



## EDITORIAL

# Gaussian process models in actuarial science

Mike Ludkovski 

Statistics and Applied Probability, UC Santa Barbara, USA

Email: [ludkovski@pstat.ucsb.edu](mailto:ludkovski@pstat.ucsb.edu)

(Received 19 January 2026; revised 19 January 2026; accepted 22 January 2026; first published online 26 February 2026)

### Abstract

Gaussian Process (GP) modeling is a probabilistic, non-parametric framework for describing spatio-temporal dependence that is well-suited for fitting risk-related surfaces. I summarize the main emerging actuarial use cases of GPs, including their applications in longevity modeling, insurance contract valuation, and loss development. The editorial also discusses further contexts with potential for GP-based approaches.

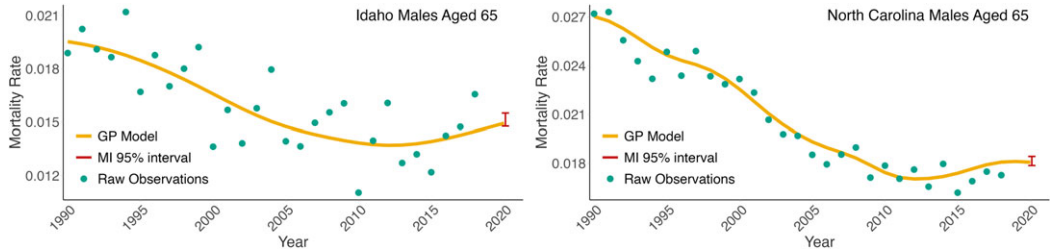
**Keywords:** Gaussian processes; longevity modeling; stochastic loss development; variable annuities

Gaussian Processes (GPs) (Ludkovski and Risk, 2025, 2026) offer a rich statistical learning ecosystem for probabilistic modeling of input–output relationships. GP regression is a supervised machine learning technique that fits a function to a set of training data. The GP quantifies the idea that one can fit an infinite number of such functions: along with the best prediction, the GP outputs the corresponding posterior distribution over functions, yielding data-adaptive uncertainty bands while providing an expressive non-parametric fit. Rather than attempting to overparameterize like in the deep learning paradigm, GPs need only a handful of hyperparameters, helping with interpretability. In contrast to neural networks that prioritize settings with enormous quantities of training data, GPs excel in low-data environments (arguably more common in actuarial contexts), where they moreover offer a transparent way to inject expert views.

Methodologically, GPs extend the multivariate normal distribution, being fully characterized by its mean function  $m(\mathbf{x}; \vartheta)$  and kernel function  $k(\mathbf{x}, \mathbf{x}'; \vartheta)$ . GP predictions are closed-form linear-algebraic expressions tied to the inverse of the covariance matrix  $\mathbf{K}$ , with the fitting effort devoted to learning the hyperparameters  $\vartheta$ . Conceptually, GPs engage the functional viewpoint, treating the underlying actuarial quantities not as a vector/matrix of discrete entries (say, loss development factors by year), but as a smooth curve, which removes the issue of data frequency. For instance, in mortality modeling, the GP representation naturally connects intra-annual and annual scales. Moreover, GPs subsume the temporal aspect, traditionally handled via time-series methods, dropping the Markovian assumption. If necessary, the latter can be recovered through certain kernels (e.g., the Exponential kernel is equivalent to the AR(1) process), and more generally, one can argue that Markovianity has been advanced for tractability rather than fundamental reasons.

### Longevity analysis

To manage longevity risk, actuaries construct dynamic life tables that reflect the dependence of population mortality rates on the two fundamental covariates of Age  $x_{ag}$  and calendar Year  $x_{yr}$ .



**Figure 1.** Male mortality rate at age 65 over time in Idaho (left) and North Carolina (right) fitted with (1). The vertical error bars indicate the 95% range of annual mortality change in 2020, cf. Ludkovski and Padilla (2025) for full details.

