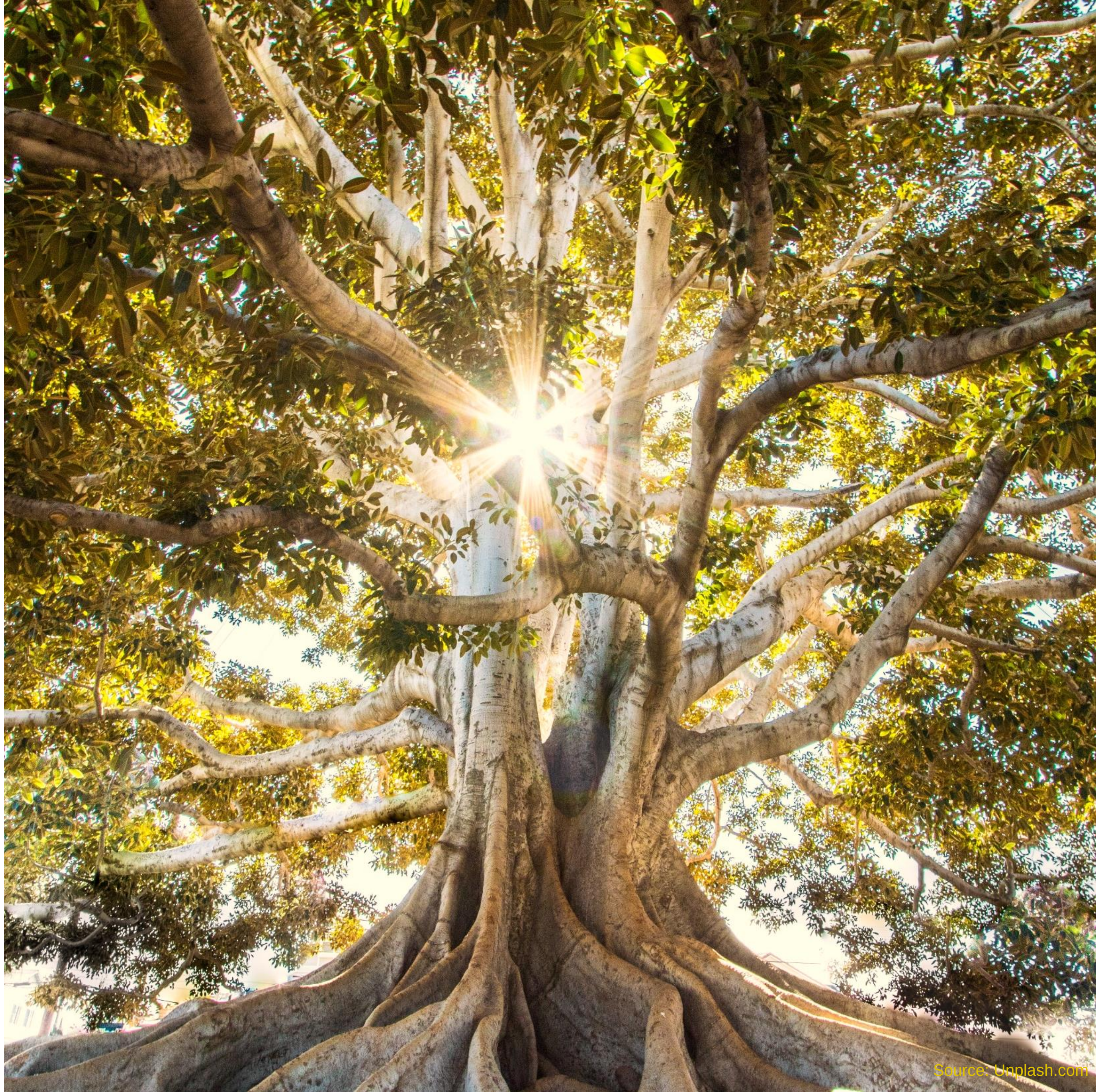
Modern approaches in mortality modeling considering the impact of COVID-19

Asmik Nalmpatian | LMU Munich - Department of Statistics, Allianz SE | Prof. Dr. Christian Heumann, Dr. Stefan Pilz

