Reply to reviewer 2

We thank the reviewer for the careful reading and sensible suggestions to improve the manuscript. Follows are point-by-point reply to the comments. Replies are in italic font.

Versions and references for the modeling systems should be identified in Table 1.

References have been added

Maps of the European, North American, and combined domains should be added.

Map added as new Figure 1 (all other Figure numbers incremented by 1).

Please explain what is meant in lines 194-196 on page 7 regarding "past use in policy-relevant emissions scenario simulation, with changes in emissions policies that may affect the deposition".

The has now been changed into : "The NA years were selected due their policy-relevance; the years 2010 and 2016 have featured in policy-relevant emissions scenario simulations by governments in the continent." Hopefully this clarifies the intent of the sentence.

It would be helpful to know which emissions models were used, not only the source of the data, i.e., 2011v6.3 and 2016 beta, in the unified approach for forest fire emissions in North America and Canada.

We've revised the section 2.2.1 slightly to mention the sources of information in more detail, with references for the biomass burning emissions data. The overarching 2010 and 2016 emission modeling platform URLs shown in section 2.2.1 (https://www.epa.gov/air-emissions-modeling/2011-version-63-platform, https://www.epa.gov/air-emissions-modeling/2016v72-beta-and-regional-haze-platform, and http://views.cira.colostate.edu/wiki/wiki/10197) provide more detailed information on the emissions inventory and processing approaches taken to generate these platforms. This includes information on the generation and processing of wildfire emissions with the exception of Canada 2010. Specifically, SMARTFIRE2 was used for U.S. wildfire emissions in both 2010 (https://www.epa.gov/sites/default/files/2016-09/documents/2011v6_3_2017_emismod_tsd_aug2016_final.pdf) and 2016 (http://views.cira.colostate.edu/wiki/Attachments/Inventory%20Collaborative/Documentation/2016beta_0311/National-Emissions-Collaborative_2016beta_point-fire_11Mar2019.pdf), while a combination of FireWork and FINN was used for Canada in 2016 (http://views.cira.colostate.edu/wiki/Attachments/Inventory%20Collaborative/Documentation/2016beta_0311/National-Emissions-Collaborative_2016beta_canada-mexico-ptfire_07Mar2019.pdf). The generation of the 2010 wildfire emissions for Canada is described in Chen et al. (2013) and this reference has been added to the revised manuscript.

In lines 338-346 on page 13, consider summarizing and briefly explaining some of the key motivating factors that have led to the development of different resistance frameworks. For example, is it the evolution of measurement systems, availability of observational data, inclusion of missing deposition pathways, etc.

This portion of text has been added:

"Several motivating factors likely led to the development of a diversity of resistance frameworks. In the intervening years subsequent to Wesely's introduction of the resistance framework concept, new measurement capabilities (for higher time resolution information, for greater chemical speciation, higher precision measurements) allowed the original algorithms to be tested and modified. Developments in plant physiology understanding have also resulted in improved stomatal resistance parameterizations. Examples include the observation-based introduction of bidirectional fluxes for ammonia gas, and improved understanding of the role of CO2 fluxes in the deposition of other gases. Also, some divergence in approaches is likely due to algorithm developments having been made in the context of specific regional models – each of which encompasses a diverse range of process representation algorithms, vertical resolutions, horizontal resolutions, etc.. An algorithm which provided good performance relative to surface concentration observations within the context of one regional model thus may not have resulted in as good performance in another model, further spurring model-specific development. These factors have resulted in the variety of approaches for gas-phase deposition in current regional models, and provide the part of the motivation for this first attempt at cross-comparing the results of the models' deposition algorithms in detail – to show and explain the causes for these differences."

Please describe the characteristics that are used to guide the mapping to each generic land use/land cover category. For example, what are the definitions of mixed forest or herbaceous cover?

We thank the reviewer for the request for clarification - there is a whole discipline behind this issue that has its pinnacle research in remote sensing using satellite data. In order to carry out a meaningful comparison on a land-use category basis, common land-use categories had to be devised. However, the assignment of land use categories from the native model categories to the AQMEII4 set is of necessity a source of uncertainty. The need for cross-comparison of the land use category assignments is one of the reasons why both the assigned AQMEII-4 land use categories and the original model land use categories are included in the reported information from the participants. The additional text is as follows:

"We also note that the mapping of LULC types from the individual model land use classifications to the AQMEII-4 land use classifications is an unavoidable source of uncertainty in the land-use specific diagnostics. The 15 AQMEII-4 land use types themselves were based on a survey of land-use classifications used in 17 regional models. For example, while "Herbaceous" is available as an AQMEII-4 land use category, its intent is for use for moors and heathlands, while AQMEII-4 land use category "Wetlands" encompasses wetlands which are diversely described in individual model land use categories as herbaceous, wooded, and permanent wetlands, as well as swamps, and peatbogs. However, some categories were held in common by most models (e.g. Evergreen needleleaf forest, Deciduous broadleaf forest, snow and ice, mixed forest (usually taken as a combination of needleleaf and deciduous forests)), while others could easily be classified according to the broader landscape type of which they were a member (e.g. different types of Tundra were recommended to be classified as the AQMEII-4 Tundra

classification). Both the AQMEII-4 and "native model" land use types were reported by participants – with the aim using both sets of information to determine the extent to which land use database variation may be a factor in estimating deposition velocities, and to provide information on specific land use types used by specific models when these differences appear to be large."

Check the references as well as the definitions of variables in Appendix B (which should be at the first instance) for completeness.

Done. All Tables and Figures in the Appendix were reviewed for consistency. A number of corrections were made to the formulae in the Tables for the CMAQ-STAGE and LOTOS-EUROS model, and the description of Zhang's model has also been corrected. The CHIMERE model has been removed from the list of models (original submission Figure B7 and Table B7) since that participant stated that they will be unable to submit model results to AQMEII-4 in the intervening time since the Technical Note was submitted. Our other reviewer requested that we use fractions rather than (xxxx)-1 notation to make the formulae more clear and we've followed through with this. Another round of updates to these and the references may also result from ACP's copyediting and typesetting processes, to improve clarity.