

Nat. Hazards Earth Syst. Sci. Discuss., author comment AC2
<https://doi.org/10.5194/nhess-2021-389-AC2>, 2022
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Reply on RC2

Lennart Marien et al.

Author comment on "Machine learning models to predict myocardial infarctions from past climatic and environmental conditions" by Lennart Marien et al., Nat. Hazards Earth Syst. Sci. Discuss., <https://doi.org/10.5194/nhess-2021-389-AC2>, 2022

Comment 1:

This paper evaluate different machine algorithm to predict myocardial infarctions (MI) using environmental exposures variables. The paper is well structured and clear. The authors found a poor performance on predicting daily or weekly MI levels, but a good performance on predicting the yearly values. I think the analysis and the interpretation of the results could be improved considering the following aspects: }

Response: We thank the reviewer for the review and these positive comments. We address the comments made, below.

Comment 2:

The authors considered a linear representation of the exposure (features). This is highly questionable especially for meteorological variables. It is well known that the effect of ambient temperature on several health outcomes is non linear with higher health impact for coldest and hottest temperature. This could be modelled using splines reparametrisation or in a simpler way using 3-piece linear spline (segmented linear) function (e.g. see Armstrong, Ben Models for the Relationship Between Ambient Temperature and Daily Mortality, *Epidemiology*: November 2006 - Volume 17 - Issue 6 - p 624-631). My suggestion is to consider a reparametrisation of temperature related variables considering both the heat and cold effects.

Response: We thank the reviewer for bringing up this important issue. Unlike more traditional methods in epidemiology we do indeed not make any prior assumption on the form of the exposure-response relationship between variables (meteorological or

otherwise). This is part of the data-driven study design, where little to no preconceptions of underlying mechanisms are assumed beforehand. Moreover, of the methods we have applied in this paper only ridge regression is a linear method. Decision Trees, Random Forests, Gradient Boosting as well as Multi-Layer Perceptron can tackle highly-nonlinear problems. The exposure-response models are very well suited to time series modelling as displayed in the Armstrong paper. Instead of a time series modelling approach, we use an approach based on multivariate machine learning regression models. These models do not require the presupposition of a known exposure-response relationship. We also point out that this study is aimed towards developing models to make long-term tendency projections at climate-timescales (i.e., 30 years). At such timescales underlying statistical properties may change gradually which would not be reflected by any prescribed exposure-response function derived from historical or current data. Instead, we hope that by letting the models pick these up based on the provided data alone, the application to ensemble data of climate simulations will provide an improved generalization. We also believe that it is of legitimate scientific interest to apply and evaluate models other than those that have been traditionally used, especially in the light of the success ML models have seen in other branches of science and technology.

We propose to address this in the revised paper by clarifying the reasoning behind our study design and giving a more concise overview of the differences to time series modelling, clearly pointing out limitations such as this one.

Comment 3:

The authors considered a lag of 3 days, and this would be enough for the heat effect of temperature and for most pollutants, but it has been shown that the cold effect could have a longer delayed effects (up to 3 or 4 weeks). I suggest to increase the lag up to at least 21 days.}

Response: We thank the reviewer for this excellent suggestion. However, increasing lag to 21 days for all variables would greatly increase the computational effort required to run the simulations and especially to conduct the hyperparameter tuning. We therefore propose to address this issue by increasing lag of only the temperature variables to 21 days and repeating the experiments, by adding another predictor, that estimates cold exposure during the past 21 days (e.g., overall minimum temperature during those 21 days).

Comment 4:

I think that the good performance at yearly level is totally expected as you are considering year and day of the year as features in your models, so environmental features doesn't seem having any role here. To see if the environmental features have a role on predicting the yearly values models without time variables should be tested.

Response: Here, the reviewer is raising an important issue. Adding the year as a feature understandably casts doubt on the relevance of environmental and other predictors in this study design, suggesting the model may simply learn the desired answers based on temporal predictors, effectively acting as a lookup-table. Before proposing a solution to this weakness, we would like to point out two factors here. First, the analysis of the variable importance does not support the notion that predictors other than year and day of the year are irrelevant. In fact, environmental predictors come out on top of the relevant predictors across all experiments carried out. Specifically, the year is shown in Figure A7 to have only low to medium relevance.

Second, we have taken great care to make sure that the models do not use any of the test (validation) data during training, i.e., to prevent data leakage. The correct responses could therefore not just be memorized by the models, because they have never seen them during the training step.

However, to accommodate this comment we propose to remove the year as a predictor entirely, repeating all experiments. While this comes with substantial effort, we believe that rectifying this issue would increase the robustness of the study results, and we will describe any differences.

Comment 5:

Some demographic features were considered and I think they are relatively stable over time, so I think they shouldn't contain any information at daily or weekly level. Changes on the demographic structure should be captured by the trend (year) variable. If the objective was to standardise the outcome the authors could consider to use as outcome the daily incidence of MI, perhaps considering logarithm values in order to have a more symmetric distribution, without considering demographic predictors.

Response: The reviewer is correct that demographic predictors generally do not undergo large changes in a matter of days or weeks. However, as explained in the methods section our models by construction require daily input for all predictors. This is a technical limitation that we can not change or alleviate. We expect changing to annual values for all days in the year would not change the ML predictions, and moreover, we likely are closer to the 'true' value of population and other demographic changes as they happen gradually within the year in the exposed population.

Comment 6:

I disagree with sentences in line 90-95. Actually it is possible considering case-controls design nested within time-series using case-crossover design, that is each case is matched with days before and after the case day and the association measured conditioning on

those risk set.

Response: We thank the reviewer for this important correction. We will change the sentence accordingly.