Insurance Data Science Conference | Advances in mortality modelling

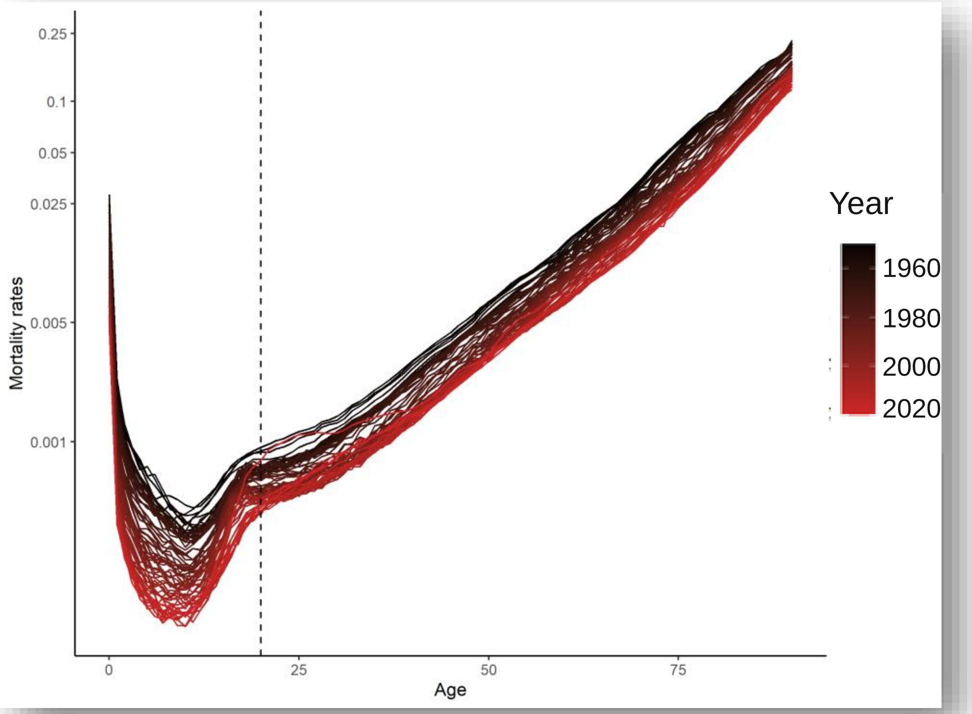
June 15, 2023


Source: Unplash.com

Navigating Mortality Trend Forecasting Amidst Event Uncertainty

Over the past two centuries, there has been a significant and consistent decrease in mortality rates across all age groups.

Evolution of mortality of the US female population aged 0–90 for the years 1950–2020



Challenge

- Uncertainty about the future mortality trends intensified by the COVID-19 pandemic
- Advanced modeling techniques are imperative to effectively capture the impact on mortality

Business goal

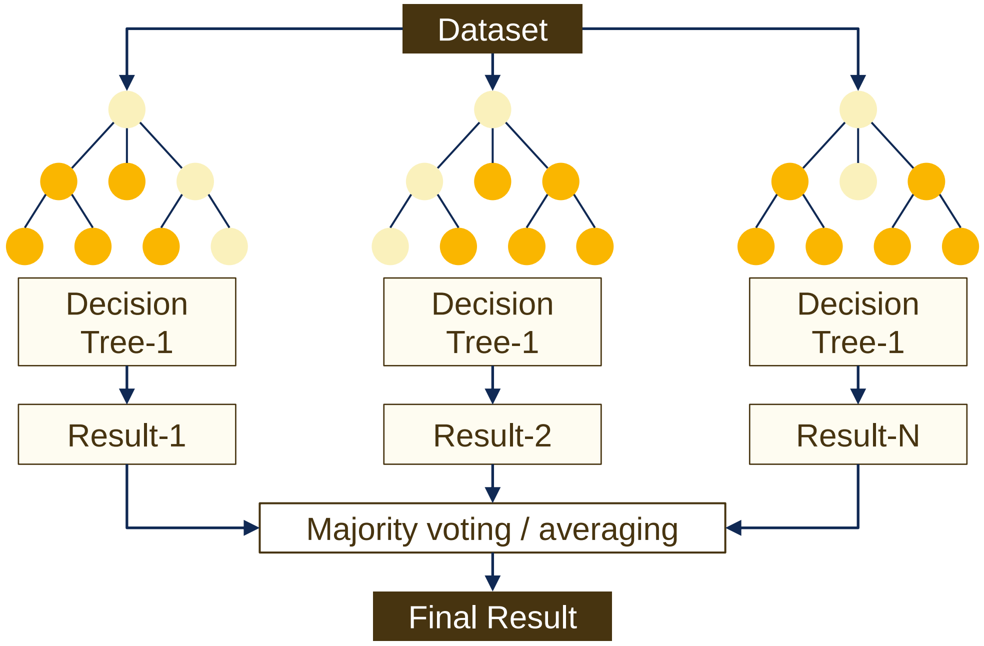
- Effective liability management and informed decision making for pension schemes
- Balancing premiums for optimal coverage and competitiveness for life insurers

