We would like to thank the referees for all the comments. The authors' responses are listed as follows with corresponding changes made in the manuscript.

Anonymous Referee #1

In this study, the authors assessed the water-soluble oxidative potential of PM2.5 collected in Southeastern U.S. based on both DTT and AA assays, and compared the results of two assays in the view of their association with chemical components, sources and emergency department (ED) visits. In my opinion, this is an important and careful study with large database, providing essential information on the origin and potential health outcome of the water-soluble oxidative potentials of PM2.5. In addition, the result could help future studies to better interpret the data based on those assays. However, there are several issues that should be addressed in the manuscript. A major concern is on the oxidative potential of PM2.5. Actually I'm afraid that oxidative potential from water-soluble components is far from enough for the evaluation of PM2.5's toxicity or health effects, given that a series of studies have suggested that some hydrophobic components (e.g. represented by PAHs) from vehicles are the major toxic components on human health (Delfino et al., Environ Health Perspect 2010, 118: 756-762; etc.). Why not measure the oxidative potential of hydrophobic components if the authors wanted to link the oxidative potential of PM2.5 with some health outcomes?

Authors' response: This paper does focus only on the water-soluble PM components contributing to oxidative potential, and doesn't include the water-insoluble fraction. There is no doubt that water-insoluble components also have potential adverse health impacts. There is a logical reason to look separately at soluble vs insoluble components, however, since they are likely to have differing physiological effects once deposited onto the aqueous environment of the air-lung interface. Although we decided to focus solely on soluble species, we are not implying that insoluble species are un-important, it is just not the focus of this work. Besides, there is no standard protocol so far for measuring the water-insoluble ROS activities.

Changes in manuscript: Added to the Introduction section "Although in-soluble components are important, since there is no current standard protocols for measuring the water-insoluble oxidative potential, we focus solely on the water-soluble AA and DTT activities."

The following are some specific comments:

1. Page 30615, section 2.2.1: Could sonication for half an hour in the water phase generate OH radical, which could result in great oxidative potential?

   Authors' response: There is some concern on the alteration of the sample due to OH formation during sonication [*Miljevic et al.*, 2014]. We have conducted a limited number of experiment on the effect of sonication on the measured oxidative potential by comparing water-soluble DTT activity from the same extracts after sonication vs shaking for 3hrs. Our findings suggest little difference between the two. Therefore, although sonication may generate OH radicals, it doesn't seem to have significant effects on our DTT assay results. A manuscript on the results is in preparation.

Changes in manuscript: page 30617 line 14, added "Although OH may form during sonication (Miljevic et al., 2014), it has little effect on our ROS measurement since we compared the water-soluble ROS activities from the same sample that had been extracted by shaking for 3 hours vs sonication and found no significant differences, average ratio is 1.08 ± 0.20, n = 7."

2. Page 30625, line 15: The spatio-temporal analysis could not draw the conclusion that the oxidative potential is influenced by different components from different sources, because there are no evidences in this part showed the similar trend for chemical components and sources.

Authors' response: We have removed the reference to chemical components.

Changes in manuscript: Page 30625 line 15: changed "These results indicate that water-soluble DTT and AA activities of PM2.5 were influenced by different aerosol components from differing sources that varied with season." to "These results indicate that there are differences in the sources for water-soluble AAv and DTTv, with traffic emissions apparently a more significant source for AAv. Correlation analysis with specific aerosol components provides further insights."

3. Page 30626, line1: r2 or r? In case of r2, it's better to convert it into r, since r is used throughout the manuscript.

Changes in manuscript: page 30626 line 1: changed "r2>0.5" to "r>0.7".

4. Page 30626, line 25: Since Pearson's r is used. Please provide information on the normality of the data and on whether data was log-transformed.

Authors' response: As is found for almost all atmospheric constituents, their frequency distributions are log-normally distributed. However, the correlation analyses were performed using measured concentrations rather than log-transformed concentrations as this is a straightforward interpretation method for these data. Since interpretation of the correlation analysis results is similar whether done on a linear or log basis, we prefer to keep the former here.

Changes in manuscript: No changes.

5. Page 30630, section 3.2.2: It is better to provide detailed data of the risk ratio (together with 95% CI) in the text instead of in Fig. 4 only.

Authors' response: The risk ratios for AA for asthma/wheeze and congestive heart failure are added to the text. The detailed data for other health outcomes are provided as supplemental material.

Changes in manuscript: Page 30630, line 13, changed "For asthma/wheeze and congestive heart failure, the risk ratios for an increase of an interquartile range for DTT estimated from

the model (Eq. 6) were above 1 (Fig. 4), and the 95% confidence intervals exclude 1, indicating a statistically significant positive association between estimated DTT and the ED visits for these health outcomes." to "For asthma/wheeze and congestive heart failure, although the risk ratios for an increase of an interquartile range for AA_v^e were above 1 [1.005 and 1.003 for Asthma/wheeze and CHF, respectively, Fig. 5(a)], the 95 % confidence intervals crossed 1 (0.994-1.015 and 0.986-1.020 for Asthma/wheeze and congestive heart failure, respectively), indicating a non-statistically significant association between AA_v^e and the ED visits for these health outcomes."

6. Page 30631, line 3: The authors claimed that the epidemiology analysis "support aerosol particle oxidative potential as a mechanism contributing to these PM-induced adverse health effects", which I think might not be true. A more robust association doesn't necessarily mean a possible mechanism. For example, sulfate is considered as "benign" (page 30612, line 13), but actually sulfate is strongly associated with adverse health effects in epidemiology studies, probably because sulfate was co-emitted with toxic pollutants (Grahame, EHP, 2012). Personally, I think the association of DTT could also possibly be attributed to co-emission with toxic pollutants, especially given that (1). DTT is sensitive to organic species; (2). There is a lack of toxicology studies showing the health effects of oxidative potential; (3). No significant association was observed for AA. Although the authors attributed it to the different uncertainties (page 30630, line 19), it is not convincing since r values of 0.60 and 0.68 are not that different.

