows
```

We can use the following code (based on `tidyverse` and "piping" principles) to calculate the population of the 1st of January 2021 by NUTS3 regions, sex and age:

```
pop %>% group_by(NUTS3, sex, age) %>%
       summarise(pop = as.vector(coef(lm(pop ~ year)) %*% c(1, 2021))) %>%
       mutate(year = 2021) -> pop2021
```

We acknowledge that the linear trend in the population is a rather simplistic assumption. In subsequent analyses in Switzerland, we proposed a spatio-temporal approach similar with equation (1) to model the population counts had the pandemic not occurred (Riou et al., 2023). The code for that analysis is also online available online (https://github.com/jriou/covid19_ascertain_deaths). Once we obtained the year 2021 we performed an additional linear interpolation to calculate weekly number of population as shown on Figure 2, panel B.

3.3 Covariates

We used covariates related to ambient temperature, national holidays, and year of death, to improve the model's predictions. Data on air temperature during 2015-2020 in Italy at 2m above the surface of land were retrieved from the ERA5 reanalysis data set of the Copernicus climate change program
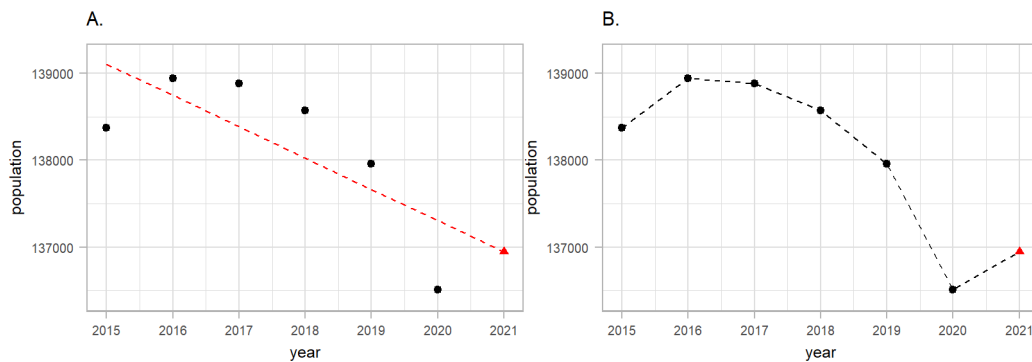

Figure 2: A schematic representation of the weekly population estimation procedure focusing on females aged 40-49 in Venice during 2015-2019 as an example. On panel A we show how we used the historical data (black points) and fit a linear regression (red dashed line) to predict 2021 (red triangle). On the panel B we show how we predicted weekly population by drawing lines between the years.
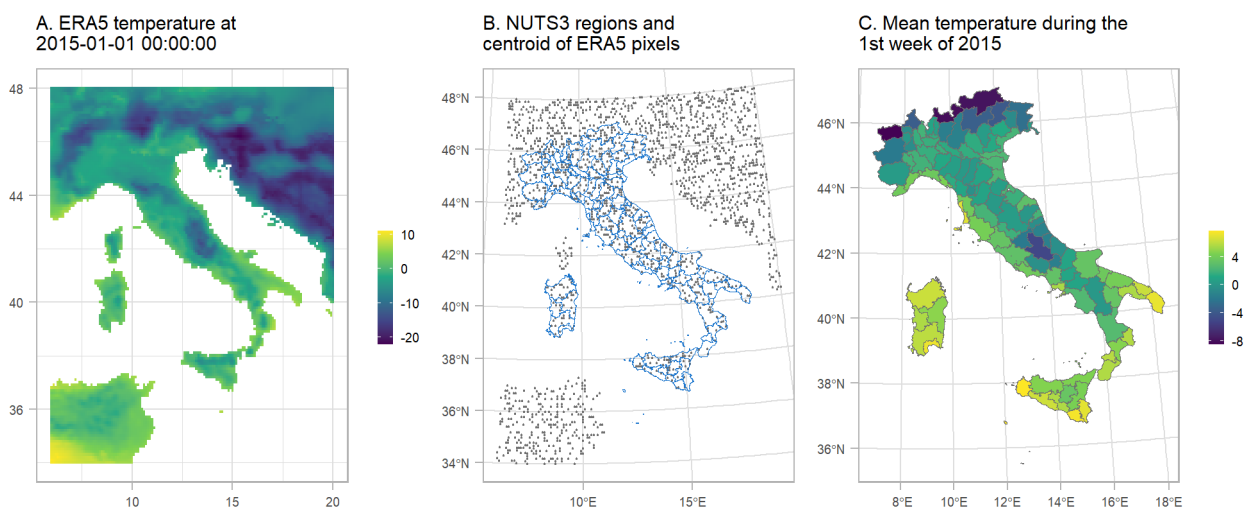

Figure 3: Schematic representation of the temperature misalignment procedure. A) The temperature obtained by ERA5 at 2015-01-01 00:00:00. B) NUTS3 regions in blue and a sample of the centroids of the pixels from the ERA5 raster. C. Mean temperature per NUTS3 region during the 1st week of 2015.

(Hersbach et al., 2020). The geographical resolution of the ERA5 estimates is $0.25° \times 0.25°$ (panel A of Figure 3). We calculated the weekly mean by the centroids of the $0.25° \times 0.25°$ grid (panel B of Figure 3) and then averaged the weekly temperature over the ERA5 centroids that overlay with the NUTS3 regions (panels B and C of Figure 3).

3.4 Fitting the model

The modelling process in `INLA` consists of three main steps: (1) the selection of priors, (2) definition of the model "formula" (which sets out the expression for the generalised linear predictor), and (3) the call to the main function `inla`, which computes the estimates.

In particular, we constructed the PC priors for $\sigma_\epsilon = \sqrt{1/\tau_\epsilon}, \sigma_z = \sqrt{1/\tau_z}, \sigma_b = \sqrt{1/\tau_b}$ and $\sigma_w = \sqrt{1/\tau_w}$ based on the assumption that it is unlikely to have a relative risk higher than $\exp(2)$ based solely on spatial, yearly and seasonal variation, Figure 4, panel A. For the mixing parameter ϕ, we set $\Pr(\phi < 0.5) = 0.5$ reflecting our lack of knowledge about whether overdispersion or strong spatial autocorrelation should dominate the field b, Figure 4, panel B.

These assumptions can be encoded using the following code:

```
# Defines the priors
hyper.bym <- list(
        theta1 = list('PCprior', c(1, 0.01)),
        theta2 = list('PCprior', c(0.5, 0.5))
)
hyper.iid <- list(theta = list(prior="pc.prec", param=c(1, 0.01)))

# Defines the model "formula"
formula =
        deaths ~ 1 + offset(log(population)) + hol + id.year +
        f(id.tmp, model = 'rw2', hyper = hyper.iid, constr = TRUE, scale.model = TRUE) +
        f(id.wkes, model = 'iid', hyper = hyper.iid, constr = TRUE) +
        f(id.time, model = 'rw1', hyper = hyper.iid, constr = TRUE, scale.model = TRUE,
         cyclic = TRUE) +
        f(id.space, model = 'bym2', graph = "W.adj", scale.model = TRUE, constr = TRUE,
        hyper = hyper.bym)

control.family = inla.set.control.family.default()

# Calls INLA to fit the model
inla.mod = inla(formula,
        data = dat,
        family = "Poisson",
        verbose = TRUE,
        control.family = control.family,
        control.compute = list(config = TRUE),
        control.mode = list(restart = T),
        num.threads = round(parallel::detectCores()*.8),
        control.predictor = list(link = 1))
```

After fitting the model, we take 1000 samples from the (approximated) posterior distribution of the linear predictor and we use each drawn sample as the mean of a Poisson distribution to retrieve the predicted mortality counts:

```
post.samples <- inla.posterior.sample(n = 1000, result = inla.mod)
predlist <- do.call(cbind, lapply(post.samples, function(X)
exp(X$latent[startsWith(rownames(X$latent), "Pred")])))

pois.samples <- apply(predlist, 2, function(Z) rpois(n = length(Z), lambda = Z))
```

This allows us to estimate the entire predictive posterior distribution of the mortality counts, incorporating both the sampling and the linear predictor uncertainty.

3.5 Model validation

To examine model validity, we performed a cross validation, leaving out one historical year at a time and predicting the weekly number of deaths by NUTS3 regions for the year left out. As part
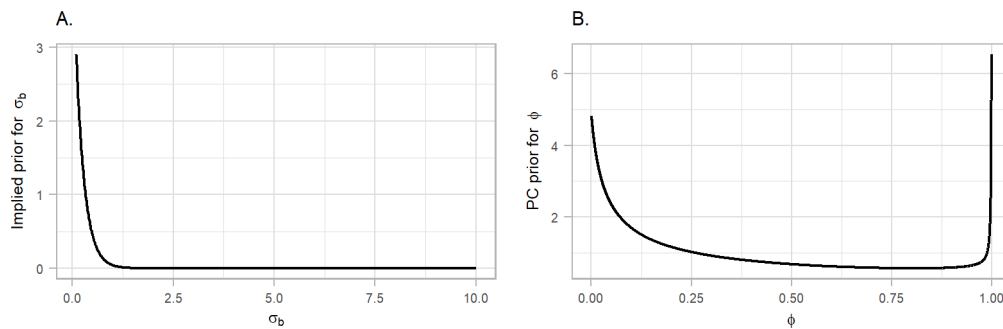
Figure 4: Penalised complexity (PC) priors for the hyperparameters of the spatial field. A) The implied PC prior for the standard deviation (as the original scale of the prior is the precision). B) The PC prior for the mixing parameter ϕ.

of this validation we used future years to train predictions for past, but in a previous study we also showed that the model selected was performing well when other types of cross-validations were used (Riou et al., 2023). For each stratum, we calculated the correlation between observed and fitted and a coverage probability, i.e. the probability that the observed death fall into the 95% credible interval (95% CrI) of the predicted.

4 Results

4.1 Cross-validation

Overall, we found that the model performed well in predicting the expected number of deaths. The correlation between true and expected number of deaths varied from 0.39 (95% CrI: 0.37, 0.40) in females 40< to 0.95 (95% CrI: 0.94, 0.95) in females 80> and coverage probability from 0.92 in females 40< to 0.96 in males 60-69, 70-79:

```
#              Correlation Coverage
# less40F 0.39 (0.37, 0.40)    0.92
# 40-59F  0.78 (0.77, 0.78)    0.95
# 60-69F  0.83 (0.82, 0.83)    0.95
# 70-79F  0.91 (0.90, 0.91)    0.95
# 80plusF 0.95 (0.94, 0.95)    0.95
# less40M 0.51 (0.50, 0.52)    0.93
# 40-59M  0.83 (0.83, 0.84)    0.95
# 60-69M  0.87 (0.86, 0.87)    0.96
# 70-79M  0.92 (0.91, 0.92)    0.96
# 80plusM 0.94 (0.94, 0.94)    0.95
```

4.2 Expected number of deaths

The object `pois.samples.list` contains 1000 samples from the posterior predictive distribution (2), i.e. 1000 samples of the expected number of deaths by age, sex, NUTS3 regions and week, had the pandemic not occurred. We can access the different age-sex groups as follows:

```
names(pois.samples.list)
# "F_less40" "F_40_59"  "F_60_69"  "F_70_79"  "F_80plus"
# "M_less40" "M_40_59"  "M_60_69"  "M_70_79"  "M_80plus"
```

where F stands for females and M for males across the different age groups. To get an idea about the structure of the data, we can use the `head()` function for females 60-69 and check the first 10 samples (V1 through to V10) of excess deaths for the first 6 weeks of 2020 in the 001 region (Torino)