The typical modeled quantity is log-mortality,  $y(\mathbf{x}) = \log(D_{x_{ag}, x_{yr}}/E_{x_{ag}, x_{yr}})$  where  $D, E$  is the number of deceased/exposed in the given Year and Age bracket  $\mathbf{x} = (x_{ag}, x_{yr})$ , yielding a *mortality surface*  $\mathbf{x} \mapsto f(\mathbf{x})$  that captures the idealized hazard rate of input cell  $\mathbf{x}$ . Building upon the logic that a mortality of a 61 year old in 2019 is more similar to that of a 59 year old in 2018 than that of a 73 year old in 2025, a GP approach describes  $f(\cdot)$  through its covariance structure plus the observation likelihood  $y(\mathbf{x})|f(\mathbf{x})$  (Debón et al., 2010). Predictions of  $f(\cdot)$  correspond to de-noised forecasts, and after adding back the observation, noise can be converted into projections of future deceased counts, life expectancies, etc.

A common choice for the needed bivariate covariance kernel  $k(\mathbf{x}, \mathbf{x}')$  is the separable anisotropic Age-Period product  $k_{ag}(x_{ag}, x'_{ag}) \cdot k_{yr}(x_{yr}, x'_{yr})$ . For example, we may employ the Squared Exponential (SE) kernel in Age and a Matérn-5/2 (M52) kernel in Year:

$$k_{ag}^{(SE)}(x_{ag}, x'_{ag}) \cdot k_{yr}^{(M52)}(x_{yr}, x'_{yr}) = \eta^2 \exp\left(-\frac{(x_{ag} - x'_{ag})^2}{2\ell_{ag}^2}\right) \left(1 + \frac{\sqrt{5}r_{yr}}{\ell_{yr}} + \frac{5r_{yr}^2}{3\ell_{yr}^2}\right) \exp\left(-\frac{\sqrt{5}r_{yr}}{\ell_{yr}}\right), \quad (1)$$

with  $r_{yr} = |x_{yr} - x'_{yr}|$ , which necessitates fitting just three hyperparameters:  $\ell_{ag}, \ell_{yr}, \eta^2$ . The lengthscales  $\ell_{ag}, \ell_{yr}$  are interpreted as the strength of mortality dependence across ages and years. A typical  $\ell_{ag} \in [10, 20]$  captures demographic generations: two ages that are more than  $2\ell_{ag}$  apart are effectively uncorrelated in their evolution. We opt for different kernel families in (1) because mortality rates are smooth as a function of Age, but have a rougher structure in Year (to obtain a random-walk-like behavior, substitute a Minimum kernel for  $k_{yr}$ ). The construction is highly flexible; for example, to add a birth cohort effect, one augments the GP kernel (1) by multiplying with a third term  $k_{coh}(x_{yr} - x_{ag}, x'_{yr} - x'_{ag})$ . Similarly, to account for the Gompertz law whereby mortality is (nearly) exponential in Age, one includes a linear trend into the prior mean function  $m(x_{ag}, x_{yr}) = \beta_0 + \beta_1 x_{ag}$ , so that the GP is effectively modeling the de-trended residuals.

A key quantity of interest for actuaries is the *mortality improvement factors* (MI) that capture the annual rate of change of mortality over time. Figure 1 shows the raw mortality data of two U.S. states and their respective (separately obtained) de-noised GP estimates. As of 2020, the GP models predicted conclusive mortality deterioration in Idaho and a statistically insignificant improvement in North Carolina. Raw Year-over-Year ratios are too noisy, especially in the smaller Idaho dataset, and a probabilistic framework is essential to quantify the significance of such MI estimates. The analytic differentiation property of GPs leads to a tractable and consistent joint treatment of a mortality surface and its MI within a single GP model.

GPs are also fruitful in relational models that impose a hierarchical relationship across two mortality surfaces. This is essential for small population analysis, where stand-alone models lack credibility as there are insufficient death counts. One solution is deflator-based frameworks

(Hardy & Panjer, 1998) that take in an external mortality table  $m_{x_{ag}, x_{yr}}^{ref}$  and estimate the deflators  $\theta_{x_{ag}, x_{yr}}$  in

$$D_{x_{ag}, x_{yr}} \sim \text{NegBin}(e^{\theta_{x_{ag}, x_{yr}}} \cdot m_{x_{ag}, x_{yr}}^{ref} \cdot E_{x_{ag}, x_{yr}}; \omega). \quad (2)$$

In the recent contribution (de Melo et al., 2025), we used GPs to capture Age- and Year-specific deflators, successfully bridging between the national Brazilian life tables and the experience of a small (under 50K total exposed) Brazilian corporate pension fund. Deflators help with coherent projections: it is common to assume that the long-term MI of the modeled population matches that of the reference, which corresponds to making  $\theta_{x_{ag}, x_{yr}}$  asymptotically constant in  $x_{yr}$ .

