Annals of Actuarial Science (2026), **20**, pp. 22–53
doi:10.1017/S1748499525000089

ORIGINAL RESEARCH PAPER

Analyzing state-level longevity trends with the U.S. mortality database

Mike Ludkovski and Doris Padilla

Department of Statistics and Applied Probability, University of California, Santa Barbara, CA, USA
Corresponding author: Mike Ludkovski; Email: ludkovski@pstat.ucsb.edu

(Received 09 August 2024; revised 21 April 2025; accepted 21 April 2025; first published online 6 August 2025)

Abstract

We investigate state-level age-specific mortality trends based on the United States Mortality Database (USMDB) published by the Human Mortality Database. In tandem with looking at the longevity experience across all the states, we also consider a collection of socio-demographic, economic, and educational covariates that correlate with mortality trends. To obtain smoothed mortality surfaces for each state, we implement the machine learning framework of Multi-Output Gaussian Process regression (Huynh & Ludkovski, AAS, 2021) on targeted groupings of 3–6 states. Our detailed exploratory analysis shows that the mortality experience is highly inhomogeneous across states in terms of respective Age structures. We moreover document multiple divergent trends between best and worst states, between Females and Males, and between younger and older Ages. The comparisons across the 50+ fitted models offer opportunities for rich insights about drivers of mortality in the U.S. and are visualized through numerous figures and an online interactive dashboard.

Keywords: Mortality modeling; multi-output Gaussian Processes; U.S. state mortality; U.S. mortality database

1. Introduction

The United States Mortality Database United States Mortality Database (USMDB) provides a high-quality complete dataset regarding Age- and Year-indexed mortality experience across the 50 U.S. states and the District of Columbia. First published in 2019, USMDB offers a novel opportunity for actuarial and statistical insights at sub-national granularity. Comparative analysis of the U.S. states presents an interesting case study: while there are shared economic, cultural, health and to some extent demographic characteristics, there is also a lot of *heterogeneity*. To handle this situation, we seek a modeling framework that captures and leverages similarities but does not lump all states together. We moreover emphasize the need for a statistical approach. First, many of the U.S. states are small: 15 states have population less than 2 million and six are under one million, meaning that respective yearly deceased counts for a single Age are in the low hundreds. As a result, single-population models for such states lack credibility. Second, in order to make meaningful comparisons across states, joint modeling is necessary, otherwise, the inter-state differences are not statistically coherent. Third, many of our goals are about deeper patterns of mortality, including mortality improvement factors (i.e., annual changes in mortality rates) or mortality Age structures. The respective raw quantities, such as raw year-over-year changes in mortality rates, are much too volatile, and smoothing techniques are imperative.

With the above in mind, we design a custom statistical workflow for studying the USMDB. Our approach is based on creating targeted groupings of a few states at a time. In the first step, we build

© The Author(s), 2025. Published by Cambridge University Press on behalf of Institute and Faculty of Actuaries. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/S1748499525000089 Published online by Cambridge University Press

collections of *similar* states, using a range of auxiliary state-level covariates that reflect economic, demographic, and geographical state characteristics. In the second step, we employ data pooling, modeling a group of similar states as a joint dataset that treats State as a factor level. To do so, we use a (multi-output) Gaussian Process-driven stochastic mortality model. The resulting setup "patches" together localized models to offer an overall view across 51 states; (To simplify notation, we treat District of Columbia as the 51st state) it highlights regional similarities while sharpening national inequalities. Methodologically, this offers a novel middle ground between a massive joint model for all 51 populations at once and a collection of 51 single-population models.

After constructing state groupings and fitting a stochastic mortality model, we embark on a detailed exploratory analysis, supplemented by an online interactive `RShiny` dashboard (Ludkovski & Padilla, 2023). Among others, we visualize and discuss (i) rankings of states in terms of their Age-linked mortality evolution in time; (ii) their recent mortality improvement factors; (iii) their Age-structure of mortality rates; (iv) Age-structure of mortality improvement factors; (v) patterns of above vis-a-vis state characteristics. We highlight several take-aways that are likely to be new to the actuarial audience in terms of the aggregate behavior of the 51 states and respective "outliers." For example, we document that there is a wide range of mortality improvement factors (i.e., annual changes in mortality rates) among states, with some improving and others experiencing rising mortality. Similarly, we document a diversity of Age structures of state mortality relative to the U.S. average and a spectrum of Age patterns in yearly mortality changes. While we do not have any insights about the causal drivers of these patterns, we cite some related literature that attempts to put together a more complete sociodemographic story. To our knowledge, this is one of the first papers to present a detailed side-by-side comparison of so many interrelated mortality models. It is also one of the few analyses fully dedicated to the USMDB state dataset.

Remark 1. The impact of the COVID-19 pandemic on U.S. mortality has been severe. Inclusion of such dramatic outliers into a stochastic mortality model is fraught since the underlying assumption is of a statistically stationary behavior across the training dataset. Moreover, it is not clear whether (or how) to upweigh or downweigh the latest experience when smoothing for mortality trends. For these reasons, we choose to exclude latest data and concentrate on pre-pandemic analysis. See also Remark 4 below.

1.1 Related literature

Our analysis of variations within U.S. mortality connects to several non-actuarial strands of extant literature. Demographers have been highlighting growing geographical disparities in mortality within U.S. since the late 20th century, see Ezzati et al (2008), Wilmoth, et al. (2011), Currie & Schwandt (2016), primarily focusing on dispersion of life expectancy (LE) at birth, e_0. These works highlight the emergence of a geographic belt whereby "the 13 worst-off states were geographically contiguous in 2004" (Fenelon, 2013), spanning Appalachia and the South. Within the economics literature, the starting point for investigating mortality disparities originated with Chetty et al. (2016) who documented strong correlation between income, geography, and LE at Age 40, e_{40}. Chetty et al. primarily worked with county-level data aggregated into commuting zones and documented gaps of 10–14 years in LE across different parts of the U.S.. Boosted by the study of "deaths of despair" in mid-life U.S. adults by Case & Deaton (2021), many researchers have documented growing gaps in adult mortality and life expectancies across both education and income (Becker et al., 2021; Bosworth, 2018; Case & Deaton, 2021; Sanzenbacher et al., 2021). We especially highlight the SOA-sponsored report by Barbieri (2020) which ranks counties by socioeconomic index scores (SIS) and then compares mortality by SIS deciles.

These macro-effects translate into state-level differences through several channels. First, states intrinsically vary in their socio-demographics, for example in the proportions of racial sub-groups

https://doi.org/10.1017/S1748499525000089 Published online by Cambridge University Press

and poverty levels. Second, states have enacted over time different policies that impact mortality. This includes anti-smoking campaigns (Couillard et al., 2021 find correlation between tobacco taxes and mortality and Fenelon, 2013 attributes cigarette smoking prevalence as explaining more than half of geographic differences in mortality), public health policies such as Medicare expansions, and environmental policies such as pollution mitigation. Third, migration patterns, for example increased concentration of college-educated persons along the coasts, have been amplifying regional differences by making states more heterogeneous over the past few decades. Moreover, migrants tend to have better underlying health characteristics, which improves the observed mortality profile of receiving states and leads to higher observed mortality in sending states Ezzati et.al (2008).

Beyond the above factors that create dependence between state characteristics and mortality rates, there is also "a portmanteau of 'place' effects," cf. Couillard et al. (2021). These include the rural/urban differences in mortality, climate effects, and other, yet to be pinpointed spatial factors. The analysis in Couillard et al. (2021) parallels in several ways our work below, with three critical differences. First, Cuillard et al. are looking at midlife mortality, studying aggregate mortality for ages 25–64. In contrast, we provide a much more granular, Age-specific analysis. Second, Cuillard et al. concentrate on the working adults, where effects such as deaths of despair (which have seen marked state-level differences Case & Deaton, 2021) are important; in contrast, we focus on the older ages 60–84. Third, Couillard et al. (2021) take raw mortality data as-is, without imposing any statistical analysis; this is sufficient for their age-aggregated approach but is inadequate for our deeper investigation, especially for mortality improvement factors. In sum, our work complements (Couillard et al., 2021) by providing an updated, age-specific, statistically smoothed analysis of state-level mortality.

Another recent analysis on the spatial disparities in U.S. mortality is by Vierboom et al. (Vierboom & Preston, 2020; Vierboom et al., 2019). In Vierboom et al. (2019), the authors use cause-specific mortality (working with 9 top-level mutually exclusive and exhaustive cause categories) across metropolitan areas and geographic regions, binning into 5-year Age intervals. The main finding is that spatial inequality rose between 2002 and 2016 and that areas that had lower mortality enjoyed larger gains. Such divergent trends were especially noticeable between large coastal metropolitan areas and rural Appalachia and South, and within lung cancer/respiratory diseases, as well as drug/alcohol abuse. Vierboom et al. focus on studying LE at birth e_0 for ages 30–85 and work with county-level data aggregated into 40 spatial units (the 9 census divisions plus Appalachia, broken out into Metro, Suburb, Small Metro, and Rural strata). The companion study (Vierboom & Preston, 2020) investigates spatial inequality in e_{65} (LE at 65) for the same 40 spatial units; see also an earlier analysis in Dwyer-Lindgren et.al (2016) using 5-year Age bins.

Li & Hyndman (2021) investigate state-level mortality through a two-level forecast reconciliation method: building single-state and national-level Lee-Carter models and then adjusting the results so that the sum of the state projections adds up to the national estimate. They primarily focus on out-of-sample performance, examining projections as much as 10 years into the future.

Several articles have recently introduced explicit consideration of spatial mortality patterns in order to handle the sparsity of death counts in small spatial units. Gibbs et al., Gibbs et al. (2020) use conditional auto-regressive priors with a county-specific linear Age trend in order to borrow information across neighboring U.S. counties. In a follow-up publication, Hartman et al. (Shull et al., 2025) construct a single multivariate spatio-temporal model to fuse data across both space and Age bins; parameter inference is done via Integrated Nested Laplace Approximations. Cupido et al. (2021, 2020) apply a spatial filtering approach to infer the latent spatial dependence in U.S. county-level mortality. Boing et al. (2020) build a hierarchical model that teases out the relative importance of states versus counties versus census tracts, showing that the latter is the primary driver of mortality variability. The consistent take-away is that taking into account spatial relationships offers better predictive power and supports the intuition that "individuals living closer

https://doi.org/10.1017/S1748499525000089 Published online by Cambridge University Press

together likely have more similar lifestyles than individuals living hundreds of miles apart" (Gibbs et al., 2020). In turn, similar lifestyle habits and environmental factors drive mortality.

1.2 Contributions

Within this landscape, our contributions can be traced along two dimensions. Methodologically, we propose a new technique for studying USMDB data, namely through creating custom, targeted groupings of a handful of states at a time. Our approach builds on Ludkovski et al. (2018), Huynh & Ludkovski (2021), Huynh et al. (2020), Huynh & Ludkovski (2024) and allows fusion of mortality data from similar states, while avoiding the need to directly model all 51 states jointly, a statistically daunting task. Instead, we advocate grouping states based on their geographical neighbors and a collection of socio-economic covariates. Such targeted groupings simultaneously improve computational efficiency and statistical efficacy. To this end, we compute a weighted Euclidean distance between state-wise principal component analysis scores obtained from the collected state covariates. Notably, we propose a nearest-neighbor-like setup, creating a separate group for each state, in contrast to a partitioning method where states are a priori clustered and multi-population models are independent fit for each cluster. Our motivation comes from spatial regression, where localized modeling (like LOESS) is generally more robust than top-down partitioning (piecewise regression). Hard partitions based, e.g., on pre-specified geographic regions, are difficult to justify or validate, and we therefore prefer data-driven, transductive (i.e., tailored to the state being modeled) groupings.

On a broader level, our approach breaks new ground in developing a meso-scopic framework for studying many inter-related subnational populations – situated between single-population methods and a macroscopic all-in joint modeling. This setup offers the ability to fuse information and draw meaningful comparisons through the built-in coherence in the mortality estimates, while avoiding the full specification of the co-dependence among dozens of populations. Our technique to adaptively group and patch similar populations would be useful for further settings, such as studying other sub-national jurisdictions (counties, provinces, federal states), or for grouping countries, e.g., within Europe.

Empirically, we provide a novel exploratory analysis about the relative experience of smoothed Age-specific mortality across U.S. states. We augment existing literature that focuses on either life expectancy or aggregate mortality (both metrics effectively averaging across many ages) with an explicit consideration of mortality as a function of Age. Moreover, we investigate the recent dynamics of mortality through inferred mortality improvement (MI) factors. We not only rank and compare states against each other but moreover correlate our projections with the collected external covariates. In sum, we confirm several previous aggregate analyses (such as strong correlation between mortality and income/obesity/geographic region) and also document several new insights, such as a heterogeneity of improvement factors as a function of Age, and the strong disparities between Male and Female MIs. These statistical findings provide a starting point for future investigations in other disciplines, such as demographics and economics, that are best equipped to identify their societal causes.

Our goal in this project is to compare mortality across U.S. states. Experiments during the writing of this paper indicate that our results are largely invariant to the specific mortality modeling framework. Therefore, one could for instance use the developed state groupings within a Li & Lee (2005) multi-population Age-Period-Cohort setup rather than with multi-output GPs. Indeed, we do not claim to provide the *best* model for USMDB and leave full benchmarking (e.g., via quantitative assessment of the quality of out-of-sample predictions) to future work. Instead, our aim is to enable meaningful smoothing and nowcasting of Age-specific mortality state-by-state, with the emphasis on pooling states for the purpose of their *relative* comparisons.

https://doi.org/10.1017/S1748499525000089 Published online by Cambridge University Press

Remark 2. The division of the US into states is not very meaningful actuarially, since many states are too large or too heterogeneous to provide direct insights into mortality disparities. For the latter purposes, one ought to concentrate on covariates such as socio-economic status or race that are known to correlate with US mortality drivers. Nevertheless, non-actuarial researchers and popular media frequently compare mortality by state, which was indeed an impetus for the creation of USMDB. Our analysis seeks to provide a stochastic mortality perspective on USMDB, documenting its features and presenting a rigorous comparative assessment of time and age structure of mortality among the states.

The rest of the paper is organized as follows. In Section 2, we summarize the raw data provided by the USMDB and introduce the stochastic model of Multi-Output Gaussian Processes and methodology used to create smoothed mortality surfaces. Section 3 describes the state grouping algorithm. Section 4 presents results regarding state-level mortality relative ranks. Sections 5 and 6 in turn analyze the respective improvement factors and correlation with state-level covariates. Section 7 concludes. Several Appendices present further plots, tables, and covariate definitions. To promote analysis of USMDB, the visualizations below are augmented with the publicly available `RShiny` tool (Ludkovski & Padilla, 2023). The dashboard replicates some of the shown figures and offers a starting point for other researchers to directly explore the outputs of our models across states, genders, and years.

2. Data and statistical model

2.1 Dataset

Built by the HMD team, the United States Mortality Database (USMDB) United States Mortality Database (USMDB) contains a complete historical set of state-level life tables for every calendar year during 1959–2019 for all 50 U.S. states and the District of Columbia (D.C.). The USMDB covers ages 0–110+ and includes separate datasets for the Male and Female populations. The raw data contain birth and death counts from the U.S. vital statistics system and incorporate the census counts and population estimates from the U.S. Census Bureau.