```
pois.samples.list$F_60_69 %>%
select(paste0("V", 1:10), EURO_LABEL, ID_space, year) %>%
head()
#    V1 V2 V3 V4 V5 V6 V7 V8 V9 V10 EURO_LABEL ID_space year
```

```
# 262 22 18 17 15 17 18 13 19 16  17   2020-W01     001 2020
# 263 18 16 21 20 16 23 13 12 23  17   2020-W02     001 2020
# 264 17 23 20 26 12 16 12 19 17  13   2020-W03     001 2020
# 265 13  8 13 30 22 17 16 22 17  19   2020-W04     001 2020
# 266 14 20 15 15 19 18 23 14 12  18   2020-W05     001 2020
# 267 17 14 14  9 14 15 17 19 19  14   2020-W06     001 2020
```

We can also calculate median and 95% CrI expected number of deaths for this specific age-sex group across all years by NUTS3 region:

```
pois.samples.list$F_60_69 %>%
select(starts_with("V"), "ID_space") %>%
        group_by(ID_space) %>%
        summarise_all(sum) %>%
        rowwise(ID_space) %>%
        mutate(median = median(c_across(V1:V1000)),
LL = quantile(c_across(V1:V1000), probs= 0.025),
UL = quantile(c_across(V1:V1000), probs= 0.975)) %>%
        select(ID_space, median, LL, UL) %>%
        head()
```

```
# A tibble: 107 × 4
# Rowwise:   ID_space
# ID_space median    LL    UL
# <chr>     <dbl> <dbl> <dbl>
# 1 001        814   748   885.
# 2 002         72    55    90
# 3 003        139   116   167
# 4 004        212   183.  243.
# 5 005         86    67   107
# 6 006        178   150   208
# 7 007         46    33    62
# 8 008         86    68   107
# 9 009        113    93   135.
# 10 010       341   301.  380
# ... with 97 more rows
```

4.3 Excess mortality

The above results can be combined in different ways using the functions `get2020data()` and `get2020weeklydata()` to calculate excess mortality (in the R script functions.R). The function `get2020data()` aggregates over the entire country, NUTS2 regions, sex, age and time resulting in the object d, whereas the function `get2020weeklydata()` aggregates over the entire country, NUTS2 regions, sex and age but not over time resulting in the object d_week.

```
names(d)
# "province" "region"    "country"
names(d$province)
# "none"    "age"     "sex"     "agesex"
names(d$province$age)
# "40<"    "40-59" "60-69" "70-79" "80+"
names(d$province$sex)
# "F" "M"
names(d$province$agesex)
# "F40<"    "F40-59" "F60-69" "F70-79" "F80+"    "M40<"    "M40-59" "M60-69" "M70-79" "M80+"
```

Province stands for the NUTS3 regions (the resolution we used to fit the models), region for the NUTS2 (coarser than NUTS3, appropriate for policy making) and country for the nationwide aggregation. Within these aggregations users can select the option "none" being the total aggregation by age and sex, "age" by sex, "sex" by age and "agesex" refers to no age and sex aggregation. The objects d and d_week have similar structure and contain summary statistics for REM and NED and posterior probabilities of a positive REM or NED:

```
head(d$province$none)
# Simple feature collection with 6 features and 24 fields
```

```
# Geometry type: POLYGON
# Dimension:     XY
# Bounding box:  xmin: 6.626865 ymin: 44.06028 xmax: 9.21355 ymax: 45.95041
# CRS:           +proj=longlat +datum=WGS84
# ID_space SIGLA     DEN_UTS observed population mean.REM median.REM   sd.REM
# 1     001    TO      Torino    32478 2226317.2 21.37414   21.36545 1.116582
# 2     002    VC    Vercelli     3216  168761.7 32.27082   32.15534 3.133142
# 3     003    NO      Novara     5274  364529.4 23.62994   23.71569 2.220801
# 4     004    CN       Cuneo     8716  585479.3 20.16665   20.22069 1.742638
# 5     005    AT        Asti     3745  211437.1 26.37080   26.30691 2.660622
# 6     006    AL Alessandria     7916  416090.6 28.27706   28.21510 2.052881
# LL.REM   UL.REM exceedance.REM median.REM.cat exceedance.REM.cat median.pred
# 1 19.19396 23.50963             1          20%>         (0.95, 1]       26761
# 2 26.51331 38.74179             1          20%>         (0.95, 1]        2434
# 3 19.18442 27.97865             1          20%>         (0.95, 1]        4263
# 4 16.67966 23.54490             1          20%>         (0.95, 1]        7250
# 5 21.19545 31.77340             1          20%>         (0.95, 1]        2965
# 6 24.48449 32.33309             1          20%>         (0.95, 1]        6174
# LL.pred UL.pred mean.NED median.NED    sd.NED    LL.NED    UL.NED exceedance.NED
# 1   26296   27249 5717.153     5717.5 246.42447 5229.975 6182.075              1
# 2    2318    2543  783.265      782.5  57.49419  673.975  898.025              1
# 3    4121    4428 1006.667     1011.0  76.69947  848.925 1153.000              1
# 4    7055    7471 1461.215     1466.0 105.25269 1245.975 1661.075              1
# 5    2842    3092  780.189      780.0  62.31281  654.950  903.000              1
# 6    5982    6360 1743.404     1742.0  98.74248 1556.975 1934.125              1
# median.NED.cat exceedance.NED.cat                         geometry
# 1          1000>          (0.95, 1] POLYGON ((7.859044 45.59758...
# 2     [500, 1000)        (0.95, 1] POLYGON ((8.204465 45.93567...
# 3          1000>          (0.95, 1] POLYGON ((8.496878 45.83934...
# 4          1000>          (0.95, 1] POLYGON ((7.990897 44.82381...
# 5     [500, 1000)        (0.95, 1] POLYGON ((8.046805 45.12815...
# 6          1000>          (0.95, 1] POLYGON ((8.405489 45.20148...
```

Notice that the object d$province$none is a simple feature collection, making mapping it straightforward using the ggplot2 package and geom_sf function. In Figure 5 (plots 1A, 2A and 3A) we show the median posterior of REM for total age and sex at the national, NUTS2 and NUTS3 regional level. For these plots we used the object d with the selection "none" and plot the median REM (median.REM), for example for 3A:

```
# prov <- "Foggia"
# d$province$none %>%
#        filter(DEN_UTS == prov) %>%
#        select(geometry) %>%
#        ggplot() +
#                geom_sf(data = d$province$none, aes(fill = median.REM.cat)) +
#                geom_sf(fill = NA, col = col.highlight, size = .8) +
#                scale_fill_manual(values=colors, name = "", drop = FALSE) +
#        theme_light() + ggtitle(paste0("3A. NUTS3 regions: ", prov))
```

Overall, the REM in Italy during 2020 was between 15-20%, meaning that 15-20% more people died that year than how many would be expected based on historical data, see Figure 5, panel 1A. When the higher geographical resolution is assessed, it was revealed that north and in particular Lombardia was the region hit the worst, with the REM exceeding 20%, see Figure 5, panels 1B and 1C. Figure 6 shows a measure of uncertainty of the REM, now for the different age groups and both sexes (selection "age"). The probability of a positive excess (exceedance.REM) in older people was larger than 0.95 almost everywhere, see Figure 6.

Panels 1B and 1C of Figure 5 show the median temporal nationwide excess mortality together with 95% CrI by sex after aggregating the different age groups (using d_week and the "sex" selection). We observe a clear first pandemic wave during March and May and a second one during mid October and December in 2020. During the first pandemic wave, there were weeks when the median REM reached almost 100% in males, Figure 5.
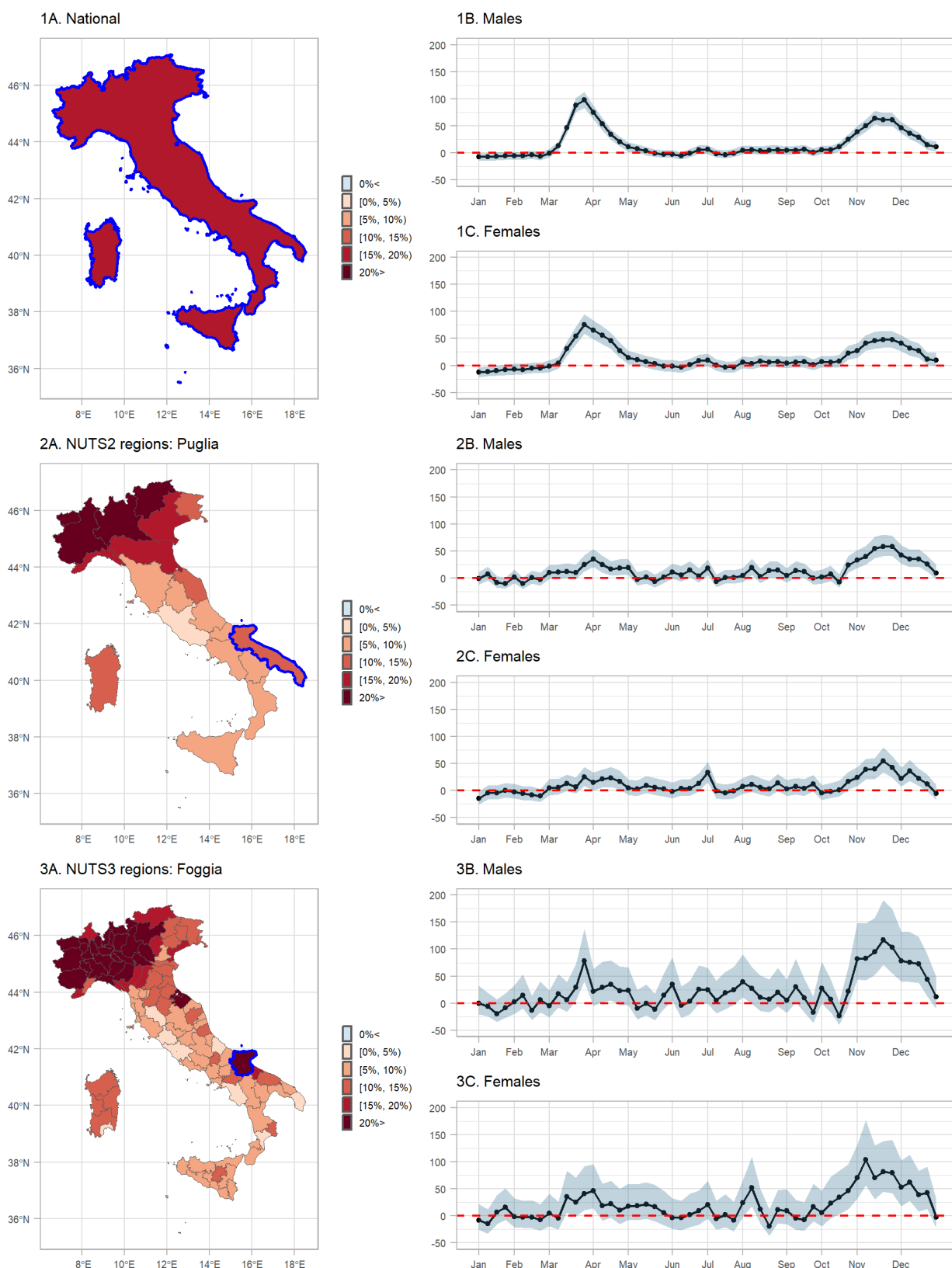
Figure 5: Median relative excess mortality by spatial region during 2020: 1A) nationwide, 2A) NUTS2 and 3A) NUTS3 level, and weekly median relative excess mortality and 95% credible intervals (95% probability that the true value lies within this interval) by sex during 2020 in: 1B) males nationwide, 1C) females nationwide, 2B) males in Puglia, 2C) females in Puglia, 3B) males in Foggia and 3C) females in Foggia.

Panels 2B, 2C, 3B and 3C of Figure 5 show the median spatio-temporal excess mortality together with 95% CrI by sex after aggregating over the different age groups. Panels 2B and 2C highlight the region of Puglia, where during the first wave of the pandemic experienced increases excess mortality. This increase follows similar trends as the nationwide excess mortality (Panels 1B and 1C). When we increase the spatial resolution in panels 3B and 3C, we highlight the province of Foggia, where there was insufficient evidence of a positive excess during the first wave, but strong during the second.
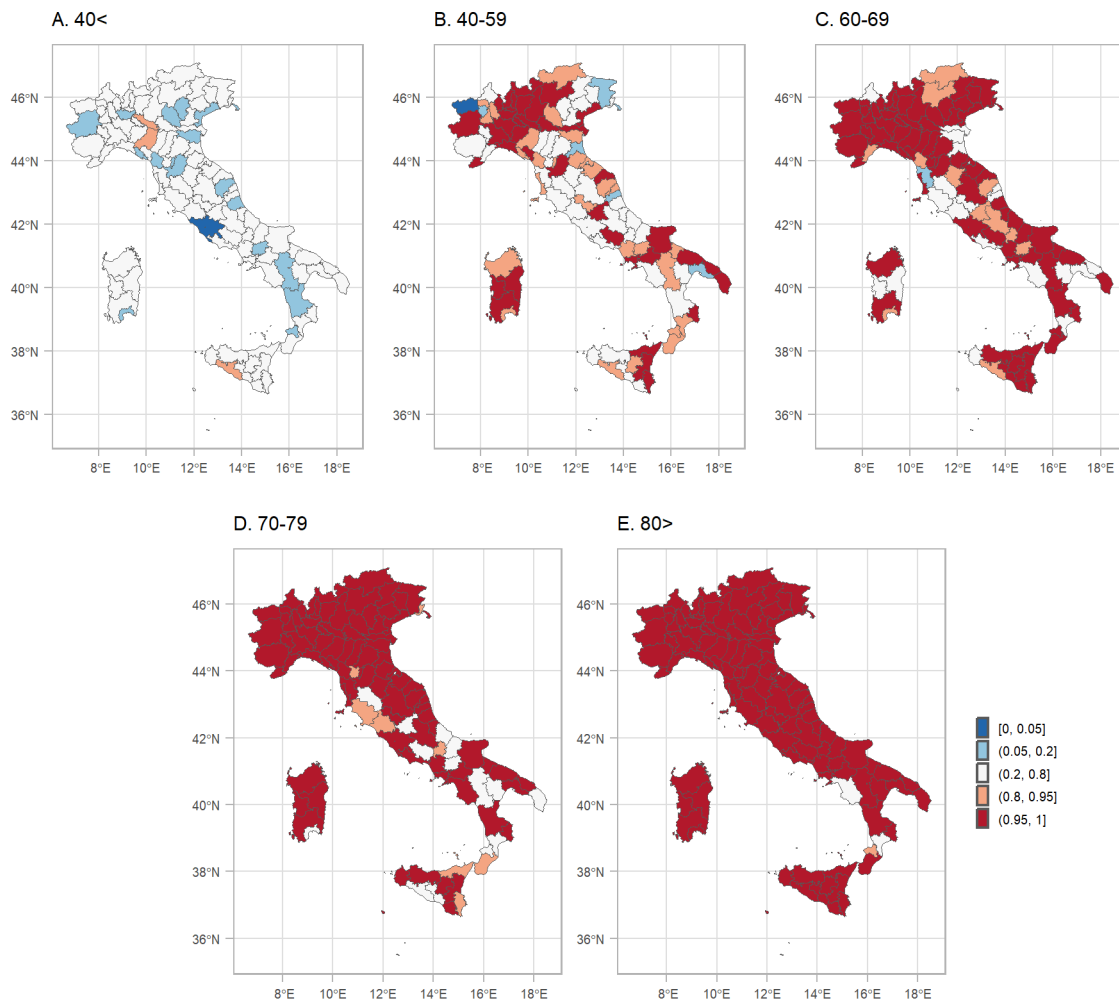
Figure 6: Posterior probability that the relative excess mortality is positive for both sexes during 2020 by age group and NUTS3 region.

4.4 Shiny Web-Application

To be able to effectively examine and communicate the different aggregation levels of the output of our modelling framework, we have also developed a Shiny Web-Application (WebApp), Figure 7. The WebApp provides spatial, temporal and spatio-temporal analysis tabs, and within each tab there are plots and summary statistics for the level of aggregation selected from the drop-down menu. Users can select across different variables (REM or NED), statistics (median or posterior probability), sex (males, females or both), age group ($40 <$, $40 - 59$, $60 - 69$, $70 - 79$, $80 >$ and all) and different geographical level (national, NUTS2 or NUTS3 regions). Summary statistics for each area are available and they are displayed in a pop-up window, which is activated by clicking on the area of interest. In addition, graphical pop-ups are provided to show the weekly estimates for each area with the `leafpop` R-package (Appelhans and Detsch, 2021), in the spatio-temporal analysis tab.