I How can the state-of-the-art mortality models' performance be improved in terms of fit and forecast, while maintaining explainability?

II How will mortality develop in the future in different countries, considering the COVID-19 impact?

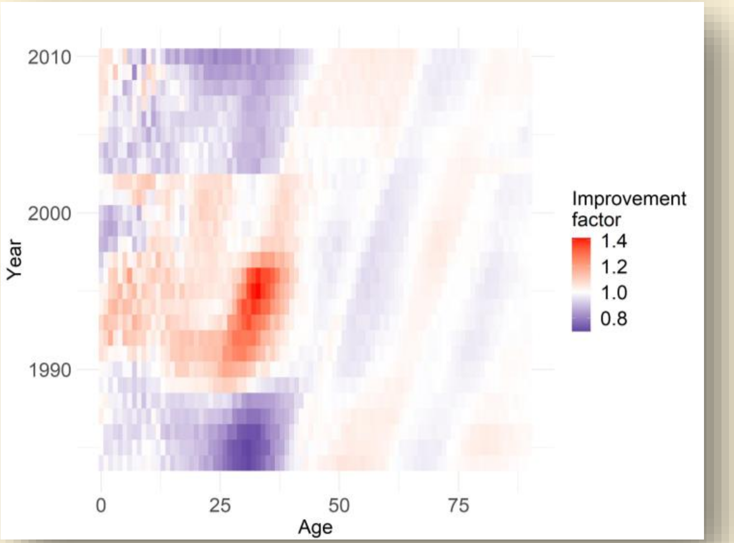
Unveiling Weaknesses and Bridging the Gap - Tree-Based ML Techniques vs. Traditional Mortality Models

Background – Random Forest



$\forall i \in \{1,...,n\}$, with n = (Age x Period x Country x Gender)

Application to the mortality context



$D_i \sim \text{Poisson}(E_i \cdot \mu_i)$

$D_i \sim \text{Poisson}(d_i \cdot q_i)$

with $d_i = E_i \cdot \mu_i^{LC}$

and $q_i \equiv 1$

Machine Learning improved mortality rates are then given by:

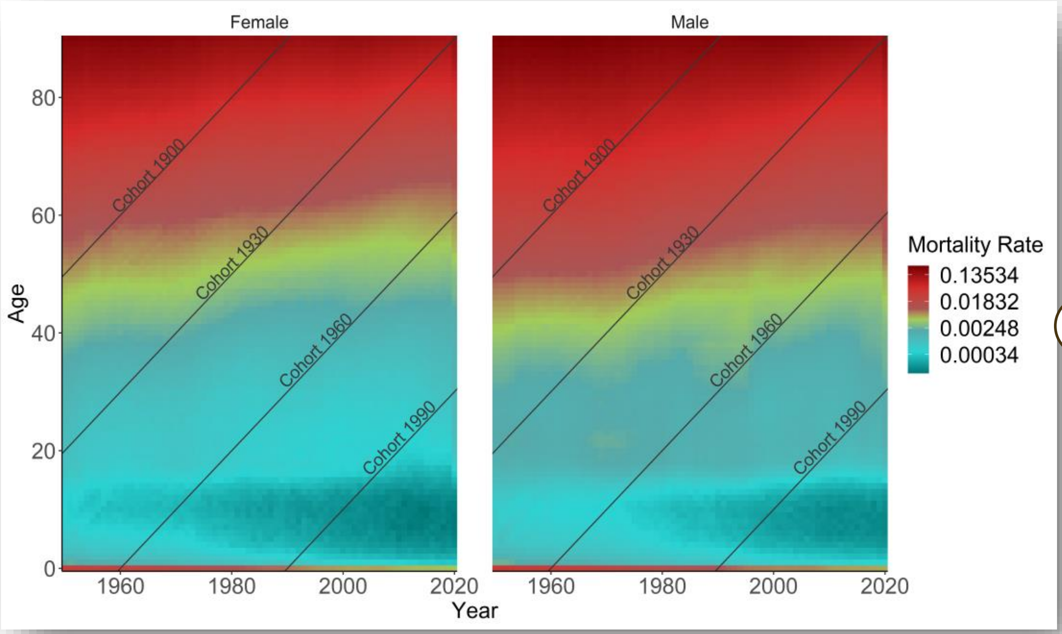
$\mu_i^{ML} = \mu_i^{LC} \cdot q_i^{ML}$

Proposal: Deprez et al. (2017), Levantesi and Pizzorusso (2019)

Idea: Utilize the original assumption of the Lee Carter (LC) model and evaluate by analyzing whether the constant factor q_i is equal to 1. Deviations indicate potential over or underestimations for specific variables and/or values.

Improving Modeling Performance - Comparing GAMs within APC Framework to Traditional Approaches

Background – Age Period Cohort Framework



∀i Є {1,...,n}, with n = (Age x Period x Country x Gender)

Application to the mortality context

- The number of deaths is still assumed to follow a Poisson distribution taking the form $D_i \sim \text{Poisson}(E_i \, . \, \mu_i)$
- Log link function establishes the relationship between the linear predictor and the expected mortality rates as follows: $\log \mu_i = \eta_i$ and the exposure $log(E_i)$ is treated as offset