Our objective is to estimate and smooth historical mortality rates and then forecast short-term calendar trends through analyzing and estimating mortality improvement factors. Throughout this paper, we thus focus on the following subset of the USMDB: (a) Calendar Years 1990–2018; (b) Males and Females, considered separately; and (c) Ages 60–84. This subset of older ages and recent years is the most relevant for actuarial applications. Omitting data from the 20th century is in line with the data-driven machine-learning framework we employ, so that long-past mortality experience is not only less relevant but potentially misleading for our model construction, in particular due to nonstationary mortality improvement patterns. We omit very old Ages since the population data underlying USMDB are only available up to an open age interval at 85+ years (Barbieri, 2020).

Due to operating on raw data, nearly all of the works cited above consider aggregated mortality across Age groups, e.g., bins of 10 years (Becker et al., 2021), or 5 years (Dwyer-Lindgren et al., 1980; Shull et al., 2025; Vierboom et al., 2019), or look at life expectancy at a given age (e_0, e_{50}, e_{65}, etc.) (Barbieri, 2020; Boing et al., 2020; Chetty et al., 2016; Vierboom & Preston, 2020; Wilmoth et al., 2011). This fundamentally obscures the non-constant *relative* impact of Age, which is well documented. Indeed, multiple studies conclude that spatial and socio-economic inequalities decline with age: "health disparities narrow with age" (Vierboom et al., 2019) and "the gap [in probabilities of dying by SIS deciles] declines progressively after age 55 years and becomes small (less than 10 percent) at ages 85 and above" (Barbieri, 2020). At the same time, most references disaggregate mortality by other, say socio-economic or cause-of-death, factors, or into smaller spatial units. In contrast, we follow the USMDB to fully disaggregate into 1-year bins by Age, but otherwise consider aggregated mortality across all individuals in the state.

https://doi.org/10.1017/S1748499525000089 Published online by Cambridge University Press

Table 1. Sample rows from the USMDB dataset $\mathcal{D}$

USMDB $\mathcal{D}$ for Males				
State (s)	Age (x_a)	Year (x_t)	Mortality Rate	Log-mortality (y)
CA	60	1990	0.014	−4.248
CA	60	1991	0.014	−4.261
⋮	⋮	⋮	⋮	⋮
CA	84	2017	0.070	−2.664
CA	84	2018	0.072	−2.631
AR	60	1990	0.014	−4.262
AR	60	1991	0.013	−4.327

Let $\mathcal{S} = (s_i)_{1 \le i \le 51}$ represent the 51 U.S. states (throughout we count D.C. as the 51st "state"). The information regarding each state $s \in \mathcal{S}$ provided by the USMDB is organized as follows:

(i) *Independent variables*: calendar year x_t and age x_a. The pair (x_a, x_t) refers to the set of persons from state s aged $a \in \{60, \ldots, 84\}$ during year $t \in \{1990, \ldots, 2018\}$.

(ii) *Dependent variables*: The death counts $D^{(a,t)}$ along with the total number of persons lived (exposed to risk) $E^{(a,t)}$ at a given age-time cell (x_a, x_t). We record the log-mortality rate

$$y^{(a,t)} := \log\left[\frac{\#\text{ of Deaths during } (x_a, x_t) \text{ age-time Interval}}{\#\text{ of Exposed to Risk during } (x_a, x_t) \text{ age-time Interval}}\right] \equiv \log\left[\frac{D^{(a,t)}}{E^{(a,t)}}\right]. \quad (1)$$

To simplify notation, when the age and calendar year are clear from the context, we drop the superscript (a, t). While death counts are generally highly accurate, exposed-to-risk are based on census estimates and systematically undercount undocumented immigrants.

In summary, the USMDB provides 5 inputs, $(s, x_a, x_t, D, E) \equiv$ (state, age, year, deaths, exposed-to-risk) and the associated output $y \equiv$ log-mortality, see Table 1. The complete dataset is denoted as $\mathcal{D}$, with $\mathcal{D}_s \subset \mathcal{D}$ representing the subset of rows associated with state $s \in \mathcal{S}$.

2.2 Multi-population Gaussian process models

Multi-population modeling aims to identify and capture mortality dependence patterns among several populations in order to fuse data and achieve coherent forecasts. We follow Huynh et al. (2020) and Huynh & Ludkovski (2021) in implementing a Multi-Output Gaussian Process (MOGP) to model USMDB longevity data. The MOGP model quantifies mortality uncertainty by probabilistically smoothing raw data and simultaneously generates stochastic out-of-sample forecasts by projecting mortality surfaces across the Age and Year dimensions. For multi-population analysis, MOGP imposes a cross-population correlation structure on top of the Age-Period pattern in each population. The data fusion employed in MOGP improves model fit, reduces model risk, and provides insights into the discrepancies among mortality trends across populations.

2.2.1 Gaussian process regression

First, we describe the mechanics of the Single-Output Gaussian Process (SOGP) model; see Ludkovski et al. (2018) for more details. Fix an arbitrary state $l \in \mathcal{S}$. We are given a sample of $n = 29 \times 25 = 725$ observed log-mortality rates. Mortality is described as a function of age and time:

- $\mathbf{x} := (x^1, \ldots, x^n)$ where $x^i := (x_a^i, x_t^i)$ for $a \in \{60, \ldots, 84\}$ and $t \in \{1990, \ldots, 2018\}$.

https://doi.org/10.1017/S1748499525000089 Published online by Cambridge University Press

- $\mathbf{y}_l(\mathbf{x}) \equiv \mathbf{y}_l := (y_l^1, \ldots, y_l^n)$ where y_l^i, $i = 1, \ldots, n$ denotes the observed log-mortality rate provided by the USMDB for a state $l \in \mathcal{S}$ at age x_a^i during year x_t^i.

Observation Likelihood. We assume that the relationship between y_l^i and x^i can be described with a latent black box function $f_l(\cdot)$ and white noise term,

$$y_l^i = f_l(x^i) + \epsilon_l^i; \qquad i = 1, \ldots, n. \tag{2}$$

The observation noise is Gaussian with $\boldsymbol{\epsilon}_l = (\epsilon_l^1, \ldots, \epsilon_l^n) \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_l := \mathrm{diag}(\sigma_l^2))$. We assume that the observation likelihood is population-dependent but constant in age and year, $\sigma_l = \mathrm{StDev}(\epsilon_l^i)\ \forall i$. The underlying function $f_l(x^i)$ intuitively represents the true mortality rate which would materialize in the absence of random idiosyncratic shocks. While a Poisson observation likelihood is appropriate for small populations, typical death counts for the considered age groups are at least in the high hundreds, making the Gaussian approximation (which is methodologically much preferred for GPs) highly accurate.

Distribution of the Latent Process. A priori, for any sample $\mathbf{x}$ of $m \geq 1$ observations, the finite dimensional distributions (fdds) of $f_l(\mathbf{x}) = (f_l(x^1), \ldots, f_l(x^m))$ are postulated to follow a multivariate Gaussian law $f_l \sim \mathcal{GP}\big(m_l, C_l\big)$ with prior (parametric) mean function, $m_l(\cdot)$, and covariance matrix $C_l(\cdot, \cdot)$,

$$m_l(\mathbf{x}) := \mathbb{E}[f_l(\mathbf{x})] = \big(\mu_l(x^1), \ldots, \mu_l(x^m)\big) \quad \text{and} \quad C_l(x, x') := \mathbb{E}\Big[\big(f_l(x) - m_l(x)\big)\big(f_l(x') - m_l(x')\big)\Big].$$

Assuming that ϵ_l^is are independent across x^is and from f, it follows that

$$\mathbf{y}_l \sim \mathcal{N}\left(m_l, C_l + \boldsymbol{\Sigma}_l\right), \qquad \boldsymbol{\Sigma}_l = \sigma_l^2 I,$$

since $\mathrm{Cov}(y_l^i, y_l^j) = \mathrm{Cov}\big(f_l(x^i), f_l(x^j)\big) + \sigma_l^2 \delta(x^i, x^j)$, where $\delta(x^i, x^j)$ is the Kronecker delta.

Mean and Covariance Structure. We assume functional representations for the mean and covariance functions $m_l(\cdot)$, $C_l(\cdot, \cdot)$ which represent the prior beliefs about the dataset. Recall that all the observed properties of a stochastic process with Gaussian fdds are characterized by $m_l(\cdot)$ and $C_l(\cdot, \cdot)$.

(i) The GP *Mean function* describes the prior trend in log-mortality rates. We use a parametric prior mean function, $m_l(x^i) = \beta_{0,l} + \sum_{j=1}^p \beta_{j,l} h_j(x)$, where $h_j(x)$s are given basis functions and the $\beta_{j,l}$s are unknown coefficients to be estimated. Letting $\boldsymbol{\beta}_l = \big(\beta_{0,l}, \ldots, \beta_{p,l}\big)^T$, $h(x) = \big(h_1(x), \ldots, h_p(x)\big)$, we use the shorthand $m_l(x) = h(x)\boldsymbol{\beta}_\ell$. Below, we postulate a linear trend in the Age dimension:

$$m_l(x^i) = \beta_{0,l} + \beta_{1,l}^a \cdot x_a^i. \tag{3}$$

The choice (3) is used to de-trend the data according to an exponential increase in mortality as a function of Age (the so-called Gompertz Law of Mortality) in our segment of interest $x_a \in \{60, \ldots, 84\}$.

(ii) The GP *Covariance kernel* captures the dependence of the response surface f_l on the varying Age and Year dimensions x_a, x_t. The GP kernel characterizes the smoothing process by quantifying the influence of inputs on the likelihood of the output. Our kernels are distance-based, capturing the logic that the mortality experience should be similar at neighboring data points, and *separable* across the Age and Period coordinates.

We concentrate on a common family of covariance functions known as the Matérn class, equipped with automatic relevance determination. The Matérn-5/2 kernel defines the covariance between arbitrary univariate inputs $x, x_* \in \mathbb{R}$ as:

$$C^{(M52)}(x, x_*; \theta) := \left(1 + \frac{\sqrt{5}}{\theta}|x - x_*| + \frac{5}{3\theta^2}|x - x_*|^2\right) \cdot \exp\left\{-\frac{\sqrt{5}}{\theta}|x - x_*|\right\}. \tag{4}$$

https://doi.org/10.1017/S1748499525000089 Published online by Cambridge University Press

This kernel is parameterized by the length scale (hyper)parameter θ, to be estimated. To construct the overall dependence structure we use a multiplicative Age-Period-Cohort (APC) structure, so that the covariance between two mortality table entries $x^i \equiv (x_a^i, x_t^i)$, $x^j \equiv (x_a^j, x_t^j)$ is

$$C_l^{(APC)}(x^i, x^j) := \eta^2 \cdot C^{(M52)}(x_a^i, x_a^j; \theta_{l,a}) \cdot C^{(M52)}(x_t^i, x_t^j; \theta_{l,t}) \cdot C^{(M52)}(x_c^i, x_c^j; \theta_{l,c}), \tag{5}$$

where $x_c := x_t - x_a$ is the Birth Cohort (i.e., year of birth of an individual who is x_a-old in year x_t), and η^2 is the process variance hyperparameter, scaling covariances to capture the typical amplitude of the response. Observe that there are three length scale hyperparameters $\theta_{l,a}, \theta_{l,t}, \theta_{l,c}$, jointly estimated together. The last cohort term in (5) is essential to capture well-known generational effects, such as the special 1918 and 1939 cohorts. The product structure is analogous to the Age-times-Year terms in the classical Lee-Carter framework. See Ludkovski & Risk (2024) for a further discussion of appropriate GP kernels for mortality.

2.2.2 GP posterior

The GP paradigm models input-output relationships by algebraically conditioning its prior distribution on the training data. The resulting posterior yields a probabilistic projection regarding the latent log-mortality surface at desired input vector $\mathbf{x}_*$, given the information in the USMDB. Note that $\mathbf{x}_*$ can refer to in-sample cells (historical smoothing) or out-of-sample cells (future forecasts), both obtained from exactly the same formulas below.

Given a prior distribution $f_l \sim \mathcal{GP}(m_l, C_l)$ and a training set $\mathcal{T} = (\mathbf{x}, \mathbf{y}_l)$, we calculate the posterior distribution $\mathbf{y}_{l,*}|\mathcal{T} \equiv \mathbf{y}_l(\mathbf{x}_*)|\mathcal{T}$ at predictive cells $\mathbf{x}_*$. Observe that $(\mathbf{y}_l, \mathbf{y}_{l,*})$ follows the Multivariate Normal distribution (MVN)

$$\begin{bmatrix} \mathbf{y}_l \\ \mathbf{y}_{l,*} \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} m_l(\mathbf{x}) \\ m_l(\mathbf{x}_*) \end{bmatrix}, \begin{bmatrix} C_l(\mathbf{x}, \mathbf{x}) + \mathbf{\Sigma}_l & C_l(\mathbf{x}, \mathbf{x}_*) \\ C_l(\mathbf{x}_*, \mathbf{x}) & C_l(\mathbf{x}_*, \mathbf{x}_*) + \mathbf{\Sigma}_{l,*} \end{bmatrix} \right). \tag{6}$$

Applying MVN conditioning expressions, the Universal Kriging equations (Rasmussen & Williams, 2006, Section 2.7) below provide both the estimated mean-function coefficients $\boldsymbol{\beta}_l = (\beta_{0,l}, \beta_{1,l})^T$ in (3) and the posterior distribution of $\mathbf{y}_{l,*}$ $p(\mathbf{y}_{l,*}|\mathbf{y}_l) \sim \mathcal{N}(m_{l,*}(\mathbf{x}_*), C_*(\mathbf{x}_*, \mathbf{x}_*))$ with the posterior mean-variance

$$m_{l,*}(\mathbf{x}_*) := \mathbb{E}[\mathbf{y}_{l,*}|\mathcal{T}] = h(\mathbf{x}_*)^T \hat{\boldsymbol{\beta}}_l + C_l(\mathbf{x}_*, \mathbf{x})(C_l(\mathbf{x}, \mathbf{x}) + \mathbf{\Sigma}_l)^{-1}(\mathbf{y}_l - H\hat{\boldsymbol{\beta}}_l); \tag{7}$$

$$\hat{\boldsymbol{\beta}}_l := \left(H^T D\right)^{-1} H^T \left(C_l(\mathbf{x}, \mathbf{x}) + \mathbf{\Sigma}_l\right)^{-1} \mathbf{y}_l; \tag{8}$$

$$\begin{aligned} C_*(\mathbf{x}_*, \mathbf{x}_*) := \mathrm{Cov}(\mathbf{y}_{l,*}|\mathcal{T}) = C_l(\mathbf{x}^*, \mathbf{x}^*) + \mathbf{\Sigma}_l + \\ + (h(\mathbf{x}_*)^T - C_l(\mathbf{x}_*, \mathbf{x})D)^T (H^T D)^{-1} \left(h(\mathbf{x}_*)^T - C_l(\mathbf{x}_*, \mathbf{x})D\right) \end{aligned} \tag{9}$$

where the matrix $C_l(\mathbf{x}, \mathbf{x}_*)_{i,j} = C_l(x_i, x_{j,*})$ represents the covariance between inputs in the training set and predictive locations $\mathbf{x}_*$, $H = \left(h(x^1), \ldots, h(x^n)\right)$ and $D := \left(C_l(\mathbf{x}, \mathbf{x}) + \mathbf{\Sigma}_l\right)^{-1} H$.

The predictive distribution of mortality rates at different age and time coordinates represented by $\mathbf{x}_*$ yield the estimated mortality surfaces. Note that (7) can be applied both in-sample (for $\mathbf{x}_*$s that are in the training set) to obtain smoothed historical experience, as well as out-of-sample, to predict mortality into the future using exactly the same algebraic expressions. In parallel, the GP model also outputs confidence intervals around $m_{l,*}(\mathbf{x}_*)$ based on the posterior covariance $C_*(\cdot, \cdot)$, depicting the confidence of the model in its own projections.