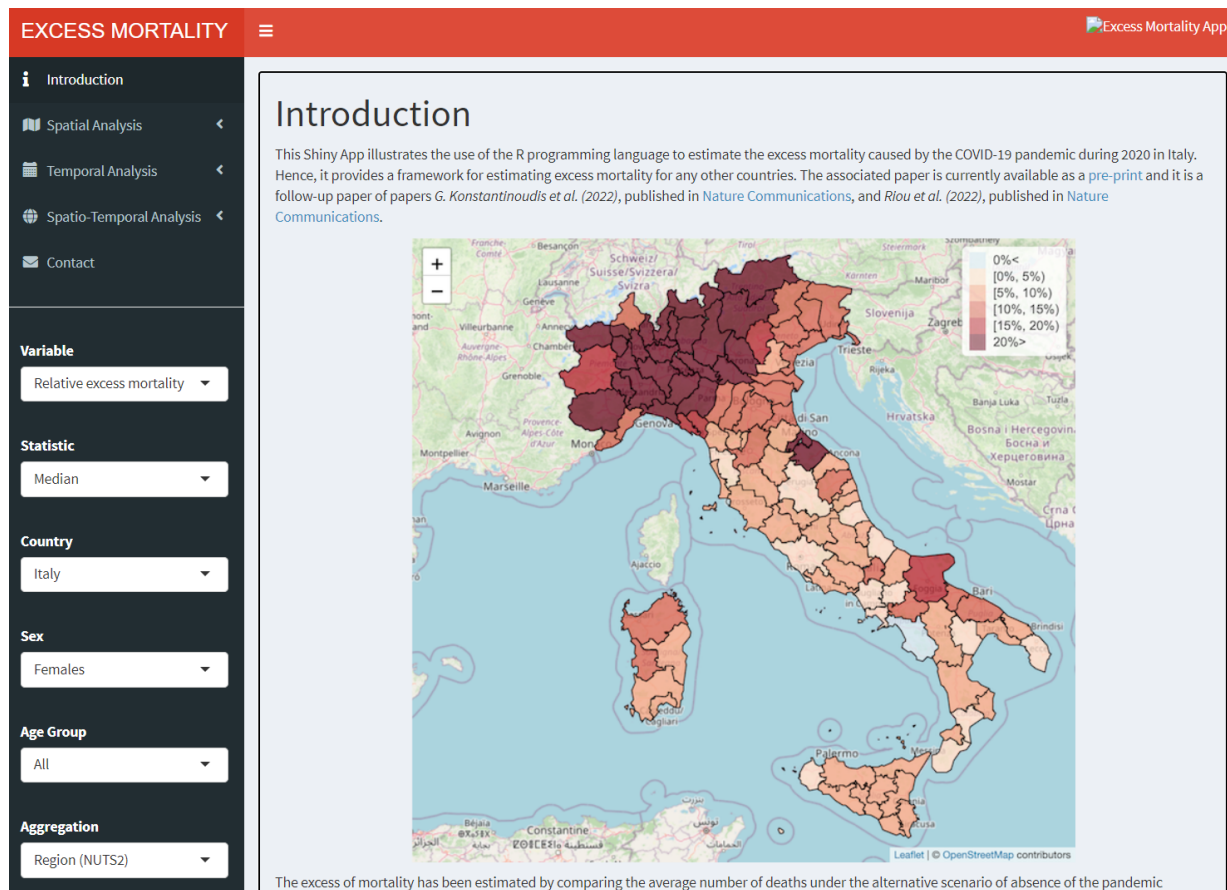
The WebApp that we have developed is hosted at http://atlasmortalidad.uclm.es/italyexcess/.

Figure 7: Shiny App developed to navigate through the excess mortality estimates during 2020 in Italy across the different aggregations.

5   Summary

This tutorial provides a detailed explanation of the workflow used previously to calculate excess mortality during the COVID-19 pandemic in 5 European regions (Konstantinoudis et al., 2022). The main model used here is slightly modified based on updated results (Riou et al., 2023). We have proposed a Bayesian workflow for estimating excess mortality and shown how to use R and INLA to retrieve fast and accurate estimates. The proposed workflow also allows for combining different models and presenting the results stratified by age, sex, spatial and temporal location. We have given a practical example of how to use the proposed framework to model the excess mortality during the 2020 COVID-19 pandemic in Italy at small area level. We also developed a Shiny App to effectively communicate the results. The methodological framework can be extended to monitor excess mortality caused by other extreme events; for instance, natural hazards such as tropical cyclones (Parks et al., 2022) or heatwaves (Konstantinoudis et al., 2023). Potential methodological extensions of the proposed framework also include modelling the younger age groups with a zero-inflated Poisson distribution. The proposed framework can also be extended to provide an automated tool for online disease surveillance and policy making.

Acknowledgements

Author contributions

V.G.R. conceived the study. M.B. supervised the study. G.K. developed the initial study protocol and discussed it with M.B., M.C., M.P. and G.B.. G.K. developed the statistical model, prepared the population and covariate data and led the acquisition of mortality data. M.C. and V.G.R. validated and modified accordingly the code. G.K. ran the analysis. G.K., V.G.R., M.B., M.C. and M.P. wrote the initial draft and all the authors contributed in modifying the paper and critically interpreting the results. V.G.R. developed the Shiny app. All authors read and approved the final version for publication.

References

J. M. Aburto, R. Kashyap, J. Schöley, C. Angus, J. Ermisch, M. C. Mills, and J. B. Dowd. Estimating the burden of the COVID-19 pandemic on mortality, life expectancy and lifespan inequality in England and Wales: A population-level analysis. *J Epidemiol Community Health*, 2021. [p89]

T. Appelhans and F. Detsch. *leafpop: Include Tables, Images and Graphs in Leaflet Pop-Ups*, 2021. URL https://CRAN.R-project.org/package=leafpop. R package version 0.1.0. [p90, 99]

D. Attali and T. Edwards. *shinyalert: Easily Create Pretty Popup Messages (Modals) in 'Shiny'*, 2021. URL https://CRAN.R-project.org/package=shinyalert. R package version 3.0.0. [p90]

J. Besag. Spatial interactions and the statistical analysis of lattice systems (with discussion). *Journal of the Royal Statistical Society: Series B*, 36:192–236, 1974. [p91]

J. Besag, J. York, and A. Mollié. Bayesian image restoration, with two applications in spatial statistics. *Annals of the Institute of Statistical Mathematics*, 43(1):1–20, 1991. [p91]

R. Bivand. *R Packages for Analyzing Spatial Data: A Comparative Case Study with Areal Data*, 2022. [p90]

R. Bivand, T. Keitt, and B. Rowlingson. *rgdal: Bindings for the 'Geospatial' Data Abstraction Library*, 2023. URL https://CRAN.R-project.org/package=rgdal. R package version 1.6-4. [p90]

M. Blangiardo, M. Cameletti, M. Pirani, G. Corsetti, M. Battaglini, and G. Baio. Estimating weekly excess mortality at sub-national level in Italy during the COVID-19 pandemic. *PLOS ONE*, 15(10): 1–15, 10 2020. doi: 10.1371/journal.pone.0240286. URL https://doi.org/10.1371/journal.pone.0240286. [p89]

W. Chang and B. Borges Ribeiro. *shinydashboard: Create Dashboards with 'Shiny'*, 2021. URL https://CRAN.R-project.org/package=shinydashboard. R package version 0.7.2. [p90]

W. Chang, J. Cheng, J. Allaire, C. Sievert, B. Schloerke, Y. Xie, J. Allen, J. McPherson, A. Dipert, and B. Borges. *shiny: Web Application Framework for R*, 2022. URL https://CRAN.R-project.org/package=shiny. R package version 1.7.4. [p90]

B. Davies, B. L. Parkes, J. Bennett, D. Fecht, M. Blangiardo, M. Ezzati, and P. Elliott. Community factors and excess mortality in first wave of the COVID-19 pandemic in England. *Nature Communications*, 12(1):1–9, 2021. [p92]

M. Dowle and A. Srinivasan. *data.table: Extension of 'data.frame'*, 2022. URL https://CRAN.R-project.org/package=data.table. R package version 1.14.6. [p90]

S. Garnier, N. Ross, R. Rudis, P. A. Camargo, M. Sciaini, and C. Scherer. *viridis - Colorblind-Friendly Color Maps for R*, 2021. URL https://sjmgarnier.github.io/viridis/. R package version 0.6.2. [p90]

G. Grolemund and H. Wickham. *Dates and Times Made Easy with lubridate*, 2011. URL https://www.jstatsoft.org/v40/i03/. [p90]

H. Hersbach, B. Bell, P. Berrisford, S. Hirahara, A. Horányi, J. Muñoz Sabater, J. Nicolas, C. Peubey, R. Radu, D. Schepers, et al. The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730):1999–2049, 2020. [p94]

R. J. Hijmans. *raster: Geographic Data Analysis and Modeling*, 2023. URL https://CRAN.R-project.org/package=raster. R package version 3.6-14. [p90]

K. Hufkens, R. Stauffer, and E. Campitelli. *The ecwmfr package: An interface to ECMWF API endpoints*, 2019. URL https://bluegreen-labs.github.io/ecmwfr/. [p90]

N. Islam, V. M. Shkolnikov, R. J. Acosta, I. Klimkin, I. Kawachi, R. A. Irizarry, G. Alicandro, K. Khunti, T. Yates, D. A. Jdanov, et al. Excess deaths associated with COVID-19 pandemic in 2020: age and sex disaggregated time series analysis in 29 high income countries. *BMJ*, 373, 2021. [p89]

J. Kaczorowski and C. Del Grande. Beyond the tip of the iceberg: direct and indirect effects of COVID-19. *The Lancet Digital Health*, 3(4):e205–e206, 2021. [p89]

G. Konstantinoudis, D. Schuhmacher, H. Rue, and B. D. Spycher. Discrete versus continuous domain models for disease mapping. *Spatial and Spatio-temporal Epidemiology*, 32:100319, 2020. ISSN 1877-5845. doi: https://doi.org/10.1016/j.sste.2019.100319. URL https://www.sciencedirect.com/science/article/pii/S1877584519301297. [p91]

G. Konstantinoudis, M. Cameletti, V. Gómez-Rubio, I. L. Gómez, M. Pirani, G. Baio, A. Larrauri, J. Riou, M. Egger, P. Vineis, et al. Regional excess mortality during the 2020 COVID-19 pandemic in five European countries. *Nature communications*, 13(1):482, 2022. [p89, 90, 100]

G. Konstantinoudis, A. Hauser, and J. Riou. Bayesian ensemble modelling to monitor excess deaths during summer 2022 in Switzerland. *arXiv preprint arXiv:2308.15251*, 2023. [p100]

V. Kontis, J. E. Bennett, T. Rashid, R. M. Parks, J. Pearson-Stuttard, M. Guillot, P. Asaria, B. Zhou, M. Battaglini, G. Corsetti, et al. Magnitude, demographics and dynamics of the effect of the first wave of the COVID-19 pandemic on all-cause mortality in 21 industrialized countries. *Nature Medicine*, pages 1–10, 2020. [p89]

V. Kontis, J. Bennett, R. Parks, T. Rashid, J. Pearson-Stuttard, P. Asaria, B. Zhou, M. Guillot, C. Mathers, Y. Khang, M. McKee, and M. Ezzati. Lessons learned and lessons missed: impact of the coronavirus disease 2019 (COVID-19) pandemic on all-cause mortality in 40 industrialised countries prior to mass vaccination [version 1; peer review: 2 approved with reservations]. *Wellcome Open Research*, 6 (279), 2021. doi: 10.12688/wellcomeopenres.17253.1. [p89]

E. Kontopantelis, M. A. Mamas, J. Deanfield, M. Asaria, and T. Doran. Excess mortality in England and Wales during the first wave of the COVID-19 pandemic. *J Epidemiol Community Health*, 75(3): 213–223, 2021. [p89]

P. Moraga. Small Area Disease Risk Estimation and Visualization Using R. *The R Journal*, 10(1):495–506, 2018. doi: 10.32614/RJ-2018-036. URL https://doi.org/10.32614/RJ-2018-036. [p89]

E. Neuwirth. *RColorBrewer: ColorBrewer Palettes*, 2022. URL https://CRAN.R-project.org/package=RColorBrewer. R package version 1.1-3. [p90]

R. M. Parks, J. Benavides, G. B. Anderson, R. C. Nethery, A. Navas-Acien, F. Dominici, M. Ezzati, and M.-A. Kioumourtzoglou. Association of tropical cyclones with county-level mortality in the US. *JAMA*, 327(10):946–955, 2022. [p100]

E. Pebesma. Simple Features for R: Standardized Support for Spatial Vector Data. *The R Journal*, 10(1): 439–446, 2018. doi: 10.32614/RJ-2018-009. URL https://doi.org/10.32614/RJ-2018-009. [p90]

T. L. Pedersen. *patchwork: The Composer of Plots*, 2022. URL https://CRAN.R-project.org/package=patchwork. R package version 1.1.2. [p90]

D. Pierce. *ncdf4: Interface to Unidata netCDF (Version 4 or Earlier) Format Data Files*, 2023. URL https://CRAN.R-project.org/package=ncdf4. R package version 1.21. [p90]

T. Plate and R. Heiberger. *abind: Combine Multidimensional Arrays*, 2016. URL https://CRAN.R-project.org/package=abind. R package version 1.4-5. [p90]

A. Riebler, S. H. Sørbye, D. Simpson, and H. Rue. An intuitive Bayesian spatial model for disease mapping that accounts for scaling. *Statistical Methods in Medical Research*, 25(4):1145–1165, 2016. [p91]

J. Riou, A. Hauser, A. Fesser, C. L. Althaus, M. Egger, and G. Konstantinoudis. Direct and indirect effects of the COVID-19 pandemic on mortality in Switzerland. *Nature communications*, 14(1):90, 2023. [p90, 93, 95, 100]

