Revision checklist - Reviewer #1

Paper “Enhancing diagnostic of stochastic mortality models leveraging contrast trees. An application on Italian data”

Submission ID: 51a8fa45-cb85-4c01-ad48-a73538820020

We would like to thank the anonymous reviewer for the helpful comments and valuable suggestions that contributed to improve our paper. We have revised the paper accordingly.
All of the reviewer’s comments and suggestions have been incorporated into the revised paper (see text in magenta).
We have provided a response to each of the comments below in blue text.

Referee:
Thanks to the authors for the efforts, explanations, and R-code. This proposal is an exciting extension for modeling mortality, but the main drawbacks remain in the paper.
On the one hand, the division into three age groups could be more or less, but it helps the interest in machine learning methods. The classical Lee-Carter method was proposed to adjust and predict all ages, which is a strong advantage.
On the other hand, authors advocate that contrast trees automatically identify the regions in which a given model provides a high error for certain combinations of ages and calendar years, which is easy to interpret. However, using traditional diagnostic methods such as residual diagnosis allows for obtaining the same.
Finally, the authors should compare models in the log scale (for example, Figure 2).

Authors:

We thank the Reviewer for his positive comment on our paper.
We also thank the reviewer for opening the floor to an interesting discussion on traditional diagnostic methods in mortality modeling. This is a crucial issue useful to mention but with any doubt too wide to address in our work and thus beyond the scope of this paper. Nevertheless, we take this chance to mention the canonical tools, such as the residual diagnosis, and the main difference with respect to the proposed method.
The following text has been included in the final manuscript:

"In mortality modeling, the objective of diagnostic checking is to ascertain whether the model fits the historical data by obeying an underlying probabilistic hypothesis. This procedure is carried out using residuals diagnosis checking with a Gaussian or more often a Poisson assumption (see, e.g., Renshaw, S. Haberman; Insurance: Mathematics and Economics 38 (2006) 556–570). Leveraging Friedman (2020), who introduces contrast trees to estimate the full conditional probability distribution without any parametric assumptions, we propose a prominent alternative, with particular regard to the intersection of Machine Learning and Mortality modeling fields. In this sense, our proposal fills the gap between mortality modeling and model diagnostics, particularly for nontraditional modeling as a machine learning framework."

We have now compared the model in the log scale (see Figure 2).