2.2.3 Shared covariance structure

Assume that we have selected a collection of $L \subset \mathcal{S}$ states and that each state-specific mortality surface f_l, $1 \le l \le L$, follows a GP $f_l \sim \mathcal{GP}(m_l, C_l)$. We proceed to create a joint model for the vector

https://doi.org/10.1017/S1748499525000089 Published online by Cambridge University Press

$f = (f_1, \ldots, f_L)$ through correlating its components. The motivation is that similar states should share alike mortality rates. Therefore, we impose a shared covariance structure which captures the dependencies between mortality rates across states.

To jointly model L outputs, we need to specify the mean and covariance kernel of the joint GP f. More specifically, let $\vec{x}^i \equiv (x^i_a, x^i_t, x^i_1, \ldots, x^i_L)$ where $x^i_l = \mathbb{I}_{\{\text{population} = l\}}$; then we take

$$f \sim \mathcal{GP}(m, \mathcal{C}) \quad \text{where } f(\mathbf{x}) = \big(f_1(\vec{\mathbf{x}}), \ldots, f_L(\vec{\mathbf{x}})\big), \tag{10}$$

$m \in \mathbb{R}^{Ln\times 1}$ is the mean vector whose elements represent the mean functions of each population $l \in L$, $\{m_l(\vec{x})\}_{l=1}^L$, and $\mathcal{C} \in \mathbb{R}^{Ln\times Ln}$ denotes the covariance matrix across the entire system.

Intrinsic Coregionalization Model (ICM). Directly specifying the cross-covariances of each output pair $f_l, f_{l'}$ becomes unwieldy for $L > 3$, so instead we rely on coregionalized kernels (Huynh & Ludkovski, 2021) which assume that each output f_l, $1 \le l \le L$ is a linear combination of Q independent latent GPs $u = \big(u_1(\mathbf{x}), \ldots, u_Q(\mathbf{x})\big)$ with shared covariance kernel $C^{(u)}(\cdot,\cdot)$. In our case, we use the $C^{(u)} = C^{(APC)}$ APC kernel from (5).

Let $a^*_q = (a_{1,q}, \ldots, a_{L,q})^T$, $1 \le q \le Q$, be the vector containing the q-th factor loadings across all populations L. Then $f(\mathbf{x}) = \sum_{q=1}^Q a^*_q u_q(\mathbf{x})$, or for the l-th population,

$$f_l(\mathbf{x}) = a_{l,1}u_1(\mathbf{x}) + \ldots + a_{l,Q}u_Q(\mathbf{x}). \tag{11}$$

The switch from L separate GPs to Q GPs $(u_1, \ldots, u_Q)$ is similar to a PCA or singular value decomposition approach and allows to reduce the number of hyperparameters in the cross-population covariance matrix from $\frac{L(L-1)}{2}$ to $Q \times L$:

$$\begin{aligned}\mathcal{C}(x,x') = \mathrm{Cov}\big(f(x), f(x')\big) &= \mathrm{Cov}\left(\sum_{q=1}^Q a^*_q u_q(x), \sum_{q=1}^Q a^*_q u_q(x')\right)\\ &= \left(\sum_{q=1}^Q a^*_q a^{*T}_q\right) \otimes \mathrm{Cov}\big(u_q(x), u_q(x')\big) \equiv B \otimes C^{(u)}(x,x').\end{aligned}$$

The $L \times L$ *coregionalization matrix* $B := AA^T$ with entries $B_{l,k} = \sum_{q=1}^Q a_{l,q}a_{k,q}$ has rank Q.

Hyperparameters. The modeling task is ultimately to learn the covariance structure, i.e., the mean and kernel functions based on the training data. The overall set of the ICM MOGP hyperparameters is $\Theta = ((\theta_j)_{j\in\{a,t,c\}}, (a_{l,q})_{l=1,\ldots,L,q=1,\ldots,Q}, (\beta_{0,l}, \beta_{1,l}, \sigma^2_l)_{l=1,\ldots,L})$. In our results below, the `R` package `kergp` Deville et.al (2015) is used to carry out the respective Maximum Likelihood estimation through Kronecker decompositions, see Huynh & Ludkovski (2021, 2024) for more details. Once Θ is estimated, ICM MOGP inference reduces to evaluating the linear-algebraic formulas in (7)–(9).

3. Model grouping

The MOGP model from the previous section works on a group of states. In this Section, we discuss how to construct such groups. The goal of grouping is to maximize data fusion and maintain computational tractability. On the one hand, joint models lead to more accurate longevity modeling than individual-state models. This is especially noticeable for the smaller states where state-level data are very noisy and trends are hard to decipher. Moreover, a joint model facilitates making comparisons about the *relative* experience of states. On the other hand, directly modeling all 51 states is computationally intractable within the MOGP framework and is unlikely to perform well anyway. The respective model would have hundreds of hyperparameters and is likely to suffer from unstable inference and identifiability issues. Furthermore, as discussed in Ludkovski et al. (2018), fusing information from different populations through a MOGP can be expected to

https://doi.org/10.1017/S1748499525000089 Published online by Cambridge University Press

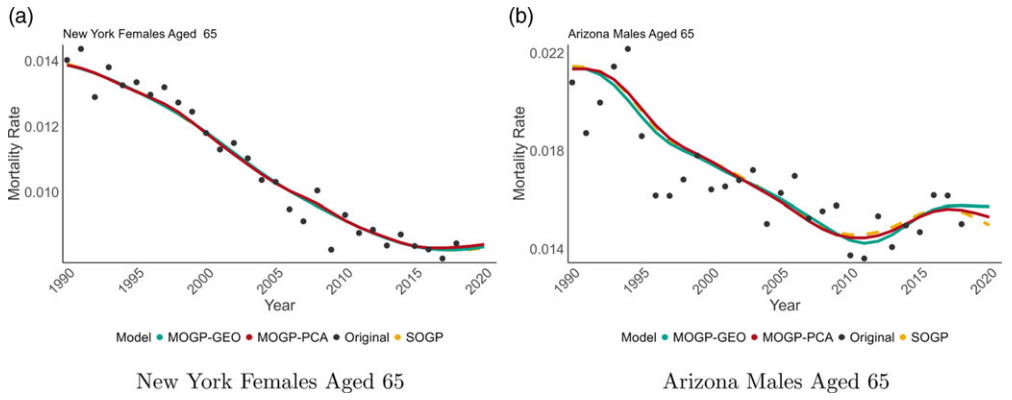


Figure 1. Raw data (black circles, covering years 1990–2018) and smoothed/projected mortality curves (for years 1990–2020) for two representative states based on three different groupings: (i) a single-state SOGP model; (ii) MOGP-GEO with geographic groupings based on U.S. Census regions in Appendix A.1; (iii) proposed MOGP-PCA.

improve predictions only when these populations are similar. Thus, it makes little sense to group, say, Massachusetts (East Coast, urbanized, wealthy state) with Alabama (South, rural, poor). To account for spatial relationships described in the Introduction, we would like whenever possible to group *neighboring* states following the maxim "Everything is related to everything else, but near things are more related than distant things" (Tobler, 1970). This operationalizes the hypothesis that neighboring states often share similar economic and demographic characteristics and thus have cognate mortality trends and experiences (see also the aforementioned spatial-based analysis of U.S. counties in Gibbs et al. (2020), Shull et al. (2025).

With the above in mind, we seek to create groups of 3–6 *similar* states, with a preference for geographic contiguity. Our groupings $\mathscr{O}_s$ are state-specific, i.e., are not mutually exclusive across ss. The grouping algorithm determines which states are alike, so as to incorporate the "right information" to provide more accurate predictions and reduce predictive uncertainty, excluding irrelevant information. A secondary concern is making sure that groupings provide enough "critical mass," namely a large enough aggregate population.

3.1 Motivation

To motivate the issue of how to group states, we briefly discuss two GP-based alternatives. First, we recall the base case of constructing a Single Output GP (SOGP) model state-by-state. This can be done using the methodology in Ludkovski, et al. (2018) and yields independently fitted GP models. Specifically, we fit 51 SOGP models utilizing the kernel (5) with hyperparameters $\theta_a, \theta_t, \theta_c, \sigma^2, \eta^2, \beta_0, \beta_1$. Second, we construct MOGP models based on a *geographic partitioning*, namely the nine U.S. Census regions, see Appendix A.1. These regions are 3–9 states in size and are purely geographically aligned. From a statistical perspective, the widely varying group size and the large size of some groups (e.g., the South Atlantic region includes 9 states with a total population of over 50 million) are challenging and slow down ICM MOGP performance.

Fig. 1 displays estimated mortality experience at Age 65 in two states as a function of year for the above two choices, as well as the proposed state-characteristic grouping from Section 3. These smoothed mortality estimates are contrasted with the raw mortality rates shown as black circles. For the latter, we observe that the observation noise is much higher in the right panel of Fig. 1 compared to the left panel. This matches the intuition that the noise in mortality data is roughly proportional to the underlying exposed population. Since Arizona's population is 7 million compared to about 20 million in New York, we expect triple the respective observation variance. The

https://doi.org/10.1017/S1748499525000089 Published online by Cambridge University Press

Table 2. The 18 selected state covariates in $\mathcal{C}$. See Appendix A.3.1 for definitions of each covariate

Economic covariates	Demographic covariates	Geographic covariates
Educational Attainment (EA)	Non-minority Pop'n (NMP)	Average Temperature (TP)
Percent Change in GDP (GDP)	Percentage Elderly (ED)	Average Rel. Humidity (RH)
Median Income (MI)	Without Health Insurance (HI)	Average Dew Point (DP)
Regional Price Parities (RPP)	Obesity Rate (OR)	Population Density (TPD)
Poverty Rate (PR)	Political Preference (PP)	
Urbanization Percentage (UP)	Percent Religious (R)	
Land in Farms (LF)	Share of Immigrant Pop'n (IP)	

estimated standard deviation of this noise is the hyperparameter $\sigma_l = 3.42\%$ for N.Y. Females, and $\sigma_l = 5.09\%$ for Ariz. Males.

All three GP forecasts are statistically unbiased and carry out viable non-parametric penalized curve fitting. In particular, the in-sample (for years 1990–2018) estimated mortality rates closely match each other. At the same time, there are non-trivial differences among the models when it comes to out-of-sample prediction, best understood as differences in mortality improvement factors. Fig. A.1 in the Appendix highlights some of the discrepancies. In general, we find that single-state models tend to produce more extreme MI and tend to over-smooth the data, while the MOGP-GEO models occasionally overfit. For example, in Fig. A.1 Vermont Males and Minnesota Females' geographic groupings lead to a negative MI estimate, while the PCA-based grouping projects a positive MI. The take-away is that better MOGP groupings can help maximize stability and interpretability. Moreover, the ideal group size is 3–6 states, enabling sufficient information borrowing from similar states without making the groups too big. Larger groups are challenging for modeling the MOGP cross-covariance described in Section 2.2.3 and affect computational time which is cubic in L.

3.2 Covariates

As documented in prior studies and confirmed in our results below, there is a strong correlation between economic and demographic variables and observed state-level mortality discrepancies. Hence we use various state characteristics to compute a customized similarity metric that drives our group selection. We work with a diverse set $\mathcal{C}$ of 18 non-mortality state-level covariates, chosen to be a representative collection of (1) economic, (2) demographic, and (3) geographic characteristics. The respective state-level data are obtained from several sources, including (FRED https://fred.stlouisfed.org; US Census Bureau https://www.census.gov, and U.S. Bureau of Economic Analysis (BEA). https://www.bea.gov). These sources do not distinguish between the Male and Female subpopulations, hence all the covariates are shared among the genders. Table 2 lists the covariates we work with; Appendix A.3.1 provides a complete description. Our covariates overlap with and are broadly similar to the 20 used in Chetty et al. (2016, Table 8) (who grouped them into Health Behaviors, Healthcare, Environmental, Labor Market, Social Cohesion, and Other Factors), the six used in Couillard et al. (2021) and the 11 covariates in Barbieri (2020).

Rather than manually combining the above covariates, many of which are highly correlated, to identify which states are similar, we opt for a statistical solution of principal component analysis (PCA). Thus, we apply PCA to the 51×18 matrix $\mathcal{C}$, to identify the main sources of differences across states. Use of PCA to summarize covariates is also advocated in Barbieri (2020). Fig. 2 and Table A.1 report the PCA results, where we focus on the first three PCA components $PC1, PC2, PC3$ that together explain over 66% of the variance in $\mathcal{C}$ and allow us to succinctly summarize the drivers of heterogeneity.

https://doi.org/10.1017/S1748499525000089 Published online by Cambridge University Press

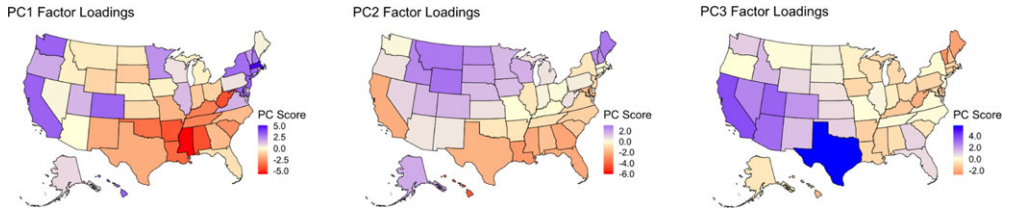


Figure 2. State-wise PCA factor loadings $\mathcal{P}_k(s)$, $k = 1, 2, 3$.

The PC1 component can be interpreted as an economic or wealth factor: the most expensive states (N.Y., Cali.) display the highest PC1 loadings, while states like Miss. and La. have lower PC1 loadings. As confirmation, cf. Table A.1 in Appendix A.3.2, most of the economic covariates from $\mathcal{C}$ are positively correlated with the first PC component. The PC2 component can be interpreted as a climate-related factor, with the Southern-most warmest states having highest PC2 loadings. Lastly, PC3 component loosely corresponds to the Sun Belt states: the Southwest plus Texas, Georgia, and Florida. The population in these states is rapidly growing due to internal migration and immigration.

Table A.1 in the Appendix shows the correlation between each covariate and LE at birth, e_0, as constructed by the CDC Center for Disease Control and Prevention (CDC). Similar statistical association between income (MI), poverty rate (PR), and age-specific mortality is documented in Li & Hyndman (2021). Other observed correlations are harder to explain, for example between state Religious Percentage and its LE. This supports our motivation to use statistical, rather than causal, ways to capture similarity. A further alternative, which however could be deemed as circular logic, would be to use past mortality trends to identify which states are similar.

3.3 Grouping by covariate similarity

The PCA factor loadings obtained above are used to construct groupings of states for the MOGP. We first define a distance metric $\mathfrak{D}(s_1, s_2)$ between any two states $s_1, s_2 \in \mathcal{S}$. We then group s with its most similar states according to $\mathfrak{D}(s, \cdot)$, aiming for geographical closeness to help with interpretability. The groups are constructed in a stepwise agglomerative manner, with 3–6 states per group. Note that since the covariates are the same across genders, all the groupings pertain both for Males and Females.

Distance between States. We use PCA components from previous section to define a distance between states. To this end, we compare the respective factor loadings $\mathcal{P}_k(s)$, generating a Euclidean metric weighted by the eigenvalues associated with each PCA component:

$$\mathfrak{D}(s_1, s_2)^2 := \lambda_1\big(\mathcal{P}_1(s_1) - \mathcal{P}_1(s_2)\big)^2 + \lambda_2\big(\mathcal{P}_2(s_1) - \mathcal{P}_2(s_2)\big)^2 + \lambda_3\big(\mathcal{P}_3(s_1) - \mathcal{P}_3(s_2)\big)^2, \qquad (12)$$

where λ_k is the eigenvalue associated with PC component $k = 1, 2, 3$, see Table 3. The intuition is to prioritize the first factor that explains the majority of the observed variability in our covariates.

