Responses to Referee 1 (R1) Comments on acp-2018-1296 “Evaluating wildfire emissions projection methods in comparisons of simulated and observed air quality”

R1 Interactive Comment: As mentioned in my initial short review of the manuscript, both the topic and the proposed methods are interesting and suitable for the ACP journal. The scope of the article is in line with other methods used to predict atmospheric emissions under foreseen climate change scenarios, taking into account population dynamics.

My main concern is with the final results presented in the article; the methods aim at providing acceptable performance as compared to NEI, but also observational data on emissions, as that should be the ultimate goal of all models. MFBs are 25% and 51% against observations. I would therefore recommend that the initial approaches are revised so that the results they yield approximate better the known emissions. There are some data related to estimates of burnt areas that clearly not suitable for that, such as the count of MODIS active fires. There are much better datasets in the USA, and also globally, for these assessments, which ultimately lead to the estimation of wildfire emissions. This could be one of the causes of the large differences between model outputs and observations.
Response: We thank the reviewer for the positive comment on the overall scope of the paper. However, the quantities predicted in this work are ambient concentrations of atmospheric pollutants, and not their emissions. Wildfire emissions were estimated in previous work cited in this paper, and are inputs to the model simulations described here. Likewise, the observational data are not those of pollutant emissions, but of their atmospheric concentrations. There are numerous factors besides the wildfire emissions that can lead to the large MFB in the predicted concentrations relative to observations. Our analyses address the uncertainties attributable to the input emissions from wildfires as well as other sources, and the atmospheric chemistry simulated by the air quality model.

We recognize that the model performance is variable for different species and time periods, and particularly poor for ozone. We analyze possible sources of these uncertainties extensively in the Results and Discussion sections. It is neither possible within the scope of this paper, nor is it our aim, to improve the wildfire emissions estimates in the National Emission Inventory, but rather to use them as a benchmark inventory for comparison with our emissions estimation methods in driving air quality simulations for 2010 over the U.S.

R1 (in reference to p. 4, line 12): Although this year may be ideal as placed latest in relation to future projections, it would be desirable that the assessment on the performance of the modeling is extended to other years.
Response: Although it is ideal to have a few years of data available for evaluating models in retrospective periods, resources available to this project limited the evaluation to one retrospective year (2010), and four future years so that the trends of fire emissions and their impacts on air quality in the future decades could also be examined in addition to assessing their adequacy for contemporary air quality. Multiple retrospective evaluations would entail processing of emissions not only for wildfire, but also all the other sectors for the three methods compared in this study, along with the meteorological and air quality model simulations for each retrospective year. This would have added considerable time and resources to the overall project, and was beyond the scope of the study.

R1 (in reference to p. 4, lines 23-28): Please, provide references for published criteria of acceptance of QA model performance.
Response: We have inserted the text citing the relevant references to the published criteria of Boylan and Russell (2006) and Emery et al. (2017) in the second sentence of this text.
Revised Text, p. 4, lines 23-24: "...we hypothesize that they will yield results within published criteria (Boylan and Russell, 2006; Emery et al, 2017) for acceptable AQ model performance with respect to observations,...."

R1 (in reference to p. 5, line 3): Please, clarify the process below, including the models used in each case, inputs and outputs. From the text below it is difficult to understand if the AAB come from Prestemon et al. from Shankar et al. or if they are estimated in the current work.

A flowchart would definitely help here!

Response: The opening sentence of Section 2 (Methods) states that the air quality (AQ) simulations used emissions estimates for wildfires from the previous work of Shankar et al. (2018). No emissions are being estimated in the current work, which only involves an evaluation of air quality model simulations. The beginning of subsection 2.1 states that the projected wildfire emission inventories from Shankar et al. (2018) were "developed using the AAB estimated by the statistical models of Prestemon et al. (2016)" with meteorological inputs from two different climate downscaling methods, labeled "statistical d-s" and "dynamical d-s". We have further clarified this in the revised text below by inserting the word "AAB" or "inventory" after those labels as appropriate, because each projected wildfire emission inventory inherits its label from the respective AAB estimation method. These and other details of the inventory development, as well as flowcharts and tables of models used in the projected inventories are already included in the primary reference, Shankar et al. (2018). To repeat those published details here would add considerably to the length of this paper, and thus we have not provided them in this section.
Revised Text, p. 5, line 10: "The statistical d-s AAB were based…".
p. 5, line 18: "Meteorological inputs for the dynamical d-s AAB estimates were provided by the Weather Research…".