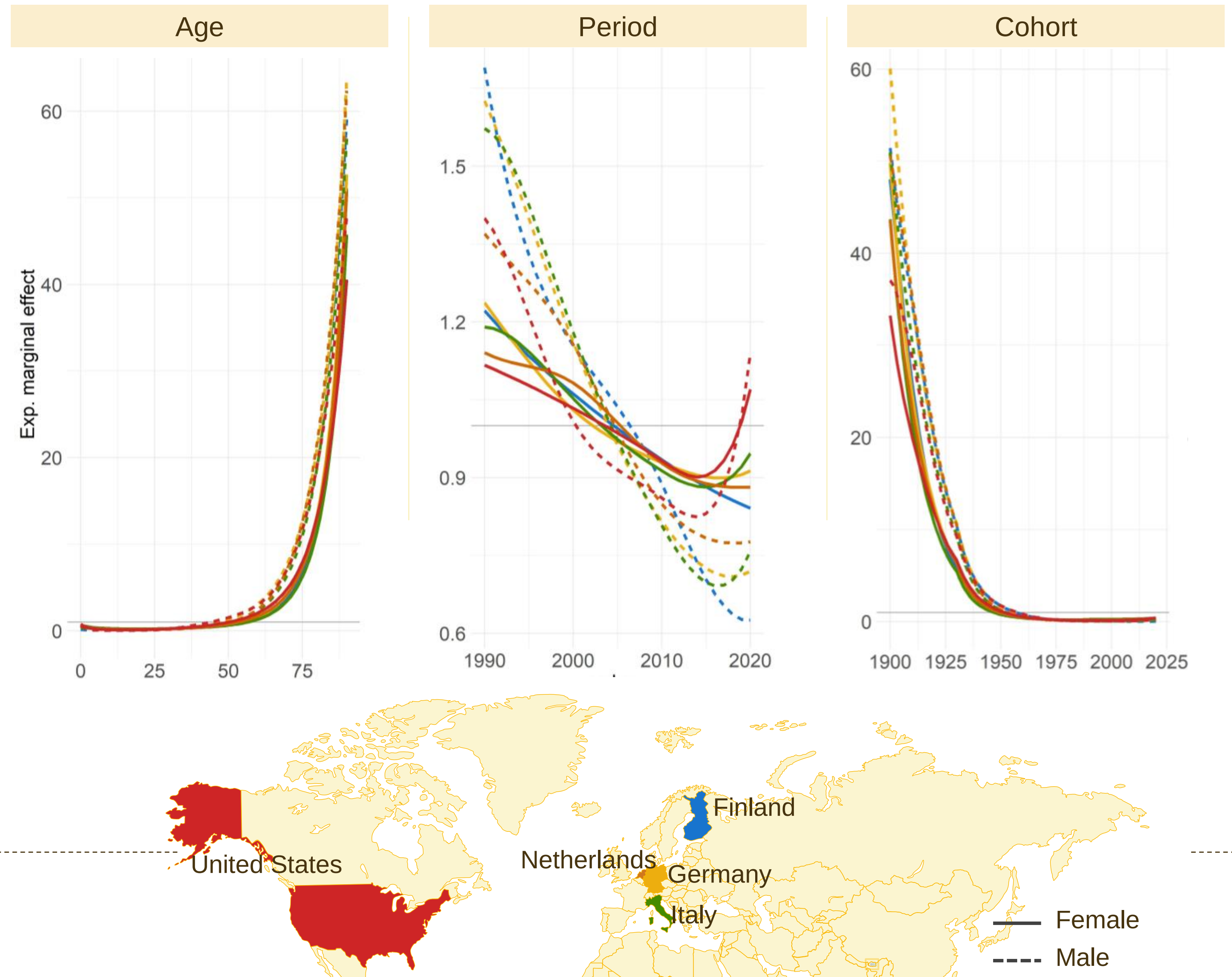
→ **Resulting model:** $\eta_i = \beta_0 + f_{country}(age_i, period_i) + log\ (E_i)$

Proposal: Clements et al. (2005)

Idea:

- Utilize GAMs for multi-populational cross-country approach
- Tackle the identifiability problem
- Implicitly capture cohort information

Unleashing the Potential of Effective Cross-Country Modeling



	ML	GAM
Explainability	✓	✓
Fit Performance Boost	✓	✓
Forecast Performance Boost	✗	✓

Projecting Trend Forecasts with GAM by Exploring 4 COVID-19 Scenarios

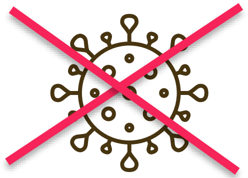
Scenario I

COVID-19 will disappear in the future

Consider 2020 and 2021 as outlier and therefore omit from training;

Assumption:

- COVID-19 is a residual, special event
- Excess mortality averages out
- No long-term effects of COVID-19 on health



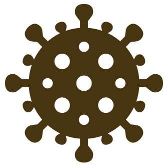
Scenario II

Expect full COVID-effect in the future

2020 and 2021 are not outlier, thus will be considered when training;

Assumption:

- COVID-related situation will continue just as it did in 2020 & 2021
- It will have the same effect on mortality over the next years

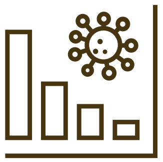


Scenario III

Flattening COVID-effect over years

Years 2020 and 2021 are not outlier, will be taken into account training but with *flattening effect* when testing;

Assumption: The effect of COVID-19 on health and mortality slowly (exponentially decay) flattens out over the next few years and fully disappears after few years



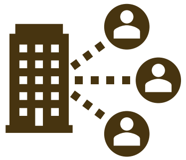
Scenario IV

Considering excess mortality

Years 2020 and 2021 are outlier, but the *excess mortality* must be taken into account;

