Atmos. Chem. Phys. Discuss., 15, C3630–C3633, 2015
www.atmos-chem-phys-discuss.net/15/C3630/2015/

Atmospheric Chemistry and Physics Discussions

Interactive comment on "Investigating the observed sensitivities of air quality extremes to meteorological drivers via quantile regression" *by* W. C. Porter et al.

Anonymous Referee #1

Received and published: 15 June 2015

This seems to me a very nice paper. It furthers an understanding of the meteorological drivers of air pollutant extremes across the U.S., both how they differ regionally and seasonally, and the differences between ozone and aerosol pollution. It also helps clarify the meteorological drivers of extreme events. I think this will be a valuable contribution.

I would recommend publication after the rather minor comments below have been addressed.

Major Comment:

1. My main concern is that the methodology is not always clear. (i) A central theme of the paper concerns quantile regression. I would guess this procedure is not widely known in the meteorological/chemical community. I would recommend adding a short section to the paper within the methodology section explaining in more detail what quantile regression is. (ii) The analysis procedure and variable selection were not clear to me. I read the relevant section several times and still did not come away with a precise understanding of the procedure. The authors need to take the time to fully explain their procedure. Maybe a schematic diagram would help (also see minor comments below).

Minor Comments:

1. P14077 L13-14: It is my understanding that measures in Beijing were not taken because of a "particularly extreme events" but because of normal high pollution levels. Paris in the last few years might be a better example of extraordinary measures taken during high pollution events.

2. P14075 L5: "fans out". I think I know what you mean, but it would be better to explain more explicitly instead of using a term in quotes.

3. P14079 : As the paper is nominally about quantile regression more background on the methodology would be appropriate as it may not be generally known. The paper goes over this in a few sentences in the introduction and provides a nice example (Fig. 1) but it would make sense to educate the community in somewhat more detail.

4. P14080, L21: "all" is a strong word. I suggest you delete it.

5. P14080, L8: "averages" – this seems to imply all variables are 3-hour averages. This does not seem consistent with some of the variable descriptions.

6. P14081: I believe the RPI as defined previously is actually equal to the ratio of vector/scale sums. This would make a low RPI (close to 0) indicative stagnant air masses.

ACPD
15, C3630–C3633, 2015

7. P14080: Variable generation. Some variables the authors averaged regionally (e.g., tke), some they do not. It would be appropriate to provide some rationalization for which variables are averaged regionally.

8. P14083, L17: “pollutant levels”. I think you said this previously – but it might be worthwhile reminding the reader here which metrics you use for ozone and pm2.5 (e.g., daily average?)

9. P14084: The procedure to select variables here is not altogether clear to me. I have read this section a number of times and am still unclear on the exact procedure. The authors should make sure it is clearly explained. Maybe a diagram would be helpful here?

10. P14085: “summed inverse rank threshold of at least 2”. This is not clear to me. Perhaps when the procedure above is explained in more detail this will also become clear.

11. P14085, L22-23: “multivariate quantile regression”. Here I assume it is linear regression?

12. P14086, L7: “frequency of appearance”. Actually the frequency of appearance is not shown, but the number of stations is shown. I would suggest showing the actual frequency would be a better metric.

13. P14087, L7: “inverse correlation”. I assume by inverse the authors mean a negative correlation.

14. Fig 1. Please state the seasonality of the measurements and how many years are used in the figure caption.

15. Table 1. Some of these variables names are not obvious and could be explained better with a footnote. What is categorical rain, best lifted index, the difference between apcp and prate, projected cloud cover? I suspect turbulent kinetic energy was generated in the boundary layer scheme – please clarify?

ACPD

15, C3630–C3633, 2015

17. Fig. 3. I had to blow this figure up to make anything out of it. I would suggest making the panel sizes bigger and possibly separating into separate figures. I always find it hard to match colors precisely. In the lower panel in each figure the individual sensitivities should be specified (the names for these variables could probably be easily shortened). In addition either in the text or the figure caption it should be specified to what extent the correlations are significant.

18. Fig 7. At what level are these slopes significant? I was struck by how similar the results were across the different quantiles. Even a 10% change seems rather small. Is this really significant? In general throughout the paper a number of correlations and regressions are made. The authors should really comment on the significance of these quantities.

Interactive comment on Atmos. Chem. Phys. Discuss., 15, 14075, 2015.