Contextualizing the Global Burden of COVID-19 Pandemic

A Historical and Geographical Exploration of Excess Mortality in France, 1901-2022

Carlo Giovanni Camarda* and Florian Bonnet

Institut national d'études démographiques (INED), Aubervilliers

Extended abstract for the EPC 2024 conference

November 1, 2023

Abstract

In this paper, we propose (1) to evaluate the global burden of COVID-19 pandemic in France in 2020 and 2021 at the finest geographic level used by Eurostat in the Nomenclature of Territorial Units for Statistics, and (2) to compare this burden with other mortality crises which occurred during the 20^{th} and 21^{st} centuries. We leverage an extensive dataset comprising a lengthy time series of mortality data sourced from reliable records, stratified by age-groups and gender, for nearly 90 French *départements*. Our analytical approach involves modeling this dataset using a non-parametric approach (P-splines) within a Composite Link Model framework, while intentionally excluding years that are unequivocally identified as crises in relevant literature. The impact and burden of each crisis are subsequently assessed by computing age-specific differences between observed mortality rates and those predicted by our models. While our approach allows for the computation of a wide range of demographic indicators to evaluate the impact of each crisis, we focus on results for life expectancy at birth (e_0) and at age 65 (e_{65}). Our results reveal geographic hot spots for each crisis as well as differences in magnitude between crises. We plan to extend these results to the period 1872-1900, compute confidences intervals around our excess mortality measures, and endogenize the choice of crisis years on the basis of our modeling.

Keywords: COVID-19 · Mortality crises · Historical demography

*Corresponding author: carlo-giovanni.camarda@ined.fr
Address: 9 cours des Humanités, 93322 Aubervilliers - France

1 Introduction

The worldwide COVID-19 pandemic has compelled researchers and policymakers to seek answers to the myriad questions it poses. Among them, the question of its global mortality burden is especially acute. To do so, scholars are mainly calculating excess mortality, defined as the difference between the number of deaths (from any cause) that occur during the pandemic and the number of deaths that would have occurred in the absence of the pandemic. These are considered to be the gold standard for estimating the global impact of the pandemic.

Comparing excess mortality due to COVID-19 with other historical mortality crises is a valuable analytical tool for placing the current pandemic in a broader historical and public health context. Such comparisons offer insights into the uniqueness and severity of COVID-19's impact by drawing parallels or distinctions with past crises such as influenza pandemics, major wars, or other epidemics. This comparative approach helps researchers and policymakers better understand the relative gravity of the pandemic, assess the efficacy of response measures, and identify patterns that might inform future preparedness efforts.

Furthermore, computing excess mortality due to COVID-19 at a fine geographic level, as opposed to a national level, is of paramount importance in understanding the pandemic's true impact. Fine-grained geographic analysis allows for a more accurate and nuanced assessment of the disparities and variations in mortality rates, which are often masked when considering national aggregates. This approach helps to identify localized hot spots, demographic factors, healthcare resource distribution, and public health interventions that are more effective in controlling the disease. The implications for different age groups are equally paramount, as understanding how various generations are affected unveils essential nuances in the societal response.

In this paper, we propose to embark on a journey through time and across geographical boundaries within a French context, juxtaposing the COVID-19 pandemic with historical crises to glean a comprehensive perspective on the challenges and adaptations that shape our world in times of emergency. To achieve this, we leverage an extensive dataset comprising a lengthy time series of mortality data, stratified by age-groups and gender, for nearly 90 French *départements*. These data are sourced from reliable records dating back to the early 20^{th} century. Our analytical approach involves modeling this dataset, intentionally excluding years that are unequivocally identified as crises in relevant literature. The impact and burden of each crisis are subsequently assessed by computing age-specific differences between observed mortality rates and those predicted by our models.