Authors' response:

First, we state that it "supports", not that it proves OP is a mechanism. We would like to address the reviewer's three comments one by one:

1) DTT is sensitive to organic species
   DTT assay is sensitive to organic species doesn't prove that OP is not a plausible mechanistic linkage. We've identified components of the organic species that are DTT active [*Verma et al.*, 2015a; *Verma et al.*, 2015b]. For example, HULIS (humic-like substances) in biomass burning or more-oxidized OA are mostly linked to DTT while biogenic secondary organics basically have no DTT activity.

2) There is a lack of toxicology studies showing the health effects of oxidative potential
   This is not true. There are numbers of studies have associated OP to PM toxicity, as we have cited in the introduction section (see page 30612 line 27).

3) No significant association was observed for AA. Although the authors attributed it to the different uncertainties (page 30630, line 19), it is not convincing since r values of 0.60 and 0.68 are not that different
   We have rearranged the text to clarify the differences in health analysis between DTT and AA are possibly due to the narrower selectivity of AA than DTT assay, rather than the uncertainties in the models, which is fairly similar (ie, r values 0.60 vs 0.68).

In summary, although a positive association does not mean a causation, the association of DTT with health effects is supported by a rational physiological mechanism (the DTT assay measures the OP of particles, which represents the ability of PM components to generate ROS and ROS can induce oxidative stress *in vivo*. It is also supported by our various findings that DTT is sensitive to specific classes of organic species and metals that could lead to health effects.

Changes in manuscript:

Page 30630 line 18: changed "Similar differences were found for estimates based…" to "The same results were found for estimates based on the two other regressions [(see Figure 5(b) and (c)]), suggesting that the null relationship of AA_v^e and positive association of DTT_v^e with these health outcomes are to some extent robust, despite the high uncertainties from the back-cast models. A possible cause for the differences in AA_v^e and DTT_v^e health associations is the more narrow selectivity of the AA assay to specific aerosol components (i.e., mostly sensitive to Cu). The AA assay may not capture the overall oxidative potential of all the various PM components as well as the DTT assay."

Referee # 2 Cort Anastasio

Introduction.

Fang and co-authors describe results from two different measurements of oxidative potential – DTT (dithiothreitol) and AA (ascorbic acid) – for ambient particles collected in the southeast U.S. as part of the SCAPE center. The DTT and AA data sets are analyzed in three different ways: (1) linear regressions between oxidative potential and various chemicals in an attempt to identify the responsible chemical species; (2) positive matrix factorization (PMF) and chemical mass balance (CMB) modeling to identify important sources of ROS-generating PM; and (3) epidemiological modeling on approximately a decadelong time series of estimated DTT and AA results to assess if either measure of oxidative potential is associated with health effects.

Authors' response: We also investigated seasonal trends and spatial distributions.

The AA results are novel and, though not fully explored, form the core of a good manuscript. With some additional pieces, this could be a very nice piece of work. On the other hand, the DTT results have all been presented previously and there is nothing that warrants spending half the manuscript on these past results. There are some new (and better) ways in which the DTT data could be treated; if this is done, it could be an important contribution to our understanding of DTT and would significantly improve the manuscript.

Authors' response: First, we stress that this is a comparison paper. Our aim is to provide a clear contrast between the AA vs DTT assays, rather than simply discuss AA results and leaving it the reader to go back and forth between multiple papers. Admittedly, the AA results do not show strong linkages to health (ie, a null result), but this is an important result. However, the real strength of this paper is to provide a direct side-by-side comparison between AA and DTT, which were both analyzed on the same filters. That is, we go through

a whole series of comparisons: spatial, seasonal, correlations to PM components, including $PM_{2.5}$ mass (important since it has been associated with adverse health effects in many studies and is regulated), source apportionment, and then health effects. We have modified the paper, including changing the title and figures, to emphasize the AA results. However, we maintain the side-by-side comparison as we view this as a highly unique and powerful result. In fact, this contrast largely addresses Major point 2 below, addressing the issue that DTT may be only responding to a few metals and that this reviewer's analytical approach to assess this assertion should be performed on this data set, as part of this paper.

Changes in manuscript:

1. We added the following in the introduction section to better illustrate our purpose, page 30614 line 21, added to the end "Throughout, we compare the AA results to our previously published DTT findings (Bates et al., 2015, Fang et al., 2015b, Verma et al., 2014) to provide a clear contrast between these two commonly utilized assays to assess aerosol oxidative potential and possible associations with health endpoints."
2. Changed to the title to "Oxidative Potential of Ambient Water-Soluble PM2.5 in the Southeastern United States: Contrasts in Sources and Health Associations between Ascorbic Acid (AA) and Dithiothreitol (DTT) Assays".
3. Moved Figure S2 (Protocol schematics for conducting Ascorbic Acid assay) to the main text.
4. Modified Figure 4 by putting AA health results first. To emphasize the robustness of epi results from different models, Figure S7 (health results from the other two models) were moved to the main text and combined with Figure 4.
5. Changes were made throughout the manuscript to increase the discussions of AA results (see the complete manuscript with marked changes).