**Bayesian lens:** The intrinsic connection between GPs and the Bayesian paradigm offers ample opportunities for probabilistic model assessment. First, one may impose hyperpriors on GP hyperparameters to inject expert opinion into the fitting procedure, which is especially convenient for long-term forecasting. Second, running MCMC, one can integrate over the closed-form GP predictive equations, providing hyperparameter posterior distributions to capture model uncertainty. de Melo et al. (2025) used this method to show the inadequacy of assuming age-constant deflators, or of using the Poisson rather than the Negative Binomial likelihood in (2). Third, the technique of Bayes Factors (BF) allows comparing the fit of different kernels, enabling compositional *kernel search*. Leveraging the fact that kernels can be multiplied and added and combining with the building blocks of Age-Period-Cohort (APC) kernels  $k_{ag}, k_{yr}, k_{coh}$  as in (1), offers an analog of the “standard” APC construction (Dowd et al., 2020). In Risk and Ludkovski (2024), we developed a genetic algorithm that evolves kernels according to hoist and mutation operators and then ranks them by BF fitness to explore thousands of compound kernels across a collection of national mortality surfaces. The following insights were obtained: (i) multiplicative structure is generally sufficient; (ii) cohort effects are heterogeneous: clearly needed in some countries, but effectively absent in others; (iii) the complexity of mortality surfaces (in terms of the presence of additive and nonstationary kernels) varies markedly, hence the number of APC blocks ought to be picked adaptively.

**Joint modeling:** A major theme in the past decade has been to encompass multiple populations simultaneously, such as capturing the joint male–female mortality trends, or doing comparative analysis across countries or sub-national areas (such as U.S. states, Ludkovski and Padilla (2025)). GPs offer an elegant representation of multiple mortality surfaces via the Multi-Output GP framework (Huynh & Ludkovski, 2021, 2024). In their essence, MOGPs embed the joint mortality dynamics through one-hot-encoding of the population index  $l$ , such that the overall covariance is separated into the product of the within-population (APC-style) structure and the cross-population covariance matrix. A parsimonious low-rank representation is sought for the latter, e.g., via a Coregionalization approach. In Huynh & Ludkovski (2024), we built a Kronecker linear-algebraic structure that splits populations across multiple axes, successfully treating up to 30 mortality surfaces aggregated across Gender-Cause-Country coordinates.

## Variable annuities modeling

Variable Annuities (VAs) blend investment vehicles and life insurance products. The most common VAs are of the Guaranteed Minimum Death Benefit (GMDB) type, offering a minimum payment to the beneficiary upon death, or the GMAB type (Guaranteed Minimum Accumulation Benefit), offering a minimum account value after a specified period. VA plans may have further riders, such as Return of Premium (RP), Annual Roll-Up (RU), Annual Ratchet, etc., that are combined with the above. For instance, in a DBRU plan, the death benefit (DB) is tied to stock and bond performance, with the benefit rate increasing annually (RU). Calculating the Fair Market Value (FMV) of VAs requires simultaneously accounting for the stochastic future value of market-based investments, embedded insurance guarantees, contract types, and

policyholder characteristics, and is usually handled through computationally intensive nested Monte Carlo simulations. To price  $N = 10^5$  policies each requiring  $M = 10,000$  market samples, the total simulation effort amounts to  $M \cdot N = 10^9$ .