2 Data and Methods

To examine the range of mortality crises over the past century, we rely on data spanning from 1901 to 2022. Our analysis involves death counts categorized by age-groups, which have evolved over time. In the initial years, various age-grouping structures were employed. From 1968 onwards, we obtained death counts by individual ages. Additionally, we used population counts by individual ages for the 1st of January of each year, estimated by (Bonnet, 2020) using methodologies akin to those employed by the Human Mortality Database.

From a methodological standpoint, there are two challenges in modeling these data. First, we require a versatile model to estimate the underlying force of mortality across both age and time, enabling the removal of certain years from the fitting process without affecting the remaining estimated trends. The use of P-splines enables us to achieve smoothness across both age and time in a regression context, without the need to impose a predefined model structure. This flexibility allows for the adjustment of estimated trends to align with the observed data (Currie et al., 2004). Additionally, the years of crises can be conveniently excluded from the estimation process by incorporating regression weights. Given that our data encompasses age 0, we have customized the approach outlined in the work by Camarda (2019), which introduces a specialized coefficient for infant mortality.

Secondly, we need a framework to accommodate the time-varying age-grouping structure in our mortality data. In this study, we opt to incorporate a non-parametric approach like P-splines within a Composite Link Model framework. Building upon the seminal research by Eilers (2007), recent advancements have been introduced to model mortality based on coarsely grouped data within the framework of Composite Link Models (Rizzi et al., 2018). In our study, we extend this approach to accommodate varying grouping structures in a two-dimensional context, resulting in the ultimate estimates for all years, broken down into single years of ages.

By integrating these two methodologies, we can accurately calculate excess mortality for every year of crisis and across all age groups by taking the difference between the estimated and observed levels. While our approach allows for the computation of a wide range of demographic indicators to evaluate the impact of each crisis, due to space limitations, we will focus on presenting results for life expectancy at birth (e_0) and at age 65 (e_{65}).

In this proposal, we have defined mortality crises based on the existing literature; we have identified four major mortality crises that have strongly affected France over a short period since 1901, in addition to the COVID-19 pandemic of 2020 and 2021 (Bonnet and Camarda, 2022). These were the heatwave of 1911 (Rollet, 2010), the Spanish flu of 1918 and 1919, the Hong Kong flu of 1968 and 1969 (Meslé, 2010), and the heatwave of 2003 (Robine et al., 2008). Furthermore, without considering them as mortality crisis years, we did not consider the war years 1914–1917 and 1939–1945 in our modeling.

Finally, we have restricted our analysis for this proposal to the 84 *départements* for which the dataset begins in 1901 and whose geographical boundaries have not changed between 1901 and 2022. In particular, we have excluded the *départements* of north-eastern France, for which data is available only since 1921, and the *départements* of the Paris region, for which the boundaries were reshaped in 1968. For now, we have also restricted our analysis to female mortality.

3 Preliminary results

For illustrative purposes, Figure 1 displays both observed and estimated values of life expectancy at birth (e_0) and at age 65 (e_{65}) for three specific French *departements*: Aveyron, Mayenne, and Rhône. It is evident that years identified as crises exhibit a lower life expectancy, although each geographical unit responds differently to each type of crisis. Furthermore, each of these crises predominantly affected specific age groups. Therefore, declines in life expectancy at birth may obscure specific mortality shocks that primarily impacted higher age groups. Conversely, disparities in e_{65} can serve as indicators of the impact of crises on older age groups. Specifically, e_0 in Aveyron and Mayenne were profoundly affected by the heatwave of 1911, which was not true in Rhône. In 1918, Rhône was hugely affected by the influenza: loss in e_0 was around 15 years. Interestingly, for these three *départements*, losses in life expectancy were higher at age 0 than at age 65 in 1911 and 1918-1919, but higher at age 65 than at age 0 in 1967-1968, 2003 and 2020-2021.

Figure 2 presents the losses in e_0 and e_{65} for all the *départements* in 1911, 1918 and 2020. For clarity reasons, a positive value means that fitted life expectancy was higher than observed life expectancy. Regarding e_0, it is particularly interesting to note that the impact of COVID-19 was not comparable to that of the heatwave of 1911 or the Spanish flu of 1918: all *départements* suffered a loss of e_0 ranging from 0 to 2 years, whereas the loss of e_0 ranged from 8 to 14 years for the Spanish flu in 1918 (with the highest values in the south of France), and from 2 to 8 years for the *départements* in the South and far North of France during the heatwave of 1911. Regarding e_{65}, the results are much more comparable for the three mortality crises, with losses in e_{65} ranging in most cases between 0.5 and 1.5 years. Losses were the highest during the Spanish flu of 1918, and very different from one territory to another during the COVID-19 pandemic.

4 Future research

During the next months, we plan to extend our results in several directions. First, we plan to present our results for both men and women, and for all French *départements*, even if some of them do not have mortality and population data for the entire 1901-2022 period. In addition, we plan to extend our analysis to the period 1872-1900, for which we have digitized data but of lesser quality.

Second, we plan to compute confidence intervals around our excess mortality measures. Calculating these confidence intervals is crucial for robust and reliable data interpretation: fine-grained analyses often involve small populations, making data susceptible to increased variability. By establishing confidence intervals, researchers can quantify the uncertainty associated with their estimates, providing a range within which the true excess mortality likely falls. This not only aids in statistical rigor but also allows for meaningful comparisons between different geographic areas and years, helping to identify significant differences and trends while accounting for inherent fluctuations.

Third, we plan to change the way we identify crisis years. At this stage, these are chosen exogenously, in line with what has been revealed in the literature, based on analyses at national level. However, such analyses may mask even larger mortality shocks, but contained within small regions. We therefore plan to endogenize the choice of crisis years on the basis of our modeling.

References

Bonnet, F. (2020). Computations of French lifetables by department, 1901–2014. *Demographic Research*, 42:741–762.

Bonnet, F. and Camarda, C. G. (2022). Estimating Subnational Excess Mortality in Times of Pandemic. An application to French *départements* in 2020. Pre-print available on medRxiv: `https://www.medrxiv.org/content/10.1101/2022.12.12.22283346v1`.

Camarda, C. G. (2019). Smooth Constrained Mortality Forecasting. *Demographic Research*, 41(38):1091–1130.

Currie, I. D., Durbán, M., and Eilers, P. H. C. (2004). Smoothing and Forecasting Mortality Rates. *Statistical Modelling*, 4:279–298.

Eilers, P. H. C. (2007). Ill-posed Problems with Counts, the Composite Link Model and Penalized Likelihood. *Statistical Modelling*, 7:239–254.

Meslé, F. (2010). A spectacular decline in influenza mortality: the role of vaccination. *Population Societies*, 470(8):1–4.

Rizzi, S., Halekoh, U., Thinggaard, M., Engholm, G., Christensen, N., Johannesen, T. B., and Lindahl-Jacobsen, R. (2018). How to estimate mortality trends from grouped vital statistics. *International Journal of Epidemiology*, 48:571–582.

Robine, J.-M., Cheung, S. L. K., Le Roy, S., Van Oyen, H., Griffiths, C., Michel, J.-P., and Herrmann, F. R. (2008). Death toll exceeded 70,000 in europe during the summer of 2003. *Comptes rendus biologies*, 331(2):171–178.

Rollet, C. (2010). La canicule de 1911. observations démographiques et médicales et réactions politiques 1. In *Annales de démographie historique*, number 2, pages 105–130. Cairn/Softwin.

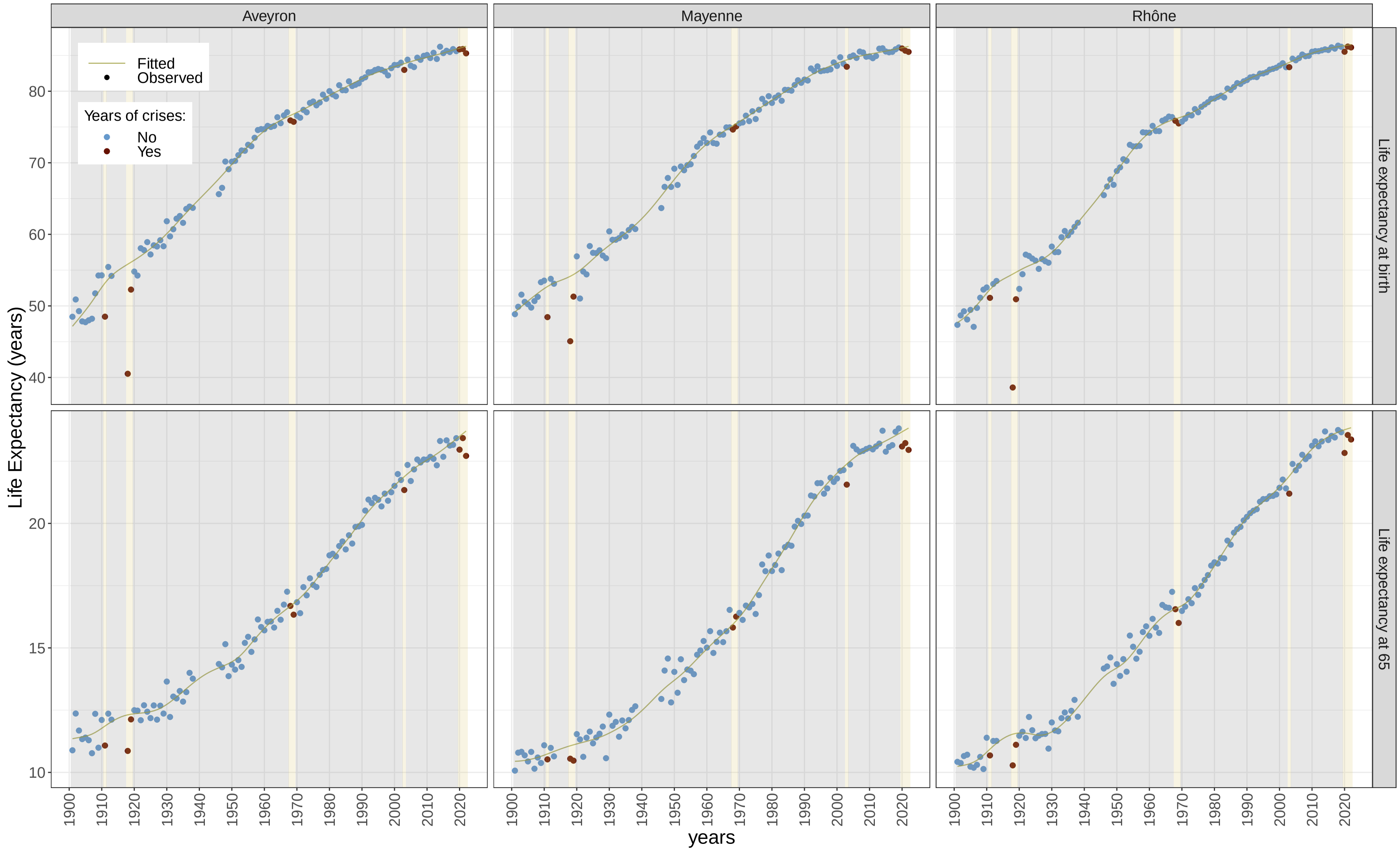

Figure 1. Observed and fitted life expectancy at birth and age 65 in three specific French *départements*: Aveyron, Mayenne, and Rhône. Fitted values obtained by CP-splines over age(-groups) and years, omitting years identified as experiencing known mortality crises (green dots) as well as war years.

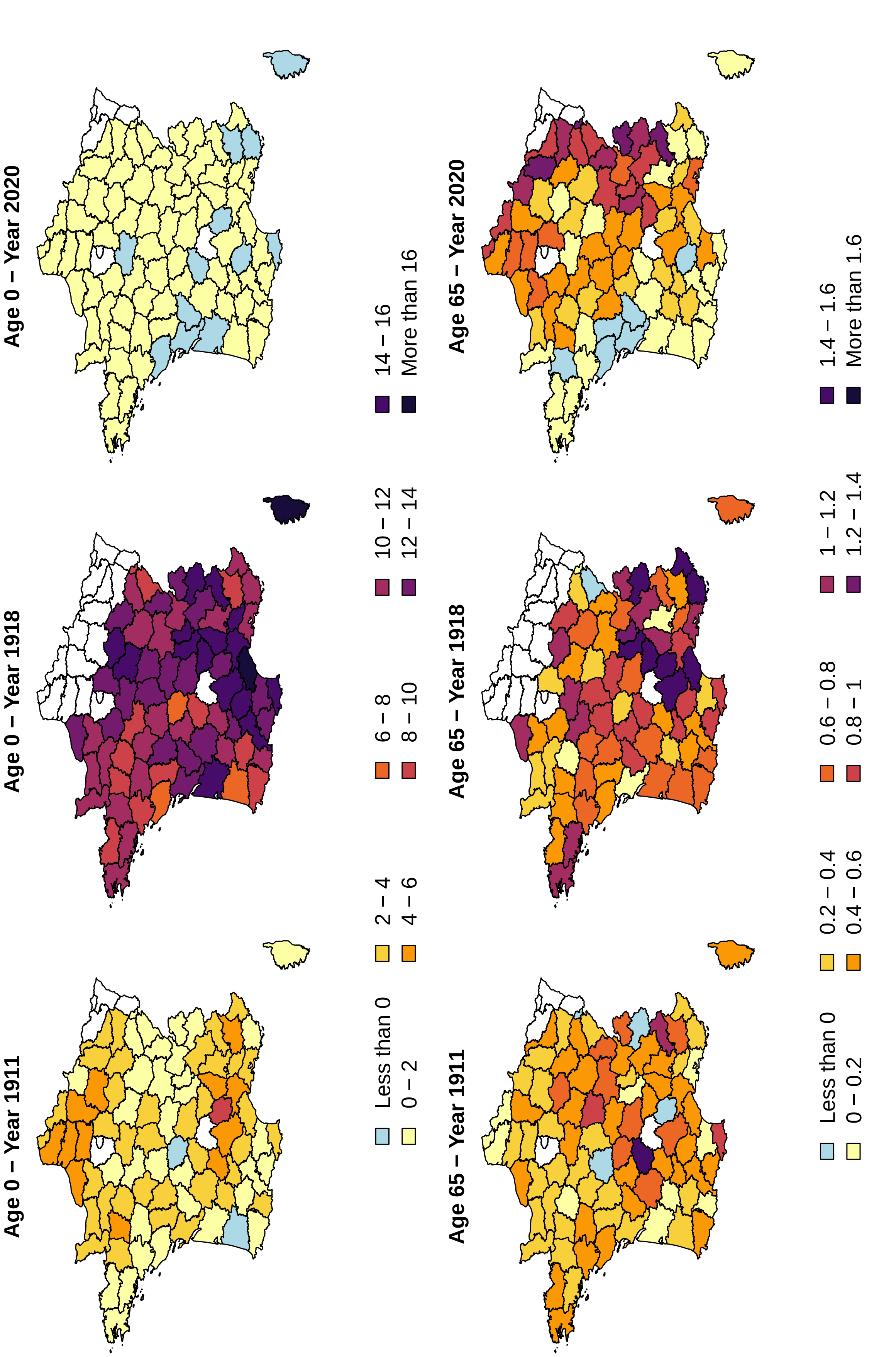

Figure 2. Excess mortality (in years) computed using life expectancy at birth and at age 65 for 3 years of mortality crises (1911, 1918, 2020).