Response to referee #1

We would like to thank the reviewer for providing constructive comments that have helped us improve the manuscript. The reviewer comments are shown in regular font, our responses are shown in bold fond, and a summary of the changes made to the manuscript in response to each comment is shown in italics.

Major comments

1 The content in results section is organized with model pairs, but without clearly starting at the beginning of each subsection with what we can learn from the two different setups, which makes the following discussion confusing.

Response: We agree with the reviewer that adding a statement about what can be learned from a given model pair at the beginning at each subsection would make the paper easier to follow.

Changes in manuscript: Such statements have been added at the beginning of the comparison between different model pairs. This study aims at understanding the model variability originating from the physical processes in CTMs and its impact on the inert tracers of lateral boundary ozone reaching the surface. The first model pair (WRF/CMAQ versus WRF/CAMx) used the same meteorological inputs for CTMs (and therefore the same representation of advection) but different representations for the other important physical processes that the inert tracers undergo (e.g. vertical mixing and dry deposition). The situation is the opposite for the second model pair (WRF/CMAQ versus COSMO-CLM/CMAQ). Therefore, the results from the two model pairs serve as a good example for demonstrating the relative importance of 3D advection and the other physical processes in CTMs on the model variability in inert tracers at surface.

2 The discussions in results section are not focused on the major conclusions (and maybe too lengthy for subsections). Many descriptions and discussions of the minor temporal and spatial features. In my opinion, should be removed or reorganized, especially those cannot be readily explained by current model pairs.

Response: We agree that many descriptions and discussions were lengthy and not sufficiently focused on the key messages. Moreover, due to lack of extra sensitivity simulations, some discussions were not supported firmly by the simulations presented in the original manuscript. Therefore, three sensitivity simulations have been conducted for WRF/CMAQ to investigate the impact of different physical processes on the inert tracers at the surface. These physical processes are important processes that the inert tracers undergo in CTMs, including wet and dry deposition and sub-grid cloud mixing. In addition, one more sensitivity simulation has been carried out to investigate the impact of vertical grid structure on inert tracers at the surface. With the help of these sensitivity simulations for WRF/CMAQ, the model differences in inert tracers between WRF/CMAQ and the other models can be tied to specific processes in the revised manuscript. Ideally, the sensitivity simulations conducted for WRF/CMAQ would also have been performed for the other three models. Unfortunately, however, since these sensitivity simulations were not part of the AQMEII3 protocol most sensitivity simulations are not available for the other three models.

Changes in manuscript: The results from these sensitivity simulations have been included in the revised paper. The results have also been reorganized so that the results are discussed with respect to the impact of each physical process on the total inert tracers at the surface and on the relative contributions of inert tracers from different altitude ranges, respectively.

3 The multi-panel color-coded maps are difficult for a reader to follow (especially there are 7 of them) and sometimes misleading (because the color scales are different). Since the paper mainly focuses on the difference between mountain west and eastern US, the authors may consider to summarize the results in two regions and present it to the main points of their discussion.

Response: We agree that the panel plots showing the entire domain may have made the discussions hard to follow. As suggested by the reviewer, results are shown for sub-regions in the revised manuscript.

Changes in manuscript: In the revised paper, seven sub-regions were selected from the domain based on their proximity to the lateral boundaries, elevations, and climate, including WB (region close to western boundary), NB (region close to northern boundary), MT (west mountain area), GP (great plain area), NE (northeast), SE (southeast), and ATL (the Atlantic Ocean). When explaining the impact of physical processes on the inert tracers at surface, the results (including the vertical profiles and diurnal cycles of inert tracers) are averaged over sub-regions, so that readers can clearly see the different/similar roles that certain physical process plays in a variety of regions over the U.S. Some of the panel plots of the entire domain are still kept and reorganized to show the general differences between different simulations and two tables have been added to summarize the averaged results over land. The color scales are kept the same when showing the same metric in different sub-sections in the results.

4 Discussions are mostly descriptive and qualitative. Readers may be interested in quantitative estimates on the uncertainty associated with the model configurations and physical processes.

Response: Due to the limited data available to us at the time the original manuscript was prepared, it had not been possible for us to provide more quantitative estimates regarding the uncertainty associated with the model configurations and physical processes. However, we agree with the reviewer that providing more quantitative estimates would be of interest to the reader. Therefore, we have carried out extra sensitivity simulations for WRF/CMAQ to quantify the effects of different process representations on the fate of inert boundary condition tracers. Ideally, the sensitivity simulations conducted for WRF/CMAQ would also have been performed for the other three models. Unfortunately, however, since these sensitivity simulations were not part of the AQMEII3 protocol, , most sensitivity simulations are not available for the other three models.

Changes in manuscript: In the revised paper, the uncertainty of inert tracers at the surface associated with the model configurations and physical processes has been quantitatively estimated for WRF/CMAQ. In addition, with the help of the extra sensitivity simulations for WRF/CMAQ, more quantitative estimates are also provided explaining the roles of model configurations and physical processes in the differences in inert tracers between WRF/CMAQ and the other models.

Minor comments

1 Page 7 Line 18-20; Line 31-32: I am not convinced that the higher BC2 (lower BC1) in the mountain west is due to vertical mixing between mid-troposphere and the PBL. More likely, the existence of high mountains can divert the wind in BC1-enriched lower troposphere, so there is less BC1 advected to mountain regions, which is very different from vertical mixing.

Response: We agree that this is a plausible alternative explanation for this behavior.

Changes in manuscript: This statement has been removed in the revised paper. With the current data, no further analysis can be made to determine whether the higher BC2 and lower BC1 over the mountain is due to vertical mixing or the diversion of the wind. In addition, in the revised paper, we removed the discussions built on the results of WRF/CMAQ alone. Instead, we focused on the comparison between WRF/CMAQ and the extra sensitivity simulations for WRF/CMAQ to quantitatively estimate the uncertainty related to model configurations and physical processes, so that our points are much better supported by the results and presented to the readers more clearly.

2 Page 7 Line 21: "One possible explanation is that the impact from vertical mixing and the impact from LBCs cancel out in the mountain west region." Are all these coming from lateral boundary conditions as the authors are analyzing lateral boundary inert tracer? What exactly is the "impact from LBC" here referring to?

Response: Yes, they all come from lateral boundary as we are analyzing the lateral boundary inert tracers reaching the surface of the U.S. The levels of inert tracers at the surface are affected by 3D advection, vertical turbulent mixing and some other physical processes. Here, the "impact from LBC" refers to the inert tracers coming to the mountain area through horizontal advection. As explained in the following sentence, since the western boundary is predominantly an inflow boundary for lateral boundary inert tracers, it should be expected that BC1 in the west is higher than that in the east due to horizontal advection. We agree that this statement is confusing and should be modified.

Changes in manuscript: In the revised paper, we removed such discussions built on the results of WRF/CMAQ alone. Instead, we focused on the comparison between WRF/CMAQ and the extra sensitivity simulations for WRF/CMAQ to quantitatively estimate the uncertainty related to model configurations and physical processes, so that our points are much better supported by the results and presented to the readers more clearly.

3 Table 1: Do the differences in "gas phase chemistry", "wet deposition for tracers", "impact of sub-grid clouds on radiation", and "impact of sub-grid clouds on ozone photochemistry" affect BC1 and BC2 results? What are the vertical turbulent mixing schemes for different model, which are discussed in the paper?

Response: The differences in "gas phase chemistry", "impact of sub-grid clouds on radiation", and "impact of sub-grid clouds on ozone photochemistry" do not affect BC1 and BC2 since they are chemically inert tracers. The differences in "wet deposition for tracers" might affect BC1 and BC2 since it is one of the physical processes that the inert tracers would undergo. The vertical turbulent mixing schemes used in the CTMs are ACM2 for WRF/CMAQ and COSMO-CLM/CMAQ, and K-theory for WRF/CAMx WRF/DEHM.

Changes in manuscript: Since our paper focuses on the physical treatment in CTMs and its impact on inert tracers of lateral boundary ozone, the content in Table 1 has been revised to only include the physical processes that the inert tracers undergo, including dry and wet deposition, sub-grid cloud mixing and vertical turbulent mixing in CTMs. The processes related to ozone chemistry have been removed.

4 Figure 10: The authors may consider to move figure 10 to main results. It shows nicely why this study is important.

Response: In the original paper, Figure 10 was placed in "summary and discussion" to demonstrate the relative importance of physical processes versus the chemical processes on the lateral boundary ozone. As the reviewer suggested, we have moved it to main results to show the important impacts of physical processes on the lateral boundary ozone reaching the surface.

Changes in manuscript: This Figure has been moved to the section of results, under the sub-section of WRF/CMAQ versus WRF/CAMx.

Response to referee #2

We would like to thank the reviewer for providing constructive comments that have helped us improve the manuscript. The reviewer comments are shown in regular font, our responses are shown in bold fond, and a summary of the changes made to the manuscript in response to each comment is shown in italics.

1 The simulated tracers, which are meant to indicate the impact of BC ozone on surface ozone, are not compared in any way to the simulated surface ozone values of the models.

Response: No, the results of inert tracers for lateral boundary ozone are not compared with the simulated ozone values, because quantifying model performance for ozone is beyond the scope of this study. As the reviewer subsequently pointed out, the title of the original manuscript may have been misleading, leading readers to expect a quantitative estimate of the contributions of lateral boundary ozone to ozone levels at the surface. However, this is not the goal of our paper.

Changes in manuscript: To avoid the confusion, the title has been changed to "Attributing Differences in the Fate of Lateral Boundary Ozone in AQMEII3 Models to Physical Process Representations". In addition, the introduction has been modified to emphasize that this paper focuses on quantifying the model variability originating from the physical processes in CTMs, and aims at understanding how different representation of physical processes in CTMs may lead to the differences in the LB ozone that eventually reaches the surface across the U.S.

2 The tracer results are not juxtaposed to observations (or model errors) in any way despite the surprising fact that the tracer results are shown only for the location of the ozone observations.

Response: We agree with the reviewer that since the modeled results had not been compared with observations in the original manuscript, it did not make sense that the results were shown only at the locations where ozone observations are available. The reason for displaying the model data in this fashion in the original manuscript was that such data extracted at monitoring locations was readily available while gridded model fields were not. However, such gridded model fields have now been obtained from all participating groups and are used in the revised manuscript.

Changes in manuscript: Instead of using the AQMEII3 dataset at ozone receptors, gridded model fields have been obtained and are used in the revised paper. To generate this gridded data set, participating groups re-gridded the modeled hourly values for inert tracers at the surface to a common domain for analysis and comparison, covering the area from 23.5^o N/-130.0^o W to 58.5^o N/ -59.5^o W (green shaded area in Figure 1a in the revised paper) with grid spacing of 0.25^o by 0.25^o. Spatially continuous maps of the model differences in inert tracers at surface are shown in the revised paper.

3 A BC tracer which is meant to mimic ozone should be subject to an equivalent chemical loss, which is not the case in this study.

Response: The reviewer is correct that a BC tracer aimed at quantitatively attributing surface ozone to BC should include the chemical loss of ozone. Several studies have used reactive tracers to estimate the contributions of lateral boundary ozone to the ozone level at surface over the U.S. However, few studies have investigated the uncertainty in the impact of lateral boundary ozone on the ozone levels at the surface. The goal of this paper is to focus on the model variability originating from the physical processes in CTMs, and it aims at understanding how different representation of physical processes in

CTMs may lead to the differences in the LB ozone that eventually reaches the surface across the U.S. This research goal can be fulfilled by using chemically inert tracers for lateral boundary ozone.

Changes in manuscript: Modifications have been made in the introduction section to emphasize the research goal of this study, and we also clarified that "it is necessary to include the chemical loss of LB ozone when quantitatively estimating the impact of LB ozone, as shown in the comparison between inert and reactive LB ozone tracers by Baker et al. (2015)" to avoid any misunderstanding in the use of inert tracers.

4 The impact of the top boundary is not discussed despite that fact that stratospheric ozone can influence surface ozone in certain situations.

Response: We agree that stratospheric ozone can influence surface ozone in certain situations. However, quantifying the contributions of stratospheric ozone to surface ozone across the U.S. is outside the scope of this paper. In terms of the influence of the top boundary on simulation results, we also refer the reviewer to our response to comment #11.

Changes in manuscript: Modifications have been made in the title and in the introduction section to emphasize the research goal of this study, so that the readers do not expect a quantitative estimate of the contributions of ozone from lateral boundaries or stratospheric intrusion to the ozone levels at the surface.

5 The choice of the DMA8 average needs to be better motivated as it excludes night time values, when differences in mixing can be large between different models.

Response: We agree that it may be also important to look at the model differences during the night time.

Changes in manuscript: In the revised paper, the diurnal cycles of the model differences in the inert tracers at surface are also investigated, and the results are shown by averaging over sub-regions on a seasonal basis.

6 The difference between the models are discussed in a lot of detail, which is overwhelming for the reader.

Response: We agree that the original manuscript included a lot of details that sometimes may have distracted the reader from the main points.

Changes in manuscript: Some minor features in the temporal and spatial differences in inert tracers between different models have been removed. Instead, in the revised paper, when comparing the model differences in inert tracers, the discussion is reorganized so that the impact of different physical treatment in CTMs on inert tracers are discussed individually, including vertical resolution, wet scavenging, dry deposition, sub-grid cloud mixing, and vertical turbulent mixing.

7 The title of the study suggest that the paper will give an estimate of the impact of the BC on simulated surface ozone. I did not find any such estimate in the paper. The simulated ozone surface concentrations are not compared to the BC tracer results in anyway. I strongly suggest to change the title of the paper to avoid giving a false impression of its content.

Response: We agree that the title of the original manuscript may have been misleading and may have led the reader to expect an estimate of the impact of the BC on simulated surface ozone, which is not the goal of this study.

Changes in manuscript: The title has been changed to 'Attributing Differences in the Fate of Lateral Boundary Ozone in AQMEII3 Models to Physical Process Representations".

8 Although artificial tracers cannot be directly evaluated with observations, an analysis of the model errors for ozone in relation to the tracer simulation should have been an aspect of the paper.

Response: We agree that an analysis of the model errors for ozone in relation to the tracer simulation could be interesting. However, such analysis would not provide information about the model variability associated with the model configurations and physical processes in the lateral boundary ozone reaching the surface, which is the goal of this research. Therefore, such analysis is not included in the paper. However, we agree that the differences in tracers can sometimes explain differences in ozone concentrations, highlighting the importance of this study. In particular, it was found that the differences in DM8A O_3 and DM8A BCT between WRF/CMAQ and WRF/CAMx show strong spatial correlations with similar magnitudes except for summer, as shown in Figure 10 in the original paper (now Figure 7 in the revised paper).

Changes in manuscript: The Figure 10 in the original paper has been moved to the main results section under sub-section WRF/CMAQ versus WRF/CAMx as part of Figure 7.

9 On the other hand, the paper does give the impression that observations are considered because the observations are discussed (figure 1, section 2.3) and all plots of tracer simulation (figures 3-4, 6-10) are provided at the sampling location of the observations. But there is no use of this, which is confusing. Producing continuous maps of the tracers simulations would have been more appropriate. The only figure (11) showing a comparison with observation is a time series of mean differences between rural and urban observations. This could be an interesting figure but it is just mentioned in passing with no detailed discussion.

Response: We agree with the reviewer. Please see our response to comment #2 regarding the use of gridded model fields in the revised manuscript.

Changes in manuscript: Instead of using the AQMEII3 dataset at ozone receptors, gridded model fields have been obtained and are used in the revised paper. To generate this gridded data set, participating groups re-gridded the modeled hourly values for inert tracers at the surface to a common domain for analysis and comparison, covering the area from 23.5^o N/-130.0^o W to 58.5^o N/ -59.5^o W (green shaded area in Figure 1a in the revised paper) with grid spacing of 0.25^o by 0.25^o. Spatially continuous maps of the model differences in inert tracers at surface are shown in the revised paper.As to Figure 11 in the original paper, it has been removed in the revised paper as such discussion is not closely related to the main points of this paper.

10 Getting a better knowledge of the impact of boundary ozone concentrations, both from a modelling or more factual point of view, on surface ozone over North-America is an interesting question. Using dedicated tracer simulations is an appropriate method to investigate the problem. The authors apply tracers that are only subject to dry and wet deposition but not the chemical ozone loss. It is debatable – as also pointed out by the authors in the conclusion - if this approach does full justice to the problem of ozone because chemical ozone loss could be an important factor for the importance of BC on surface ozone. A quantitative comparison of the contribution of (i) chemical loss, (ii) loss by dry deposition and (iii) by wet deposition to the ozone budget in the study area for the different models would have been

needed to demonstrate if the used approach is reasonable or not. (It would probably also show that wet deposition of ozone is of minor importance). Truly inert BC tracers, i.e. also without deposition, give some indication of the transport processes only. They would be not specific to the characteristics of ozone (apart from the BC values) but could still give some information. Hence, a better discussion of the runs without dry deposition could be helpful.

Response: We agree that to quantitatively estimate the contributions of lateral boundary ozone to the ozone levels at surface, the chemical loss of ozone must be included. This has also been confirmed by comparing the inert and reactive lateral boundary ozone tracers at the surface across the U.S. (Baker et al.,2015). However, the goal of this paper is not to provide such an estimate of the contributions of lateral boundary ozone. Instead, this paper focuses on quantifying the model variability originating from the physical processes in CTMs, and aims at understanding how different representation of physical processes in CTMs may lead to the differences in the LB ozone that eventually reaches the surface across the U.S. Using chemically inert tracers is an appropriate approach for accomplishing this goal. We agree with the reviewer that a quantitative comparison of the effects of different processes that the inert tracers undergo would be valuable. Therefore, extra sensitivity simulations for WRF/CMAQ have been carried out to provide quantitative estimates of the uncertainty of inert tracers at surface associated with the model configurations and physical processes, including vertical resolution, wet scavenging, dry deposition and sub-grid cloud mixing. In addition, with the help of these sensitivity simulations, more quantitative estimates are also provided at the process level when explaining the differences in inert tracers between WRF/CMAQ and the other models.

Changes in manuscript: The introduction section has been modified to clarify the research goal and the reasons that inert tracers were used. The results section is reorganized so that the impact of different physical treatment in CTMs on inert tracers are discussed one by one.

11 The authors find very little transport from the lateral BC to the surface for tracer BC3 (< 250 hPa). It needs to be discussed in more detail, if this means that stratospheric ozone does not impact surface ozone in the study area at all. However, the top boundary condition would need to be considered for such an investigation. Considering only lateral influx with BC3, given that the models had only 1, 3 or 5 model levels above 250 hPa, seems not sufficient. A realistic top boundary is required to simulate the impact of upper tropospheric and stratospheric ozone on surface ozone.

Response: First, we would like to emphasize that the BC3 results shown in this study are seasonal averages over broad regions. Therefore, they would not be expected to show the impact of episodic stratospheric intrusion events affecting specific monitors within the domain, and the relatively low seasonally and regionally averaged BC3 values do not indicate that such events did not happen. Second, neither CMAQ nor CAMx consider downward fluxes through the top of the modeling domain. Therefore, the only direct estimates of stratospheric influences within the domain are from the BC3 tracers which represent stratospheric ozone at the location of the lateral boundaries and are then advected and potentially mixed downward within the regional domain. Finally, we would like to emphasize that on a hemispheric-to-global scale, there is a large contribution of stratospheric ozone to free tropospheric ozone. Therefore, while the BC2 tracer represents free tropospheric ozone at the location of the lateral boundaries for the regional scale models, it implicitly includes stratospheric contributions occurring at larger space and time scales, and the important contribution of BC2 to surface tracer concentrations

across the regional domain therefore implies that the stratospheric ozone at hemispheric-to-global scale impacts surface ozone.

Changes in manuscript: We do not believe any changes to the manuscript are needed in response to this comment.

12 A clarification of the setup of the BC tracers would be helpful. Did the values of the ozone BC from the global model agree with observations? Where the BC placed at exactly the same position for all models or did the location of the model boundaries vary from model to model? This could have had an impact on the results.

Response: We agree that adding this information may be helpful for readers. The reviewer also made a good point that the location of the model boundaries may vary from model to model and may have had an impact on the results. WRF/CMAQ, WRF/CAMx and COSMO-CLM/CMAQ have the same (or very similar) modeling domains, making their results of inert tracers comparable to each other. However, WRF/DEHM has a very different domain coverage. Therefore, in the revised manuscript WRF/DEHM results are only used in a relative sense to assess the effects of dry deposition on simulated tracer concentrations.

Changes in manuscript: More details have been added in the section of "chemically inert tracers" regarding the setup of the tracers. Given that the models use different vertical resolution, Figure 2 has also been modified to better demonstrate how each model attributes lateral boundary ozone at different altitude ranges to BC1, BC2 and BC3. The lateral boundary conditions of ozone for all models are derived from C-IFS and have been evaluated against observations (Hogrefe et al., 2018). This statement is added in the section of "model description". The modeling domains and the analysis domain have also been added in the revised paper and shown are in Figure 1a.

13 The main finding of the authors is that vertical mixing differs between the models. It would be interesting to elaborate on the reasons for this. The authors could provide a confirmation of this finding by studying profiles of primary species emitted over the domain. Also, a statement on which of the different mixing scheme leads to the best results would give useful information to the reader.

Response: The differences in vertical turbulent mixing are either due to different parameterizations used in different CTMs (e.g. WRF/CMAQ and WRF/CAMx in this study), or due to the different meteorological input fields (e.g. WRF/CMAQ and COSMO-CLM/CMAQ in this study). It is true that profiles of primary species emitted over the domain may help to confirm the model differences in vertical turbulent mixing. However, the vertical profiles are not available for WRF/CAMx and WRF/DEHM as such data were not required to be archived in AQMEII3. The vertical profiles of inert tracers for WRF/CMAQ and COSMO-CLM are compared to show their potential difference in vertical mixing. Identifying the scheme that leads to the best results is beyond the scope of the paper. The goal of this analysis is to demonstrate that the process of vertical turbulent mixing is an important source of uncertainty when estimating the impact of lateral boundary ozone on the ozone levels at the surface.

Changes in manuscript: The reasons for different vertical turbulent mixing have been clearly mentioned when comparing each model pair. The importance of vertical mixing in causing the tracer differences has been supported by the sensitivity simulations of WRF/CMAQ for the model pair of WRF/CMAQ and

WRF/CAMx, and by the analysis of PBL height and vertical profiles of inert tracers for the model pair of WRF/CMAQ and COSMO-CLM/CMAQ.

14 DMA8 is the only time aggregation of the discussed BC results. The choice of DMA8 for the given applications needs to be better motivated. As the focus of the study is to identify differences in vertical mixing, night-time values should not be disregarded as it is most likely the case for DMA8 of ozone. It is during night time that different vertical mixing (or the lack thereof) can have the largest impact on surface values. Likewise it would be interesting to check if the time period for the calculation of DM8A was the same for the models. Differences in the diurnal cycle can be an indication of different simulation of vertical mixing.

Response: We agree that the choice of DM8A may not describe the complete picture of the model difference in vertical mixing. Therefore, at the revised manuscript also presents an analysis of the diurnal cycles in the model differences. As the reviewer pointed out, we do find that in some cases, the differences in the diurnal cycle can be an indication of the differences in vertical mixing between different models. The time period for calculating the DM8A was not exactly the same across the models, but the difference is usually within two hours.

Changes in manuscript: In the revised paper, the diurnal cycles of the model differences in the inert tracers at the surface have been added, and the results are shown as averages over sub-regions on a seasonal basis.

15 The colour scale for absolute values could be improved. In my printout the magenta (e.g. 25-30 ppb range, top, Fig 3) looks very similar to the red colors indicating high values.

Response: We apologize for the problems in the color scales.

Changes in manuscript: The color scales have been adjusted for easier readability.

~~Multi-Model Comparison~~Attributing Differences in the ~~Impact~~Fate of Lateral Boundary ~~Conditions on Simulated Surface~~ Ozone ~~across the United States Using Chemically Inert Tracers~~in AQMEII3 Models to Physical Process Representations

Peng Liu[1], Christian Hogrefe[2], Ulas Im[3], Jesper H. Christensen[3], Johannes Bieser[4], Uarporn Nopmongcol[5], Greg Yarwood[5], Rohit Mathur[2], Shawn Roselle[2], Tanya Spero[2]

[1] NRC Research Associate, in the National Exposure Research Laboratory, U.S. Environmental Protection Agency, Research Triangle Park, NC, 27711, USA
[2] National Exposure Research Laboratory, U.S. Environmental Protection Agency, Research Triangle Park, NC, 27711, USA
[3] Aarhus University, Department of Environmental Science, Frederiksborgvej 399, 4000, Roskilde, Denmark
[4] Helmholtz-Zentrum Geesthacht, Institute of Coastal Research, Max-Planck-str. 1 21502 Geesthacht, Germany
[5] Ramboll, ~~773 San Marin Drive~~7250 Redwood Boulevard, Suite ~~2115~~105, Novato, CA 94945, USA

Correspondence to: ~~Peng Liu (liu.peng@epa.gov)~~Christian Hogrefe (Hogrefe.Christian@epa.gov)

Abstract. ~~This~~ Increasing emphasis has been placed on characterizing the contributions and the uncertainties of ozone imported from outside the U.S. In chemical transport models (CTMs), the ozone transported through lateral boundaries (referred to as LB ozone hereafter) undergoes a series of physical and chemical processes in CTMs, which are important sources of the uncertainty in estimating the impact of LB ozone on ozone levels at the surface. By implementing inert tracers for LB ozone, the study ~~represents an inter- comparison of four~~ seeks to better understand how differing representations of physical processes in regional~~-scale air quality simulations, focused on understanding similarities and~~ CTMs may lead to differences in the simulated ~~impact of~~LB ozone ~~lateral boundary conditions (LBCs) on~~that eventually reaches the ~~ground-level ozone predictions~~surface across the U.S. ~~The chemically inert tracers were implemented in the simulations as a diagnostic tool to understand the similarities and differences between models at process level. For all simulations,~~For all the simulations in this study (including WRF/CMAQ, WRF/CAMx, COSMO-CLM/CMAQ, and WRF/DEHM), three chemically inert tracers ~~(BC1 BC2 and BC3) are used to track the impact of ozone specified at different altitudes along the lateral boundaries of the modeling domain encompassing the contiguous U.S. The~~ which generally represent the altitude ranges ~~specified for BC1, BC2, and BC3 broadly represent the~~ of the planetary boundary layer (~~PBL~~BC1), free troposphere~~,~~ (BC2), and upper troposphere-lower stratosphere~~, respectively~~ (BC3), are tracked to assess the simulated impact of LB specification.

~~The four simulations, namely WRF/CMAQ, WRF/CAMx, WRF/DEHM and COSMO-CLM/CMAQ, can have considerable differences in the simulated inert tracers at surface, indicating their different estimates in the impact of lateral boundary on surface ozone within the U.S. due to the physical processes alone in chemical transport models. WRF/CMAQ is used as a base case, and the differences between WRF/CMAQ and the other three models are examined, respectively. The model pair of COSMO-CLM/CMAQ and WRF/CMAQ shows the smallest differences in inert tracers, with the difference in BCT (sum of~~

BC1, BC2 and BC3) peak in winter to be 1.6 ppb averaged across all sites. The model pair of WRF/DEHM and WRF/CMAQ shows the largest differences, with difference in BCT peak in summer to be 8.1 ppb averaged across all sites.

Furthermore, the model differences in inert tracers are discussed with respect to the physical processes that inert tracers undergo. It is found that the process of vertical turbulent mixing between the PBL and the free troposphere is the main cause of the model differences in the simulated inert tracers, especially the relative contributions of BC1 and BC2 to the total inert tracers, in most seasons and regions of the U.S., although the processes of sub-grid cloud mixing and dry deposition can also be important drivers for specific regions and seasons. Comparing WRF/CAMx with WRF/CMAQ, their differences in vertical grid structure explain 10%-60% of their seasonally averaged differences in inert tracers at the surface. Vertical turbulent mixing is the primary contributor to the remaining differences in inert tracers across the U.S. in all seasons. Stronger vertical mixing in WRF/CAMx brings more BC2 downward, leading to higher BCT (BCT=BC1+BC2+BC3) and BC2/BCT at surface in WRF/CAMx. Meanwhile, the differences in inert tracers due to vertical mixing is partially counteracted by their difference in sub-grid cloud mixing over the southeastern U.S. and the Gulf coast region during summer. The process of dry deposition adds extra gradients to the spatial distribution of the differences in DM8A BCT by 5 - 10 ppb during winter and summer.

COSMO-CLM/CMAQ and WRF/CMAQ show similar performance in inert tracers both at the surface and aloft through most seasons, which suggests the similarity between the two models at process level. The largest difference is found in summer. Sub-grid cloud mixing plays a primary role in their differences in inert tracers over the southeastern U.S. and the oceans in summer. Our analysis on the vertical profiles of inert tracers also suggests that the model differences in dry deposition over certain regions are offset by the model differences in vertical turbulent mixing, leading to small differences in inert tracers at surface in these regions.

1 Introduction

Studies based on chemical transport models (CTMs) have shown that air quality in the U.S. can be considerably influenced by pollutants beyond the U.S. boundaries, such as through intercontinental transport, and through stratosphere-to-troposphere exchange (Zhang et al., 2011; Lin et al., 2012; Nopmongcol et al., 2016; Langford et al., 2017; Lin et al., 2017; Hogrefe et al., 2017; 2018). Mathur et al., 2017). Similar findings have also been reported based on routine observations and field campaign measurements (e.g. Cooper et al., 2012; Gratz et al., 2015; Langford et al., 2015), especially at the rural and elevated locations in the western U.S. Recent revisions to the National Ambient Air Quality Standards (NAAQS) for ground-based ozone (Federal Register, 2015), which further lowered both the primary (health-based) and secondary (welfare-based) standards from 75 ppb to 70 ppb, have placed for ground-based ozone (Federal Register, 2015). Therefore, increasing emphasis has been placed on the need to characterize the contributions and the uncertainties of ozone imported from outside the U.S.

To quantify the impact The contributions of emissions and pollutants ozone from outside the region of interest, U.S. to the surface ozone within the U.S. has been estimated by several studies with different approaches have been implemented in CTMs and broadly fall into two categories. The first category is called, including "source sensitivity approach", " (such as "bruteforce" method and the decoupled direct method in 3-dimensions (DDM-3D) (Dunker, 1984; Yang et al., 1997; (e.g. Dolwick et al., 2015)), the Path-Integral Method (Dunker et al., 2002). The second category is called "source apportionment2017), and the tagged species approach". For example, pollutants from different sources can be tracked directly by adding chemically reactive tracers to CTMs. Based on the approaches above, a variety of model-specific tools have been developed and enhanced for different CTMs, (such as DDM-3D for the Community Multiscale Air Quality (CMAQ) model (Cohan et al., 2005) and for the Comprehensive Air quality Model with extentions (CAMx) (Yarwood et al., 1996; Dunker et al., 2002), respectively, the Integrated Source Apportionment Method (ISAM) for CMAQ (Kwok et al.,2015), the Ozone Source Apportionment Technology (OSAT) for CAMx (ENVIRON, 2015), the adjoint model for GEOS-Chem (Henze et al., 2007), and a tagging scheme for ozone production from NO sources in the Model for Ozone and Related chemical Tracers (MOZART-4) (Emmons et al.,Ramboll, 2018), and the chemically reactive tracers (2012). With the help of these tools, the impact of ozone from outside the U.S. on the surface ozone within the U.S. has been quantitatively estimated (e.g. Baker et al., 2015; Dolwick et al., 2015; Nopmongcol et al., 2017).)).

Any single model realization of CTMs with the implementation of these approaches, however, may not be enough to describe the complete picture of interest due to the uncertainties in air quality modeling, which stem from the model inputs (e.g. Hanna et al., 2001) and the CTMs themselves (as reviewedThe simulated ozone levels by Russell and Dennis, 2000). Multi-model ensembles can help with uncertainty quantification in air quality modeling (Mallet and Sportisse, 2006; Vautard et al., 2006). In addition, the prediction by a multi-model ensemble may outperform any individual member (e.g. van Loon et al., 2007; Langner et al., 2012). Creating multi-model ensembles is however resource intensive and often reliant on extensive collaborations across multiple participating organizations. For example, about 40 institutions from the U.S. and Europe have participated in the three phases of theregional CTMs can be influenced by uncertainties in specification of lateral boundary (LB) conditions. For example, in the phase 3 of Air Quality Model Evaluation International Initiative (AQMEII) (Galmarini et al. 2012, 2015, 2017), enabling the inter-comparison across a variety of CTMs and the examination of the multi-model ensembles for both the U.S. and the Europe. Substantial effort in the past multi-model studies has focused either on the operational evaluation (Dennis et al., 2010) across models, in which modeled results are compared against corresponding measured data (e.g. Solazzo et al., 2012; Im et al, 2015), or on the generation and analysis of multi-model ensembles (e.g. Delle Monache and Stull, 2003; Mallet et al., 2009; Galmarini et al. 2013; Kioutsioukis et al., 2016). Only limited efforts AQMEII3), Hogrefe et al. (2018) analysed the impact of LB ozone derived from four global or hemispheric CTMs on the ozone predictions over the U.S. using CMAQ and found significantly varying impacts of LB conditions on predicted surface ozone levels. Furthermore, LB ozone undergoes a series of physical and chemical processes in CTMs, which may be represented differently due to different model configurations and parameterizations chosen by the models (Russell and Dennis, 2000). Limited efforts, however, have been devoted towards elucidating the reasons at the process level for the noted similarities/differences in surface ozone prediction among the models. Campbell et al. (2015), for example, compared the indicators of model predictions in surface ozone sensitivity across regionaland the impact of LB ozone, though studies have suggested the important role that the processes in CTMs in Phase 2 of AQMEII. In Phase 3 of AQMEII (play in explaining the model differences. For example, also in AQMEII3~~)~~,, Solazzo et al. (2017) ~~examined how~~compared the model errors in surface ozone ~~prediction can be apportioned to different~~ predictions over the U.S. and the Europe from several regional CTMs and showed that errors across a series of time scales ~~and how the errors~~ could be ~~related~~attributed to ~~the~~different chemical and physical processes in ~~CTMs. Such process-level diagnostic inter-comparison of~~the CTMs.

Understanding how the differences in model predictions can be attributed to scientific processes in CTMs is important for several reasons. First, comparison in fundamental processes can help to ~~achieve~~mitigate the reducible error in air quality models, which can be achieved by the scientific ~~advances~~improvements in the representations of the physical and chemical processes in CTMs, so that model prediction from a single CTM can be improved~~, as reviewed by~~ (e.g. Zhang et al., 2012). ~~. (2012). Second, process-level multi-model comparisons can help to guide the research directions to reduce model uncertainty by~~Second, identifying the major process(es) that contributes to the variability across models~~.~~ can help to guide the research directions to reduce model uncertainty and error. Last, a better understanding of the model similarities and differences at the process level could improve multi-model ~~ensembles~~ensemble by increasing the independence of ensemble members.

~~This study, performed under AQMEII3, aims to investigate how different representations of physical processes and different model configurations in CTMs may lead to the differences in the estimated impact of ozone imported from outside the U.S. on the surface ozone within the U.S. To enable such diagnosis, all participating groups in AQMEII3 implemented chemically inert tracers to track the transport, scavenging and deposition of ozone imported from the lateral boundaries into the simulation domain, which covers the COntiguous United States (CONUS) and southern Canada (Figure 1). It should be noted that since chemical loss is not considered, the inert tracers should only be regarded as a diagnostic tool to help understand the similarities and differences between models, rather than a tool for the long-range-transport attribution. The inert tracer method has its advantage in two aspects. First, compared with the “source sensitivity” and “source apportionment” approaches mentioned above, the method of inert tracers is much easier to be consistently implemented in different CTMs by all AQMEII3 participants. Second, since the representation of chemistry differs across the CTMs, the inert tracer method isolates the differences related to the representation of 3-D transport (advection, turbulent mixing and cloud mixing) and the physical sinks from the differences in chemical processes in CTMs.~~This study, therefore, focuses on examining the impact of physical treatments in CTMs on LB ozone and aims at a better understanding of how different representations of physical processes in CTMs may lead to the differences in the LB ozone that eventually reaches the surface across the U.S. To keep track of the LB ozone, chemically inert tracers for LB ozone have been implemented in all participating models in this study, in which chemical loss of LB ozone is excluded. The important thing to clarify is that it is necessary to include the chemical loss of LB ozone when quantitatively estimating the impact of LB ozone, as shown in the comparison between inert and reactive LB ozone tracers by Baker et al. (2015). This study, instead of providing such quantitative estimate, aims at understanding the model variability that originates from the physical treatments in CTMs and its impact on the LB ozone reaching the surface. The implementation of chemically inert tracers enables us to completely focus on the impact of physical treatments in CTMs. Otherwise, it would be very difficult to disentangle the impact of chemical processes from the impact of physical processes if chemically reactive tracers for LB ozone are employed, as chemical and physical processes are intricately coupled in CTMs.

This paper is organized as follows. Section 2 describes the model configurations and how the chemically inert tracers are implemented. In section 3, the seasonal impact of lateral boundary ozone due to the representations of physical processes and the model configurationstreatment in CTMs on inert tracers at surface is compared among modelsexamined by season. The model comparing WRF/CMAQ to several sensitivity simulations. Then WRF/CMAQ is used as a base case and the differences in inert tracers between WRF/CMAQ and three other models are investigated and discussed with respect to the physical processes that inert tracers are involved. Section 4 discusses how the model differences noted in section 3 may change if chemical loss of the ozone from lateral boundaries was also consideredFinally, the findings are summarized in Section 4.

2 Methods

2.1 Model description

This study, performed as part of AQMEII3, investigates simulations conducted by four different research institutes/groups from the U.S. and Europe, using state-of-the-art regional CTMs. Based on The four simulations are named using the combination of the regional CTMs and the models used for theto generate their meteorological driving fields and CTMs, the four simulations are named asinputs: WRF/CMAQ, WRF/CAMx, COSMO-CLM/CMAQ, and WRF/DEHM in this study. Important model details are summarized in Table 1, with additional model . A description of the model features availableand emissions can be found in Solazzothe technical note by Galmarini et al. (2017). Two additional sensitivity simulations were also available, namely WRF/CMAQ_NODDEP and WRF/DEHM_NODDEP, in which all settings were the same as WRF/CMAQ and WRF/DEHM, respectively, except that the dry deposition of inert tracers was turned off. The simulation period consists ofis the entire year of 2010, which was determined by AQMEII3 based on the data availability of emissions and observations data. The four simulations share the same chemical boundary conditions for all simulations were derived from the Composition Integrated Forecasting System (C-IFS) global modelingmodelling system (Flemming et al., 2015) by the European Centre for Medium-Range Weather Forecast (ECMWF). The anthropogenic emissions were harmonized to be consistent across the models under the effort of AQMEII3. For WRF/DEHMForecasts (ECMWF). The LB ozone derived from C-IFS has been evaluated against observations (Hogrefe et al., 2018). WRF/CMAQ and WRF/CAMx share the same modelling domain (Figure 1a). The size of the modelling domain for COSMO-CLM/CMAQ is like that for WRF/CMAQ, but was shifted westward by 48km. WRF/DEHM, however, has a very different domain coverage than other models (Figure 1a). Therefore, the results of inert tracers for LB ozone are directly comparable among WRF/CMAQ, WRF/CAMx and COSMO-CLM/CMAQ, the anthropogenic emissions were prepared on a gridded and monthly basis using HTAP_v2.2 database (Janssens-Maenhout et al, 2015); while for the other simulations, the anthropogenic emissions were provided by U.S. EPA (Pouliot et al., 2015) on a gridded and hourly basis. The North American emissions in the HTAP_v2.2 database were compiled using the U.S. EPA emissions and therefore, the two sets of emission data in North America have the same annual total for each species in each sector. Non-anthropogenic emissions were estimated separately by each individual research group/institute. For example, biogenic emissions for WRF/CMAQ were estimated using the Biogenic Emissions Inventory

System (BEIS) v3.14 (Schwede et al., 2005; Pouliot et al., 2015) with USGS land use data, while the biogenic emissions for WRF/CAMx were determined by the Model of Emissions of Gases and Aerosols from Nature (MEGAN) v2.1 (Guenther and Wiedinmyer, 2007; Sakulyanontvittaya et al., 2008) with land use data from the North American Land Change (NALC) Monitoring System. A comprehensive description of the model activity and emissions related to AQMEII3 can be found in the technical note by Galmarini et al. (2017). In short, with the current modeling setup, the results from the four simulations mainly reflect the variability due to different meteorological driving fields, model configurations and the representation of processes in CTMs, all of which are important components of the uncertainty in air quality modeling (Fox, 1984; Solazzo and Galmarini, 2016).but not WRF/DEHM.

2.2 Chemically inert tracers

For each simulation, three chemically inert tracers were added specifically at the lateral boundaries to track ozone from outside of the modeling domain. The tracers undergo the same physical processes as ozone, including advection, diffusion, cloud mixing/transport (if represented in CTMs, as summarized in Table 1), scavenging and deposition, with no emissions or chemical formation/destruction. The only exception is that WRF/CAMx does not include any wet deposition for the inert tracers (Nopmongcol et al., 2017).at different altitudes from outside the modelling domain. The three tracers, representing the LB In all models, the deposition velocity of tracers was set to be the same as that of ozone.

The three tracers, representing the lateral boundary ozone from the planetary boundary layer (PBL), the free troposphere, and the upper troposphere-lower stratosphere respectively, are defined as follows: BC1 for vertical layers below 750 mbhPa (~ 2.5 km); BC2 for layers between 750 mbhPa (~2.5 km) and 250 mbhPa (~10 km); and BC3 for layers above 250 mbhPa. Initial conditions for all tracers were set to be zero and a ten-day spin-up period was used in the simulations. The lateral boundary conditions (LBCs) of the tracers were set to be the same values as the LBCs of ozone at the corresponding vertical layers, with zero values assigned in other layers. Therefore, these tracers are able to provide information on how specified ozone LBCs at different altitude ranges may eventually influence the simulated surface ozone in the U.SFor example, for WRF/CMAQ, BC1 is the LB ozone from layer 1 to 21, BC2 from layer 22 to 31, and BC3 from layer 32 to 35 (Figure 2). Therefore, these tracers can provide information on the altitude ranges which the LB ozone reaching the surface originates from. Due to the different vertical grid structure used by each model, differences occur in the attributions of LB ozone to inert tracers across models. For example, Figure 2 shows the typical pressure at each vertical level for the four models. In WRF/CMAQ, BC2 starts from layer 22, with the pressure at the bottom of the layer about 725 hPa; while in WRF/CAMx, BC2 starts from layer 17, with the pressure at the bottom of the layer about 755 hPa. Such differences may result in differences in the relative contributions of BC1 and BC2 to the total inert tracer at the surface, but are not expected to significantly change the total amount of BC1 and BC2 reaching the surface, which has also been confirmed later in section 3. BC3 starts from very similar pressure levels for WRF/CMAQ, WRF/CAMx and WRF/DEHM, but is different for COSMO-CLM/CMAQ due to its very coarse vertical resolution in the upper troposphere-lower stratosphere. The impact of such differences on inert tracers at the surface is also found to be small in general as the seasonal averaged contribution of BC3 at the surface is usually very small (less than 1.5 ppb) relative to BC1 and BC2 across the U.S. except for summer.

The tracers undergo the same physical processes as ozone, including 3D advective transport, vertical turbulent mixing, sub-grid cloud mixing (if represented in CTMs), scavenging and deposition. In all models, the deposition velocity of tracers was set to be the same as that of ozone. The physical processes that the inert tracers undergo in each model have been summarized in Table 1. To better distinguish the impact of each physical processes, a series of sensitivity simulations has been conducted for WRF/CMAQ, including WRF/CMAQ_noddry, WRF/CMAQ_nodwet, WRF/CMAQ_nocldmix (as described in Table 2). Ideally, the sensitivity simulations conducted for WRF/CMAQ are also desired for the other three models. However, since these sensitivity simulations were not part of the original design for AQMEII3 and entail additional non-trivial resource commitments from each participating organization, most sensitivity simulations are not available except for WRF/DEHM_noddry (Table 2). In addition, the vertical resolution, especially in the free troposphere, has been shown to be important for air quality models (e.g. Mathur et al., 2017; Eastham and Jacob, 2017). To investigate the impact of this physical treatment on LB ozone, a sensitivity simulation WRF/CMAQ_27aL was conducted. This simulation is the same as WRF/CMAQ except for using the same vertical grid structure as WRF/CAMx, as WRF/CAMx has the coarsest vertical resolution in free troposphere among the four models.

2.3 Data for analysis

As required by AQMEII3, the modeled hourly values for ozone, tracers and meteorological fields were uploaded to the ENSEMBLE data system developed by the EC Joint Research Center (Galmarini et al., 2001) for the purpose of sharing results between AQMEII3 participants. For ozone, the inert boundary condition tracers, and other trace gases, AQMEII3 participants were required to provide the modeled hourly values at the surface at the locations where ozone monitors are available. The observation networks constituting this set of monitor locations include the Clean Air Status and Trends Network (CASTNET), the Air Quality System (AQS), the SEARCH network for the U.S. and the National Air Pollution Surveillance (NAPS) network for Canada. In the studied area, there are 278, 614 and 587 monitors for urban, rural and suburban locations respectively (Figure 1). All multi-model comparison in section 3 was conducted based on the surface data at these monitors, unless otherwise specified. The fact that only surface level concentrations were available for this study poses a limitation to disentangling the impact of lateral boundary ozone on surface ozone at process-level.

Due to the different modelling domain and horizontal resolution across the models, the participating groups followed the AQMEII3 protocols and re-gridded the modeled hourly values for inert tracers at the surface to a common domain for analysis and comparison, covering the area from 23.5º N/-130.0º W to 58.5º N/ -59.5º W (green shaded area in Figure 1a) with grid spacing of 0.25º × 0.25º. In addition to the surface data, 3D data for inert tracers are also available for WRF/CMAQ, its sensitivity simulations, and COSMO-CLM/CMAQ, and have been interpolated to the same elevation levels so that the vertical profiles of inert tracers can be compared. The corresponding 3D data for WRF/CAMx and WRF/DEHM are not available, as 3D data were not included in the data archival protocols of AQMEII3.

Seven sub-regions are selected across the analysis domain (Figure 1b), based on their proximity to the lateral boundaries, elevations, and climate (Karl and Koss, 1984), including WB (region close to western boundary), NB (region close to northern boundary), MT (west mountain area), GP (great plain area), NE (northeast), SE (southeast), and ATL (the Atlantic Ocean). When calculating the statistical metrics for each sub-region, only the grid cells over land will be used for analysis except for the ATL sub-region.

3 Results

For all simulations involved inIn this study, it has been found thatsection, the seasonally averaged contribution of BC3 tomodel results for the mixing ratios of total inert tracerstracer (namely the sum of BC1, BC2 and BC3, hereafter referred to as BCT) is in general small at the surface (less than %5 on a seasonal average basis). Similar findings have also been reported by Nopmongcol et al. (2017). Furthermore, the differences in seasonally averaged BC3/BCT at the surface among models are also usually small (within ± 2.5%), compared with the differences in (BC1/BCT) and in (BC2/BCT). Therefore, only the resultsrelative contributions of BC1, BC2 and BCT will be shown in the subsequent analysis.

3.1 Multi-model comparison in the impact of lateral boundaries usingeach inert tracers

The model results for tracer to BCT are examined first, as BCT represents the impact of lateral boundary ozone from all altitude ranges due to the physical processes in CTMs.. First, WRF/CMAQ is used as a base case, and the differences inimpact of a variety of physical processes on the inert tracers at the surface is investigated by comparing WRF/CMAQ with sensitivity simulations. Then, the model differences are investigated and attributed to different physical treatment in CTMs for the model pairs of WRF/CMAQ versus WRF/CAMx, WRF/CMAQ versus COSMO-CLM/CMAQ, and WRF/CMAQ versus WRF/DEHM, respectively. All analysis was conducted on a seasonal basis.

The metrics examined for BCT and the relative contributions of each tracer include the daily maximum 8-hour average (DM8A) BCT between WRF/CMAQ and the other 3 models are shown on a seasonal basis (Figure 3). For each model,values and the diurnal cycles. The DM8A BCT is and relative contributions are calculated at the grid cells where the monitors are located as follows. First, forFor each daymodel, the eight-hour window when the modeled DM8A ozone occurs is found. for each day at each $0.25^{\circ} \times 0.25^{\circ}$ grid cell across the analysis domain. Then the averagedaverage mixing ratios of BCTeach tracer during that eight-hour window are calculated using the modeled hourly data at the surface, whichand these are referred to as DM8A BCT. Finally, the DM8A BCT are averaged for each season.

Considerable differences are seen between the models both spatially and seasonally, indicating that significant model differences in the impact of lateral boundary ozone can result from the physical processes in CTMs alone. For example, BCT from WRF/CAMx can be higher than that from WRF/CMAQ by as much as about 5 ~ 7.5 ppb for some regions during winter and summer; BCT from WRF/DEHM can be lower than that from WRF/CMAQ by as much as about 12.5 ~ 17.5 ppb in the southeastern U.S. during summer.

Meanwhile, in some instances, similar results in DM8A BCT (differences smaller than 2.5 ppb) between models are also found. WRF/CMAQ and WRF/CAMx show similar BCT results during fall, as do WRF/CMAQ and COSMO-CLM/CMAQ in all seasons at most locations. The similar model results for BCT, however, do not necessarily indicate model similarity in the physical processes that inert tracers undergo, because the four simulations share the same chemical boundary conditions, including the LBCs for the inert tracers. Hence, it is needed to confirm that the inert tracers at the surface are primarily affected by the physical processes in the CTMs per se, instead of by the LBCs.

This issue is investigated by examining the relative contributions of BC1BC1, DM8A BC2 and BC2 to BCT. If the inert tracers at the surface are determined by LBCs, little spatial variability would be expected in the relative contributions of tracers. For all simulations, it is found that the relative contributions of BC1 and BC2 vary significantly across the modeling domain. The ratio of (DM8A BC1) to (DM8A BC3. Then DM8A BCT (in ppb), DM8A BCT) from WRF/CMAQ is shown as an example (Figure 4, first row). The ratio of (BC1/BCT, DM8A BC2) to (/BCT and DM8A BCT) is not shown, as the ratio of (BC1+BC2) to BCT is almost 100%. The values of DM8A BC1 and DM8A BC2BC3/BCT (in percentage) are calculated following the same steps. Finally, the daily metrics are averaged for each season. For the seasonal averaged diurnal cycle for inert tracers, at each hour, the daily values for DM8A BCT. It is also found that (DM8A BC1)/(DM8A BCT) along the western and northern U.S. boundaries is usually higher than the inner U.S., due to higher DM8A BC1 (Figure 4, second row) and lower DM8A BC2 (Figure 4, third row) along the two boundaries. This is because the western and northern boundaries of the inert tracers at that hour are averaged over the season. The subsequent analysis mainly focuses on the direct differences in the metrics above between two simulations (e.g. DM8A BC1/BCT from simulation domain are the predominantly inflow boundaries of the modeling domain, especially in winter and spring. In addition, since the LBCs of BC2 are zero below 750 mbar by definition, grid cells close to these two boundaries tend to have higher DM8A BC1 and lower DM8A BC2 at the surface than those in the interior portions of the domain. In other words, the relative impact of LBCs and physical processes on inert tracers at the surface varies spatially, with larger impact of LBCs on the grid cells close to these two boundaries than the grid cells in the inner domain. Such features are also noticed in other models.A minus DM8A BC1/BCT from simulation B).

In short, the spatial distribution of the relative contributions of tracers clearly shows that the 3.1 WRF/CMAQ

The physical processes in the CTMs are involved in determining the inert tracers at the surface for the studied area. Hence, the model inter-comparison in inert tracers can be used to indicate similarities and differences between models regarding these physical processes.

3.2 Multi-model comparison in the impact of physical processes on inert tracers

A variety of physical processes are involved in determining the mixing ratio of inert tracers at the surface, including 3-D advection, turbulent mixing, of sub-grid cloud mixing, and wet scavenging, and dry deposition. The model differences and similarities in BCT for each model pair shown in the previous section will be explained and discussed in context of these physical processes by comparing model predictions in BC1, BC2 and their relative contributions to BCT.

Formatted: Font color: Red

3.2.1 WRF/CMAQ

The results from WRF/CMAQ are first examined, since WRF/CMAQ is used as the base case for the model inter-comparison. The most are important feature for BCT in WRF/CMAQ is that for all seasons, higher DM8A BCT is seen in the mountain west region than the eastern U.S. (Figure 3, first row) processes that the inert tracers undergo, and may be treated differently by CTMs due to the higher BC2 in the west (Figure 4, third row). This spatial contrast in inert tracers is not surprising because the higher elevation in the mountain region tends to generate more mixing between the mid-troposphere and the PBL, resulting in higher BC2 at surface in the mountain west U.S.

It is also expected that stronger mixing between the mid-troposphere and the PBL in the mountain region helps to dilute BC1 at surface. However, compared with the eastern U.S., lower levels in BC1 is not found in the mountain west region except for winter (Figure 4, second row). One possible explanation is that the impact from vertical mixing and the impact from LBCs cancel out in the mountain west region. Since we are examining surface concentrations and since the western boundary is predominantly an inflow boundary, it should be expected that BC1 in the west is higher than that in the east due to the impact of LBCs alone. Another possible explanation is that physical processes other than vertical mixing also playdifferences in parameterization methods, the meteorological inputs (Table 1), and/or the discrete grid structures. With a role in determining tracers at surface, such as dry deposition, as a clear spatial contrast in the ozone dry deposition velocity occurs between the western and the eastern U.S. (Figure 5, first row), especially in spring and summer. This is further illustrated through the series of sensitivity simulationsimulations for WRF/CMAQ_NODDEP. In, how the absence of dry deposition,LB ozone reaching the surface level mixing ratios of BC1 and BC2 increase across the domain, especially during spring and summer (Figure 4), but the increase is not uniform across the domain. Without the impact of dry deposition, lower BC1 is seen in the mountain west region than the eastern U.S. for all seasons.U.S. is modified by these processes is investigated in this model.

The results from WRF/CMAQ and WRF/CMAQ_NODDEP clearly show that vertical mixing and dry deposition interact to determine the spatial contrast in BCT at the surface. Due to the higher elevation in the mountain western U.S., BC1 in the west tends to be lower than the east through vertical transport. Meanwhile, more BC1 is removed from the atmosphere in the eastern U.S. than the mountain west region due to dry deposition, except for winter, when the impact of dry deposition is almost uniform spatially. As a result, BC1 levels at surface is relatively uniform across the interior of the U.S. compared with BC2. For BC2, on the contrary, the difference between the mountain western and the eastern U.S. is magnified when the two processes come into play together. In short, the differences in vertical transport (due to elevation difference) and dry deposition between the eastern and the mountain western U.S. mainly contribute to the spatial contrast of inert tracers at surface.

Other physical processes may also play a role in determining the mixing ratio of inert tracers regionally during certain season. For example, DM8A BC2 in eastern U.S. from WRF/CMAQ_NODDEP (Figure 4) is comparable to that in the west during summer, indicating that other physical process(es) may exist to enhance the mixing of BC2 from mid-troposphere downward into PBL in the eastern U.S., such as sub-grid cloud mixing. However, with the current data available, it is not possible to further diagnose the cause of the differences at the process level.

Last, results of (DM8A BC1/DM8A BCT) clearly show that throughout the interior U.S., BC2 plays a dominant role in BCT. Additionally, although dry deposition significantly changes the absolute mixing ratios of the simulated tracers, it changes little in the relative contributions of the tracers (Figure 4), since the dry deposition flux is in general proportional to the absolute mixing ratio of the species.

3.2.For DM8A BCT, it is not surprising to find that dry deposition significantly reduces DM8A BCT for all seasons by as much as about 10 ppb averaged over the U.S. (Table 3). Sub-grid cloud mixing in general slightly increases DM8A BCT (Table 3), because the sub-grid cloud mixing in CMAQ tends to mix the air aloft (e.g. above PBL), which is richer in BCT (especially BC2), downward into the PBL. This is later confirmed by results in the relative contributions of tracers. The largest impact of sub-grid cloud mixing is found in summer, with increases in DM8A BCT over 2.5 ppb across the eastern U.S. and the Atlantic Ocean. In spring and fall, the impact is generally smaller but not negligible, as increases in DM8A BCT still exceed 1ppb regionally. In winter, the impact is less than 1ppb across the U.S. Lastly, for wet scavenging, little change in DM8A BCT (less than 0.1 ppb for domain average) is found, so that the impact of this process on the simulated inert tracers is negligible for WRF/CMAQ and not shown. In addition, the impact of the three processes are relatively uniform across the U.S., as small deviations are found (summarized in Table 3).

For the relative contributions of inert tracers, only the differences in DM8A BC1/BCT between WRF/CMAQ and its sensitivity simulations are shown (Figure 3) to illustrate the changes in the relative contributions of inert tracers at the surface, as the changes in DM8A BC1/BCT and in DM8A BC2/BCT are usually the same in magnitude with opposite sign. The changes in DM8A BC3/BCT are less than 0.5% domain-wide in all seasons, except for the differences between WRF/CMAQ and WRF/CMAQ_nocldmix in summer, which will be discussed later. The impact of dry deposition is usually within ±5% and the direction of the change varies with space and time. At the surface, the removal of inert tracers is proportional to the absolute mixing ratios of each tracer, which in turn will be updated through vertical mixing in the PBL. In other words, the process of dry deposition does not modify the relative contributions of inert tracers directly, but through vertical turbulent mixing. Therefore, in regions where the vertical gradient of the tracer is steeper within the PBL, a larger impact of dry deposition on the DM8A BC1/BCT is expected. To confirm this hypothesis, the seasonal averaged vertical profiles of BC1/(BC1+BC2) and the maximum daytime PBL height in WRF/CMAQ are examined in each sub-region at 2pm (local standard time) (Figure 4). For example, in WB and NE sub-regions, the change in BC1/(BC1+BC2) from the surface to the top of PBL is larger in summer than winter, which is consistent with the larger differences seen in DM8A BC1/BCT between WRF/CMAQ_noddry and WRF/CMAQ during summer than winter (Figure 3, first row). In contrast, the change in BC1/(BC1+BC2) from the surface to the top of PBL is larger in winter than summer in SE and ATL, which is consistent with the larger impact of dry deposition over these two regions in winter. In addition to the vertical gradient of BC1/(BC1+BC2), the magnitude and direction for the change in BC1/(BC1+BC2) at the surface also depends on the amount of air exchanged between the surface and the aloft. Therefore, the impact of dry deposition on BC1/(BC1+BC2) varies in season and space.

The sensitivity towards sub-grid cloud mixing shows that it always tends to decrease DM8A (BC1/BCT) and increase DM8A (BC2/BCT) (Figure 3, second row), leading to slightly higher DM8A BCT in WRF/CMAQ than WRF/CMAQ_nocldmix

(Table 3). The largest impact is found in summer, when convection is most active and frequent, especially over the gulf coast area and the Atlantic Ocean with change in DM8A BC1/BCT of about 10% and in DM8A BCT of about 2.5 ~ 5 ppb. In other seasons, sub-grid cloud mixing mainly affects the western coast area and the oceans with its impact on the other areas across the U.S. usually less than 1.0 ppb in DM8A BCT and less than 5% in DM8A BC1/BCT. For wet scavenging, it is found that its impact on the relative contributions of tracers is also negligible, with the differences in DM8A BC1/BCT less than 0.1% domain wide in all seasons (not shown). In addition to these three processes, vertical grid structure is also an important model configuration in CTMs as it affects the vertical transport of inert tracers and the attribution of LB ozone to inert tracers. Comparing WRF/CMAQ with WRF/CMAQ_27aL shows that the coarser vertical structure in the free troposphere only slightly increases DM8A BCT (Table 3) but significantly modifies the relative contributions of BC1 and BC2 at surface (Figure 3, third row).

As mentioned, the impact of a given physical process in CTMs on LB ozone at the surface may not be isolated from other physical processes. In this study, vertical turbulent mixing could also be involved in determining the differences in inert tracers at surface between WRF/CMAQ and its sensitivity simulations discussed above. To investigate the impact of vertical turbulent mixing in conjunction with the physical treatment discussed above, the diurnal cycles of the differences in BC1/BCT and BC2/BCT at the surface are examined between WRF/CMAQ and its sensitivity simulations. The results for summer when the largest diurnal variance usually occurs, are shown in Figure 5. The diurnal change in the differences in BC1/BCT is generally much smaller than the diurnally averaged differences in BC1/BCT between WRF/CMAQ and WRF/CMAQ_nocldmix, and between WRF/CMAQ and WRF/CMAQ_27aL over all sub-regions (Table 5), suggesting that the impact of sub-grid cloud mixing and vertical resolution on the relative contributions of inert tracers at surface is in general much stronger than the impact of diurnal variability in vertical mixing. In contrast, the magnitude of the diurnal variance in the difference in BC1/BCT exceeds the diurnally averaged difference in BC1/BCT in the sub-regions of WB and NB between WRF/CMAQ and WRF/CMAQ_noddry (Table 5), suggesting the stronger dependence of the dry deposition process on the vertical turbulent mixing in determining the inert tracers at the surface. Similar results are found for BC2/BCT (not shown). There is no obvious pattern in the diurnal variance in the differences in the relative contributions of inert tracers, except for the model pair of WRF/CMAQ and WRF/CMAQ_nocldmix. In summer, for example, their differences in BC2/BCT and BC1/BCT (WRF/CMAQ_nocldmix minus WRF/CMAQ) always decrease during daytime (Figure 5). This is because sub-grid cloud mixing becomes less effective in reducing the vertical gradient of inert tracers in daytime due to the stronger turbulent mixing in daytime than in nighttime.

Lastly, the sum of the differences in (BC1/BCT) and (BC2/BCT) at the surface is approximately zero between WRF/CMAQ and WRF/CMAQ_noddry, and between WRF/CMAQ and WRF/CMAQ_27aL over all sub-regions in all seasons. For WRF/CMAQ_nocldmix and WRF/CMAQ, especially in summer, the differences in BC1/BCT and BC2/BCT do not add up to zero in the MT, GP, NE and SE sub-regions (Figure 5) due to the negative differences in BC3/BCT. This result suggests that sub-grid cloud mixing also transport more BC3 downward through deep convection at high altitude. For example, the vertical profiles of (BC2+BC3) from the two simulations clearly show that the mixing ratio of (BC2+BC3) in WRF/CMAQ is higher than that in WRF/CMAQ_nocldmix from the altitude ~ 3-4 km in MT, GP and NE regions, and from ~ 5-6 km in SE (Figure 6a). In WB and NB, however, the mixing ratio of (BC2+BC3) in WRF/CMAQ does not exceed that in WRF/CMAQ_nocldmix until about 2km (Figure 6a), so that the sub-grid cloud mixing has little impact on the vertical transport of BC3 and the differences in BC1/BCT and BC2/BCT almost add up to zero (Figure 5).

3.2 WRF/CAMx vs. WRF/CMAQ

Formatted: Heading 2

This model pair has some important features in common for air quality modeling, which the other model pairs do not have. The two models used the same meteorological inputs for CTMs and were configured with the same horizontal resolution. However, it is interesting to find that the differences in BCT at the surface for this model pair are still comparable to and sometimes even larger than those in the other model pairs (Figure 3, so that there should be little differences in 3D advection. Meanwhile, the two models use different representations for the other important physical processes that the inert tracers undergo, including vertical turbulent mixing, dry and wet deposition, and sub-grid cloud mixing (Table 1).

Comparing WRF/CAMx with WRF/CMAQ, for all seasons, WRF/CAMx shows consistently lower DM8A BC1 and higher DM8A BC2 across the U.S., leading to lower (DM8A BC1/DM8A BCT) for all seasons (Figure 6) and higher DM8A BCT except for fall (Figure 3). In fall, the differences in BC1 and BC2 are comparable in magnitude with the differences in other seasons, but tend to cancel out each other, leading to smaller difference in BCT. The results suggest that a systematic difference exist between the two models, and such systematic difference overwhelms other differences in the model configurations and the representations of physical processes, leading to consistent differences in the simulated tracers across season and space. The two simulations were driven by the same meteorological inputs and utilized the same horizontal resolution, so that the horizontal transport and vertical advection should not in principle be the dominant contributor to these modeled differences in inert tracers.

As to the vertical turbulent mixing, WRF/CMAQ used the parameterization ACM2 (Pleim 2007) for vertical diffusion, while WRF/CAMx used "K-theory". Compared with ACM2, K-theory may be less efficient in the mixing of the convective boundary layer during the deep vertical convection (ENVIRON, 2015). During the neutral and stable conditions, both parameterizations are able to adequately characterize the vertical mixing (ENVIRON, 2015). Therefore, if the parameterizations for vertical turbulent mixing dominate the differences in simulated inert tracers, larger differences are expected to exist during summer, when deep convection is most frequent, rather than the consistent differences across season.

Another important systematic difference for this model pair is the vertical grid resolution, which may result in significant difference in the vertical diffusion between the two models. Though the two models have similar vertical structure from surface to about 900 mbar (~ 1km), WRF/CAMx used coarser vertical resolution from 900 to 300 mbar (Figure 2), which may lead to higher vertical mixing between the PBL and the free troposphere and eventually tend to dilute the BC1 levels at surface and entrain greater amounts of BC2 downward to the surface.

In addition, sub-grid cloud mixing is also different in the two models (Table 1). In general, clouds enhance the vertical mixing between PBL and free troposphere, thus decreasing BC1 and increasing BC2 within the PBL. Lacking the representation of this process, WRF/CAMx would be expected to have higher BC1 and lower BC2 than WRF/CMAQ if all other physical processes were similar. However, the results show the opposite, indicating that the processes other than sub-grid cloud mixing dictate differences in tracers, though locally the impact of this process can be as important as the impact of vertical grid resolution. As an example, along the gulf coast, where convection is usually frequent and strong (especially during spring and summer), the differences in BC1 and BC2 between the two models (within ±2.5 ppb) are smaller than the inner regions (Figure 6). A possible explanation is that the impact of sub-grid cloud mixing and the impact of vertical turbulent mixing cancel out. Though WRF/CAMx tends to mix more BC2 downward into the PBL and mix more BC1 into the free troposphere than WRF/CMAQ due to its coarser vertical grid resolution in the free troposphere, the differences in tracers between the two models decrease due to the process of sub-grid cloud mixing in WRF/CMAQ along the gulf coast.

As to the physical sinks for inert tracers, WRF/CAMx did not include wet deposition of the inert tracers, the concentrations of tracers are expected to be higher than those in WRF/CMAQ. However, the opposite was found for BC1. Furthermore, the impact of wet deposition on pollutants usually shows a strong seasonality, with peak in summer or spring, but no seasonal variation was found in the differences between the two models. Hence, wet deposition is unlikely to dominate the systematic model differences in inert tracers either. Last, the impact of dry deposition is investigated. Significant difference is found regarding the dry deposition velocity of ozone for this model pair (Figure 5, second row), resulting from the differences in the dry deposition scheme (Table 1) and the land use data. However, the difference in dry deposition velocity is not uniform spatially and temporally, which is opposite to the uniform spatial and temporal distributions in the differences in inert tracers. Furthermore, the differences in BC1 and BC2 are opposite in sign (Figure 6), and thus, it is impossible that the difference in dry deposition between the two models dominates the differences in BC1 and BC2 at the same time. Therefore, we believe that dry deposition process does not dominate the systematic model differences in inert tracers. In addition, as shown in the comparison between WRF/CMAQ and WRF/CMAQ_NODDEP (Figure 4), dry deposition changes little on the relative contributions of inert tracers. Therefore, the significant difference in (DM8A BC1/DM8A BCT) for this model pair does not result from the process of dry deposition.

To summarize, the model difference in vertical mixing due to the different vertical grid resolutions likely overwhelms other differences in the model configurations and the representations of physical processes throughout all seasons across the entire U.S. domain, though other process(es), such as sub-grid cloud mixing, can also be important locally during specific season.

3.2.3 WRF/DEHM vs WRF/CMAQ

Comparing WRF/DEHM with WRF/CMAQ, the difference in BCT (Figure 3) is not dominated by the difference in BC2 alone, as found in WRF/CAMx and WRF/CMAQ. Instead, whether BC1 or BC2 dominates the differences in BCT depends on season and location (Figure 7). In addition, considerable spatial variability is also found in the difference in BC1, BC2 and (DM8A BC1/DM8A BCT) (Figure 7). For example, large positive and negative differences (more than 15%) in (DM8A BC1/DM8A BCT) coexist within the U.S. except for winter, and this feature is not seen in the other model pairs. One possible explanation is that the differences in the modeled inert tracers for this model pair were determined by multiple processes at the same time. It is also possible that certain model representation of physical processes or model configuration mainly contributes to the differences in inert tracers, however, the impact of this factor on inert tracers changes spatially and temporally. Our analysis shows that the latter is likely to be the case.

The impact of dry deposition on the differences in simulated inert tracers is first investigated using WRF/CMAQ_NODDEP and WRF/DEHM_NODDEP sensitivity simulations. Without dry deposition, the model difference in DM8A BCT significantly decreases (Figure 3, last row) across all seasons, especially for spring and summer. This is because the differences in BC1 and BC2 are usually similar in magnitude but opposite in sign across most of the U.S. (Figure 7), leading to smaller differences in BCT. And by comparing the (DM8A BC1/DM8A BCT) from the model pairs with and without dry deposition (Figure 7), it is reconfirmed that the process of dry deposition has little impact on the relative contributions of inert tracers at surface. Then, the model pair of WRF/CMAQ_NODDEP and WRF/DEHM_NODDEP is used to investigate the impact of other physical processes on the differences in inert tracers. The impact of wet deposition on the differences in inert tracers for this model pair is believed to be negligible for two reasons. First, the differences in BC1 and BC2 between WRF/CMAQ_NODDEP and WRF/DEHM_NODDEP are found to be in opposite sign across the majority of U.S. Hence, this physical sink is unlikely to contribute and dominate the differences in BC1 and BC2 simultaneously. Secondly, the model difference in precipitation is in general small, except that significant higher precipitation is found in the southeastern U.S. in WRF/DEHM during summer (not shown). Though BC2 in WRF/DEHM_NODDEP is lower than that in WRF/CMAQ_NODDEP for this region in summer, which could result from wet deposition, the higher BC1 in WRF/DEHM_NODDEP is not consistent with its higher precipitation.

Given that the physical sinks are likely to play a minor role in the differences in the simulated inert tracers between WRF/CMAQ_NODDEP and WRF/DEHM_NODDEP, the representation of 3-D transport (advection, turbulent mixing and cloud mixing) should primarily contribute to the differences in two ways, namely (1) the meteorological inputs of CTMs to drive 3D advection and (2) the parameterizations for turbulent mixing and cloud mixing in CTMs.

WRF/DEHM was run at a much coarser horizontal resolution than WRF/CMAQ (Table1) and the meteorological inputs for the two models were generated from different global climate models, so that considerable differences in the meteorological inputs for CTMs could exist between the two models. The model differences in horizontal transport has been suggested by a previous study in AQMEII3 (Solazzo et al.,2017), in which the surface wind is compared over southeastern, northeastern U.S. and the region of California with noticeable differences found for all seasons. However, the feature of the opposite signs for the differences in BC1 and BC2 in most regions of the U.S. between WRF/CMAQ_NODDEP and WRF/DEHM_NODDEP (Figure 7) suggests that the model differences in horizontal transport, especially within PBL, should not be the dominant reason behind the model difference in inert tracers. A possible argument for the opposite signs in the differences in BC1 and BC2 is that the model differences in horizontal transport change with altitude range. For example, for spring, the lower BC1 in the central U.S. (Figure 7, fifth row) could result from the horizontal transport within the PBL; while the higher BC2 in this region (Figure 7, sixth row) could be initiated by higher BC2 in the free troposphere due to the opposite difference in horizontal transport in the free troposphere. Though it is not possible for this study to further investigate the model differences aloft due to the limitation of the availability of surface data only, we do not think the model differences in the horizontal transport of inert tracers in the free troposphere could be the main reason for the differences in the modeled inert tracers at the surface. This is because the mixing ratios of BC2 in the free troposphere should be in general uniform across the modeling domain, especially on a seasonal basis.

Therefore, we believe that the differences in vertical transport between the two models are very likely to be mainly responsible for the differences in inert tracers at the surface. With limited meteorological data at the surface, the difference in PBL height during daytime between the two models is used to indicate the potential model differences in vertical turbulent mixing. It should be noted that the definition of PBL height depends on the parameterizations of PBL (as summarized by Banks et al., 2016). The meteorological fields of the two models were generated using different PBL schemes. For WRF/DEHM, MYJ scheme (Mellor and Yamada, 1974, 1982; Janjić, 2002) was used, in which PBL height is defined by a prescribed threshold of turbulent kinetic energy. For WRF/CMAQ, ACM2 scheme (Pleim, 2007) was used, in which PBL height is determined by the Richardson number calculated above neutral buoyancy level. Therefore, the difference in PBL height between the two models can partially stem from the different definitions of PBL height, instead of resulting from different vertical turbulent mixing alone. However, compared with other meteorological data available, such as wind at 10 m height and temperature at 2 m height, PBL height is still a better indicator for vertical turbulent mixing.

It is found that the differences in inert traces in some regions are consistent with the differences in PBL. For example, in winter, WRF/DEHM shows higher PBL than WRF/CMAQ in the eastern U.S. (Figure 8, first row), which may explain the lower BC1 and higher BC2 at surface (Figure 7) in this region. In summer, WRF/DEHM shows much lower PBL than WRF/CMAQ in the southeastern U.S., which may explain the higher BC1 and lower BC2 in this region. In spring, however, the differences in inert tracers over this region do not match the difference in PBL height, especially in the area of Florida. The results suggest that the model differences in other physical process(es), such as sub-grid cloud mixing could also influence the vertical mixing of inert tracers, so that the model differences in vertical turbulent mixing fail to dominate the model difference in inert tracers over this region. Unfortunately, with the current data available, it is impossible to further separate the impact of these processes.

It is interesting to note that the model differences in vertical transport mainly contribute to the differences in the modeled inert tracers for the model pair of WRF/CMAQ_NODDEP-WRF/DEHM_NODDEP and the model pair of WRF/CMAQ-WRF/CAMx as well. However, consistent differences in inert tracers are only found in WRF/CMAQ-WRF/CAMx, but not in WRF/CMAQ_NODDEP-WRF/DEHM_NODDEP. This is because for WRF/CMAQ-WRF/CAMx, the coarser vertical resolution in the free troposphere in WRF/CAMx always tends to dilute the BC1 levels at surface and entrain greater amounts of BC2 downward to the surface, which is affected little by space and time. However, for WRF/CMAQ-WRF/DEHM, the mixing between PBL and the free troposphere in WRF/DEHM can be stronger or weaker than that in WRF/CMAQ depending on the meteorological conditions and the responses of the parameterization of turbulent mixing to the meteorological conditions. As a results, though the differences in BC1 and BC2 between WRF/CMAQ_NODDEP-WRF/DEHM_NODDEP are usually in opposite sign, as what is also found in the model pair of WRF/CMAQ-WRF/CAMx, the differences show large spatial and temporal variability.

3.2.4 COSMO-CLM/CMAQ vs WRF/CMAQ

Comparing COSMO-CLM/CMAQ with WRF/CMAQ, it is found that this model pair has smaller differences in DM8A BCT than the other model pairs. Furthermore, the two models also show similar BC1 and BC2, with differences less than 2.5 ppb for most of the locations. The model differences in the relative contributions of BC1 and BC2 (Figure 9) are also much smaller than other model pairs, except over the southeastern U.S. during summer.

Given that the two models are driven by the same global meteorology and nudging is also applied above the PBL, the model differences in horizontal transport and vertical advection are expected to be small, especially at the spatial and temporal scales considered in this study. For dry deposition, it has been shown that this process barely changes the relative contributions of inert tracers. For wet deposition, the model difference in precipitation is also small (not shown). For sub-grid cloud mixing, its impact, if any, is not expected to be important across the whole U.S. for all seasons. Therefore, the similar model performance for this model pair is mainly due to the similar representation of turbulent vertical mixing. Similar to the previous model pair, the PBL height in daytime is used to suggest the potential model differences in vertical turbulent mixing. For COSMO-CLM/CMAQ, an extended MYJ scheme was used (Doms et al., 2011). As mentioned before, the PBL height can only partially suggest the potential differences in vertical turbulent mixing, as PBL height is defined in different ways in different parameterizations.

It is interesting to find that COSMO-CLM/CMAQ has consistently higher PBL than WRF/CMAQ (Figure 8, second row), suggesting that more BC2 aloft is expected to be mixed downward into PBL and more BC1 to be diluted in PBL. However, the stronger vertical turbulent mixing in COSMO-CLM/CMAQ is not strong enough to overwhelm the other differences for this model pair, so that COSMO-CLM/CMAQ does not have consistent higher DM8A BC2 and lower DM8A BC1 than WRF/CMAQ (Figure 9). Such results are in opposite to the results seen for the model pair of WRF/CAMx and WRF/CMAQ, in which stronger vertical mixing in WRF/CAMx has a dominant impact on the model difference in inert tracers. The stronger vertical mixing in COSMO-CLM/CMAQ is mainly due to different meteorological fields, because COSMO-CLM/CMAQ and WRF/CMAQ have very similar vertical structure from surface to until 400 mbar (Figure 2). The stronger mixing in WRF/CAMx is due to its coarser vertical structure, especially from 900 to 300 mbar. The contrasting results of the differences in inert tracers in these two model pairs further suggest the important role that vertical grid resolution (both within the PBL and in the mid-troposphere) plays on the eventual simulated boundary tracer levels at the surface. In addition, though different PBL schemes were used by COSMO-CLM/CMAQ and WRF/CMAQ when the meteorological fields were generated, the same vertical turbulent mixing scheme was used for the CTMs, which may also contribute to the small model differences in the inert tracers.

During the summer over the southeastern U.S., the model differences in other processes may become important as well. COSMO-CLM/CMAQ simulates much less precipitation (sum of convective and non-convective rain) than WRF/CMAQ, which could lead to less scavenging and higher inert tracer levels. Meanwhile, this region typically experiences frequent and strong convective activity during summertime. Less precipitation in COSMO-CLM/CMAQ could indicate weaker convection and weaker sub-grid cloud mixing, leading to less BC2 brought downward from the mid-troposphere to PBL. The impact of less sub-grid cloud mixing may be stronger than the impact of less scavenging and finally result in lower BC2 in the COSMO-CLM/CMAQ simulations.

4 Summary and Discussions

In summary, the inert tracer method is a good balance of the needs for simplicity and the ability to provide diagnostic information for the model inter-comparison with respect to the physical processes that the tracers undergo. The four simulations show considerable differences in their estimates on the impact of lateral boundary ozone due to the model configurations and the representations of physical processes, though the magnitude of the differences varies from season to season, and from model to model. For all the processes that the inert tracers undergo, vertical turbulent mixing stands out as a primary contributor to the model differences in inert tracers, especially in the relative contributions of BC1 and BC2 to BCT. For the simulated mixing between PBL and the free troposphere, the differences in vertical grid structures, not only within the PBL but also in the free troposphere employed among models, can be as important as the differences in the representation of vertical turbulent mixing.

The results of the modelThe model differences in DM8A BCT are relatively small (within ±5 ppb) in spring and fall. In winter and summer, however, the differences can reach as much as 7.5~10 ppb regionally (Figure 7, first row). The results demonstrate that the physical treatment in CTMs serves as an important source of uncertainty when estimating the impact of LB ozone on ozone level at the surface aside from the meteorological inputs and the lateral boundary conditions. Differences in DM8A O_3 and DM8A BCT between WRF/CMAQ and WRF/CAMx show strong spatial correlations with similar magnitudes except for summer (Figure 7, second row). Two processes lead to the weaker agreement between the difference in inert tracers and the difference in ozone in summer. First, the chemical decay due to photolysis of LB ozone is the strongest in summer and not represented by inert tracers. Second, the chemical formation of ozone peaks in summer. The results suggest that the impact of physical treatment can compete or even overwhelm the impact of chemistry on the LB ozone reaching the surface in some cases. There are significant differences in the relative contributions of BC1 and BC2 at the surface in all seasons, which are usually much larger than the differences found in BC3/BCT (within in 2.5%) in all seasons. Hence, only the results in DM8A BC1/BCT are shown for this model pair (Figure 8, first row) to illustrate the model differences in the relative contributions of inert tracers at the surface.

The impact of vertical grid structure on the model differences is first examined by comparing (WRF/CAMx minus WRF/CMAQ) with (WRF/CAMx minus WRF/CMAQ_27aL). For DM8A BCT, about 10%, 60%, 20% and 40% of the difference between WRF/CAMx and WRF/CMAQ over land can be attributed to the difference in vertical resolution in winter, spring, summer and fall, respectively (Table 3). For the relative contributions of inert tracers, about 60% of the differences in DM8A BC1/BCT between WRF/CAMx and WRF/CMAQ over the land can be attributed to their differences in the vertical resolution in all seasons (Figure 8, second row; Table 4).

As to the impact of other physical processes on the model differences, while wet deposition of the inert tracers is not represented in WRF/CAMx, our analysis for WRF/CMAQ and its sensitivity simulations in the previous section has shown that the impact of wet deposition is negligible in WRF/CMAQ, and therefore the absence of this process in WRF/CAMx should not be a significant contributor to the model differences in inert tracers. The impact of sub-grid cloud mixing in WRF/CMAQ is usually pronounced over ocean and coastal regions with an averaged change of less than 1 ppb in DM8A BCT over land except during summer (Table 3). Sub-grid cloud mixing in WRF/CMAQ also always decreases (BC1/BCT) and increases (BC2/BCT) at the surface. Although WRF/CAMx does not represent this process, the DM8A BC1/BCT in WRF/CAMx is usually lower than that in WRF/CMAQ_27aL (Figure 8, second row) over land. Therefore, the remaining differences between WRF/CAMx and WRF/CMAQ_27aL do not result from wet scavenging or sub-grid cloud mixing. To investigate the impact of dry deposition on DM8A BCT, the seasonal averaged dry deposition velocity at surface is compared between the two models, and a correlation is seen between the differences in BCT and the differences in dry deposition velocity. The differences in BCT (WRF/CAMx minus WRF/CMAQ, Figure 7, first row) tends to increase when and where the dry deposition velocity in WRF/CAMx is smaller (Figure 7, third row). For example, in winter, the difference in DM8A BCT in the north is about 5 ppb higher than that in the south with lower dry deposition velocity found in WRF/CAMx over the northern part of the domain. Similar results also are seen in summer. In addition, the spatial distributions of the difference in BCT are more uniform in spring and fall than those in winter and summer, and correspondingly, the spatial distributions of the ratio in dry deposition velocity are also more uniform in spring and fall. Therefore, we believe that the large spatial gradient of the difference in DM8A BCT in winter and summer between the two models is primarily due to their differences in dry deposition.

However, dry deposition does not explain the higher DM8A BCT in WRF/CAMx when the dry deposition velocity in WRF/CAMx is also faster, such as in spring. In addition, the remaining model difference in the relative contributions of inert tracers at the surface (as shown in Figure 8, second row) cannot be explained by the difference in dry deposition alone, because as mentioned above, the dry deposition modifies the relative contributions of inert tracers at the surface through vertical turbulent mixing with relatively small changes in the relative contributions of inert tracers at the surface. Therefore, the remaining differences in inert tracers at the surface can only be explained by their difference in vertical turbulent mixing. WRF/CMAQ used the parameterization ACM2 (Pleim 2007), while WRF/CAMx used “K-theory” (Table 1). Under neutral and stable conditions, both parameterizations can adequately characterize the vertical mixing (Ramboll, 2018); while during periods of deep vertical convection, K-theory is less efficient in the mixing of the convective boundary layer (Ramboll, 2018). However, our results indicate that WRF/CAMx always tends to have stronger vertical turbulent mixing than WRF/CMAQ. On one hand, as shown in Figure 7, DM8A BCT in WRF/CAMx is higher than that in WRF/CMAQ even when dry deposition velocity in WRF/CAMx is faster, indicating that more air aloft (with richer BCT) is brought downward to compensate the loss of inert tracers. On the other hand, the DM8A BC1/BCT in WRF/CAMx is always lower than that in WRF/CMAQ over land (Figure 8, second row) with correspondingly higher DM8A BC2/BCT (not shown), suggesting again that WRF/CAMx mixes more air from aloft (with lower BC1/BCT and higher BC2/BCT) downward than WRF/CMAQ. In addition, the NB sub-region usually shows larger differences in DM8A BC1/BCT than other regions (Figure 8, second row). This is because the vertical gradients of (BC2+BC3) and BC1 in NB from surface to 3km are usually steeper than the gradients in other regions in WRF/CMAQ (e.g. in summer as shown in Figure 6a, 6b), so that the stronger vertical mixing in WRF/CAMx tends to have a larger impact on inert tracers at the surface in this region. The stronger vertical mixing in WRF/CAMx also compensates for the lack of sub-grid cloud mixing to a certain extent, leading to smaller differences in DM8A BCT and DM8A BC1/BCT especially over the SE and the Gulf coast region during summer.

To further illustrate the role of differences in vertical mixing between the two models, the diurnal cycles of the differences in BC1/BCT and BC2/BCT between WRF/CAMx and WRF/CMAQ_27aL are examined over the sub-regions. Little variance (less than 1%) is found over most of the sub-regions, except that clear diurnal change is noticed over WB and NB in most seasons. Over these two regions, the differences in BC2/BCT and BC1/BCT (Figure 9a, 9b) grow from night to daytime, as the vertical turbulent mixing becomes stronger.

To summarize, 10% to 60% of the seasonal averaged differences in inert tracers between WRF/CMAQ and WRF/CAMx at the surface can be attributed to their difference in the vertical grid structure in the free troposphere. The vertical turbulent mixing primary contributes to the remaining differences across the entire land in all seasons. Stronger vertical mixing in WRF/CAMx brings more BC2 downward, leading to higher DM8A BCT and BC2/BCT at the surface in WRF/CAMx. The differences in inert tracers due to vertical mixing is partially counteracted by their difference in sub-grid cloud mixing over the SE and the Gulf coast region during summer. The process of dry deposition adds extra gradients to the spatial distribution of the differences in DM8A BCT by about 5 ~ 10 ppb during winter and summer. Unfortunately, it is impossible to further quantitatively attribute the model differences in inert tracers to the processes of dry deposition and vertical turbulent mixing with the sensitivity simulations available in this study.

3.3 COSMO-CLM/CMAQ vs WRF/CMAQ

Unlike the model pair in the previous section, COSMO-CLM/CMAQ and WRF/CMAQ do not share the same meteorological inputs; however, the same physical parameterizations are used in CMAQ to represent the processes inert tracers undergo (Table 1). Furthermore, the two models have similar vertical resolution from the surface up to about 400 hPa (Figure 2), which covers the majority of the pressure range for BC1 and BC2. The differences in DM8A BCT and the relative contributions of inert tracers for this model pair are usually much smaller than the differences between WRF/CAMx and WRF/CMAQ. For example, the differences in DM8A BCT are within 2.5 ppb (Table 3; Figure 10, first row) across most of the U.S. The results indicate that the uncertainty stemming from the physical treatment in CTMs may rival or exceed the uncertainty from meteorological inputs, especially when nudging is applied to generate the meteorological fields (Table 1) with constrains above the PBL at synoptic scales. The largest differences at the surface occur in the summer over the SE and ATL sub-regions with lower DM8A BCT in COSMO-CLM/CMAQ by about 5 ppb and 10 ppb, respectively. This is because the large difference in BC2 is not offset by the difference in BC1 or BC3 as in other regions (Table 6). The physical process(es) that contributes to the large differences over the two areas will be discussed later.

The difference in the relative contributions of inert tracers at the surface is dominated by the difference in DM8A BC1/BCT and DM8A BC2/BCT over most regions in all seasons, except MT, GP and NE in summer, where the differences in DM8A BC2/BCT and in DM8A BC3/BCT dominate (Table 6). Therefore, the model differences in DM8A BC2/BCT are shown to demonstrate the model differences in the relative contributions of inert tracers (Figure 10, second row). The difference in DM8A BC2/BCT is in general small (within 5%) over most of the U.S., except during the summer. In summer, the large differences in DM8A BC2/BCT over MT and GP result from their differences in BC2 and BC3 at the surface (Table 6), due to their difference in vertical resolution above 400 hPa. Over SE and ATL, however, the large differences in BC2/BCT result from their difference in BC2 alone at the surface, suggesting the impact of other physical processes than vertical resolution in these two regions.

In SE and ATL during summer, COSMO-CLM/CMAQ shows much larger vertical gradients in BC1 and in (BC2+BC3) than WRF/CMAQ from 5km to 3km (Figure 6a, 6b). Of all the physical processes inert tracers undergo, sub-grid cloud mixing and horizontal transport may contribute to the large difference in vertical gradient at this altitude range. Separate analysis reveals that the precipitation over the southeastern U.S., which is mainly convective rain in summer given the horizontal resolution of the simulations, is smaller in COSMO-CLM/CMAQ than in WRF/CMAQ (not shown), suggesting weaker sub-grid cloud mixing in COSMO-CLM/CMAQ. By comparing the vertical profiles of WRF/CMAQ_nocldmix and COSMO-CLM/CMAQ, similar vertical gradients in (BC2+BC3) and BC1 are found between these two simulations in ATL. The results confirm that much less BC2 is mixed downward from 3 ~ 5km into the PBL in COSMO-CLM/CMAQ and the difference in BC2 cannot be compensated by the differences in BC1 and BC3, leading to the large negative differences in BCT at the surface between the two models. In SE, however, the vertical gradient in COSMO-CLM/CMAQ from 5km to 3km is still larger than that in WRF/CMAQ_nocldmix, which is likely due to the differences in horizontal advection between COSMO-CLM/CMAQ and WRF/CMAQ.

In general, the two models show similar vertical profiles of BC1 and (BC2+BC3) over the sub-regions through all seasons, which again suggests the similarity between the two models at the process level. The vertical profiles in summer are shown (Figure 6a, 6b) and discussed as the largest differences in inert tracers are found in summer both at the surface and aloft. Furthermore, the vertical profiles do suggest the potential compensation between different physical processes in this season over certain sub-regions, leading to small differences in DM8A BCT at the surface. In WB and NB, (BC2+BC3) in COSMO-CLM/CMAQ begins to exceed that in WRF/CMAQ from about 2km, which can be due to its stronger vertical turbulent mixing as suggested by PBL height (Figure 6a). However, BC1 in COSMO-CLM/CMAQ is lower than WRF/CMAQ at any altitude (Figure 6b), suggesting that the difference in BC1 at the surface (about 2.3 ppb) is not dominated by their difference in vertical turbulent mixing, but by their difference in dry deposition. Though COSMO-CLM/CMAQ also tends to remove more (BC2+BC3) at the surface by dry deposition as well, the (BC2+BC3) at the surface is compensated by mixing more air aloft downward through its stronger vertical turbulent mixing. One thing to point out is that different parameterizations are used to diagnose PBL height in their simulations for meteorology, with ACM2 in WRF and an extended MYJ scheme (Doms et al., 2011) in COSMO-CLM, so that the PBL height can be defined differently. However, as very large difference in PBL height is seen here between the two models, PBL height is a reasonable factor to suggest the potential difference in their vertical turbulent mixing. Similarly, in GP, the BC1 in COSMO-CLM/CMAQ is always higher than that in WRF/CMAQ (Figure 6a), suggesting weaker dry deposition. At the same time, the less efficient removal at the surface helps to decrease the difference in (BC2+BC3) at surface, since less (BC2+BC3) in COSMO-CLM/CMAQ is available aloft (Figure 6b) in this region. As a result, the difference in (BC2+BC3) decreases from 3km to surface. Conversely, in MT, (BC2+BC3) in COSMO-CLM/CMAQ is lower than WRF/CMAQ until about 2km; while BC1 shows the opposite. The results suggest that the differences in inert tracers over MT are dominated by the stronger vertical mixing in COSMO-CLM/CMAQ alone rather than by other physical processes. The MT region may also influence its neighbouring region through horizontal advection, leading to the slightly increase in the difference in (BC2+BC3) and in BC1 over GP from 7km to 3km.

To summarize, the two models show similar vertical profiles of BC1 and (BC2+BC3) over sub-regions across most seasons, which suggests the similarity between the two models at process level. The largest differences are noted during summer. The different attributions of LB ozone to BC2 and BC3 between the two models in general have small impact on DM8A BCT at the surface, with the largest difference in DM8A (BC2+BC3) about 2.0 ppb found in summer over MT, GP and NE. At the same time, the different attributions of BC2 and BC3 significantly change the relative contributions of BC2 and BC3 at the surface over the three regions. The results are similar to what is found between WRF/CMAQ and WRF/CMAQ_27aL where the attributions of LB ozone to BC1 and BC2 are significantly different. The model differences in sub-grid cloud mixing play a primary role in their large differences in DM8A BCT and DM8A BC2/BCT over ATL and SE in summer. Our analysis also suggests the model differences in vertical turbulent mixing over most of the domain and in dry deposition over certain sub-regions in summer. However, the impact of different dry deposition on inert tracers at the surface is almost offset by the model difference in vertical turbulent mixing on inert tracers.

3.4 WRF/DEHM vs. WRF/CMAQ

Given the different simulation domains between the two models, the results cannot be compared directly to investigate the impact of physical treatment on inert tracers. However, the model sensitivity of inert tracers at the surface to the process of dry deposition can be compared between WRF/CMAQ and WRF/DEHM using the sensitivity simulations denoted WRF/CMAQ_noddry and WRF/DEHM_noddry in Table 2. The impact of dry deposition on DM8A BCT in WRF/DEHM is about 50% higher than that in WRF/CMAQ except during winter (Table 3). Such large differences are not surprising given that neither the meteorological inputs nor the parameterizations are the same for the process of dry deposition between the two models (Table 1). However, both models show similar magnitudes in their changes in the relative contributions of inert tracers at the surface. For both WRF/CMAQ and WRF/DEHM, since little change is found in DM8A BC3/BCT (less than 0.5% across the entire U.S. in all seasons), only the results in DM8A BC1/BCT are shown to illustrate the model sensitivity of the relative contributions of inert tracers at surface to the process of dry deposition. For both models, the sensitivity of DM8A BC1/BCT is in general small (Table 4), with the change less than 5% in all seasons across the U.S. However, the spatial distributions of the change in DM8A BC1/BCT are very different between the two models. The change in DM8A BC1/BCT in WRF/CMAQ

(Figure 3, first row) shows much more spatial variance than that in WRF/DEHM (Figure 11), suggesting the differences in the vertical profiles of inert tracers and the differences in the process of turbulent mixing between the two models.

4 Summary and Discussion

This study investigated the impact of physical treatment in CTMs on lateral boundary (LB) ozone reaching the surface across the U.S. with the implementation of inert tracers for LB ozone. The differences in inert tracers at the surface between different models are attributed to model differences at the process level.

The analysis focused on inter-comparing three models with each other, namely WRF/CMAQ, WRF/CAMx and COSMO-CLM/CMAQ. WRF/CMAQ and WRF/CAMx share the same meteorological inputs, but the physical processes that the inert tracers undergo (other than 3D advection) are represented differently. On the other hand, the WRF/CMAQ and COSMO-CLM/CMAQ simulations are driven by different meteorological fields but share the same CTM. The model differences in DM8A BCT between WRF/CMAQ and COSMO-CLM/CMAQ is found to be usually much smaller than that between WRF/CMAQ and WRF/CAMx across the U.S. in all seasons. The results indicate that the uncertainty stemming from the physical treatment in CTMs may compete or exceed the uncertainty from meteorological inputs, especially when nudging is applied to constrain the synoptic-scale meteorology above the PBL. Furthermore, the model differences in inert tracers are investigated at the process level. Different vertical resolutions and discretization are used by the three models, leading to differences in the attributions of LB ozone to BC1, BC2 and BC3. The impact of vertical grid structure on DM8A BCT at the surface is usually small (within 1 ppb) across the U.S., but not negligible regionally with the seasonal averaged changes in DM8A BCT exceeding 1 ppb. At the same time, the vertical grid structure significantly modifies the relative contributions of inert tracers at the surface. These findings suggest a need for finer vertical resolution in both the free troposphere and the lower stratosphere to better represent the impact of intercontinental transport of ozone and the ozone intrusion on ozone levels at the surface. Dry deposition strongly affects the DM8A BCT at the surface in all seasons. However, its impact on the relative contributions of inert tracers is usually small even when the process is represented by different parameterization and driven by different meteorology. Sub-grid cloud mixing is found to be important in the western coastal U.S. during winter, spring and fall. In summer, its impact extends to the majority of U.S. with significant impact on both DM8A BCT and the relative contributions of inert tracers at the surface. Wet scavenging is found to have little impact on the inert tracers at the surface.

Our analysis also indicates that there are significant differences in vertical turbulent mixing among the three models. Both WRF/CAMx and COSMO-CLM/CMAQ are very likely to have stronger vertical mixing than WRF/CMAQ, with the same meteorology driving the turbulent mixing but represented by different parameterization in WRF/CAMx and WRF/CMAQ, and with different meteorology driving the turbulent vertical mixing but represented by the same parametrization method in COSMO-CLM/CMAQ and WRF/CMAQ. As to the relative contributions of inert tracers at the surface, in winter, spring and fall, when the impact of other processes (especially the sub-grid cloud mixing) on the relative contributions of inert tracers is weak, the differences in DM8A BC1/BCT and DM8A BC2/BCT between COSMO-CLM/CMAQ and WRF/CMAQ are within

5% across the majority of the U.S. regions and 5 ~ 10% in the remaining regions; while the differences between WRF/CAMx and WRF/CMAQ_27aL are larger than 5% in the majority of the U.S. regions. The results indicate that the differences in vertical turbulent mixing between WRF/CAMx and WRF/CMAQ_27aL could also be greater than the differences between COSMO-CLM/CMAQ and WRF/CMAQ. As to the DM8A BCT at the surface, the process of dry deposition often interacts with the vertical mixing in determining the inert tracers at the surface. For example, in summer, for COSMO-CLM/CMAQ and WRF/CMAQ, the impact of different dry deposition on inert tracers at the surface is almost compensated by the opposite impact of the model difference in vertical turbulent mixing on inert tracers over the sub-region of WB and NB; while for WRF/CAMx and WRF/CMAQ, larger difference in DM8A BCT is noted in regions where dry deposition in WRF/CAMx is weaker so that the difference in the simulated BCT due to vertical mixing is further enlarged.

The inter-comparison in inert tracers also suggest that it should be essential to understandsimulated by different models also suggests that when similar estimates on the impact of lateral boundary ozone are found between different simulations, the results do not necessarily meanimply that the agreement has been reached for the same reason, unless a careful comparison is performed at the process level to rule out the possibility of cancelling process contributions. To carry out such analysis, process analysis (PA) (Jeffries and Tonnesen, 1994) is desired for all simulations involved. Unfortunately, the PA tool is either not available, or it was not turned on duringinvoked in the simulation, assince PA iswas not a standard design protocol for the AQMEII3 participants.. We recommend the future model inter-comparison studies to include the PA tool as a standard design. The relatively higher contributions ofprotocol to enable consistent process-level comparison. Additionally, given the important role that turbulent mixing and sub-grid cloud mixing can play in determining the BC2 tracer to BCT also suggest further examination of inert tracers at the surface, aloft data would be extremely valuable in understanding the model concentrations and difference/similarity in these processes in the free troposphere and how those differences across models influence the surface predictions. Unfortunately, only minimal aloft data is available from the models participating in AQMEII3, precluding such an analysis here. Future. Therefore, future model inter-comparison studies should consider more detailed and standard archiving of 3D model information, to enable a more complete diagnostic assessment of role of differing 3D transport representation on differences in surface level predictions across these modeling systems.

Though the four simulations have shown significant differences in the physical processes that the inert tracers undergo due to the model configurations and representations of physical processes, it could be speculated that if chemical degradation of O_3 imported into the domain was also considered, the differences in physical processes would become less important.

We think that whether the differences in physical processes would still be important for the impact of LBC ozone compared with chemical loss is case dependent. For example, it is found that the differences in simulated DM8A O_3 and DM8A BCT between WRF/CMAQ and WRF/CAMx (Figure 10) show strong spatial correlations and similar magnitudes (especially during winter and spring), suggesting that differences in representation of the 3D transport of LBC O_3 influence the noted differences in surface O_3 predictions between the two models on a seasonal average basis. During summer however the agreement is weaker and this is likely associated with the lack of representation of LBC O_3 chemical decay due to photolysis. During winter, spring and fall, the modeled differences in BCT are dominated by the difference in BC2. However, during summer, lateral boundary ozone from the mid-troposphere could decrease as the photolysis of ozone is most active in summer. In addition, the chemical formation of ozone also peaks in summer. Therefore, different from the other seasons, the impact of lateral boundary ozone in summer due to physical processes become less important than chemical processes, so that the model difference in O_3 is not consistent with the difference in BCT anymore. Similar findings have also been reported by Baker et al. (2015), in which the chemically reactive and inert tracers were compared by season.

LBC O_3 could also be lost due to titration by NOx especially within the boundary layer. The contrast in DM8A O_3 between rural and urban sites is compared amongst models to indicate the model similarity/difference in chemical environment at surface. WRF/CMAQ and WRF/CAMx is found to have very similar contrast between rural and urban sites (Figure 11), suggesting that the impact of lateral boundary ozone depletion at surface could be similar between the two models as well.

Acknowledgements and Disclaimer

We gratefully acknowledge the Air Quality Model Evaluation International Initiative (AQMEII) for facilitating the analysis described in the manuscript by designing and coordinating internally consistent regional-scale air quality model simulations. During the conduct of this work, P. Liu held National Research Council post-doctoral fellowships. The views expressed in this article are those of the authors and do not necessarily represent the views or policies of the U.S. Environmental Protection Agency. Aarhus University gratefully acknowledges the 604 NordicWelfAir project funded by the NordForsk's Nordic Programme on Health and Welfare 605 (grant agreement no. 75007), the REEEM project funded by the H2020-LCE Research and 606 Innovation Action (grant agreement no.: 691739), and the Danish Centre for Environment 607 and Energy (AU-DCE). The simulation of WRF/CAMx was supported by the Coordinating Research Council Atmospheric Impacts Committee.

References

Baker, K. R., Emery, *C.*, Dolwick, P., and Yarwood, G.: Photochemical grid model estimates of lateral boundary contributions to ozone and particulate matter across the continental United States, Atmos. Environ., 123, 49–62, 2015.

Banks, R. F., Tiana-Alsina, J., Baldasano, J. M., Rocadenbosch, F., Papayannis, A., Solomos, S., and Tzanis, C. G.: Sensitivity of boundary-layer variables to PBL schemes in the WRF model based on surface meteorological observations, lidar, and radiosondes during the HygrA-CD campaign, Atmospheric Research, 176–177, 185-201, 2016.

Brandt, J., Silver, J. D., Frohn, L. M., Geels, C., Gross, A., Hansen, A. B., Hansen, K. M., Hedegaard, G.B., Skjøth, C. A., Villadsen, H., Zare, A., and Christensen, J. H.: An integrated model study for Europe and North America using the H.: The Danish Eulerian Hemispheric Model with focus on intercontinental transport, Atmos. Environ., 53, 156-176, 2012.

Campbell, P., Zhang, Y., Yahya, K., Wang, K., Hogrefe, C., Pouliot, G., Knote, C., Hodzic, A., San Jose, R., Pérez, J. L., Jiménez-Guerrero, P., Baró, R., and Makar, P.: A multi-model assessment for the 2006 and 2010 simulations under the Air

Formatted: English (United States)

Formatted: English (United States)

Quality Model Evaluation International Initiative (AQMEII) phase 2 over North America: Part I. Indicators of the sensitivity of O_3 and $PM_{2.5}$ formation regimes, Atmos. Environ., 115, 569-586, 2015.

Cohan, D., Hu, Y., Hakami, A., and Russell, A. G.: Nonlinear response of ozone to emissions: source apportionment and sensitivity analysis, Environ. Sci. Technol., 39, 6739-6748, 2005 — a three-dimensional air pollution model used for the Arctic. Atmospheric Environment, 31, 4169-4191, 1997.

Cooper, O.R., Gao, R. S., Tarasick, D., Leblanc, T., Sweeney, C.: Long-term ozone trends at rural ozone monitoring sites across the United States, 1990-2010. J. Geophys. Res. Atmos. 1984-2012, 117, 2012.

Delle Monache, L., and Stull, R.: An ensemble air quality forecast over western Europe during an ozone episode, *Atmos. Environ.*, 37, 3469–3474, 2003.

Dennis, R., Fox, T., Fuentes, M., Gilliland, A., Hanna, S., Hogrefe, C., et al.: A framework for evaluating regional-scale numerical photochemical modeling systems, Environ. fluid Mech., 10, 471-489, 10.1007/s10652-009-9163-2, 2010.

Dolwick, P., Akhtar, F., Baker, K. R., Possiel, N., Simon, H., and Tonnesen, G.: Comparison of background ozone estimates over the western United States based on two separate model methodologies, Atmos. Environ., 109, 282-296, 2015.

Doms, G., Főerstner, J., Heise, E., Herzog, H.-J., Mironov, D., Raschendorfer, M., Reinhardt, T., Ritter, B., Schrodin, R., Schulz, J.-P., and Vogel, G.: A Description of the Nonhydrostatic Regional COSMO Model Part II: Physical Parameterization. German weather service, Offenbach, Germany, 2011. Available online at http://www.cosmo-model.org/content/model/documentation/core/cosmoPhysParamtr.pdf

Dunker, A. M.: The Decoupled Direct Method for Calculating Sensitivity Coefficients in Chemical Kinetics, J. Chem. Phys., 81,2385–2393, 1984.

Dunker, A. M., Yarwood, G., Ortmann, J. P., and Wilson, G. M.: The Decoupled Direct Method for Sensitivity Analysis in a Three-Dimensional Air Quality Model Implementation, Accuracy, and Efficiency, Environ. Sci. Technol., 36, 2965-2976, doi: 10.1021/es0112691, 2002.

Emmons, L. K., Hess, P. G., Lamarque, J.-F., and Pfister, G. G.: Tagged ozone mechanism for MOZART-4, CAM-chem and other chemical transport models, Geosci. Model Dev., 5, 1531-1542, doi:10.5194/gmd-5-1531-2012, 2012.

ENVIRON: The Comprehensive Air quality Model with extensions (CAMx), version 6.20. (Available from http://www.camx.com), 2015.

Dunker, A.M., Koo, B. and Yarwood, G.: Contributions of foreign, domestic and natural emissions to US ozone estimated using the path-integral method in CAMx nested within GEOS-Chem. Atmospheric Chemistry and Physics, 17(20), p.12553, 2017.

Eastham, S. D. and Jacob, D. J.: Limits on the ability of global Eulerian models to resolve intercontinental transport of chemical plumes, Atmos. Chem. Phys., 17, 2543-2553, https://doi.org/10.5194/acp-17-2543-2017, 2017.

Federal Register: U.S. Environmental Protection Agency, National Ambient Air Quality Standards for Ozone – Final Rule, Federal Register 80, available at: https://www.gpo.gov/fdsys/pkg/FR-2015-10-26/pdf/2015-26594.pdf, 2015.

Formatted: English (United States)

Formatted: Font: Not Italic

Flemming, J., Huijnen, V., Arteta, J., Bechtold, P., Beljaars, A., Blechschmidt, A.-M., Diamantakis, M., Engelen, R. J., Gaudel, A., Inness, A., Jones, L., Josse, B., Katragkou, E., Marecal, V., Peuch, V.-H., Richter, A., Schultz, M. G., Stein, O., and Tsikerdekis, A.: Tropospheric chemistry in the Integrated Forecasting System of ECMWF, Geosci. Model Dev., 8, 975–1003, doi:10.5194/gmd-8-975-2015, 2015.

Fox, D. G.: Uncertainty in air quality modeling, Bull. Amer. Meteor. Soc., 65, 27-36, 1984.

Galmarini, S., Bianconi, R., Bellasio, R., and Graziani, G.: Forecasting the consequences of accidental releases of radionuclides in the atmosphere from ensemble dispersion modelling, J. Environ. Radioact., 57,203-19, 2001.

Galmarini, S., Rao, S. T., and Steyn, D. G.: Preface for the Special Issue -- AQMEII: An International Initiative for the Evaluation of Regional-Scale Air Quality Models – Phase 1, Atmos. Environ., 53, 1–3, 2012.

Galmarini, S., Kioutsioukis, I., and Solazzo, E.: E pluribus unum*: ensemble air quality predictions, Atmos. Chem. Phys., 13, 7153-7182, 2013.

Galmarini, S., Hogrefe, C., Brunner, D., Baklanov, A., and Makar, P.: Preface Article for the Atmospheric Environment Special Issue on AQMEII Phase 2, Atmos. Environ., 115, 340-344, 2015.

Galmarini, S., Koffi, B., Solazzo, E., Keating, T., Hogrefe, C., Schulz, M., Benedictow, A., Griesfeller, J. J., Janssens-Maenhout, G., Carmichael, G., Fu, J., and Dentener, F.: Technical note: Coordination and harmonization of the multi-scale, multi-model activities HTAP2, AQMEII3, and MICS-Asia3: simulations, emission inventories, boundary conditions, and model output formats, Atmos. Chem. Phys., 17, 1543-1555, doi:10.5194/acp-17-1543-2017, 2017.

Gratz, L. E., Jaffe, D. A., and Hee, J. R.: Causes of increasing ozone and decreasing carbon monoxide in springtime at the Mt. Bachelor Observatory from 2004 to 2013, Atmos. Environ., 109, 323–330, doi: 10.1016/j.atmosenv.2014.05.076, 2015.

Guenther, A., and Wiedinmyer, C.: User's guide to the Model of Emissions of Gases and Aerosols from Nature (MEGAN), Version 2.01, 2007.

Hanna, S. R., Lu, Z., Frey, H. C., Wheeler, N., Vukovich, J., Arunachalam, S., Fernau, M., and Hansen, D. A.: Uncertainties in predicted ozone concentrations due to input uncertainties for the UAM-V photochemical grid model applied to the July 1995 OTAG domain, Atmos. Environ., 35, 891–903, 2001.

Henze, D. K., Hakami, A., and Seinfeld, J. H.: Development of the adjoint of GEOS-Chem, Atmos. Chem. Phys., 7, 2413-2433, https://doi.org/10.5194/acp-7-2413-2007, 2007.

Hogrefe, C., Liu, P., Pouliot, G., Mathur, R., Roselle, S., Flemming, J., Lin, M., and Park, R. J.: Impacts of different characterizations of large-scale background on simulated regional-scale ozone over the continental United States, Atmos. Chem. Phys. Discuss.,., 18, 3839-3864, https://doi.org/10.5194/acp-2017-676, in review, 201718-3839-2018, 2018.

Im, U., Bianconi, R., Solazzo, E., Kioutsioukis, I., Badia, A., Balzarini, A., Baró, R., Bellasio, R., Brunner, D., Chemel, C., Curci, G., Flemming, J., Forkel, R., Giordano, L., Jiménez-Guerrero, P., Hirtl, M., Hodzic, A., Honzak, L., Jorba, O., Knote, C., Kuenen, J. J. P., Makar, P. A., Manders-Groot, A., Neal, L., Pérez, J. L., Pirovano, G., Pouliot, G., San Jose, R., Savage, N., Schroder, W., Sokhi, R. S., Syrakov, D., Torian, A., Tuccella, P., Werhahn, J., Wolke, R., Yahya, K., Zabkar, R., Zhang, Y., Zhang, J., Hogrefe, C., Galmarini, S.: Evaluation of operational on-line-coupled regional air quality models over Europe and North America in the context of AQMEII phase 2. Part I: Ozone, *Atmos. Environ.*, *11*5, 404-420, 2015.

Janjic, Z.: Nonsingular implementation of the Mellor–Yamada level 2.5 scheme in the NCEP meso model, NCEP Office Note, No. 437, National Centres for Environmental Prediction, 61, 2002. Available at http://www.emc.ncep.noaa.gov/officenotes/newernotes/on437.pdf

Janssens-Maenhout, G., Crippa, M., Guizzardi, D., Dentener, F., Muntean, M., Pouliot, G., Keating, T., Zhang, Q., Kurokawa, J., Wankmüller, R., Denier van der Gon, H., Kuenen, J. J. P., Klimont, Z., Frost, G., Darras, S., Koffi, B., and Li, M.: HTAP_v2.2: a mosaic of regional and global emission grid maps for 2008 and 2010 to study hemispheric transport of air pollution, Atmos. Chem. Phys., 15, 11411–11432, doi:10.5194/acp-15-11411-2015, 2015.

Jeffries, H. E., and Tonnesen, S. A.: comparison of two photochemical reation mechanisms using mass balance and process analysis. Atmos. Environ.,28, 2991-3003, 1994.

Kioutsioukis, I., Im, U., Solazzo, E., Bianconi, R., Badia, A., Balzarini, A., Baró, R., Bellasio, R., Brunner, D., Chemel, C., Curci, G., van der Gon, H. D., Flemming, J., Forkel, R., Giordano, L., Jiménez-Guerrero, P., Hirtl, M., Jorba, O., Manders-Groot, A., Neal, L., Pérez, J. L., Pirovano, G., San Jose, R., Savage, N., Schroder, W., Sokhi, R. S., Syrakov, D., Tuccella, P., Werhahn, J., Wolke, R., Hogrefe, C., and Galmarini, S.: Insights into the deterministic skill of air quality ensembles from the analysis of AQMEII data, Atmos. Chem. Phys., 16, 15629-15652, doi:10.5194/acp-16-15629-2016, 2016.

Karl, T.R., and Koss, W.J.: Regional and National Monthly, Seasonal, and Annual Temperature Weighted by Area, 1895-1983. Historical Climatology Series 4-3, National Climatic Data Center, Asheville, NC, 38, 1984.

Kwok, R.H.F., Baker, K.R., Napelenok, S. L., and Tonnesen, G. S.: Photochemical grid model implementation and application of VOC, NOx, and O3 source apportionment, Geosci. Model Dev., 8, 99-114, doi.org/10.5194/gmd-8-99-2015, 2015.

Langford, A. O., Senff, C. J., Alvarez II, R. J., Cooper, O. R., Holloway, J. S., Lin, M. Y., Marchbanks, R. D., Pierce, R. B., Sandberg, S. P., Weickmann, A. M., and Williams, E. J.: An overview of the 2013 Las Vegas Ozone Study (LVOS): Impact of stratospheric intrusions and long-range transport on surface air quality, Atmos. Environ., 109, 305–322, 2015.

Langford, A. O., Alvarez II, R. J., Brioude, J., Fine, R., Gustin, M. S., Lin, M. Y., Marchbanks, R. D., Pierce, R. B., Sandberg, S. P., Senff, C. J., Weickmann, A. M., and Williams, E. J.: Entrainment of stratospheric air and Asian pollution by the convective boundary layer in the southwestern U.S., J. Geophys. Res. Atmos., 122, 1312–1337, 2017.

Langner, J., Engardt, M., Baklanov, A., Christensen, J. H., Gauss, M., Geels, C., Hedegaard, G. B., Nuterman, R., Simpson, D., Soares, J., Sofiev, M., Wind, P., and Zakey, A.: A multi-model study of impacts of climate change on surface ozone in Europe, Atmos. Chem. