Response to reviewer comments: "Temporal and spatial analysis of ozone concentrations in Europe based on time scale decomposition and a multi-clustering approach"

Eirini Boleti, Christoph Hueglin, Stuart Grange, Andre Prevot, Satoshi Takahama

Response to reviewer 2

The manuscript presented by Boleti et al. examines trends on surface ozone concentrations across a number of stations over Europe for the period 2000-2015. They use a time-scale decomposition to analyse long-term (LT), seasonal (S) and short-term (W) variations. Then, they apply a clustering technique and they finally calculate the trends in the clusters obtained. In addition, they analyse the ozone-temperature relationship over the different clusters and sub-periods. Their classification is consistent with previous studies and their results show a general decreasing in the ozone concentrations, mostly in the ozone peaks. In addition they find a reduced sensitivity in the ozone- temperature relationship over most of the clusters defined. Overall, I found the manuscript very interesting and complete. The methodology applied is robust and consistent, as well as the results presented. However, I also think that there are some parts in the current version that should be improved in order to be published, in particular the methods sections (see my comments below). In my opinion the manuscript might be a good contribution to Atmos. Chem. Phys and the scientific community. Therefore, I would be happy to support the publication of the present manuscript after addressing some comments, which I consider that would be useful to improve it. I have a few general comments and some specific comments:

Main Comments

There are some parts in the methods section that are not very clear, and in my opinion this section is essential to follow the manuscript. Therefore I have some comments and questions that I would like to ask the authors:

1. Section 3.1. Time scale decomposition (page5): The authors should explain in more detail the IMF. How the number of coefficients (c j) is selected? The authors say "By adding together IMFs with frequencies around 40 days and 3 years we obtain the seasonal variation of O3 (c 1 +...c 10)", but why 40 days and 3 years? Is this based on the previous study from Boleti et al.2018? I think that this information should be included in order to help the reader to better understand the methodology.

The selection is discussed in Boleti et al. (2018) and for the reader that is more interested on that matter we refer to the above publication. We have added the sentence: "A more detailed discussion on the choice of the IMFs for the seasonal and short term variations can be found in Boleti et al. (2018)."

2. Sections 3.3-3.4 Long-term trends (page 6, 7): I understand that for the peaks of O3 metrics the method explain in section 3.3 cannot be applied. But, it would possible to use the same method, i.e. GAMs models also for daily mean and MDA8 O3, wouldn't it?

   Yes, that would be another good approach to estimate meteorological influence of daily mean and MDA8 O_3.

3. Why the authors define the warm season as May-September? Why April is not included? I think this should be further clarified, since usually ozone season ranges from April to September (e.g. EEA, 2019, Fleming et al. 2018)

   This is based on the estimation of the MTDM which In previous reports by the EEA refers to the period between May and September. Thus, for consistency and to compare our study with the EEA studies, we keep this period of the year as representative for occurrences of peak O_3. We now mention the above in the manuscript: "The models are fitted for the warm season May-September as by definition the MTDM refers to this period of the year."

4. Reading the modelling part (section 3.3, 3.4, page 6) is not clear the input data to calculate the trends, e.g. the GAM models are fitted to each cluster that contains a number of stations, so the models are applied individually to each station, am I right?

   A GAM model is applied to each station separately and the trends are calculated for the individual stations as well.

5. Regarding the analysis of seasonal cycle of O3, why do the authors chose the mean of O3 and not the MDA8 O3?

   The reason we use daily mean O3 in our analysis is to be consistent throughout the manuscript and be able to make comparisons between the different O3 metrics and clusters.

Specific comments

1. L26-30 of page 3. The authors applied a filter to obtain the time series, and only those with a maximum of 15% of missing values and maximum of 120 consecutive days are used. Is this 15% applied to whole period (16 years) or each year? And the consecutive days? I assume that they refer those 120 consecutive days in one year, is that correct? Can the authors clarify this?

   15% missing data and the 120 consecutive days refer to the whole period of measurements. Thus, we clarify this with the following addition: "time series with a maximum of 15% of missing values, and a maximum of 120 consecutive days with missing values are used for the whole period of measurements, leaving the study with 291 sites across the European domain."

2. L1 of page 5. I would add that the clusters are identified by using each component L(t), S(t), W(t) separately to the algorithm.

The following sentence is already mentioned in the manuscript: "For identification of the clusters the LT(t), S(t) and W(t) of the daily mean and MDA8 O_3 were used as input time series in the PAM algorithm."

3. L4 of page 8. What are the temperature ranges considered?

   The temperature ranges from 7 to 35 degrees Celsius in intervals of four degrees; (7,11] (11,15] (15,19] (19,23] (23,27] (27,31] (31,35]

4. L4 of page 9. Why do the authors leave the results of MDA8O3 in the supplement and the results if the O3 in the main text? Wouldn't it be more interesting to see the results for MDA8O3?

   We refer to our answer of the main comment 1 from referee 1. In summary, O3 daily mean and MDA8 clusters are rather similar, the choice of clusters between the two metrics does not affect the conclusions. In addition, for comparison with other studies and consistency throughout the manuscript, we believe it is more interesting to present the daily mean O3 clusters.

5. L19-23 of page 9. Please refer to figure 3.

   Done.

6. L25 of page9. Just a comment regarding Fig.4. the colours for "Po Valley" and "Central South" maybe could be changed, they are quite similar and it is hardly to distinguish the stations that belong to each cluster.

   This is a valid remark, we changed the color of the Po Valley cluster to a darker blue color.

7. L4-L20 of page 10. Should Figure 5 be referred here? I couldn't find any reference to Fig.5.

   This is true, the figure is now moved to the section 4.4. O3 seasonal cycle trends, after a similar suggestion of referee 1.

8. L3 of page 11. "decreasing O3 trends", maybe it should be specified "decreasing daily O3 means".

   That is right, it is corrected to "decreasing daily O3 means".

9. L4 of page 11. In my opinion, the table 1 with the number of stations should be introduced before (e.g. when presenting the clusters).

   Indeed, we have moved this table in section 4.1. Cluster analysis.

10. L9 of page 14. Where these percentages 62% and 18% came from exactly? Is there any figure to support this? This question is in the line of one my previous comment (i.e. how the models are fitted).

    We do not have a figure to support this. These percentages refer to the sites, where negative trends were estimated from all studied sites in our data set, i.e. ratio: sites with negative trends/number of all sites, 62% before and 18% after meteo-adjustment.

11. L10 of page 19. In the North sites the variability of temperature is lower and O3 is also more influenced by transport.

This is indeed a useful addition to the manuscript. "In addition, in these northern regions the variability of temperature is lower compared to the central and southern parts of Europe, while O_3 concentrations are more influenced by intercontinental transport mechanisms."

12. L24-26 of page 21. The authors attributed the decreasing O3-Temperature to NOx reductions, and this is likely one of reasons (but it is not showing here) and then, they mention that "changes in the sensitivity across sites are mainly driven by regional meteorological conditions", so what about the NOx emission reductions just mentioned? I think this last paragraph is important and it must be rewritten.

We rewrote this paragraph to the following one: "Finally, the sensitivity of O_3 to temperature has weakened since 2000 with a rate of around 0.04 ppb/K/year, i.e. formation of O_3 became weaker at high temperature conditions, that can be attributed to the decrease of NO_x concentrations. It was shown that the trend of the sensitivity differs across sites that are influenced by different meteorological conditions."

References

Boleti, E., Hueglin, C., Takahama, S., 2018. Ozone time scale decomposition and trend assessment from surface observations in switzerland. Atmospheric Environment 191, 440–451.