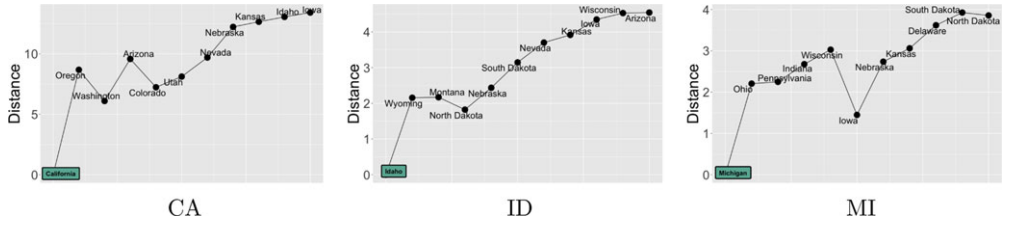
Identifying Similar States. Fix a state $s \in \mathcal{S}$. We construct the set $\mathcal{N}_s$ of nearest neighbors of s as follows:

(i) Compute the geographical neighbors of s, $O_1(s) = \{s_*: s \text{ and } s_* \text{ are geographically contiguous}\}$. That is, $O_1 \equiv O_1(s)$ is the collection of states which share a physical boundary with s.
(ii) Add to $\mathcal{N}_s$ the state s_1 that minimizes the PCA-based distance to s in O_1, $s_1 = \arg\min_{s_* \in O_1} \mathfrak{D}(s, s_*)$.

https://doi.org/10.1017/S1748499525000089 Published online by Cambridge University Press

Table 3. Summary of principal component analysis for state covariates

	PC1	PC2	PC3
Standard deviation	2.53	1.70	1.59
Proportion of variance	0.36	0.16	0.14
Eigenvalue λ_k	6.486	3.424	2.548



CA ID MI

Figure 3. The ten most similar states $\mathcal{N}_s$ for s =Cali. (*left*), Idaho (*middle*), and Mich. (*right*).

(iii) Update the neighborhood definition to include s_1:

$$O_2 \equiv O_2(s \cup s_1) = \{s_* : s_* \text{ is geographically contiguous with either } s \text{ or } s_1\}.$$

(iv) Repeat steps (ii) and (iii) with $s_n = \arg\min_{s_* \in O_n} \mathfrak{D}(s, s_*)$ for $n = 2, \ldots, 10$.

Fig. 3 visualizes the results from running the above algorithm on a few representative states and until $|\mathcal{N}_s| = 10$. The y-axis shows the distance $\mathfrak{D}(s, s_n)$ between the target state $s \in \mathcal{S}$ and the "neighbor" added in round n. We observe the common "zig-zaggging" pattern, indicating that there are states which are further apart geographically, yet closer with respect to $\mathfrak{D}$; see, for example, Michigan and Iowa, or Idaho and North Dakota. Indeed, N.Dak. is the closest to Idaho in terms of $\mathfrak{D}(s, \cdot)$ even though they are not geographically neighboring.

Proposed Grouping $\mathscr{O}_s$ for $s \in \mathcal{S}$. By construction, the neighborhoods $\mathcal{N}_s = \{s, s_1, s_2, \ldots\}$ are contiguous; given the above non-monotonicity in Fig. 3, we adjust them to take into account $\mathfrak{D}$-similarity. We start with $s, s_1, s_2 \in \mathscr{O}_s$; that is, the first three states from $\mathcal{N}_s$ are in the group of s. Next, let $s_3' = \arg\min_{s^* \in \mathcal{N}_s \setminus (s_1, s_2)} \mathfrak{D}(s^*, s)$ be the next closest state to s in terms of $\mathfrak{D}(\cdot, s)$. We include $s_3' \in \mathscr{O}_s$, if and only if $\mathfrak{D}(s_3', s) < \max_{i=1,2} \mathfrak{D}(s_i, s)$; so that we create a group of 4 states if s_3' is a better (closer) fit in distance to s than either of the two neighbors s_1, s_2 already in the group. The above procedure enforces a high degree of geographic contiguity (at least two states are guaranteed to be contiguous with s) but also allows to add one more state that is not geographically close but is very similar to s in terms of the PCA loadings. As an example, Iowa is added to Mich.'s group (Fig. 4, right), and N. Dak. is added to Idaho's group (middle panel), while Cali. stays in a group of 3 (no further state is closer than s_3 =Wash. in that case, left panel).

As a final step, we ensure that all groups have sufficient underlying population to yield credible estimation. To this end, we augment additional states (in order of their distance $\mathfrak{D}$) until the group $\mathscr{O}_s$ has a total population of at least 5 million. This is particularly relevant for the Mountain West region, where Idaho, Montana, Wyoming, North and South Dakota (all with populations under 1.5 M) tend to group together.

Remark 3. Alaska and Hawaii lack natural geographical neighbors. See Appendix A.4 for the methodology (and results) used to calculate AK and HI groupings.

A sample of the resulting 51 groupings $\mathscr{O}_s$ (one for each s) are shown in Fig. 4; the full list of $\mathscr{O}_s$s is in Appendix A.4. The colors correspond to $\mathfrak{D}(s_*, s)$, i.e., how close a given state is to its selected

https://doi.org/10.1017/S1748499525000089 Published online by Cambridge University Press

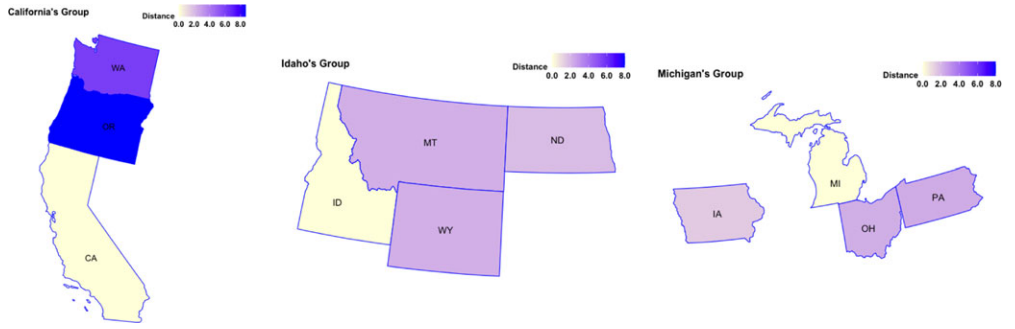


Figure 4. Selected state groupings $\mathscr{O}_s$, $s =$ Cali. (*left panel*), Idaho (*middle*), and Mich. (*right*).

neighbors. We note that closeness in terms of $\mathfrak{D}$ varies; there are many very similar (in terms of PCA factor loadings) states in the Midwest, while California's neighbors are all much less similar to it.

Fig. A.3 in Appendix A.4 visualizes the size of the constructed groups. Recall that by default $|\mathscr{O}_s| = 3$; groups of four arise when one of the two most similar (in terms of $\mathfrak{D}$) states is not contiguous with s, but farther away. Most of these cases arise in the Midwest and Mid-Atlantic regions. In addition, in Mountain West we have groups of 5 or 6 due to low state population counts.

The groups $\mathscr{O}_s$ are constructed separately for each state. To get a sense of how similar are the groups for different target states, we define the concept of reciprocity. States A and B are in positive reciprocity (PR) if $B \in \mathscr{O}_A$ and $A \in \mathscr{O}_B$, i.e., they are mutual members in the respective groups. The states that do not experience any PR are La., Miss., R.I., S.Dak., Texas, and W.Va.. This means that these states are quite different from all their geographical neighbors and are not getting selected for their neighbors' groups. The triple

$$\big(\text{Arizona, Nevada, Utah}\big)$$

is the resulting $\mathscr{O}_s$ for these three states which are both contiguous geographically and are very similar across covariates, forming a mini-cluster. Similarly, the following pairs of states result in the same constructed groups $\mathscr{O}_s$:

$$\big(\text{Montana, Wyoming}\big); \quad \big(\text{Connecticut, New York}\big).$$

4. State mortality predictions

Following the grouping method in the previous section, we construct 51 MOGP models for the 51 states. For each state s, its fitted mortality surface is based on its group $\mathscr{O}_s$ and hence borrows strength from several other similar states. Since the various $\mathscr{O}_s$ overlap, there is interdependence throughout rather than a fixed regional dependence. For all MOGP, we use the ICM covariance structure with rank $Q = 3$. We then use (7) to generate the MOGP-predicted mortality estimates for years 1990–2020. Recall that the training USMDB data is up to 2018; thus, we report both in-sample retrospective analysis of U.S. state mortality up to 2018 and *prospective* predictions for 2019 and 2020. Since our training set consists of pre-pandemic data, the 2020 predictions can be viewed as a statistical baseline for how 2020 would have looked like without COVID-19.

In this Section, we analyze the relative mortality projections across calendar Years (Section 4.1) and then individual Ages (Section 4.2). Section 5 then investigates the *mortality improvement factors* (MI), obtained as the time-gradient of these mortality surfaces. While we primarily address the MOGP posterior means, we also comment about the posterior credible intervals, see e.g. Fig. 11.

https://doi.org/10.1017/S1748499525000089 Published online by Cambridge University Press

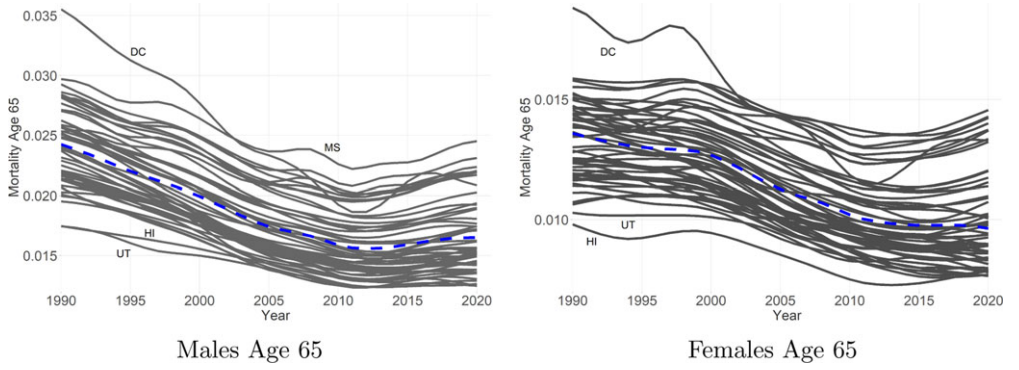


Figure 5. MOGP-PCA mortality rates for age 65 Males (*left* panel) and Females (*right*) and years 1990–2020. U.S. national SOGP-smoothed rate is shown as dashed blue.

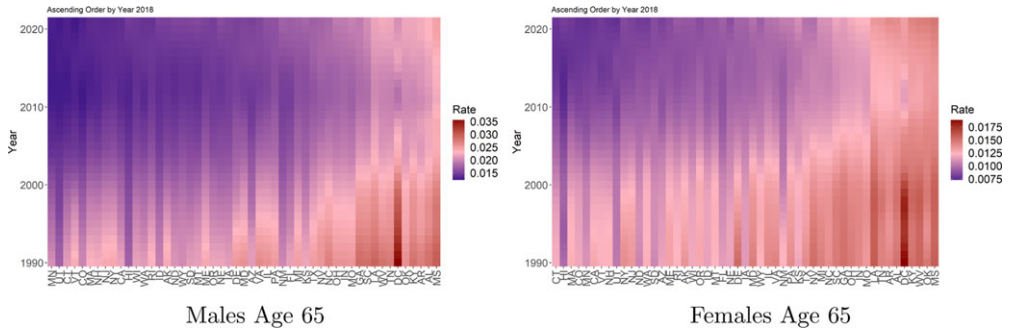


Figure 6. MOGP-PCA smoothed mortality rates for age 65 Males (*left* panel) and Females (*right*) and years 1990–2021. States are ordered from left to right by their mortality in 2018.

4.1 Time structure

Fig. 5 shows the bulk behavior of state mortality rates across years 1990–2020 fixing the Age, namely at age 65. As a baseline, we also compute and show the fitted national-level U.S. mortality rate. This curve (dark blue dashed lines in the plot) is generated by fitting a SOGP model to the aggregated U.S. mortality experience. As expected, it lies in the core of the state curves and is close to the population-weighted average of the state-level mortality projections. See also Fig. A.4 in Appendix A.5 for Age 75 counterpart, and the interactive `RShiny` widget (Ludkovski & Padilla, 2023) where users can select any other desired Age. Notice that the mortality trend before 2010 is universally positive, while in the past decade mortality has been either stagnating or deteriorating. We also observe that although there is a strong common trend, mortality evolution has been rather variable across states. There are several outlying curves in Fig. 5, notably D.C. and Hawaii, as well as many "cross-overs" where states change relative ranks over time. This heterogeneity increased during the 2010s compared to 2000s, and there are also more cross-overs in the Male populations compared to Females. Moreover, the spread among states is very substantial: e.g., from just over 0.75% inferred mortality in 2020 for 65-year old Females in N.Dak. and Conn., to 1.45% in Miss. and Okla. – a ratio of 1.88 at the right edge of the right panel.

A more in-depth visualization is provided by the heatmaps in Fig. 6 where columns denote states and rows denote years; darker gradient corresponds to lower mortality. The states are sorted in increasing order of their mortality as of 2018. The expected behavior as we move up across a column is a smooth transition from red to blue/purple corresponding to improving mortality

https://doi.org/10.1017/S1748499525000089 Published online by Cambridge University Press

Table 4. Top-5 and bottom-5 states ranked by MOGP-PCA-projected mortality rate in 2020 at ages 65 and 75. The best and worst states are in the first rows

Best States				Worst States			
Males		Females		Males		Females	
Age 65	Age 75	Age 65	Age 75	Age 65	Age 75	Age 65	Age 75
Vt.	Colo.	Minn.	Hawaii	Miss.	Miss.	Miss.	Ky.
Utah	Hawaii	Colo.	Cali.	Ala.	Ala.	Okla.	W.Va.
Minn.	Vt.	Conn.	Ariz.	Ark.	Ky.	W.Va.	Miss.
Conn.	Conn.	Utah	Fla.	W.Va.	Okla.	Ky.	Okla.
Colo.	Cali.	Cali.	Conn.	Ky.	Tenn.	Ala.	Ala.

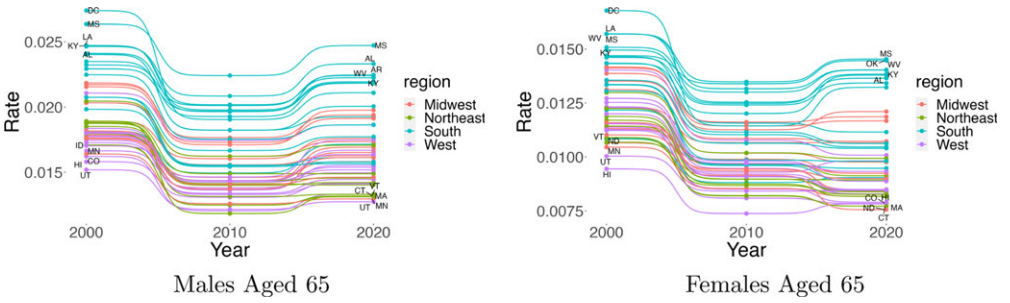


Figure 7. State rankings in terms of age 65 mortality rate across years 2000, 2010, and 2020.

throughout 1990–2020. While this occurs for some states (notably D.C. is displaying this pattern, having gone from one of the biggest laggards to being middle-ranked by 2020), for many columns the pattern is much more checkered. Note the colored stripes that indicate different historical mortality paths for states that nowadays have similar rates. Only a few states, such as Utah, Colo., and Hawaii show a consistent positive mortality trend throughout the past three decades.