In the spirit of machine learning, GPs can be employed for FMV assessment. Representing a policy as  $(\mathbf{x}, y)$ , the GP acts as a surrogate that provides a statistical prediction  $f(\mathbf{x})$  of a VA's market value  $y$  as a function of input features  $\mathbf{x}$ . Concurrently, *scenario reduction* is used to find a set of representative contracts that expedite GP learning without sacrificing the integrity of the analysis. One advantage of the GP surrogate is organically embedding feature selection. Given a  $d$ -dim. feature space  $\mathcal{X}$ , one employs a separable SE product kernel with Automatic Relevance Determination (ARD):

$$k(\mathbf{x}, \mathbf{x}') = \eta^2 \prod_{j=1}^d \exp(-(x_j - x'_j)^2 / (2\ell_j^2)). \quad (3)$$

ARD assigns a unique lengthscale parameter,  $\ell_j$ , to the base kernel  $k^{(SE)}(x_j, x'_j)$  associated to the  $j$ -th input. During optimization, feature relevance is assessed by the magnitude of  $1/\ell_j$ . Features that end up with a large fitted  $\ell_j$  make  $k^{(SE)}(x_j, x'_j)$  almost constant, thereby diminishing the influence of  $x_j$  in (3). Goudenège et al. (2021) use GPs to model GMDB contracts in the context of a Heston-Hull-White model for stochastic volatility and interest rates, studying the sensitivity of FMV to market model parameters. In their setting, the primary factors  $x_j$  are the account value, the respective volatility, and the interest rate, and among their goals is learning the fair fee structure for a given premium  $P$ , and estimating the Delta of the account value.

## Loss reserving

Another actuarial setting where probabilistic modeling of surfaces arises naturally is stochastic loss reserving. Estimation of unpaid loss liabilities requires ingesting the upper triangle of historical claims to project the distribution of the lower triangle of future claims. These are naturally indexed by the accident year  $x_{ay}$  and development lag  $x_{dl}$ . GPs provide a two-stage non-parametric framework to do so, first inferring a latent  $f(x_{ay}, x_{dl})$  and then constructing the predictive distribution of future realized losses  $Y$ 's. For instance, in Chain Ladder, the inferred logarithmic loss development factors are combined together to inflate to the ultimate lognormal development factor.

GPs allow modelers to obtain the full joint distribution of the lower triangle cells, useful, for example, for solvency reserving based on quantiles of ultimate paid claims. Following the pioneering Lally and Hartman (2018), who directly modeled the cumulative losses as a GP, several recent works applied GPs to represent incremental loss ratios. Ludkovski and Zail (2022) employed lag-dependent hurdle observation likelihood to capture the zero-inflated loss distributions; De Virgilis and Carnevale (2025) explored multiple kernel families in a related setting. Griffith (2025) used MOGPs to simultaneously model the paid and incurred loss triangles. All these setups worked with fully Bayesian GPs implemented in Stan in order to capture hyperparameter uncertainty. It is worth mentioning the nonstationary dependence structure whereby losses get more correlated for larger lags, which is resolved through the Gibbs nonstationary kernel in Griffith (2025) and through input warping in Lally and Hartman (2018).

## Outlook

There are plenty of open challenges where GPs have strong potential. In longevity modeling, the current elephant in the room is the impact of COVID. The global pandemic has highlighted the inherent unknowns of studying mortality: the possibility of never-before-seen events and the potential for structural breaks. The extreme spike in 2020 and 2021 mortality was a “5-sigma” event for pretty much any stochastic model, showing the limits of risk management planning and

rendering much of traditional goodness-of-fit metrics moot. The open question is whether mortality trends are back to “normal” or whether there is a new post-COVID trend. GPs equipped with the Changeoint kernel could be used to test whether COVID created a regime change in mortality dynamics. A related development is the newly available intra-annual mortality datasets (Jdanov et al., 2021) that unlock the potential to investigate the transient effects of environmental factors. Open questions concern disentangling epidemics (such as seasonal flu) from longer-term patterns, studying the impact of climate effects (heat waves, natural catastrophes), and the changing seasonal features. Through their continuous “spatial” lens, GPs are well-suited to consistently downscale to sub-annual frequency; respective seasonal patterns may be handled through the Periodic kernel that allows to separate out a (non-sinusoidal, non-parametric) repeating pattern from an overall trend. Coherence across long-term mortality forecasts is another ongoing challenge. GPs offer a framework that consistently blends a priori postulates (such as two mortality surfaces being asymptotically parallel) with data-driven projections.

Demographers aim to treat mortality in conjunction with fertility, looking for the full picture of population dynamics. Stochastic fertility models (which can be multi-output, such as separately modeling the birth rates of the first, second, and so forth child of a mother) remain in their infancy and would benefit from the above GP features. More investigation into custom observation likelihoods that best align with actuarial contexts (for example, input-dependent noise via heteroskedastic GPs for VA simulations) is also warranted.

Before concluding, let me highlight uses of GPs for yield curve modeling (Cousin et al., 2016) and financial hedging. There are undoubtedly many further actuarial problems where GP modeling will prove fruitful. Curious readers may consult the book Ludkovski and Risk (2025) and the student-oriented chapter Ludkovski and Risk (2026) for a more in-depth treatment of GPs.

**Data availability statement.** Not applicable as no new data were created or analyzed in this study.

**Funding statement.** There was no external funding.

**Competing interests.** None.

## References

- Cousin, A., Maatouk, H. & Rulli  re, D. (2016). Kriging of financial term-structures. *European Journal of Operational Research*, 255(2), 631–648.
- De Virgilis, M. & Carnevale, G. (2025). Applications of Gaussian-inverse Wishart process regression models in claims reserving. *Variance*, 18.
- Deb  n, A. M. A., Mart  nez Ruiz, F. & Montes, F. (2010). A geostatistical approach for dynamic life tables: The effect of mortality on remaining lifetime and annuities. *Insurance: Mathematics and Economics*, 47(3), 327–336.
- de Melo, E. F., Ludkovski, M. & Targino, R. S. (2025). Analyzing pension fund mortality with Gaussian processes in a sub population framework, arXiv preprint arXiv: 2506.03584.
- Dowd, K., Cairns, A. J. & Blake, D. (2020). CBDX: a workhorse mortality model from the Cairns–Blake–Dowd family. *Annals of Actuarial Science*, 14(2), 445–460.
- Gouden  ge, L., Molent, A. & Zanette, A. (2021). Gaussian process regression for pricing variable annuities with stochastic volatility and interest rate. *Decisions in Economics and Finance*, 44(1), 57–72.
- Griffith, D. (2025). Multioutput gaussian processes for loss ratio development. *Variance*, 18.
- Hardy, M. R. & Panjer, H. H. (1998). A credibility approach to mortality risk. *ASTIN Bulletin: The Journal of the IAA*, 28(2), 269–283.
- Huynh, N. & Ludkovski, M. (2021). Multi-output Gaussian processes for multi-population longevity modelling. *Annals of Actuarial Science*, 15(2), 318–345.
- Huynh, N. & Ludkovski, M. (2024). Joint models for cause-of-death mortality in multiple populations. *Annals of Actuarial Science*, 18(1), 51–77.
- Jdanov, D. A., Galarza, A. A., Shkolnikov, V. M., Jasilionis, D., N  meth, L., Leon, D. A., Boe, C. & Barbieri, M. (2021). The short-term mortality fluctuation data series, monitoring mortality shocks across time and space. *Scientific Data*, 8(1), 235.

- Lally, N. & Hartman, B.** (2018). Estimating loss reserves using hierarchical Bayesian Gaussian process regression with input warping. *Insurance: Mathematics and Economics*, **82**, 124–140.
- Ludkovski, M. & Risk, J.** (2025). *Gaussian process models for quantitative finance*. Springer Cham.
- Ludkovski, M. & Risk, J.** (2026). Gaussian processes for statistical learning in actuarial science. In Cohen, A. & Loke, S. H. (eds.), *An Invitation to Undergraduate Research in Risk Management. Foundations for Undergraduate Research in Mathematics*. Cham: Birkhäuser.
- Ludkovski, M. & Zail, H.** (2022). Gaussian process models for incremental loss ratios. *Variance*, **15**(1).
- Ludkovski, M. & Padilla, D.** (2025). Analyzing state-level longevity trends with the U.S. mortality database. *Annals of Actuarial Science*, 1–32.
- Risk, J. & Ludkovski, M.** (2024). Expressive mortality models through Gaussian process kernels. *ASTIN Bulletin: The Journal of the IAA*, **54**(2), 327–359.