Responses to the comments of the reviewer

(1) The manuscript has been improved, and the authors have addressed many of the issues identified.

We appreciate the positive assessment.

(2) The evaluation is still mischaracterizing how they classify performance. The Ramboll work should not be used to say that the annual performance is excellent/good/average (They state: "Based on these criteria, the ability of the model to reproduce the annual average concentrations of the sites is excellent for OA, good to excellent for PM2.5, EC, and ammonium, good for sulfate, and average for nitrate." The Emery et al. metrics were developed purely for assessing a model's ability to simulate daily quantitates. As the authors clearly show, one might expect much better performance using annual values, but that also can cover up major problems (e.g., large, but off-setting biases by season). This paragraph must be changed. They can present the specific values, but should not provide an evaluation of how good they are unless they have a scientific basis for such, e.g., a peer-reviewed paper or agency report that has laid out an evaluation scale for annual performance.

We have followed the suggestion of the reviewer and removed the characterizations (excellent, good, average, problematic) of the annual performance of the model both from the text and Table 1. We have also removed the corresponding column from Table S2 in the Supplementary Information of the paper. These characterizations are only used now for the daily performance. The performance metrics are presented for both the annual and daily averages.