Relative ranks of states are of great interest. By looking at the first few/last columns of the heatmaps in Fig. 6, we can read off the states with the worst/best mortality. Table 4 summarizes which states are projected to have the highest (worst) and lowest (best) mortality in 2020. Northeast and Pacific states have the overall best mortality, and the Southern states are at the bottom. There is a lot of consistency between the genders as well, with Miss., Ala., and Ky. being among the bottom-5 in all four columns on the left of Table 4. We caution against reading too much into the precise rankings: our fitted models yield predictive standard deviations of latent log-mortality on the order of $\sqrt{C_*(x_*, x_*)} \in [0.04, 0.1]$ (depending on size of the state), which corresponds to standard errors of about 0.2%–0.3% on the original scale of mortality rates. For example, Colo.'s Male mortality in 2020 is projected to be 1.304% (ranked fifth) but with a 95% predictive interval of [1.141%, 1.489%] which would be anywhere within the top-20. There is a complex correlation between the projections of different states which makes relative ranks less volatile, but the upshot is that it is not statistically possible to decide whether a given state is in top-5 or top-10. Nevertheless, these rankings echo Chetty et al. (2016) regarding life expectancy for individuals in the bottom income quartile, who have singled out Tenn., Ark., Okla. as the worst performers and Cali. and Vt. as the top-ranked states.

To better visualize the relative ranks of states across time, Fig. 7 shows a bump chart ranking states by their fitted/projected mortality rates in 2000, 2010, and 2020 at Age 65. We observe that all states experienced a decrease in their mortality rates between 2000 and 2010. However, for the 2010s the picture is largely reversed. In fact for Males, only Maine and New York are projected to

https://doi.org/10.1017/S1748499525000089 Published online by Cambridge University Press

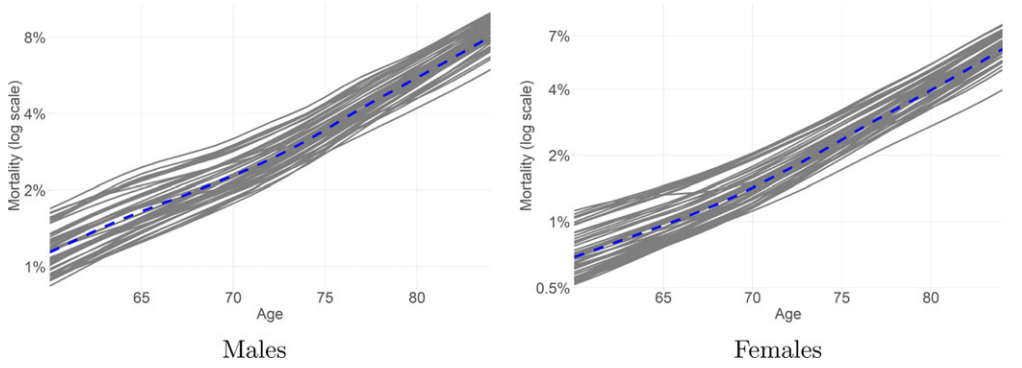


Figure 8. MOGP-PCA projected mortality rates (on the log scale) across ages 60–85 for year 2020. The dashed blue line shows the fitted SOGP model for the U.S. nationwide mortality.

improve between 2010 and 2020. For Females, the picture is mixed; 16 states (AL, AR, DC, FL, HI, IN, KY, LA, MO, MS, NM, OH, OK, TN, UT, WV) are projected to have a worse Female Age 65 mortality in 2020 compared to 2010, while the rest are improving. Our analysis corroborates the county-level findings in Shull et.al (2025) who report an improvement in mortality between 2000 and (approximately) 2014, followed by a recent deterioration in county-level rates popularly attributed to *deaths of despair*.

Moreover, Fig. 7 indicates a certain split among Female mortality: laggard states, primarily in the South, have deteriorating Female mortality in 2020 compared to 2010, while the best-performing states in the Northeast and West continue to experience improving mortality. In other words, there is a *divergent* pattern where states with low (Female) mortality are doing relatively better compared to states with higher mortality, with this pattern aligning with regional partitions. Similar age-aggregated conclusions appear in Fenelon (2013), Chetty et al. (2016).

Remark 4. The COVID-19 pandemic caused a major spike in US mortality in 2020 and 2021, as well as generated lingering mortality "aftershocks" (including possibly initiating a new structural break in future mortality improvements). It requires several years of post-pandemic data to disentangle the respective short-term and long-term impacts on mortality. In the considered MOGP models, the range of year-dependence is 3-6 years, so such analysis would become feasible in the near future, though not quite yet. Specialized GP kernels, such as those that incorporate change points (Saatçi et al., 2010) could be applied, but are beyond the scope of our work. As it stands, including 2020–2023 in our analysis would degrade all the model fits and "bake in" pandemic excess deaths trend for future projections.

4.2 Age structure

To complement our discussion of the temporal pattern of mortality, we next discuss its structure in Age. Fig. 8 presents the MOGP-smoothed age structure of mortality for Males and Females. The online `RShiny` app (Ludkovski & Padilla, 2023) shows an interactive version of this Figure for any user-specified Year. Given that mortality increases exponentially in Age, we plot log-mortality rates, which are roughly linear, matching the prior mean function in (3). We observe that in bulk the Age structure is very consistent across states, meaning that there is a high correlation between relative mortality experience at different Ages. Consequently, state-level mortality rankings are largely age invariant. For example, Connecticut Male mortality is the lowest in the nation at Age 60 and is 4th, 5th, and 4th lowest at Ages 65, 70, 75, respectively (Females similarly rank 3rd, 3rd, 4th, and 5th at those ages). At the other end of the spectrum, Mississippi has the highest Male mortality across *all* ages and is in the worst-5 for all ages for Females as well.

https://doi.org/10.1017/S1748499525000089 Published online by Cambridge University Press

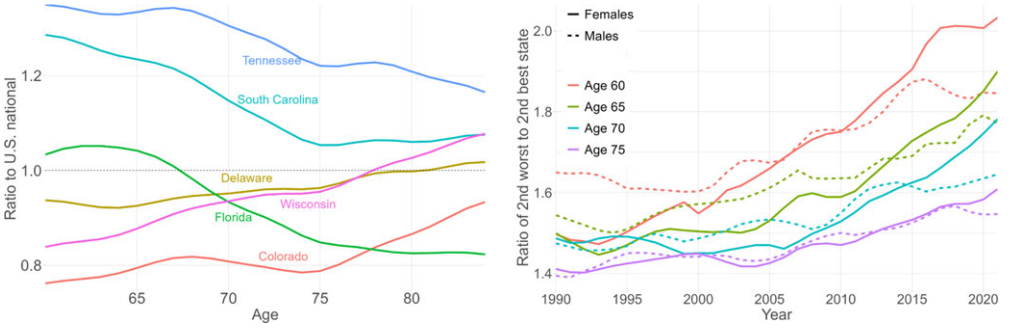


Figure 9. *Left:* mortality rates for Males in 2020 expressed as a ratio of U.S. national average, as a function of age for 6 representative states. *right:* ratio of second-worst to second-best state mortality at four different ages, as a function of year.

Although the overall shape in Fig. 8 is cylindrical, implying a fixed spread between log-mortality of the best and worst state, as a function of Age, the individual behavior of states relative to the U.S. average exhibits highly heterogeneous patterns. The right panel of Fig. 9 (also available interactively) shows those ratios for Males in 2018 for a selection of representative states. For some states, younger Ages are doing better than average, while older Ages are worse off, see e.g., Wisconsin where Male mortality at Age 60 is 84% of national average, while at Age 80 it is 103% of national average. For other states, younger Ages do relatively worse than old Ages; this is the pattern for Tenn. and S.C. in Fig. 9. Male Tennesseans aged 60 have mortality that is 37% above U.S. average, while those aged 80 are only 21% above average. Yet other states have no discernible trend: Delaware Males are within 8% of U.S. average rates across all ages.

In all, we observe distinct *clusters* of state mortality age structure: the increasing pattern of Wisc. is repeated for AK, ID, IA, ME, NE, NH, RI, UT, VT, VA; the decreasing pattern of S.C. and Tenn. is repeated in AR, DC, KY, LA, MS, NV, OK, and WV. CA, CO, CT, MA, MN, MT, NJ, NY, ND, OR, SD, WA, and WY have a largely Age-invariant negative gap to national average (doing better at all ages), while IN, KS, MO, NC, OH, PA, and TX have a largely Age-invariant positive gap to national average (doing a bit worse at all Ages). Finally, there are a couple of idiosyncratic outliers like Florida, where younger ages are worse than average but older ages are much better than average. The above reveals a novel structural affinity among state mortalities.

We moreover document a widening gap between mortality rates of best/worst states across time. Thus, the spread between states in Fig. 8 has been increasing over the past two decades, something already observed in Fig. 5 at Age 65. The right panel of Fig. 9 shows the ratio of mortality of the second best state vs. the second worst state (we remove the best and the worst to stabilize this metric, akin to winsorizing) at four representative Ages and across years. As expected, the relative spread in mortality is lower for older ages because the underlying rates themselves increase in Age (so the absolute spread is in fact growing in Age). However, the noteworthy pattern is that the dispersion dramatically increased since ∼2005. In 2020, the Female mortality in the worst states at Age 60 is estimated to be more than double compared to the best states, see the right edge of the right panel in Fig. 9, while thirty years ago it used to be only 50% higher. At Age 70 the ratio is more than 75% in 2020 compared to 45% in 2000.

Fig. A.5 in Appendix A.6 provides another visualization that addresses two further aspects. First, the figure shows the relative ranks of states at two different Ages, complementing the Year structure in Fig. 7. We observe that several states have drastically different ranks at Age 65 versus Age 75. For example, D.C. Female mortality in 2018 ranks 47th at Age 65 but 10th at Age 75; Ariz. moves from 15th to 7th and Cali. from 8th to 3rd at those ages. In the opposite direction, Vt. drops from 5th to 24th and Maine drops from 20th at Age 65 to 31st for Age 75. Second, Fig. A.5 compares the rankings based on the MOGP setup versus those from the single-population SOGP

https://doi.org/10.1017/S1748499525000089 Published online by Cambridge University Press

models. While for most states the rankings are very stable across the models (showcasing that all our results are indeed data-driven), for about a dozen states there are sizable changes. In particular, the SOGP rankings for some of the smaller states are often significantly different. For example, at Age 65, Vt. Females rank 5th best across the MOGP models, but are only 9th based on SOGP smoothing; Mont. ranks 22nd according to MOGP but 12th according to SOGP. At Age 75, SOGP suggests that R.I. is 26th rather than 15th according to MOGP. This illustrates our previous point about credibility and the need to borrow information from "neighbors" to achieve model fit. As discussed in Section 3.1, the SOGP estimates can be unstable when the raw data is very noisy.

5. Mortality improvement

In this section, we study mortality improvement (MI) factors, i.e., the relative change in mortality rates across years. The most common metric is a year-over-year change expressed in annualized percentage terms (such that lower mortality corresponds to a positive improvement),

$$MI_s(x_a, x_t) = 1 - \exp\big(m_{*,s}(x_a, x_t) - m_{*,s}(x_a, x_{t-1})\big) \equiv \frac{e^{m_{*,s}(x_a,x_t-1)} - e^{m_{*,s}(x_a,x_t)}}{e^{m_{*,s}(x_a,x_t-1)}}, \tag{13}$$

where the posterior mean $m_{*,s}(\cdot)$ is defined in (7).

We find that the Matérn-5/2 kernel in (4) is not ideal for MI analysis. As can be seen in Appendix A.7, the respective surface tends to have high-level fluctuations (see spurious "bands" around age 70–75, and many "speckles" in the heatmap), which are not noticeable when looking at $m_{*,s}(\cdot)$ but become significant when considering MI. This feature can be understood by recalling that the smoothness of a GP as determined by the behavior of its kernel $C(x, x')$ as $x' \to x$. Matérn-ν kernels yield fits that are $\nu - 1/2$ times differentiable, so the $M52$ kernel in (4) leads to a predictive surface that is exactly twice differentiable.

To remedy this issue, we refit our MOGP models using the following separable Squared-Exponential kernel (15) across (Age, Year, Cohort):

$$C_l(x^i, x^i_*) := \eta^2 C^{(SqExp)}(x^i_a, x^i_{a,*}; \theta_{l,a}) \cdot C^{(SqExp)}(x^i_t, x^i_{t,*}; \theta_{l,t}) \cdot C^{(SqExp)}(x^i_c, x^i_{c,*}; \theta_{l,c}) \tag{14}$$

where

$$C^{(SqExp)}(x, x'; \theta) := \exp\Big\{ - \frac{(x - x')^2}{2\theta^2} \Big\}. \tag{15}$$

The SqExp kernel yields infinitely differentiable fitted mortality surfaces, which are thus smoother temporally and yield more interpretable time gradients.

Remark 5. The conceptual aim of MI is to understand the Year trend in mortality. In practice, there are period effects, i.e., "common shocks," such as heat waves, that correlate observed mortality at different Ages and same Year. By "over-smoothing" through the SqExp kernel, we remove short-term temporal fluctuations, focusing on the more durable time structure of mortality. Fig. A.2 in the Appendix further illustrates how the SqExp kernel smoothes the temporal trends relative to Matérn-5/2.

Fig. 10 visualizes the model-based MIs at Age 65 and Year 2020 across the 51 states. Recall that the training USMDB data is up to 2018, so that we show the MOGP-PCA forecast for the MI trend two years into the future, on the eve of the pandemic. The results display a lot of heterogeneity in state-level MIs, with some states experiencing improvements (blue shades) and others stagnation (gray) or deterioration (negative improvement rate, orange). Notably, the statistical significance of estimated MIs is not high; the MOGP provides standard errors of about $\pm 1\%$, so for many states and Ages it is impossible to conclusively decide whether its MI is positive or negative, see right panel of Fig. 11. At the 95% posterior significance level, only AZ, NV and UT definitely have

https://doi.org/10.1017/S1748499525000089 Published online by Cambridge University Press

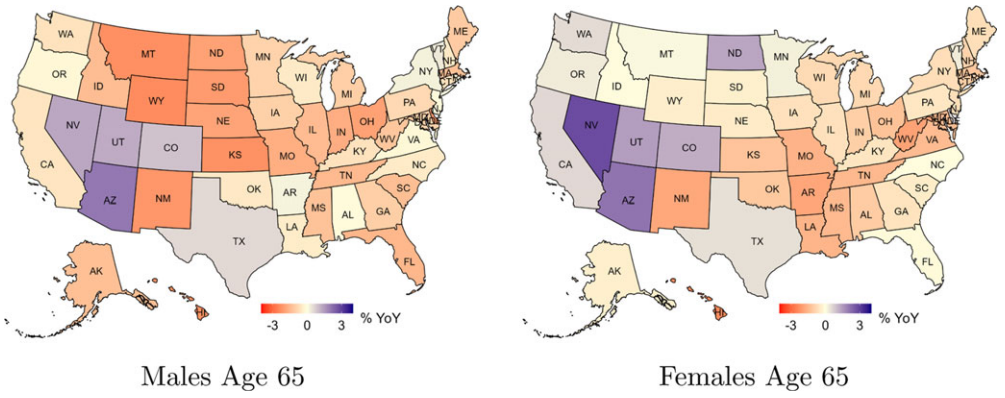


Figure 10. MOGP-based annualized mortality improvement factors in 2020 and age 65.

positive MI for Females (and only AZ for Males) at Age 65, while Females in 9 states (and Males in 14 states) have conclusively negative MI at Age 65.