R1 (in reference to p. 5, line 29): Not clear from where the AAB come from. From above, line 5, it seems that AAB used are those of Prestemon et al. 2016. Why are AAB estimated again, or are these the ones of Prestemon et al.?
Response: Yes, the only AAB estimates used in this work are by Prestemon et al. (2016); no new ones have been estimated in this paper. This section on the emissions inventories briefly describes inputs being used in the air quality modeling. We have clarified that these are not new estimates in the revised lines below.
Revised Text, p. 5, line 3: "We used two projected wildfire emission inventories…hereafter "NEI benchmark"; we highlight their main features here".
p. 5, line 27: "Each set of AAB estimated as described in Prestemon et al. (2016) and Shankar et al. (2018) was used to calculate daily wildfire emissions with …".

R1 (in reference to P.6, line 8): Please, elaborate on how burnt areas are estimated in SMARTFIRE, since this step is critical for the estimation of forest fire emissions. Detail the

"references therein" to clarify how AB are estimated. Errors in this process are bound to be propagated in the next steps.
Response: The purpose of this paper is not to evaluate the NEI wildfire inventory method. It is, rather, to evaluate the inventory projection methods we have developed by comparing simulations using those estimates against historical observations. We also compare these simulation results against those using the EPA-developed NEI as the benchmark, because it is considered the standard inventory for AQ modeling. Thus we do not elaborate on the details of how the NEI wildfire inventory was developed. However, per the review comment, we have changed the text slightly to cite the relevant references (Raffuse et al., 2009; Pouliot et al., 2012). In the later sections under Results and Discussion, we explain differences seen among these methods in species and time periods that are significant for wildfire emissions, and discuss the relevant aspects of the NEI, from both fire and non-fire sectors, that could contribute to those differences.
Revised Text, p. 6, line 4: "… SMARTFIRE -- Raffuse et al., 2009; Pouliot et al., 2012)".

R1 (in reference to p. 8, lines 1-4): This seems to indicate that the models do not perform properly. If that is not the conclusion, please, explain!
Response: It is true that ozone model performance is not in the good-to-acceptable range for any of the simulations for August - November. However, differences in AQ simulation results for ozone among the three inventories used are virtually negligible. This indicates that the ozone performance issues are most likely not a result of wildfire emissions, which are the only source of difference among these inventories, i.e., all three simulations use the same emissions for all other sectors throughout the year. PM model performance is more in the acceptable range of the performance metrics as a result of compensating errors in the component species. The severe overpredictions in nitrate and unspeciated dust, and sulfate to some extent, are offset by underpredictions in organic carbon (OC). Some of these biases are related to the model chemistry and persist to varying degrees beyond the fire season, while some others are related to missing non-wildfire sources (e.g., residential wood combustion in the case of OC). On the other hand, ammonium from ammonia, and elemental carbon, both of which are primarily emitted in wildfires, show good-to-acceptable model performance. These points are discussed in detail in the Results and Discussion sections.

R1 (in reference to p. 12, lines 14-16): MODIS fire counts cannot be used to estimate AB; any assumption of this in relation to emissions is thus wrong! The comparison of emissions based on these estimates is thus meaningless. There are many other sources that can be used to estimate burnt areas as well as direct MODIS products that are freely available in FIRMS and LANCE NASA sites.
Response: The NEI wildfire emissions estimates for 2010 represent the EPA's first use of satellite data for this purpose. It is true that the EPA spent considerable effort to improve its satellite-based estimation methods in later inventories of wildfires, but we were obliged to use the appropriate NEI inventory year for the simulations that we conducted. The burned-area estimate in the 2010 NEI wildfire inventory is based on MODIS fire counts available from the NOAA Hazard Mapping System, and assumes that the area burned at the location of each fire detect is the same as the pixel size, 1 km^2. This can lead to over- or underestimates as discussed in Soja et al. (2009), which provides details of how the method underestimates burned areas for small fires. We have slightly reworded the text to make this point. While there are more

advanced data products for estimating area burned, they were not deployed by the EPA in compiling its 2010 wildfire inventory. Our stated aim is the comparative evaluation of our methods, and not the improvement of the EPA inventory, but we use it here as a benchmark towards identifying potential areas for improvement in all three inventories.
Revised text, p. 12, line 20: "…due to the difficulty of under-canopy detection of small fires by the MODIS instrument..."

Literature added to the References:
Raffuse, S. M., Pryden, D. A., Sullivan, D. C., Larkin, N. K., Strand, T., and Solomon, R.: SMARTFIRE Algorithm Description. Paper prepared for the U.S. Environmental Protection Agency, Research Triangle Park, NC, by Sonoma Technology, Inc., Petaluma, CA, and the U.S. Forest Service, AirFire Team, Pacific Northwest Research Laboratory, Seattle, WA STI-905517e3719, 2009.