



Contents lists available at ScienceDirect

Environmental Pollution

journal homepage: www.elsevier.com/locate/envpol

Response to: "Letter to editor on 'impact of air pollution on symptom severity during the birch, grass and ragweed pollen period in Vienna, Austria: Importance of O₃ in 2010–2018' [environ. Pollut. 263PA (2020) 114526]"[☆]

Vicendese and Erbas (2020) commented on our recent publication examining the influence of selected air pollutants on symptom data in Vienna (Austria) during the birch, grass and ragweed pollen seasons from 2010 until 2018 (Berger et al., 2020). They acknowledged the impact of this study since it is one of the first contributions in this field working with crowd sourced symptom data. Additionally, they commented on general gaps in knowledge concerning the analysis of such data in general and the methods used in our study in particular.

The first comment addressed the use of a linear regression model in our methodology. They point to the fact that non-linear threshold effects of pollen and pollution can be found (e.g. Vicendese et al., 2013) and explain that Splines via Generalized Additive Models GAMs can be used to reveal linear effects of the exposure variable too and that they can have a higher prediction accuracy. However, our main goal in this study was to answer the question if a dependence of selected air pollutants and symptoms exist, which was shown by a significant effect in the linear regression model. The effect size was provided only in addition, so in difference to GAM, prediction is not the main goal in our setup. The transformation usually does not provide perfect linear data, but the possible relationships are made more linear. Ravindra et al. (2019) stated in their review that besides all the advantages of GAM such flexible regression models number more degrees of freedom and is required than for a regular linear regression model. Moreover, there is concern about the lag value for GAM. We thus decided to apply in this first contribution a simpler approach. However, first preliminary experiments (using the R package mgcv (Wood 2017) show that the explained deviance does not drastically increase when introducing a smoothing function to our model setup, e.g. for our more complex model for POAC pollen and ozone the explained deviance was 89.8% and increases by about 1 percentage point with the same covariates.

Hastie and Tibshirani (1987) proposed the possible application of GAM as "diagnostic tool for identifying parametric transformations of the covariates in a standard linear analysis". This might be in fact a very interesting way forward.

However, this does not imply that we reject the application of GAM. In fact, we would be interested in trying to apply such tools

to our datasets to check if they are suited for these methods in future approaches as we specify just above and also below at the end of our reply of correspondence.

The second comment related to auto-correlation and skewness of the data. To tackle the second, we transformed our data on a sqrt scale for pollen and on a log scale for temperature to make distances between values more even (close to symmetry for the purpose of variance correction), which was comprehended and justified by Vicendese and Erbas (2020). To avoid problems with auto-correlated residuals we worked with two approaches, differenced values and built a more complex model to remove this auto-correlation.

Exposure variables such as pollen and air pollutants are time varying by nature (e.g. the seasonal dependence of pollen occurrence) and Vicendese and Erbas (2020) correctly stated that we treated the different years and the different days within a pollen season as if the effect would be independent of it. We agree completely that in future studies this could and should be an improvement, however given the limited number of days in a single pollen season and in general the burden to collect data, the data availability might not always allow it.

The implementation of GAM is proposed to provide a visual exploration to necessitate interpretability. We are aware of concerns with our analyses and discussed them or considered them (e.g. auto-correlation) in a separate "Limitations" section. Therein, we stated that the statistical models and methods used cannot imply causation and that spurious correlation due to hidden factors might exist or could be reported in future studies (using other models and methods). Additionally, we agree with the statement that the untransformed variables do not have a linear association hence the introduction of the transformations was performed. A possible threshold effect could be modeled explicitly in an extended linear model, but of course also implicitly in a GAM. Both approaches need further analysis.

However, after this first big analysis of our data there is of course a range of tools that could be applied and we are grateful for those concrete suggestions. Our model can be tested and developed more with the focus on prediction and effect size, which was not the first and foremost goal in our study. It needs to be awaited to have results from an analysis with GAM if this approach works (better) for our datasets.

Finally, Vicendese and Erbas (2020) emphasize the importance of allergic respiratory disease as a public health problem causing

DOI of original article: <https://doi.org/10.1016/j.envpol.2020.115790>.

[☆] This paper has been recommended for acceptance by Eddy Y. Zeng.

a continually increasing economic burden. We completely agree with them on the critical effects of allergies and the gaps in knowledge that still exist concerning the time periods of greatest risk for exposure to these environmental factors and the complexity of modelling such factors due to collinear relationships and the seasonality, cyclicity and non-linearity. There are still many lessons that will have to be learned now at the beginning of the era of exploiting symptom data to understand the true burden and the development and onset of symptoms of persons concerned. Therefore, we are highly interested in exploring the predictive performance in terms of accuracy and reliability of different models among them GAM as they have proposed in future studies and want to stimulate the use of innovative modelling frameworks.

We are therefore thankful for the important hints made by [Vicendese and Erbas \(2020\)](#), that provide directions for future studies and aim to examine those options with our datasets and the proposed new tools.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- [Berger, M., Bastl, K., Bastl, M., Dirr, L., Hutter, H.-P., Moshhammer, H., Gstöttner, W., 2020. Impact of air pollution on symptom severity during the birch, grass and ragweed pollen period in Vienna, Austria: importance of O₃ in 2010–2018. *Environ. Pollut.* 263, 114526.](#)
- [Hastie, T., Tibshirani, R., 1987. Generalized additive models: some applications. *J. Am. Stat. Assoc.* 82, 371–386.](#)
- [Ravindra, K., Rattan, P., Mor, S., Aggarwal, A.N., 2019. Generalized additive models: building evidence of air pollution, climate change and human health. *Environ. Int.* 132, 104987.](#)
- [Vicendese, D., Erbas, B., 2020. Correspondence. *Environmental Pollution*.](#)
- [Vicendese, D., Olenko, A., Dharmage, S., Tang, M., Abramson, M., Erbas, B., 2013. Modelling and predicting low count child asthma hospital readmissions using General Additive Models. *Open J. Epidemiol.* 3 \(3\), 125–134.](#)
- [Wood, S.N., 2017. Generalized Additive Models: an Introduction with R, second ed. Chapman and Hall/CRC.](#)
- Markus Berger*
Paracelsus Medizinische Privatuniversität, Strubergasse 21, 5020, Salzburg, Austria
- Research Group Aerobiology and Pollen Information, Department of Oto-Rhino-Laryngology, Medical University of Vienna, Währinger Gürtel 18–20, 1090, Vienna, Austria
- Katharina Bastl
Research Group Aerobiology and Pollen Information, Department of Oto-Rhino-Laryngology, Medical University of Vienna, Währinger Gürtel 18–20, 1090, Vienna, Austria
E-mail address: katharina.bastl@meduniwien.ac.at.
- * Corresponding author. Research Group Aerobiology and Pollen Information, Department of Oto-Rhino-Laryngology, Medical University of Vienna, Währinger Gürtel 18–20, 1090, Vienna, Austria.
E-mail address: markus.berger@stud.pmu.ac.at (M. Berger).

1 October 2020
Available online xxx