Comment on acp-2021-376
Anonymous Referee #2

Referee comment on "Spatiotemporal Variability in the Oxidative Potential of Ambient Fine Particulate Matter in Midwestern United States" by Haoran Yu et al., Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2021-376-RC1, 2021

Yu et al report on extensive measurements of PM2.5 OP (oxidative potential) based on an analysis involving 5 different acellular approaches. The analysis was performed on samples collected at a number of sites in the midwestern US and the paper reports on comparisons between the assays and PM2.5 mass. It is stated that a second paper will focus on the PM2.5 chemical components driving these results. The paper is based on a substantial amount of work and provides more insights into the utility of current ways to characterize OP, and it also sheds light on the potential usefulness of using these assays in health studies.

A major conclusion is that the poor correlation between all the various assays, when compared at one site, (and this is largely true for all the sites), implies all these types of OP assays are needed for health studies. One could also conclude, that all of these assays (except possibly one) are each deficient, and no ideal assay exists. It may also even suggest that if no comprehensive OP assay is available, then maybe the approach is flawed since the goal of using these assays was to develop a comprehensive single measure of aerosol toxicity. Since this group of assays appears to fail in demonstrating this goal, instead maybe one should focus on the specific species that drive OP and not use these assays? How does one know if even more assays are needed to fully characterize PM2.5 OP? Furthermore, how would all these various OP measurements, even if available to health researchers, be utilized in a health study, ie how would they be combined to give an overall better indicator of PM2.5 OP? These questions are important and should likely be considered; a discussion beyond the conclusion that all these assays should be utilized, is warranted

The data do support other studies showing variability between various OP measures and PM2.5 mass, suggesting PM2.5 mass is a poor predictor of the ability of particles to cause oxidative stress (assuming these assays are good measures of OP). This is an important finding.
Comparisons between sites using different samplers operating at the same time depends on some level of measurement precision to argue that observed differences (poor correlations) are really due to differences in aerosol particles at the sites. This applies to the gravimetric measurement of PM2.5 mass and the various OP measurements. The authors do discuss variability in the negative and positive controls, but the data shown in Table 1 is only the precision of the analysis and does not consider sampling, filter storage or extraction. Can it be stated that this precession for all the species measured and PM2.5 mass is significantly better (lower variability) than that of the comparisons between sites. It would be especially interesting to know the precession of the methanol extracts, which based on the extraction approach is likely the most imprecise measurement (curiously it also shows the least variability between assay results from various sites). A more comprehensive discussion is warranted that includes specifically addressing if the differences seen are real or just noise.

One conclusion that may be drawn from this work and which is consistent with past studies is that the DTT assay is the most comprehensive measurement of OP (see, for example, discussion in lines 289-407). This may be because DTT includes electron transfer reactions from both organic species and metals, whereas AA, GSH and production of OH in the various assays is likely largely driven by metals. One could actually discuss an interpretation of the data in which the most assay meets the goal of being the most comprehensive. For example, maybe instead of arguing that all assays in their various forms are needed, one could try to assess which is best?

Specific Comments.

Line 20-21, not sure how higher site to site correlations proves methanol extracts includes more insoluble species? The idea that methanol extracts a greater fraction of OP than water is well known.

Lines 142 to 148, Charrier et al (2016) suggest a mass concentration for measurement of OP to limit nonlinear effects of 10ug PM/mL, here the authors use 100 ug/mL, why and what is the effect of doing this, ie does it solve the nonlinear problem?

It would be useful to provide the composition of the simulated lung fluid.

One issue with current measurements of OP by the various methods is that there is a range of approaches used for each of the methods. This makes comparisons between this work and other studies complicated. It would be valuable to know exactly how these various methods compare to what has been utilized in other studies. For example, maybe a table in the supplement could provide more details on the methods used here links to past studies that used the exact same approach.
Line 238-239, this statement should be supported with data.

Line 274, typo, change “into” to “in”?

How do the authors explain the data where OP in water extracts is greater than OP methanol when it is established that methanol extracts water soluble species plus organic species? Seems this result demonstrates the lack of precision of the methanol method. Or are the authors implying that some water soluble species that contribute to OP are not extracted and detected in the methanol method?

What is the difference between methanol soluble OP and methods that attempt to measure all OP, e.g., that associated with surfaces of solid particles?