There are notable regional patterns in Fig. 10; for example, the Sun Belt states experience a positive MI while most of the South experiences deteriorating mortality. There are several "outlier" states in the Southwest: Nevada and Arizona have exceptionally positive mortality improvement, while New Mexico has negative MI across both genders. Table 5 summarizes the top- and bottom-5 states in terms of MI at $x_* = (65, 2020)$ across genders. Comparing with the previous section, the MI pattern implies that the Sun Belt is catching up to the lower-mortality Northeast states; at the same time, Midwest and Southern states are falling behind, amplifying the national discrepancies. This can be linked to the causal analysis in Chetty et al. (2016) who report that (based on raw data and aggregating all working ages) "Hawaii, Maine, and Massachusetts had the largest gains in LE [between 2001 and 2014] (gaining $>$ 0.19 years annually) when men and women in the bottom income quartile were averaged. The states in which low-income individuals experienced the largest losses in LE (losing $>$ 0.09 years annually) were Alaska, Iowa, and Wyoming," and an earlier similar conclusion for changes in LE between 1983 and 1999 in Ezzati et al. (2008).

As seen in Fig. 10, MI for Females tends to be higher compared to Males. At Age 65 (resp. 75) Females have higher MI than Males in 32 (resp. 37) out of 51 states, see the left panel of Fig. 11 and the `RShiny` dashboard. This predicts a widening gap in Male-Female mortality. Moreover, the mortality trends are often flipped between genders: for instance, at Age 65, only in 8 states is Male MI positive, while $MI > 0$ for Females in 13 states. This difference is even more stark at Age 70 where 39 states have positive Female MI (of which 9 are positive with more than 95% posterior probability), but only 20 states (and just 2 with 95% credibility) have positive Male MI. That being said, the association between MIs across genders is not very strong, cf. the right panel of Fig. 11 which compares Male and Female MIs at Age 65, furthermore showing the respective posterior 90% credible intervals.

Next, we investigate the age structure of annualized mortality improvement rates. Fig. 12 displays MI for Males and Females between ages 60–84, sorted according to MI at Age 65. Looking across Ages, we see a mixed pattern of improvement/deterioration for nearly all columns, exceptions being Females in the Southwest and Pacific (UT, NV, AZ, CA, WA, TX) who improve at all ages, as well as a few states (KS, NE, SD, TN, and MA) where Males have negative MI for all ages.

The most common pattern in Fig. 12 is a deterioration of mortality (orange gradient) for ages below 70 and positive MI (blue gradient) for ages above 75. This means that the Age-slope of the mortality curves is flattening. One explanation could be generational (i.e., a cohort effect), reflecting better cumulative health of the older baby boomers (who are 65–75 in 2020) compared

https://doi.org/10.1017/S1748499525000089 Published online by Cambridge University Press

Table 5. Top-5 and bottom-5 states in terms of MOGP PCA-SqExp-based projected annual mortality improvement at age 65 in year 2020. The best and worst states are in the first row

Best MI				Worst MI			
Males		Females		Males		Females	
Age 65	Age 75	Age 65	Age 75	Age 65	Age 75	Age 65	Age 75
Ariz.	Vt.	Nev.	Colo.	D.C.	D.C.	D.C	D.C.
Nev.	Utah	Ariz.	Nev.	Wyo.	N.Mex.	Hawaii	S.Dak.
Utah	Colo.	Utah	Alaska	Mont.	Hawaii	W.Va.	Hawaii
Colo.	Ky.	N.Dak.	Or.	Kan.	Md.	Ark.	N.Mex.
Texas	Ariz.	Colo.	Ariz.	Hawaii	Va.	N.Mex.	Kan.

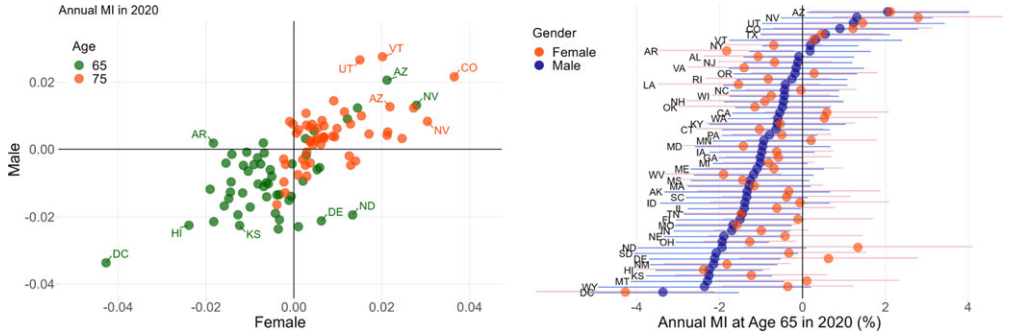


Figure 11. *Left panel*: state-wise mortality improvement factors across genders at ages 65 and 75. MIs are computed based on the SqExp MOGP-PCA model (14). *Right: Male vs Female* MIs at age 65, together with the respective 90% posterior credible intervals. The states are sorted according to the Male MIs.

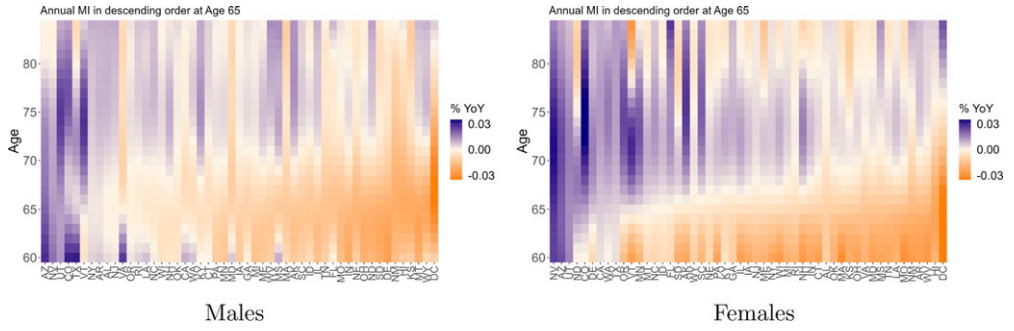


Figure 12. MOGP-based mortality improvement rates in 2020. States are sorted by MI at age 65.

to their younger counterparts. For Females, we also often observe a convex shape, with Ages ≤ 67 and ≥ 80 deteriorating, and Ages around 70–75 improving.

Some states exhibit idiosyncratic age structure of MI. For example, Males in Ariz., Nev., Va., and Texas are experiencing strong mortality improvement at younger ages, and mortality deterioration at older ages, the opposite pattern to above. That could reflect inter-state migration patterns between the working-age and retiree populations. In a few states (OR, OK, WI, RI, CT Males, NE Females), the MI factors are very close to zero throughout, indicating static mortality experience. It is an open problem for demographers and social scientists to discuss and identify the full context and causes of all these new findings.

https://doi.org/10.1017/S1748499525000089 Published online by Cambridge University Press

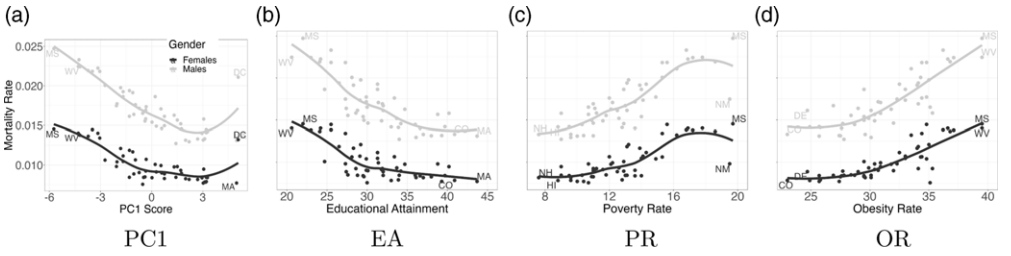


Figure 13. MOGP-PCA smoothed mortality rates at age 65 for both genders in year 2018 against four selected state covariates.

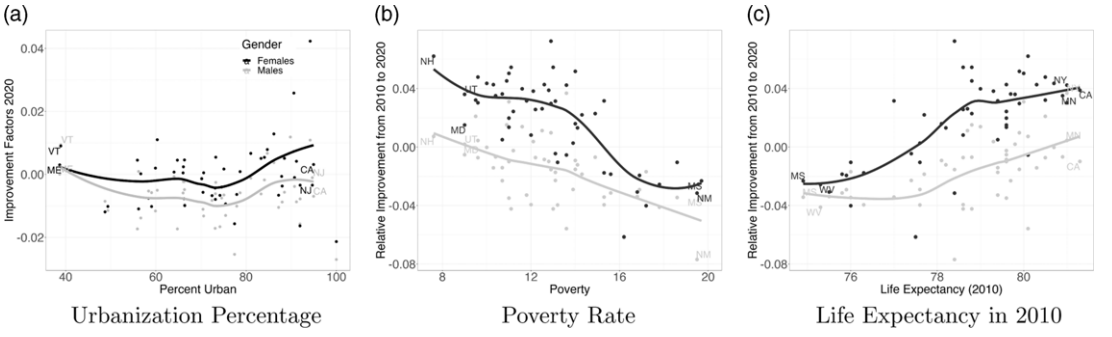


Figure 14. MOGP PCA-SqExp MI factors at age 65 for both genders in year 2018 against three selected state covariates.

6. Explanatory covariates

To gain further insights into the drivers of mortality, we document the relationship between our predicted mortality rates/improvement factors and state characteristics from Section 3.2. To this end, we compare 22 variables (18 covariates, 3 PC factors, and life expectancy at birth from Center for Disease Control and Prevention (CDC)) against three different MOGP-PCA model outputs: (a) 2020 mortality predictions, (b) 2020 MI factors, (c) 2020 vs 2010 improvement ratio. We use the SqExp kernel (14) for computing improvement factors and the M52 kernel (4) for computing mortality rates.

Fig. 13 displays the relationship between predicted mortality rates in 2020 and several covariates, namely educational attainment, poverty rate, obesity rate, and PC1 scores. To visualize the nonlinear dependence, we include a regression curve estimated using LOESS regression. The latter excludes several outliers at the extreme left/right of the plots: D.C. often has markedly different covariates (due to it being a 100% urban region), see e.g., Figs. 13a and 14a, and so does Mississippi (Figs. 13b,13c and 14b).

First, we observe that the smoothed curves are essentially parallel for the Male and Female populations, implying that the impact of different state covariates (which are shared across genders) is very similar for both genders. Second, we observe a strong correlation between economic variables and mortality rates. Recall, as suggested in Section 3.2, that the states' PC1 factor loadings are correlated with economic prosperity. Therefore, the relationship observed in Fig. 13a implies that mortality levels are positively correlated with the state's economic well-being, wealthier states having lower mortality. This matches the finding in Chetty et al. (2016) regarding the strong pattern between economic prosperity and longevity, see also the claims made in Becker et al. (2021), Sanzenbacher et al. (2021), Couillard et al. (2021) (who all fit linear relationships to age-aggregated data).

https://doi.org/10.1017/S1748499525000089 Published online by Cambridge University Press

Additional related insights are in panels (b)-(d) of Fig. 13. Fig. 13b shows that U.S. states with higher educational attainment tend to experience lower mortality; Fig. 13c–d show that higher state-wide poverty rates and higher obesity rates are associated with higher mortality. Of note, the above patterns are often *nonlinear*. For example, the positive association between state obesity rate and mortality is substantially weakened for states with obesity rate below 30% (left edge of Fig. 13d). Likewise, lower mortality associated with higher education no longer follows once over 35% of the state's population has Bachelor's degrees (right edge of Fig. 13b). These facts can be interpreted as one-sided risk drivers: states that have a lot of obese individuals (or few college-educated individuals) suffer higher mortality, but having "exceptionally" non-obese or highly educated populace is not associated with lower mortality.

Next, we analyze the relationships between our 22 variables and the annual improvement factors in 2020 along with the relative aggregate improvement in mortality rates during the decade of 2010s. We find that for most variables there is little significant correlation between them and latest state MI. We do document several interesting patterns in Fig. 14. We observe a positive relationship between MI and urbanization and a negative relationship between MI and poverty (panels a and b). This implies that rural and poorer states tend to exhibit worse mortality trends, also observed in Couillard et al. (2021). Furthermore, Fig. 14c shows that decadal MIs are positively associated with a higher life expectancy. This reinforces our discussion on divergence: states that do well (high LE) keep improving, and states that do poorly (low LE) are deteriorating. As a result, the gap between "best" and "worst" performing states increased during the final decade, see Section 4. We note that none of the bottom-7 states by LE experience a positive MI in the 2010s. Fig. 14b–c highlight that there is a bigger Female-Male MI spread for wealthier/higher-LE states, while in respective bottom states Males and Females have been doing poorly (negative MI) almost equally.

7. Conclusion

In this paper, we have developed multi-population GP models geared for an actuarial analysis of Age- and Year-specific mortality rates across the 51 U.S. states. Our work complements existing Age-aggregated studies in the economics literature, but goes much further, using the developed statistical framework to study Age-specific trends. In particular, the MOGP framework allows to analyze smoothed state-wide Mortality Improvement factors and the various age structures, patterns that are impossible to adequately capture from raw data alone.

At a basic level, we confirm the well-known and well-documented features of U.S. mortality in the late 2010s, such as deterioration of mortality for Americans in their sixties, and the vast gaps between states based on economic and health characteristics. We also observe the familiar geographic patterns for Appalachia, Southwest, the Deep South, etc. Going deeper, we uncover novel features of this heterogeneity, manifested through two channels. First, we note a lot of relative shifts, as states move up and down the rankings. These shifts are highly Age-dependent. Second, the Age structures exhibit very different behaviors, showing the limitations of using a single summary statistic like LE to explain state differences. In some states, Ages below 65 are doing relatively the best, in others those in the 70s, in yet others, the eldest.

An important insight of our exploratory analysis is the growing *divergence* across states. We document that the gaps between best and worst states are getting wider (Fig. 9), and that many (but not all) of the "laggard" states are falling further behind (due to below-average MI) while states with lowest mortality are often pulling even more ahead (Fig. 7). One exception to this is the Southwest that is catching up and moving up the ranks. We also record the divergence between Male and Female mortality trends. We emphasize that our expertise is in actuarial modeling and

https://doi.org/10.1017/S1748499525000089 Published online by Cambridge University Press

hence all proposed explanations of our findings are just educated guesses to be confirmed by subject experts, including demographers and economists.

Our analysis shows that the U.S. can be broken up into data-driven clusters of states that share similar patterns both in mortality rates and in respective trends (MI). These clusters are loosely geographical and include "Appalachia" (TN, KY, SC, AR), "Deep South" (AL, MS, WV, LA), "Pacific" (CA, OR, CO), "Southwest" (UT, AR, NV), "Upper Plains" (MN, SD, ND), "Midwest" (PA, OH, IN, MO, KS), "New England" (ME, VT, NH, RI), "North East" (MA, CT, NJ, NY). These observations validate our proposed strategy of grouping into 3–6 similar *and* neighboring states. We also identify states with truly idiosyncratic patterns, including D.C., Florida, New Mexico, Hawaii, and Texas.

To conclude, let us outline three avenues for future research. First, one may consider additional spatial scales. Looking at county-level data can help to understand further intra-state patterns, especially in large states like California or New York, where there is a lot of intra-state heterogeneity. Looking at metropolitan area data can help to isolate urban/rural effects. Analysis across these different spatial units can better tease out the impact of "place" in terms of different mortality drivers. For example, health policies (such as ACA rules and smoking regulations) are set by state, while demographics vary more in terms of urban and rural locales.

