

We thank referee #2 for their review and the various suggestions to improve the manuscript. In the following we will respond in to the different comments and explain the general changes we intend to make to the manuscript based on them. The reviewer's comments are in black italics, our responses are shown in blue. All line numbers and references refer to the originally submitted manuscript.

I really enjoyed reading this manuscript on comparing the DLNM approach with the binning approach. It is a well-performed and well-written analysis and points to a useful approach if high-resolution data is not available. In some parts it was very, very detailed, on the other hand, it showed well how the authors conceptualized their alternative approach and how they planned to show its usefulness compared to the more common approach using DLNM.

We thank the referee for this positive and encouraging assessment of the manuscript.

I only have two questions as I mostly know the DLNM approach. To my knowledge, one advantage also lies in the fact that one can in addition adjust for further environmental factors which occur at the same time as the extreme temperature exposure, e.g. adjustment for humidity or air pollution. I assume that this would not be possible for the binning approach? Especially with these stratospheric and tropospheric conditions described in the paper, maybe also the accumulation of air pollution, similar as in inversions, could contribute its share.

In principle, the binning approach can be extended to multi-dimensional binning or stratification (e.g., jointly by temperature and humidity or other covariates), yielding higher-dimensional exposure-response relationships. Such extensions, however, require substantially increased data coverage to avoid sparse-bin artefacts, particularly for aggregated (weekly/monthly) or subnational datasets. They would also necessitate additional methodological choices that go beyond the scope of the present study. We will revise the manuscript text noting this possible extension and its data requirements, and that we use the binning approach primarily as a complementary robustness check alongside the DLNM.

My second question is on the population offset that can be used in DLNMs in case you look at a population over a longer time period. The population could increase or decrease and therefore this automatically leads to more or less death counts (e.g. also the percentage of the susceptible population could change). People sometimes account for that with an offset for each year in the model. Would that also be doable with the binning approach?

We agree that, when analysing mortality over longer time periods, changes in population size or composition can influence absolute death counts and are often accounted for in DLNMs via offsets or smooth temporal terms. In principle, analogous adjustments are

also possible within a binning-based framework. For example, one could remove slowly varying baseline components (e.g. long-term trends related to population change) and apply the binning approach to detrended or anomaly-based mortality data. Such extensions are conceptually straightforward but require additional methodological choices and are most relevant for long-term analyses. In the present study, which focuses on short event-centred attribution windows, population changes are negligible, and both the DLNM and binning approaches are applied to data where long-term trends are already controlled. We therefore do not expect population offsets to affect our results. We will make sure the manuscript text reflects those points.

And one technical point: The reference on page 12, line 232 should be "The Eurowinter Group", not just "Group".

We will fix this.