L. M. Rossen, A. M. Branum, F. B. Ahmad, P. Sutton, and R. N. Anderson. Excess deaths associated with COVID-19, by age and race and ethnicity – United States, January 26–October 3, 2020. *Morbidity and Mortality Weekly Report*, 69(42):1522, 2020. [p89]

H. Rue and L. Held. *Gaussian Markov random fields: theory and applications*. CRC press, 2005. [p92]

H. Rue, S. Martino, and N. Chopin. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *Journal of the Royal Statistical Society: Series B (statistical methodology)*, 71(2):319–392, 2009. [p89, 92]

D. Simpson, H. Rue, A. Riebler, T. G. Martins, S. H. Sørbye, et al. Penalising model component complexity: A principled, practical approach to constructing priors. *Statistical Science*, 32(1):1–28, 2017. [p91]

J. Verbeeck, C. Faes, T. Neyens, N. Hens, G. Verbeke, P. Deboosere, and G. Molenberghs. A linear mixed model to estimate COVID-19-induced excess mortality. *Biometrics*, pages 1–9, 2021. doi: https://doi.org/10.1111/biom.13578. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/biom.13578. [p89]

L. A. Waller, B. P. Carlin, H. Xia, and A. E. Gelfand. Hierarchical spatio-temporal mapping of disease rates. *Journal of the American Statistical association*, 92(438):607–617, 1997. [p89]

D. M. Weinberger, J. Chen, T. Cohen, F. W. Crawford, F. Mostashari, D. Olson, V. E. Pitzer, N. G. Reich, M. Russi, L. Simonsen, et al. Estimation of excess deaths associated with the COVID-19 pandemic in the United States, March to May 2020. *JAMA Internal Medicine*, 180(10):1336–1344, 2020. [p89]

H. Wickham. *Reshaping Data with the reshape Package*, 2007. URL http://www.jstatsoft.org/v21/i12/. [p90]

H. Wickham. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York, 2016. ISBN 978-3-319-24277-4. URL https://ggplot2.tidyverse.org. [p90]

H. Wickham. *stringr: Simple, Consistent Wrappers for Common String Operations*, 2022. URL https://CRAN.R-project.org/package=stringr. R package version 1.5.0. [p90]

H. Wickham, J. Hester, and J. Bryan. *readr: Read Rectangular Text Data*, 2022. URL https://CRAN.R-project.org/package=readr. R package version 2.1.3. [p90]

H. Wickham, R. François, L. Henry, K. Müller, and D. Vaughan. *dplyr: A Grammar of Data Manipulation*, 2023a. URL https://CRAN.R-project.org/package=dplyr. R package version 1.1.0. [p90]

H. Wickham, D. Vaughan, and M. Girlich. *tidyr: Tidy Messy Data*, 2023b. URL https://CRAN.R-project.org/package=tidyr. R package version 1.3.0. [p90]

D. Wuertz, T. Setz, Y. Chalabi, and G. N. Boshnakov. *timeDate: Rmetrics - Chronological and Calendar Objects*, 2023. URL https://CRAN.R-project.org/package=timeDate. R package version 4022.108. [p90]

Garyfallos Konstantinoudis
MRC Centre for Environment and Health, Imperial College London,
St Mary's Campus, Praed St,
W2 1NY, London
United Kingdom

Virgilio Gómez-Rubio
Departamento de Matemáticas, Escuela Técnica Superior de Ingeniería Industrial-Albacete, Universidad de Castilla-La Mancha,
Albacete, Spain

Michela Cameletti
Department of Economics, University of Bergamo,

Bergamo, Italy

Monica Pirani
MRC Centre for Environment and Health, Imperial College London,
St Mary's Campus, Praed St,
W2 1NY, London
United Kingdom

Gianluca Baio
Department of Statistical Sciences, University College London,
United Kingdom

Marta Blangiardo
MRC Centre for Environment and Health, Imperial College London,
St Mary's Campus, Praed St,
W2 1NY, London
United Kingdom