Second, more in-depth analysis is warranted about Age-linked drivers of U.S. mortality. This includes the role of income, racial characteristics, education, and especially health factors. One approach would be to merge our analysis with cause-of-death data, linking to our earlier work in Huynh & Ludkovski (2024) and isolating cause-specific state differences. For example, it would offer a direct window to discuss *deaths of despair* effects, regional cancer patterns, or cardiovascular trends which in turn correlate to wealth and urban/rural discrepancies. Another approach would be along the lines of Hartman et al. (Shull et al., 2025), who apply spatial GLM methodology.

Third, additional analysis could be done to improve the MOGP models themselves. For example, in the present work we have assumed a constant observation noise σ_ℓ^2 for each state. More realistically, we expect that observation variance is proportional to underlying exposure, so that observations at older ages are more noisy since there are, e.g., many fewer 83-year-olds compared to 63-year-olds. Development of heteroscedastic MOGP variants or extending to age-specific noise terms would enhance the study's scope. In a similar vein, our analysis was done independently for Males and Females; building a joint MOGP across genders would help to achieve coherence in mortality improvement factors, cf. Huynh & Ludkovski (2021).

Acknowledgement. We thank the anonymous reviewers for helpful comments that improved the original version of this manuscript. We are also grateful to Nhan Huynh for code assistance at the early stages of the project.

Data availability statement. The primary source for the analysis is the freely available data in the US Mortality Database United States Mortality Database (USMDB) that is publicly accessible subject to creating a free USMDB account. No further pre-processing was applied to this data. Economic, demographic, and geographic state-level covariates were obtained from a variety of public websites and repositories, as listed in Appendix A.3.1. That dataset and the `R` code that supports the findings of this study are available from the corresponding author upon reasonable request.

Funding statement. This work received no specific grant from any funding agency, commercial, or not-for-profit sectors.

Competing interests. The authors declare none.

https://doi.org/10.1017/S1748499525000089 Published online by Cambridge University Press

References

Barbieri, **M.** (2020). Mortality by socioeconomic category in the United States, tech. rep., Society of Actuaries. https://www.soa.org/resources/research-reports/2020/us-mort-rate-socioeconomic/

Becker, **T.**, **Majmundar**, **M. K.**, & **Harris**, **K. M.** (2021). *High and rising mortality rates among working-age adults*. National Academies Press.

Boing, **A. F.**, **Boing**, **A. C.**, **Cordes**, **J.**, **Kim**, **R.**, & **Subramanian**, **S.** (2020). Quantifying and explaining variation in life expectancy at census tract, county, and state levels in the United States. *Proceedings of The National Academy of Sciences*, **117**(30), 17688–17694.

Bosworth, **B.** (2018). Increasing disparities in mortality by socioeconomic status. *Annual Review of Public Health*, **39**, 237–251.

Case, **A.**, & **Deaton**, **A.** (2021a). *Deaths of despair and the future of capitalism*. Princeton University Press.

Case, **A.**, & **Deaton**, **A.** (2021b). Life expectancy in adulthood is falling for those without a BA degree, but as educational gaps have widened, racial gaps have narrowed. *Proceedings of The National Academy of Sciences of The United States of America*, **118**, e2024777118.

Center for Disease Control and Prevention (CDC). https://www.cdc.gov

Chetty, **R.**, **Stepner**, **M.**, **Abraham**, **S.**, **Lin**, **S.**, **Scuderi**, **B.**, **Turner**, **N.**, **Bergeron**, **A.**, & **Cutler**, **D.** (2016). The association between income and life expectancy in the United States, 2001-2014. *Journal of the American Medical Association*, **315**, 1750–1766.

Couillard, **B. K.**, **Foote**, **C. L.**, **Gandhi**, **K.**, **Meara**, **E.**, & **Skinner**, **J.** (2021). Rising geographic disparities in US mortality. *Journal of Economic Perspectives*, **35**(4), 123–146.

Cupido, **K.**, **Fotheringham**, **A. S.**, & **Jevtic**, **P.** (2021). Local modelling of US mortality rates: A multiscale geographically weighted regression approach. *Population, Space and Place*, **27**, e2379.

Cupido, **K.**, **Jevtić**, **P.**, & **Paez**, **A.** (2020). Spatial patterns of mortality in the United States: A spatial filtering approach. *Insurance: Mathematics and Economics*, **95**, 28–38.

Currie, **J.**, & **Schwandt**, **H.** (2016). Mortality inequality: The good news from a county-level approach. *Journal of Economic Perspectives*, **30**, 29–52.

Deville, **Y.**, **Ginsbourger**, **D.**, **Durrande**, **N.**, & **Roustant**, **M. O.** (2015). Package kergp.

Dwyer-Lindgren, **L.**, **Bertozzi-Villa**, **A.**, **Stubbs**, **R. W.**, **Morozoff**, **C.**, **Kutz**, **M. J.**, **Huynh**, **C.**, **Barber**, **R. M.**, **Shackelford**, **K. A.**, **Mackenbach**, **J. P.**, **van Lenthe**, **F. J.**,. . ., **Murray**, **C. J. L.** (2016). US county-level trends in mortality rates for major causes of death. *Journal of the American Medical Association*, **316**, 2385–2401.

Ezzati, **M.**, **Friedman**, **A. B.**, **Kulkarni**, **S. C.**, & **Murray**, **C. J. L.** (2008). The reversal of fortunes: Trends in county mortality and cross-county mortality disparities in the United States. *PLoS Medicine*, **5**(4), e66.

Fenelon, **A.** (2013). Geographic divergence in mortality in the United States. *Population and Development Review*, **39**(4), 611–634.

FRED: Federal Reserve Economic Data, Federal Reserve Bank of St. Louis. https://fred.stlouisfed.org.

Gibbs, **Z.**, **Groendyke**, **C.**, **Hartman**, **B.**, & **Richardson**, **R.** (2020). Modeling county-level spatio-temporal mortality rates using dynamic linear models. *Risks*, **8**, 117.

Huynh, **N.**, & **Ludkovski**, **M.** (2021). Multi-output Gaussian processes for multi-population longevity modelling. *Annals of Actuarial Science*, **15**, 318–345.

Huynh, **N.**, & **Ludkovski**, **M.** (2024). Joint models for cause-of-death mortality in multiple populations. *Annals of Actuarial Science*, **18**, 51–77.

Huynh, **N.**, **Ludkovski**, **M.**, & **Zail**, **H.** (2020). Multi-population longevity models: A spatial random field approach. In Proceedings of the Society of Actuaries 2020 Living to 100 Symposium.

Li, **H.**, & **Hyndman**, **R. J.** (2021). Assessing mortality inequality in the US: What can be said about the future? *Insurance: Mathematics and Economics*, **99**, 152–162.

Li, **N.**, & **Lee**, **R.** (2005). Coherent mortality forecasts for a group of populations: An extension of the Lee-arter method. *Demography*, **42**(3), 575–594.

Ludkovski, **M.**, & **Padilla**, **D.** (2023). RShiny dashboard for MOGP analysis of USMDB. https://mludkov.shinyapps.io/US-state-MOGP (Accessed on Jan 2, 2024).

Ludkovski, **M.**, & **Risk**, **J.** (2024). Expressive mortality models through Gaussian process kernels. *ASTIN Bulletin: The Journal of the IAA*, **54**, 327–359.

Ludkovski, **M.**, **Risk**, **J.**, & **Zail**, **H.** (2018). Gaussian process models for mortality rates and improvement factors. *ASTIN Bulletin: The Journal of the IAA*, **48**, 1307–1347.

Rasmussen, **C. E.**, & **Williams**, **C. K. I.** (2006). *Gaussian processes for machine learning*. The MIT Press.

Saatçi, **Y.**, **Turner**, **R. D.**, & **Rasmussen**, **C. E.** (2010). Gaussian process change point models. In Proceedings of the 27th International Conference on Machine Learning (pp. 927–934).

Sanzenbacher, **G. T.**, **Webb**, **A.**, **Cosgrove**, **C. M.**, & **Orlova**, **N.** (2021). Rising inequality in life expectancy by socioeconomic status. *North American Actuarial Journal*, **25**, S566–S581.

https://doi.org/10.1017/S1748499525000089 Published online by Cambridge University Press

Shull, **M.**, **Richardson**, **R.**, **Groendyke**, **C.**, & **Hartman**, **B. M.** (2025). A multivariate spatiotemporal model for county level mortality data in the contiguous United States. *Annals of Actuarial Science, to Appear*. p. Available at SSRN 4246521.
Tobler, **W.** (1970). A computer movie simulating urban growth in the Detroit region. *Economic Geography*, **46**, 234–240.
United States Mortality Database (USMDB). https://usa.mortality.org/overview.php.
U.S. Bureau of Economic Analysis (BEA). https://www.bea.gov.
U.S. Census Bureau. https://www.census.gov.
U.S. Census Bureau (2013). Geographic regions and divisions map. Accessed on Nov 20, 2023.
Vierboom, **Y. C.**, & **Preston**, **S. H.** (2020). Life beyond 65: Changing spatial patterns of survival at older ages in the United States, 2000-2016. *The Journals of Gerontology: Series B*, **75**, 1093–1103.
Vierboom, **Y. C.**, **Preston**, **S. H.**, & **Hendi**, **A. S.** (2019). Rising geographic inequality in mortality in the United sStates. *SSM-Population Health*, **9**, 100478.
Wilmoth, **J. R.**, **Boe**, **C.**, & **Barbieri**, **M.** (2011) Geographic differences in life expectancy at age 50 in the United States compared with other high-income countries. In E. Crimmins, S. Preston & B. Cohen (Eds.), *International differences in mortality at older ages: Dimensions and sources* (pp. 333–366). The National Academies Press.

A. Appendix

A.1. Geographic Regions

We use U.S. divisions from U.S. Census Bureau (2013) when creating geographical groups for the MOGP-Geo model:

1. New England (6 states): CT, ME, MA, NH, RI, VT;
2. Mid-Atlantic (3): NJ, NY, PA;
3. East North Central (5): IN, IL, MI, OH, WI;
4. West North Central (7): IA, KS, MN, MO, NE, ND, SD;
5. South Atlantic (9): DE, DC, FL, GA, MD, NC, SC, VA, WV;
6. East South Central (4): AL, KY, MS, TN;
7. West South Central (4): AR, LA, OK, TX;
8. Mountain (8): AZ, CO, ID, MT, NM, NV, UT, WY;
9. Pacific (5): AK, CA, HI, OR, WA.

https://doi.org/10.1017/S1748499525000089 Published online by Cambridge University Press

A.2. impact of state groupings and kernel choice

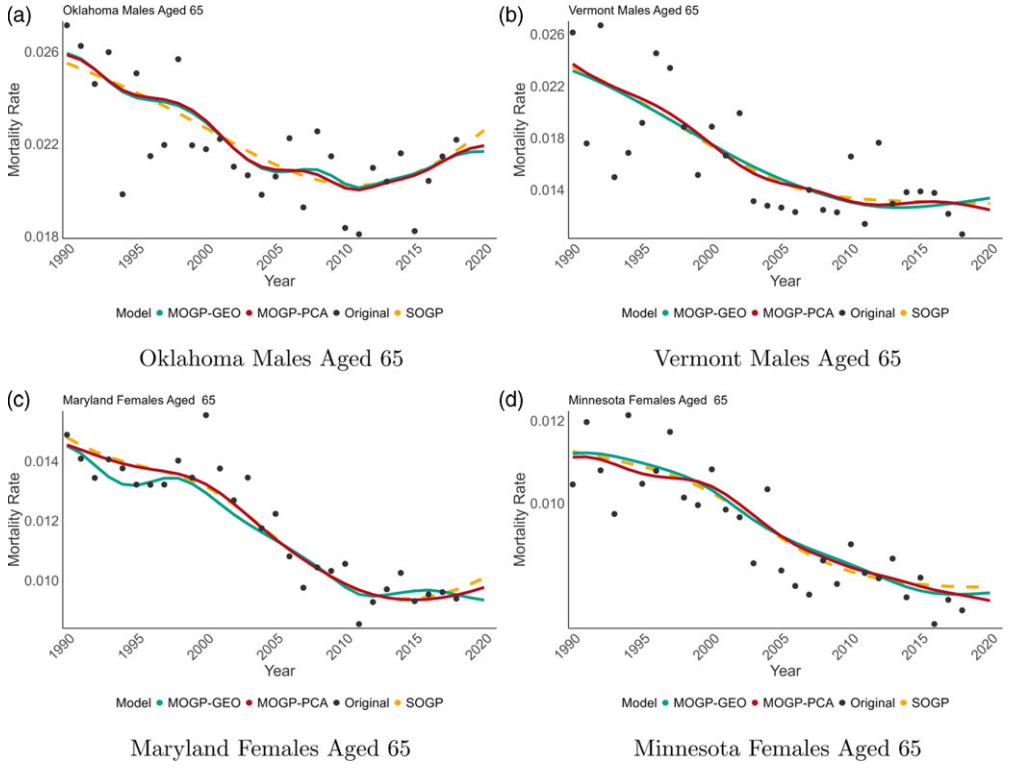


Figure A.1. Raw data (black circles) and predicted mortality rate (curves) for 4 representative states based on GP models with 3 different groupings. *Top row: Ma*les; *bottom row*: Females.

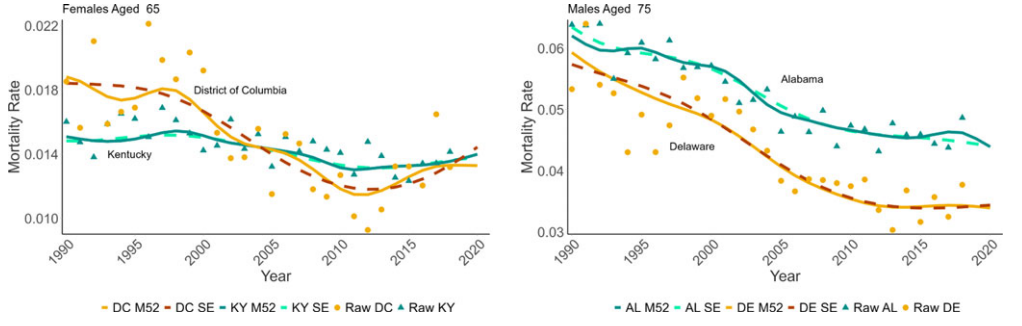


Figure A.2. Comparing the year structure of mortality at a given age across the Matérn-5/2 (4)–(5) and SqExp kernels (14)–(15) for MOGP models. We show two panels with 2 representative states each. *Left:* District of Columbia and Kentucky: Females Aged 65. *Right: Males* age 75 in Alabama and Delaware.

https://doi.org/10.1017/S1748499525000089 Published online by Cambridge University Press

A.3. State-Level Covariates

A.3.1. Data Sources

The 18 state-level covariates used in the PCA analysis described in Section 3.2 are based on 2018 data and described as follows:

Economic Covariates:

1. *Educational Attainment (EA)*: Percentage of population aged 25+ with a bachelor's degree or higher. *Source:* https://www.census.gov.
2. *Percent Change in GDP (GDP)*: Percent change in real GDP from 2017 to 2018. Here, GDP represents the inflation-adjusted market value of goods & services produced by the labor and property in the state. *Source:* https://www.bea.gov.
3. *Median Income (MI)*: Real median household income computed by the U.S. Census Bureau based on data from the Current Population Survey (CPS), the American Community Survey (ACS), and other surveys. *Source:* https://fred.stlouisfed.org.
4. *Regional Price Parities (RPP)*: Price indexes that measure the geographic price level differences. For example, an RPP of 120 means the prices within the state are on average 20 percent higher than the U.S. average. *Source:* https://www.bea.gov.
5. *Poverty Rate (PR)*: Poverty estimates are drawn from the Current Population Survey Annual Social and Economic Supplement (CPS ASEC), conducted three times per year with a sample of approximately 100,000 addresses. The Census Bureau determines poverty status by using an official poverty measure that compares pre-tax cash income against a threshold. *Source:* https://www.epi.org.
6. *Urbanization Percentage (UP)*: Percentage of state population living within urban areas. The Census Bureau classifies an urban area as "a densely settled core of census tracts and/or census blocks that meet minimum population density requirements." *Source:* https://www.census.gov.
7. *Land in Farms (LF)*: Includes (a) agricultural land used for crops, pasture, or grazing; (b) woodland and wasteland used in the farm operator's total operation, and (c) land owned and operated, as well as land rented from others. *Source:* https://www.nass.usda.gov (Page 6).