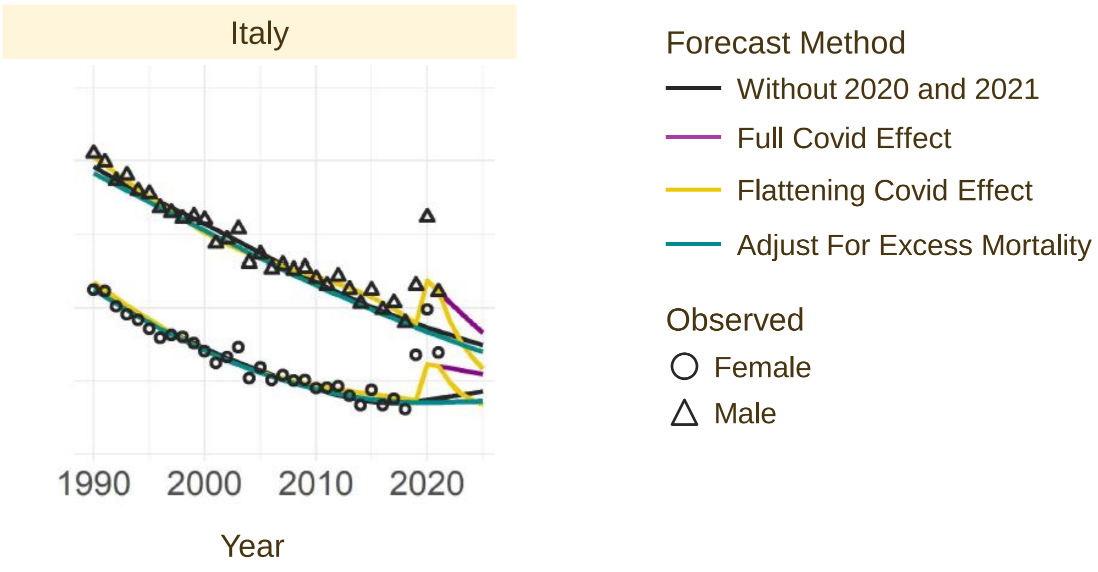
Assumption:

- Excess mortality does not average out over the coming years and must be explicitly accounted for
- Baseline mortality remains the same, so there are no behavioural changes due to COVID-19 in all age groups



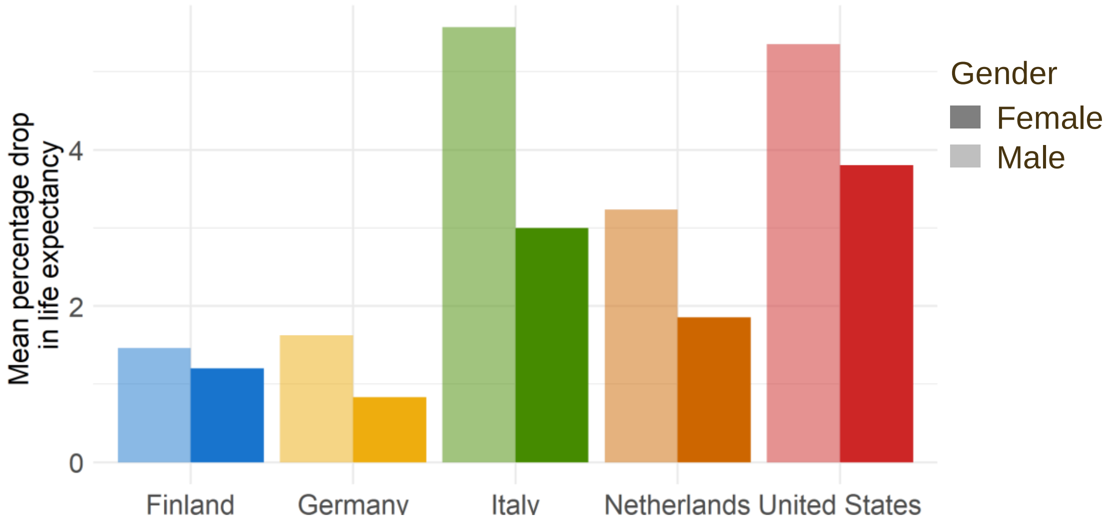
Understanding Mortality Modeling Trends under COVID-19 Impact Across Countries

Different forecast scenarios for Italian people aged 85



→ Extending the forecast until 2025, the four illustrative alternative scenarios encompass a range of expected impacts, representing varying levels of gravity from mild to severe.