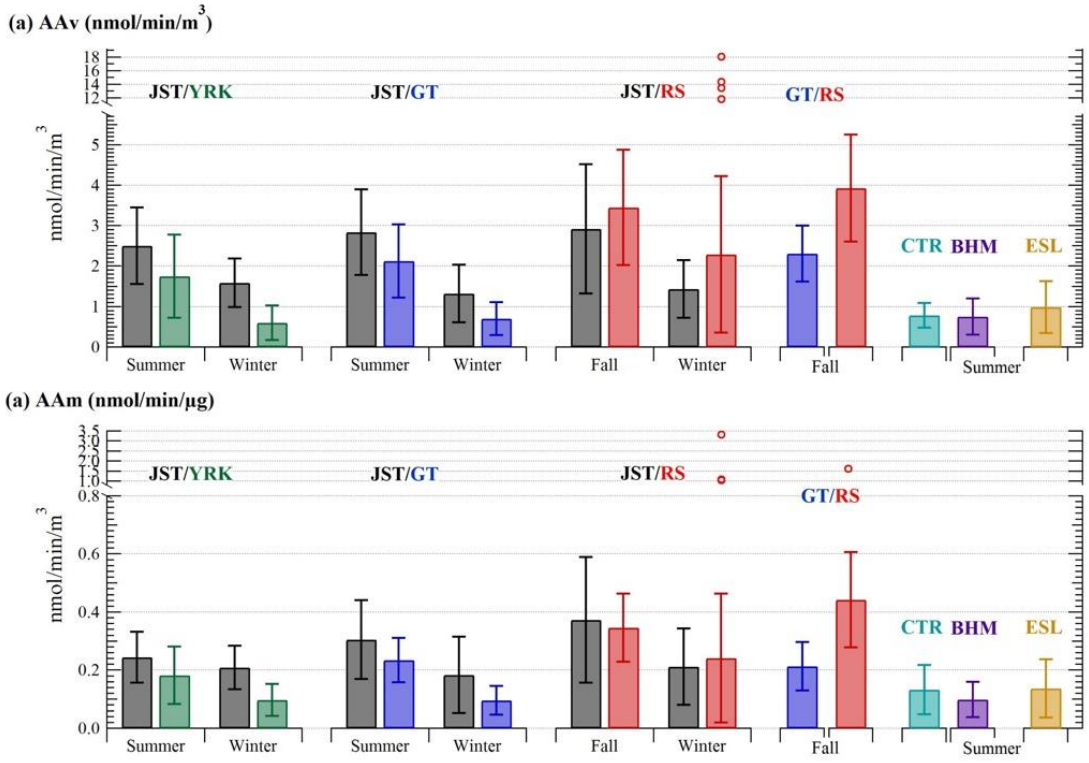
Major Points.

The DTT figures in the manuscript have all been shown (and discussed) previously: Figure 1b, the DTT results in Figure 2, and Figures 3c and c are all from Verma et al. (2014), while the DTT data in Figure 4 is from Bates et al. (2014). This previously published work represents approximately half of the data in the manuscript. I appreciate that the authors want to compare their new AA results with their old DTT results, but giving the two sets of data equal weight in the manuscript takes away from the ascorbate findings. It also makes for a repetitious experience for readers of Verma et al. (2014). My recommendation is to minimize the presentation of the previous DTT figures and the discussion of the DTT results. The comparison of the DTT and AA results is useful but could be done with a brief text discussion after each AA figure. Beyond comparisons with AA, if the authors want to present significant amounts of DTT results in the manuscript, they should be new; see the point #2 below for some suggestions on this.

The authors should more thoroughly present and discuss the AA results, as their treatment in the manuscript is often weak. For example, the authors measured hundreds of samples, but only 17 monthly averages are presented. Is there anything interesting to show from the time series data? Is there anything interesting in the mass-normalized data? What do correlation plots of DTT and

AA rates at the various sites show? What is the average value of the (DTT rate) / (AA rate) for each site/season; does this ratio say anything useful?

Authors' response: We have increased the discussions of AA, focusing the paper more on the AA results. As for the specific suggustions: Time series data essentially shows the same spatial and seasonal results as the average monthly data. We plotted the mass-normalized AA activity (in units of nmol/min/µg), see the graph below; which does not add significant new insights beyond what is already presented. DTT vs AA plots are very scattered and the ratio of DTT/AA is not very useful (this lack of correlation is imporant and discussed in the paper). Therefore, we don't find any of the above suggested analyses worthy of more detailed discussions than what is already presented in this paper.



The authors have used linear regressions to assess the significance of metals and other components in the two assays. For DTT this analysis (e.g., Figure 2) is inappropriate because (1) two of the major contributors (Cu and Mn) have non-linear responses and (2) many of the components are correlated with each other. I ranted to the authors about this in my comments to Verma et al. (2014); see http://www.atmos-chem-phys-discuss.net/14/19625/2014/acpd-14-19625-2014-discussion.html. Since these past comments were largely ignored, allow me rant again, both for the purposes of the current manuscript and more broadly as a statement to DTT/ROS users.

We repectfully disagree with the comments that our analysis is inappropriate and that we did not already address them in a previous review (more on this below).

As an alternative to linear regressions we developed a mechanistic technique to quantify the contributions of chemical species to the measured DTT (or other ROS) rate. This involves measuring (1) concentration-response curves for each species (e.g., the rate of DTT loss as a function of copper concentration), (2) concentrations of Cu and Mn in each sample, and (3) the DTT rate of loss in each sample. We recently compared results from linear regressions and the mechanistic approach for a set of samples from Fresno, California (Charrier et al., 2015). Our mechanistic approach revealed that Cu, Mn, and unknown (likely organic) species account for an average of approximately 50%, 20%, and 30%, respectively, of the measured rates of DTT loss in these samples. These percentages are approximately

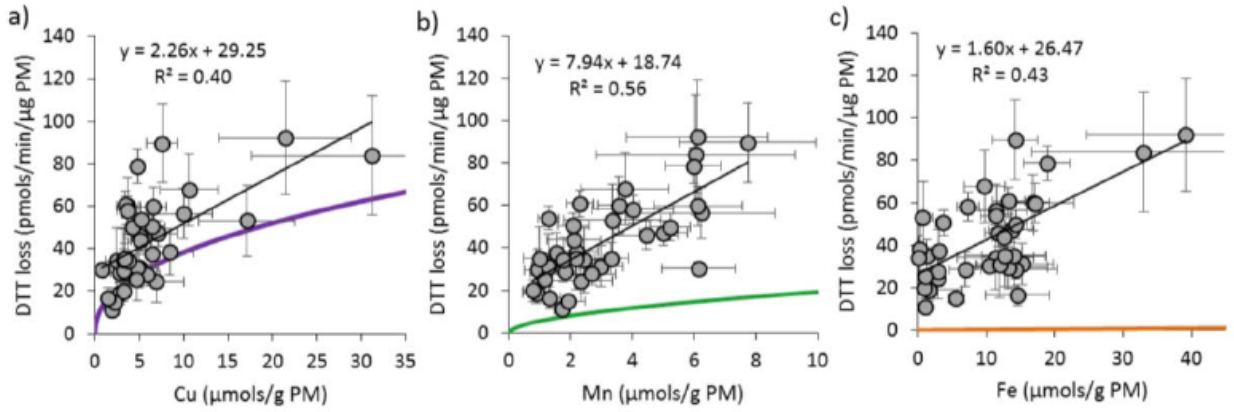


Figure A. Correlations between the measured rate of DTT loss and the concentrations of Cu, Mn, and Fe in samples from Fresno, California. Black lines represent the linear regressions. Each colored line represents the calculated rate of DTT loss from the metal based on our measured concentration-response curves (Charrier and Anastasio, 2012) and the typical PM mass loading of 10 µg mL^{-1} that was used in the assay in this study. Reproduced from Figure S8 of Charrier et al. (2015).

shown by the colored lines in Figure A. In contrast, the corresponding linear regressions for Cu, Mn, an Fe show the weakest correlation for Cu (R2 = 0.40) and the strongest for Mn (R2 = 0.56). The Fe correlation (R2 = 0.43) is as strong as the copper correlation (and has a similar slope), despite the fact that copper accounts for half of DTT loss and Fe accounts for essentially none. Clearly regressions cannot be trusted to identify the species responsible for DTT loss.

Why does this matter for the Fang et al. manuscript? Because the authors have the opportunity to use the mechanistic approach to better assess the contributions of Cu and Mn in their samples. In response to my first review of Verma et al. (2014), the authors calculated the contributions of these metals for their DTT rates with the mechanistic approach. These figures show that Cu generally makes a major contribution to the SCAPE DTT; unfortunately, the figures can only be found in the authors' second response to my comments (http://www.atmos-chem-phys-discuss.net/14/19625/2014/acpd-14-19625-2014-AR2.pdf) as they were not included in the final version of Verma et al. (2014). In a subsequent paper, Verma et al. (2015), they assessed the contributions of transition metals towards DTT but did so using linear regressions; these correlations suggest that Cu and Mn are each important in only 3 of the 7 SCAPE sample sets examined. In contrast, the mechanistic approach results show that Cu and Mn generally dominate the DTT response at every site/season, although there are some problematic samples. As an example, consider the YRK-June DTT data from Verma et al. (2015): linear regressions give R values of 0.64, 0.53, and 0.11 for Mn, Fe, and Cu, respectively. Since its regression fell below the R threshold, Cu was considered insignificant in these samples: the authors concluded that Mn

and organics each accounted for approximately half of the DTT response, while Cu did not contribute. In contrast, the mechanistic approach for YRK-June (Figure B) shows that Cu generally dominates the DTT response, Mn is significant, and unknown components sometimes contribute. The measured and calculated rates for each sample in Figure B would likely agree better if the authors measured concentration-response curves on their automated system rather than used results from the manual runs in Charrier and Anastasio (2012). I encourage the authors to pursue this for the revised manuscript.

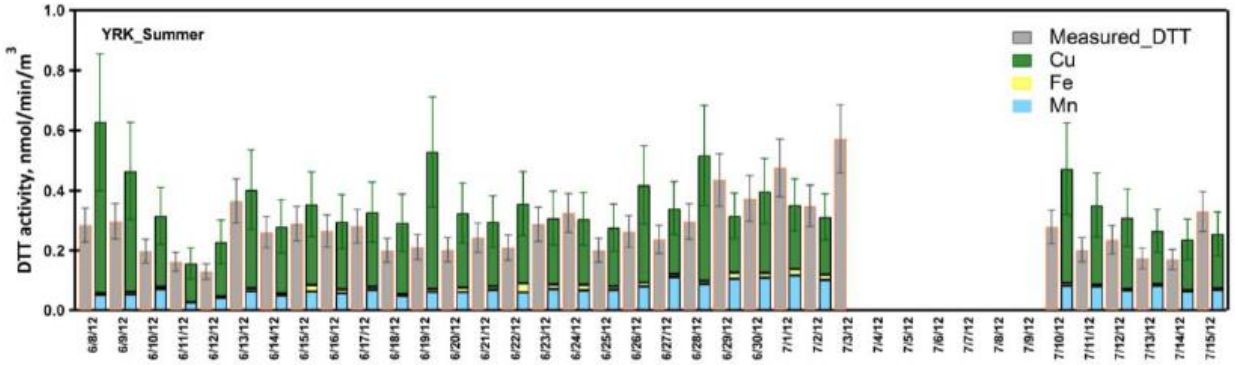
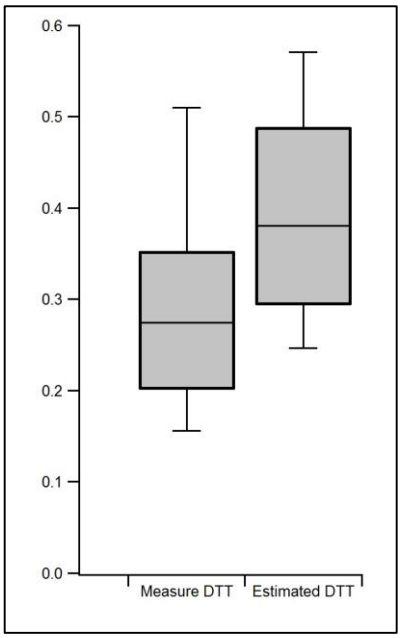


Figure B. Mechanistic assessment of the transition metal contributions to the DTT response in the YRK-June (and July) samples. From the Verma et al. response (29 October 2014) to my comments to Verma et al. (2014); see http://www.atmos-chem-phys-discuss.net/14/19625/2014/acpd-14-19625-2014-AR2.pdf

It is an open question whether Cu and Mn also have non-linear responses in the AA assay. If they do, then the linear regression assessment of which species contribute to the AA result in the manuscript might have problems. Since it would be a simple matter to measure the concentration-response curves for these metals in the authors' automated system, I recommend that they make these measurements. The authors should also use these curves to assess transition metal contributions to AA using a mechanistic approach and compare it to the regression results.

Authors' response: We have not performed the requested analysis because we do not feel it is informative, i.e., it does not definitively show what components are contributing to the observed activity), largely because the uncertainties in the reviewers mechanistic method are very large, such that it is difficult to infer definitive evidence for what species are contribution to the observed activity. The fact that this method indicates that DTT is controlled by Cu and Mn, whereas we have strong evidence that this is not the case tends to support this conclusion. Although we have addressed this before in a previous paper, we will summarize our arguments again since they are stronger, in part due to the findings presented in this paper, that DTT activity is not simply due to Cu and Mn. First, we summarize Figure B above by combining all the data. The result is shown below as a bar/whisker plot. The summation method proposed by this reviewer, based on our measurements of water soluble Mn and Cu, significantly over-predicts DTT relative to the obseved DTT, and possibly more importantly the difference is highly variable. This is also clearly demonstrated by the r2 between measured and estimated DTT of 0.22 (ie, the summation method can only explain 22% of the observed DTT variability). As we have noted before, the results appear inconclusive and the method has little predictive capability. (A similar conclusion is reached if one simply considers the uncertaintly in the method, when all uncertainties are propogated). Instead of basing an analysis on laboratory experiments using pure

compounds in isolation from the chemically complex mixtures found in ambient aerosols, we prefer to be guided by ambient data based on multiple comparisons, not all of which are correlation analyses or limited by possible nonlinear responses, ie, spatial and seasonal distributions, and mechanistic approaches, see below. Overall, our comparisions and analysis provides a consistent conclusion that DTT activity in our study was not driven solely by Mn and Cu. .



It seems, however, the main point of our paper has largely been missed by this reviewer. That is, for all the various comparisons outlined here, there are distinct contrasts between the DTT and AA assays, and the two assays are not themselves well correlated. If DTT was only responsive to Cu and Mn, (like AA), the stark contrasts between DTT and AA in the ambient data outlined here would not be expected. Again, we feel this demonstrates the strength of our papers format of a side-by-side comparison between these assays.

In further support of our approach, next we summarize the evidence that DTT is driven by more than Cu and Mn.

1) The major emission source found to be most influential from our two different source apportionment models (PMF and CMB-E) is biomass burning (Figure 3). It is known that biomass burning is not a large source for Cu and Mn, and indeed our study shows that only 16% of the Mn and 0% of Cu is associated with this source [*Fang et al.*, 2015a]. Biomass burning, however, is a known large source of OA, specifically HULIS, which we (and others) have shown is related to DTT activity and which contains compounds known to be DTT active (eg., quinones, , hydroxyquinones, etc.). Furthermore, in this paper we show AA is mostly associated with these metals and biomass burning is not a source for AA. The seasonal differences in DTT and AA, also reported in this paper, provide further support; where DTT activites are higher in winter at some sites, AA activities were not (AA actually has the opposite trend, being higher in summer), consistent with known trends in biomass burning emissions (ie, higher in winter) and confirmed by our tracers. Likewise, AA source apportionment indicated it was more associated with traffic sources, which is supported by the spatial distributions; AA was observed to be higher at urban vs rural sites, and amongst the urban sites higher at sites closer to traffic emissions. DTT does not show similar spatial trends. Thus, although the source approtionment may be affected by co-variabilties and non-linear responses, the season and spatial trends are not, and the two indepenant analyses are consistent.

2) Another important point discussed in this paper is that DTT was highly correlated with $PM_{2.5}$ mass (r=0.49-0.88 [*Fang et al.*, 2015c]). Mn and Cu do not correlate with $PM_{2.5}$ mass at most of times, (r=-0.18-0.66, two exceptions are at GT and RS winter time, r>0.75 for Mn), which is not surprising as they are small mass fractions of $PM_{2.5}$. If Cu and Mn were the main drivers of DTT as well, one wouldn't expect to see the high correlations between DTT and

PM mass. This is noteworthy since it is consistent with our epi comparisons between DTT and AA, and the many studies showing linkages between $PM_{2.5}$ mass and health. DTT is correlated with $PM_{2.5}$ and DTT shows associations with some health endpoints, AA is not correlated with $PM_{2.5}$ mass and shows less assocation with health endpoints.

3) This reivewer perfers a mechanistic analysis versus correlations. We have also done this and the results are consistent with the correlation analyses. A mechanistic approach that quantified the contributions of OA vs metals to DTT has been presented [*Verma et al.*, 2015a]. In that study we passed the water-soluble extracts, in which both water-soluble DTT activity and metals concentrations had been measured, through a C-18 column to separate hydrophilic and hydrophonic components. Both the DTT activity and metals concentrations were measured in the fraction that passed through the column (hydrophilic fraction). The hydrophobic fraction can be determined by the difference between total and hydrophibic. We found that hydrophobic DTT activity made up a large fraction of total water-soluble DTT activity and Cu and Mn were mostly excluded from the hydropholic fraction. These results indicated that the DTT activity of the water-soluble hydrophobic fraction was not driven by Cu/Mn, but another class of redox-active species, most probably organic compounds associated with HULIS. From this analyses we reported that roughly 60% of the DTT activity in that study was associated with the OA and 40% metals. As OA from biomass burning is a major source for HULIS, this is consistent with the various source apportionment and correlation analysis discussed above. Finally, we also did a mass balance analaysis, much like the reviewer has done, and found that including the OA (ie, hydrophobic DTT activity) provided better closure.

4) As further proof that DTT is not soley driven by metals we have performed a traffic tunnel/chamber study, which has shown high DTT activity for ambient aerosol containing no metals. In yet-to-be-published work we performed experiments on sample air that was drawn from a highly trafficed vehicle tunnel. Tunnel air could be analyzed directly or modified. In one experiment, the sample air was filtered to remove all aerosols (which includes all metals). The sample was then directed to a smog chamber and explosed to lights, resulting in SOA from the VOCs in the tunnel. This SOA was highly DTT active, with levels comparable to the hyrdophobic fraction of the ambient aerosols observed in SCAPE.

In summary, although this reviewer's point is well taken, that correlation is not causation, the totallity of our uniquely large data set strongly point to the role of both OA and metals to the DTT activity of water soluble components. It includes analyses that are not subject to the limitations pointed out (correlation analysis and possible nonlinear responses). The unique strength of our data is that it is large and includes long time series (ie, seasonal data), and data from multiple sites (ie, spatial data). This is in contrast to the reviewers work, from which they conclude that DTT is mainly driven by Cu and Mn. We have also shown above that the analysis proposed, does not provide a definitive indication that DT activity is driven solely by Mn and Cu, for our data set. And finally, based on this paper comparing DTT and AA, the Verma et al., (2015) paper and our tunnel data, we conclude that the approach used by this reviewer does not provide an informative way to assess the contributions of aerosol species to the water-soluble DTT activity of ambient particles measured in our study. Finally, some time back we attempted

the mechanistic approach, developing a calibration for AA response to pure Cu solutions made from copper sulfate. Although only exploratory in nature, we found the calibration curve to be somewhat unstable and the predicted AA, based on the curve and our measured water soluble Cu, greatly under predicted the AA activity relative to observed (10% of observed). For all of these reasons, we chose not to include this reviewer's mechanistic approach in the analyses of our AA data set.

4. Backcast estimates of AA and DTT activities and the epidemiological analyses. A discussion of the uncertainties in the backcast estimates of the AA response is needed. The uncertainties must be very large, as illustrated for DTT in Figure 1 of Bates et al. (2015). The equivalent figure for AA should be shown in the manuscript. As part of the discussion, how can the backcast uncertainties of AA (and DTT) activity be enormous, but the 95% confidence intervals around the RR data points in Figure 4 be quite small. Do the Figure 4 CIs include the full backcast uncertainties? How are these propagated?

Authors' response: The CIs in the figure are for the epi results and do not include uncertainty in the backcast DTT or AA. See comments to minor point 3 below. Our argument is that the AA and DTT models do roughly equally well when compared to the ambient data (similar r). How well they predict actual historical DTT and AA is not known, but we have no reason to believe that AA is predicted much worse than DTT. In fact one might expect the opposite as DTT is driven by more sources than AA. In any case, the point is that DTT shows health associations and the AA does not. This is an important conclusion despite the high uncertainties with the methods, and is consistent with the contrasts described throughout the paper. It again stresses the value of the side-by-side comparison.

Changes in manuscript: No changes.

Since AA is dominated by Cu, and there are enormous uncertainties in the backcast estimates of AA (and DTT) activity, it would be interesting to do the epidemiological modeling using measured particulate Cu rather than the predicted AA response. This might show a significant correlation with the health endpoints. Are there historical data in Atlanta that could be used for this? If this epidemiological analysis could be done relatively easily, I encourage the authors to include it in the current manuscript. If not, I hope to see it in a future manuscript.

Authors' response: We have considered this. Unfortunately, a sufficient water-soluble Cu data does (ie. Sufficient number of data above LOD) not exist for the time period of the epi study. Thus, running the epi with just water-soluble Cu would still require a backcast prediction.

Changes in manuscript: No change was made.

Minor Points.

1. "AA" is used to represent the volume-normalized rate of AA loss, but of course it's also the name of AA itself, which is confusing. Better to use something like "AAv" for the rate,

analogous to the DTT nomenclature of Verma et al. (2015). There is the same issue for DTT in the manuscript.

Authors' response: Replaced DTT and AA activity with DTTv and AAv throughout the paper. Also used DTT_v^e and AA_v^e for backcast estimates of DTTv and AAv.

2. p.30630. The authors should include the results for the other health outcomes (COPD, pneumonia, IHD) in Table S4, as the comparison with asthma and CHF would be interesting.

Authors' response: We agree and added the health outcomes of COPD, pneumonia, and IHD to Table S4.

Changes in manuscript: added the data of health risk ratios of COPD, IHD, and pneumonia to Table S4. In the main text, page 30630 line 12: changed "(results not presented)" to "(results given in Table S4)".

3. p. 30631. The authors conclude that "For the region investigated in this study, the DTT assay was a more comprehensive multi-pollutant ROS (or oxidative potential) indicator than the AA assay making DTT a potentially valuable parameter to include in future PM health-related studies." Given the very small differences in the RRs for DTT and AA, and the very large uncertainties in their backcast estimates, this conclusion is far too strong.

Authors' response: We disagree when one considers all aspects of the comparisons laid out in this paper. Admittedly, the back-cast prediction of DTT and AA activities has large uncertainties, but again it is the comparisons that are insightful; DTT shows statistically significant associations (at 95% confidence level), AA does not. And this is true for all regression models we used to backcast DTT and AA. We have now added these results to Figure 4 to emphasis the point.

Additionally, previous studies has shown good correlation between DTT and biomarkers such as hemeoxygenase expression in cells and exhaled nitric oxide fraction [*Delfino et al.*, 2013; *Li et al.*, 2003]. We believe that DTT assay is a valid accellular assay for studying the ROS-generation *in vivo* and would be valuable to include in future PM health-related studies.

Changes in manuscript: Page 30630 line 18, edited as "The same results were found for estimates based on the two other regressions [(see Figure 5(b) and (c)]), suggesting that the null relationship of AA_v^e and positive association of DTT_v^e with these health outcomes are to some extent robust, despite the high uncertainties from the back-cast models."

4. Finally, I want to apologize to the authors for taking so long to complete this review.

Anonymous Referee #3

Summary

This manuscript uses a number of important methods to bridge the continuing gap between detailed composition information and the association between PM and negative health effects. The DTT and AA assays were run on PM extracted from a number of locations across two year in the southeast United States. Correlations were made between the DTT and AA assays and PMF and CMB modeling of sources and with ED admissions from long term epidemiological data. On the whole I feel the paper has a lot of strong characteristics, particularly the new assay methodologies and their connection with source modeling. I think the paper overall is fairly strong and should be accepted with minor revisions.

Major Comments

One concern is that there is a lot of detail and information in the methods section, but that the results and discussion section feels light on new results. As noted by another reviewer, a large amount of the DTT data has been shown before in prior publications. Though comparisons with prior data are good, it feels a little lik a second bit at the apple for the DTT data. A revised manuscript that shows more data from the AA results would be compelling, though perhaps due to the strong Cu response there was concern that the AA results were not strong enough on their own? I rarely suggest making a paper longer, but at 18 pages with only 4 of results and discussion that could really be expanded and strengthened.

Authors' response: Since this work has combined field campaign, air quality modeling and epidemiologic analysis, it involves extensive methodological descriptions. We have eliminated the replicate method descriptions, now providing just references, and have moved them to the supplemental materials. These include the method description of DTT assay from Fang et al. (2015b) and water-soluble elements from Fang et al. (2015a).

We have also revised the paper, putting more emphasis on the AA results. Also see our 2nd response to the 2nd reviewer's comments.

Changes in manuscript: Moved the method description of DTT assay (page 30615 section 2.2.1) and a large fraction of "water-soluble elements" (page 30620 section 2.3.2) to the supplement and added text emphasizing the AA results.

The extensive literature citations are commendable and the introduction very nicely sets up the need for this work. Along the lines of the previous comment, at times I felt like the new findings from this study got lost in the sea of references to this group of authors previous papers, particularly those from the last 2 years and Balachandran et al. 2012. A revised manuscript that more clearly delineates the findings of this study from prior work will help readers understand the findings of this work. Though the data is admittedly abundant and at times dense, the results section feels like a list of numbers in a few paragraphs. Rewording to bring the science out from behind these numbers would be helpful.

Authors' response: We have eliminated repetitive references to previous publications and have reworded most sections to emphasize the results from AA.

In Figure 3 the amount of AA response from brake/tire wear being 44% is notable. I worry that since as the authors note that these are mechanically generated particles that they are likely solid and smaller than the 0.45 micron filter used in the methods. The still solid metals might only be dissolved when nitric acid is introduced prior to the XRF analysis. If so this would seem to suggest that this is not truly a water soluble, but rather an insoluble source that becomes soluble during analysis. Overall the possibility of solid particles less than the 0.45 um filter should be addressed in more detail. If this concern is plausible then some of the phrasing in the manuscript (and potentially even the title) should be qualified since much of the AA response might be from insoluble material.

Authors' response: If the particles are formed mechanically, they are not expected to be smaller than 0.45 μm in significant quantities. It is established that mechanically generated aerosol are mainly found in the coarse mode, but which can extend down to ~1um (eg, see Seinfeld and Pandis). Thus, some fraction of these coarse particles will be included in PM2.5, but very few below 0.45 um. (Note, we have measured the size distribution of the AA activity and there is practically no contribution to the AA activity from particles below 0.45 um). Thus, we expect few solid particle to interfere with our measurement. Regarding the nitric acid: Nitric acid is added to the extract after the extract is filtered, so it will not add to the water soluble metals signal if there are insoluble particle in the extract (ie, particles below 0.45 um).

Changes in manuscript: No changes.

AA is used as a surrogate for a more complex lung fluid and found to only really correlate with Cu. Are the authors concerned that perhaps this more simplistic model fluid is not capturing sensitivity to other species that perhaps a more complex simulated fluid might? A little bit of further discussion on whether AA is really a great choice would be helpful, a otherwise it might not be sensitive enough to mimic the responses in the lung that are observed from other types of studies.

Authors' response: Yes, the AA assay used in our work (as stated in page 30617 line 14) is a more simplistic model to a synthetic respiratory tract lining fluid model (RTLF), which contains, besides AA, urate and reduced glutathione (GSH). This ascorbate-only method has been used by other studies [*Ayres et al.*, 2008; *Mudway et al.*, 2005] to provide an alternative high throughput method to synthetic RTLF. In their work, AA-only method was performed to characterize the metal dependence of oxidative reactions, mimicking the nature of particles-antioxidant (i.e. AA) interactions at the air lung interface.

Changes in manuscript: page 30617 line 14, changed "The method in this study was based on an ascorbate-only model (Mudway et al., 2005; Ayres et al., 2008) that is a simplified approach to a synthetic respiratory ..." To "The method in this study was based on an ascorbate-only model (Mudway et al., 2005; Ayres et al., 2008) that is a simplified and alternative high throughput approach to a synthetic respiratory..."

Minor Comments/Concerns

LWCC is defined on both 30617 and 30619

Authors' response: deleted the 2nd expansion of LWCC.

Changes in manuscript: Page 30619 line 19: changed "Liquid Wave-guide Capillary Cell" to "LWCC".

700 nm was chosen as the background, but some things in the atmosphere can absorb at that wavelength. Washenfelder et al. 2015 GRL showed episodic variation of absorption for the 4 wavelength of aerosol absorption they monitored (highest was 530 nm). Have the authors done checks with full absorption spectra to determine if 700 nm is truly a safe choice for background? This would be particularly important when biomass is playing such an important role in both absorption and DTT assay activity.

Authors' response: We developed the method used to measure Brown Carbon [*Hecobian et al.*, 2010] and made the measurements (ie, it is our data) and can say definitively that there is no absorption at 700 nm from other aerosol components.

Changes in manuscript: No changes.

As a curiosity is a there a reason that XRF was used and not ICP-MS?

Authors' response: It was the method available and also easily adaptable to handle the large number of filters we had to analyze. To send out the hundreds of samples for analysis by ICP-MS would have lead to exorbitant costs (ie, over $100/sample)

Changes in manuscript: No changes.

This might just be a lack of knowledge, but ammonium bisulfate is a class listed for one of the models. How is that determined and separated from ammonium sulfate? Is that from ISOROPPIA or some other thermodynamic model?

Authors' response: The separation of ammonium sulfate and ammonium bisulfate is based on the differences in their source profiles, which were determined from an ensemble average of different CMB profiles and CMAQ results. The main distinguishing drivers in the source profiles are the average mass ratio of ammonium/sulfate. A thermodynamic model (ie, ISORROPIA) is not used.

Changes in manuscript: No changes.

Section 2.3.3 Though referencing the other work is important, I am a bit suspicious of using PM mass concentrations estimated from the sum of chemical component from Hi-Vol samples. More detail on how this has been justified (even one sentence), beyond just referencing citations would appease that concern.

Authors' response: Since total $PM_{2.5}$ measurement was not available for the RS and GT site, we had to use the summation method. Added the statement below to clarify.

Changes in manuscript: Page 30620 line 20: changed "For the RS and GT sites, the PM mass concentrations were estimated from the sum of chemical components analyzed on the same Hi-Vol" to "For the RS and GT sites, since PM2.5 mass were not available, the PM mass concentrations were estimated from the sum of chemical components analyzed on the same Hi-Vol filters (Verma et al., 2014) (Details in the Supplement)."

Are the %'s (filter sampling, etc.) chosen for propagating uncertainty on page 30632 from a reference or chosen arbitrarily? I was not clear from the text where they came from.

Authors' response: Our previous work [*Fang et al.*, 2015b] has discussed in details the uncertainties from the method, including analytical uncertainty (2-15%), calibration uncertainty (1-6%) , and blank variability. We have found that the overall precisions from co-located comparisons (1σ/slope) are 6-20%. From this, we estimated the specific uncertainties for filter sampling and extraction to be 5%.

Changes in manuscript: Page 30621 line 16, changed "which were obtained by propagating the uncertainties from filter sampling (5%), extraction (5%)" to "which were obtained by propagating the uncertainties from filter sampling (assumed to be 5%), extraction (assumed to be 5%)".

Page 30630: "the model did not observed" should be "the model did not observe"

Authors' response: edited as suggested.

Changes in manuscript: page 30630 line 31: changed "the model did not observed" to "the model did not observe".

Ayres, J. G., P. Borm, F. R. Cassee, V. Castranova, K. Donaldson, A. Ghio, R. M. Harrison, R. Hider, F. Kelly, and I. M. Kooter (2008), Evaluating the toxicity of airborne particulate matter and nanoparticles by measuring oxidative stress potential-a workshop report and consensus statement, *Inhalation toxicology*, *20*(1), 75-99.

Delfino, R. J., N. Staimer, T. Tjoa, D. L. Gillen, J. J. Schauer, and M. M. Shafer (2013), Airway inflammation and oxidative potential of air pollutant particles in a pediatric asthma panel, *J Expos Sci Environ Epidemiol*, *23*, 466-473.

Fang, T., H. Guo, V. Verma, R. E. Peltier, and R. J. Weber (2015a), PM2.5 water-soluble elements in the southeastern United States: automated analytical method development, spatiotemporal distributions, source apportionment, and implications for heath studies, *Atmos. Chem. Phys. Discuss.*, *15*(12), 17189-17227.

Fang, T., H. Guo, V. Verma, R. E. Peltier, and R. J. Weber (2015b), PM2.5 water-soluble elements in the southeastern United States: automated analytical method development, spatiotemporal distributions, source apportionment, and implications for heath studies, *Atmos. Chem. Phys.*, *15*(20), 11667-11682.

Fang, T., V. Verma, H. Guo, L. E. King, E. S. Edgerton, and R. J. Weber (2015c), A semi-automated system for quantifying the oxidative potential of ambient particles in aqueous extracts using the dithiothreitol

(DTT) assay: results from the Southeastern Center for Air Pollution and Epidemiology (SCAPE), *Atmos. Meas. Tech.*, *8*(1), 471-482.
Hecobian, A., X. Zhang, M. Zheng, N. Frank, E. S. Edgerton, and R. J. Weber (2010), Water-Soluble Organic Aerosol material and the light-absorption characteristics of aqueous extracts measured over the Southeastern United States, *Atmos. Chem. Phys.*, *10*(13), 5965-5977.
Li, N., C. Sioutas, A. K. Cho, D. Schmitz, C. Misra, J. Sempf, M. Wang, T. Oberley, J. Froines, and A. Nel (2003), Ultrafine particulate pollutants induce oxidative stress and mitochondrial damage, *Environmental Health Perspectives*, *111*(4), 455-460.
Miljevic, B., F. Hedayat, S. Stevanovic, K. E. Fairfull-Smith, S. E. Bottle, and Z. D. Ristovski (2014), To Sonicate or Not to Sonicate PM Filters: Reactive Oxygen Species Generation Upon Ultrasonic Irradiation, *Aerosol Science and Technology*, *48*(12), 1276-1284.
Mudway, I., S. Duggan, C. Venkataraman, G. Habib, F. Kelly, and J. Grigg (2005), Combustion of dried animal dung as biofuel results in the generation of highly redox active fine particulates, *Particle and Fibre Toxicology*, *2*(1), 6.
Verma, V., T. Fang, L. Xu, R. E. Peltier, A. G. Russell, N. L. Ng, and R. J. Weber (2015a), Organic aerosols associated with the generation of Reactive Oxygen Species (ROS) by water-soluble PM2.5, *Environmental Science & Technology*, *49*(7), 4646-4656.
Verma, V., Y. Wang, R. El-Afifi, T. Fang, J. Rowland, A. G. Russell, and R. J. Weber (2015b), Fractionating ambient humic-like substances (HULIS) for their reactive oxygen species activity – Assessing the importance of quinones and atmospheric aging, *Atmospheric Environment*, *120*, 351-359.