Demographic Covariates:

8. *Non-minority Population (NMP)*: Percentage of state population classified as White alone. *Source:* https://www.census.gov.
9. *Percentage Elderly (ED)*: Percentage of state population aged 65+. *Source:* https://www.census.gov.
10. *Percent Without Health Insurance (HI)*: Based on data collected for ages below 65 by the CPS ASEC and the American Community Survey (ACS). *Source:* https://www.census.gov.
11. *Obesity Rate (OR)*: Percentage of state adult ($\geq$ 18) population with body mass index of 30 or more, based on CDC Behavioral Risk Factor Surveillance System annual telephone survey. *Source:* https://obesity.procon.org.
12. *Political Preference (PP)*: Percentage of state-wide eligible voters identifying as "Democrat/Lean Democrat" in the 2017 Gallup Daily tracking dataset. *Source:* https://news.gallup.com.
13. *Religious (R)*: Percentage of religious population in the state according to a combined index based on four individual measures of religious observance. *Source:* https://www.pewresearch.org which summarizes the national 2014 Religious Landscape Study survey.
14. *Share of Immigrant Population (IP)*: Percentage of state population that are noncitizens. Based on the three-stage method by the Migration Policy Institute to assign legal status to noncitizen respondents in the U.S. Census Bureau Survey Data. *Source:* https://www.migrationpolicy.org.

https://doi.org/10.1017/S1748499525000089 Published online by Cambridge University Press

Geographic Covariates:

15. *Average Temperature (TP)*: Area-weighted state-wide averages based on climate data from the 344 continental U.S. Climate Divisions. For each division, monthly temperatures and precipitation values are calculated from daily observations. The dataset is manually augmented for AK and HI. *Source:* https://www.ncdc.noaa.gov.
16. *Average Relative Humidity (RH)*: annual historical daily average of "water vapor in the air relative to how much the air can hold," computed based on Continental U.S. Climate Divisional Dataset as for Average Temperature above. *Source:* https://www.ncei.noaa.gov.
17. *Average Dew Point (DP)*: annual historical daily average of "the minimum temperature an airmass can achieve given the amount of moisture in the air," computed based on U.S. Climate Divisional Dataset as for Average Temperature above. *Source:* https://www.ncei.noaa.gov.
18. *Population Density (PD)*: Ratio of state population divided by total geographic area of the state. *Sources:* https://www.census.gov/pop and https://en.wikipedia.org/areaofstate.

A.3.2. PCA Factor Loadings

Table A1. Factor loadings of the 18 covariates (rows, see Table 2) with respect to the first three PCA components (columns). Values are highlighted according to the PCA loadings: $|\cdot| < 0.2$, $|\cdot| > 0.7$. The last column shows the correlation between state covariates and Life Expectancy (LE) at birth in 2018 from CCenter for Disease Control and Prevention (CDC)

	Component 1	Component 2	Component 3	LE correlation
EA	**0.86**	−0.23	−0.06	0.61
GDP	0.20	−0.01	0.57	0.24
MI	**0.91**	−0.04	−0.02	**0.71**
RPP	**0.89**	−0.30	0.04	**0.71**
PR	**−0.74**	−0.38	−0.03	**−0.79**
UP	0.59	−0.45	0.43	0.48
LF	−0.24	0.14	0.64	0.04
NMP	−0.12	**0.84**	0.10	0.06
ED	−0.14	0.24	−0.53	−0.06
HI	−0.62	−0.03	0.57	−0.41
OR	**−0.78**	−0.19	−0.15	**−0.80**
PP	0.66	−0.45	−0.28	0.36
R	**−0.78**	−0.42	0.15	**−0.75**
IP	0.18	−0.32	0.60	0.28
TP	−0.30	**−0.80**	0.15	−0.30
RH	−0.06	−0.20	−0.63	−0.10
DP	−0.31	**−0.80**	−0.17	−0.35
PD	0.35	−0.51	−0.15	−0.00

https://doi.org/10.1017/S1748499525000089 Published online by Cambridge University Press

A.4. Groupings

A.4.1. Complete List of All Groupings

State s	PCA grouping $\mathscr{O}_s$
Alabama	Tennessee, Arkansas
Arizona	Nevada, Utah
Arkansas	Tennessee, Alabama
California	Oregon, Washington
Colorado	Utah, Nevada, Washington
Connecticut	New York, New Jersey
Delaware	Pennsylvania, Ohio
Florida	Georgia, North Carolina
Georgia	North Carolina, Florida
Idaho	Wyoming, Montana, North Dakota, Nebraska
Illinois	Wisconsin, Minnesota, Pennsylvania
Indiana	Ohio, Michigan, Missouri
Iowa	Wisconsin, Michigan
Kansas	Nebraska, South Dakota, Missouri
Kentucky	Tennessee, Arkansas, South Carolina
Louisiana	Arkansas, Tennessee, Alabama
Maine	New Hampshire, Vermont, Wisconsin
Maryland	Virginia, Delaware, New Jersey
Massachusetts	Connecticut, New York, New Jersey
Michigan	Ohio, Pennsylvania, Iowa
Minnesota	Wisconsin, Illinois, Oregon
Mississippi	Alabama, Arkansas
Missouri	Kansas, Tennessee, Indiana
Montana	North Dakota, Wyoming, Idaho, Nebraska
Nebraska	Kansas, Iowa, North Dakota
Nevada	Arizona, Utah
New Hampshire	Vermont, New York, Connecticut
New Jersey	New York, Connecticut, Maryland
New Mexico	Oklahoma, Missouri
New York	New Jersey, Connecticut
North Carolina	Georgia, South Carolina, Missouri
North Dakota	Montana, Idaho, Wyoming, Nebraska
Ohio	Indiana, Michigan
Oklahoma	Arkansas, Tennessee
Oregon	Washington, Nevada, Minnesota
Pennsylvania	Delaware, Ohio, Michigan
Rhode Island	Connecticut, New York, Delaware
South Carolina	North Carolina, Tennessee
South Dakota	North Dakota, Wyoming, Nebraska, Kansas
Tennessee	Kentucky, Missouri, Oklahoma
Texas	New Mexico, Arizona, Georgia
Utah	Nevada, Arizona
Vermont	New Hampshire, Maine, New York
Virginia	Maryland, Delaware, Rhode Island
Washington	Oregon, California, Colorado
West Virginia	Kentucky, Tennessee, Arkansas
Wisconsin	Iowa, Michigan
Wyoming	Montana, North Dakota, Idaho, South Dakota, Nebraska
Washington D.C.	Maryland, Virginia, New Jersey

Alaska and Hawaii.

To compute the grouping for $s \in \{\text{Alaska, Hawaii}\}$ that have no geographical neighbors, we first identify which state out of the other 50 minimizes the $\mathfrak{D}$-distance:

$$s_1 = \arg\min_{s_* \in \mathcal{S}} \mathfrak{D}(s_*, s) = \begin{cases} \text{Wisconsin} & \text{if } s = \text{Alaska;} \\ \text{Maryland} & \text{if } s = \text{Hawaii.} \end{cases}$$

https://doi.org/10.1017/S1748499525000089 Published online by Cambridge University Press

We then proceed as in Section 3.3, initializing with $\mathcal{N}_s = s_1 \cup s$. The resulting groupings are (Alaska, Iowa, Wisconsin) and (Hawaii, Maryland, District of Columbia).

A.4.2. Group Size

Fig. A.3 visualizes the sizes of $|\mathcal{O}_s|$ across the U.S.. The "default" procedure is to group each state with 2 other most similar states. A state $s \in \mathcal{S}$ ends up in a group of 4 if the latter are not geographically contiguous with it, that is a state is less similar to its neighbors than to other states that are further away. There is no particular pattern regarding regions where this occurs. Due to the sparse population of a cluster of states in Mountain West, these are in groups of 5 (6 for Wyoming) in order to achieve a total population of at least 5 million for each $\mathcal{O}_s$.

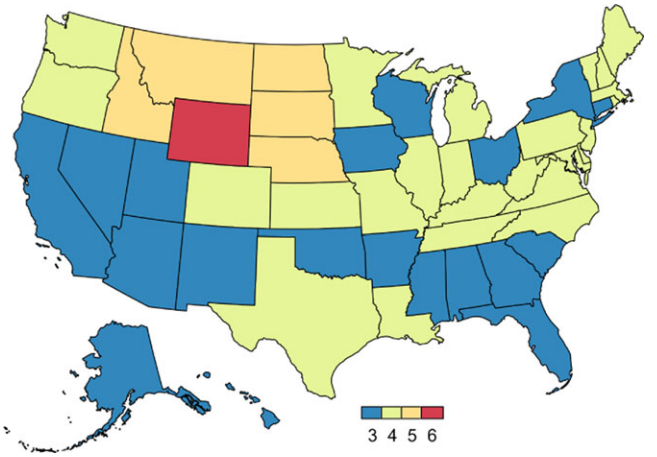


Figure A.3. Size of final state groupings, $|\mathcal{O}_s| \in \{3, 4, 5, 6\}$.

A.5. State Mortality at Age 75

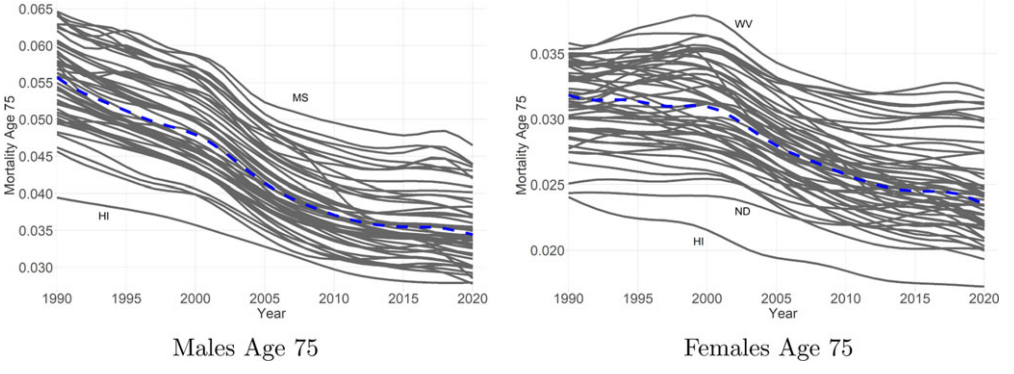


Figure A.4. MOGP-PCA mortality rates for age 75 Males (*left panel*) and Females (*right*) for years 1990–2020. SOGP-fitted U.S. nationwide mortality shown in blue. Most states continue to experience mortality improvement at this age as of 2020.

https://doi.org/10.1017/S1748499525000089 Published online by Cambridge University Press

A.6. Comparison of SOGP and MOGP State Rankings

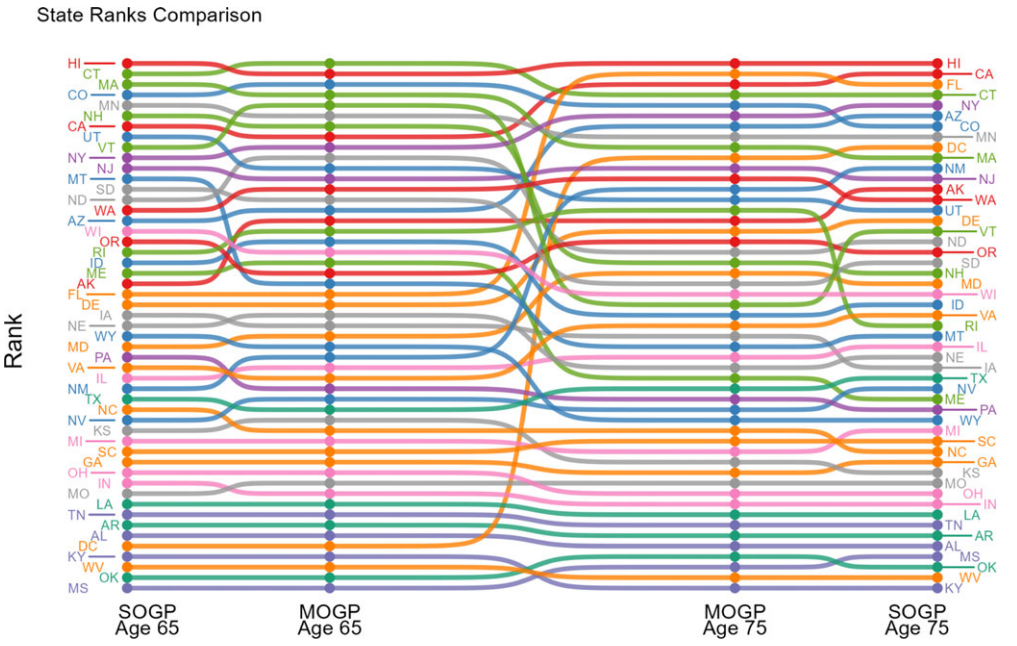


Figure A.5. Comparison of SOGP- and MOGP-based state rankings. We use Females at ages 65 and 75 in year 2018. The colors correspond to the nine geographical divisions defined in Appendix A.1. States with lowest (highest) mortality are at the top (resp. bottom).

A.7. Supplementary Plots for Mortality Improvement Factors

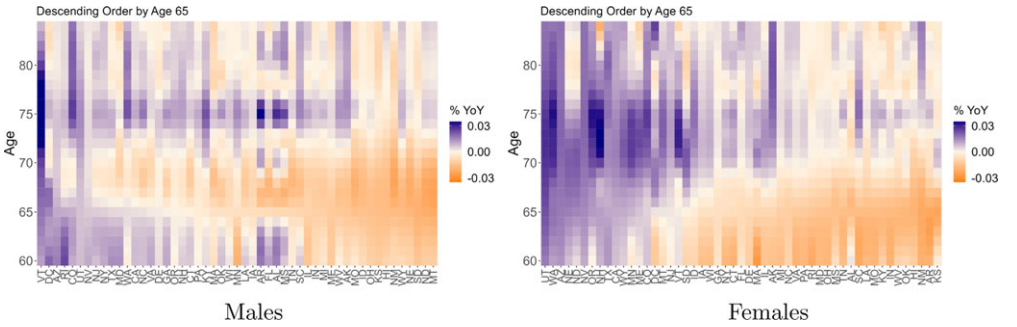


Figure A.6. MOGP-PCA mortality improvement factors in 2020 using the M52 APC kernel (5). Males (*left panel*) and Females (*right*). States are sorted by MI at age 65.

Cite this article: Ludkovski M and Padilla D (2026). Analyzing state-level longevity trends with the U.S. mortality database, *Annals of Actuarial Science*, **20**, 22–53. https://doi.org/10.1017/S1748499525000089

https://doi.org/10.1017/S1748499525000089 Published online by Cambridge University Press