The mean percentage drop is greatest for US and Italian men



→ In comparing the forecasts for life expectancy of individuals aged 50-90 between Scenarios I and II, a notable decline in life expectancy can be observed for some countries, as indicated by the predicted mean percentage drop.

Essential Insights & Key Takeaways

Improving Modeling Performance and Ensuring Explainability: Comparing Generalized Additive Models (GAM) and Tree-Based Machine Learning (ML) to Traditional Approaches

- ML techniques reveal weaknesses of traditional models, bridging the gap between standard and modern explanations
- GAM enables interpretable marginal effects
- Both techniques greatly enhance fitting performance
- GAM demonstrates superior forecast accuracy

Multipopulational Forecasting: Unleashing the Potential of Effective Cross-Country Modeling

- Achieve coherent forecasts across multiple countries
- GAM facilitate accurate and efficient cross-country fit and forecast

Embracing Expert Knowledge for Future Trend Forecasts: Navigating Extreme Events at the Edge of Time Series

- Expert knowledge crucial for trend forecasts considering extreme events
- Different countries will exhibit diverse mortality developments at various ages

Outlook

- Incorporate cause-of-death information to further explore COVID-19 impact and excess mortality
- Investigate on the relationship of socio-economic factors and causes of death

Main references

Lee, R.D., Carter, L.: Modeling and forecasting US mortality. How published of the American Statistical Association (1992)

Wood, S.N.: Generalized additive models: an introduction with R. CRC press (2017)

Hastie, J., Hastie, T., Tibshirani, R.: The Elements of Statistical Learning. New York: Springer (2016)

Reither, E., Hauser, R., Yang, Y.: Do birth cohorts matter? Age-period cohort analyses of the obesity epidemic in the United States. Social science and medicine (2009)

Girosi, F., King, G.: Demographic Forecasting. Princeton University Press (2008)Bai, J. and Perron, P. (2003). Computation and analysis of multiple structural change models, journal of applied econometrics, pp. 1–22.

Bank, W. (2021). The World Bank Group. last retrieved 30.04.2021. **URL:** databank.worldbank.org

Cairns, A. J., Blake, D. and Down, K. (2006). A two-factor model for stochastic mortality with parameter uncertainty: Theory and calibration.

Currie, I. D. (2006). Smoothing and forecasting mortality rates with P-splines. **URL:** http://www.macs.hw.ac.uk/iain/research/talks/Mortality.pdf

Clements, M. S., Armstrong, B. K., Moolgavkar, S. H. (2005). Lung cancer rate predictions using generalized additive models. Biostatistics 6(4): 576–589.

Deprez, P., Shevchenko, P. V. and Wuthrich, M. V. (2017). Machine learning techniques for mortality modeling.

Fahrmeir, L., Kneib, T., Lang, S. et al. (2007). Regression.

Lee, R. D. and Carter, L. (1992). Modeling and forecasting us mortality.

Levantesi, S. and Pizzorusso, V. (2019). Application of machine learning to mortality modeling and forecasting.

University of California, B. U. and for Demographic Research (Germany)., M. P. I. (2021). HMD: Human Mortality Database.last retrieved 30.04.2021. **URL:** www.mortality.org

Villegas, M. A. R. (2015). Mortality: Modelling, socio-economic differences and basis risk.

Weigert, M., Bauer, A., Gernert, J., Karl, M., Nalmpatian, A., Kuchenhoff, H. and Schmude, J. (2021). Semiparametric apc analysis of destination choice patterns: Using generalized additive models to quantify the impact of age, period, and cohort on travel distances.

Wood, S. (2017). Generalized additive models: An introduction with r.

Zeileis, A., Kleiber, C., Kramer, W. and Hornik, K. (2003). Testing and dating of structural changes in practice, computational statistics and data analysis, pp. 109–123.

Thank you!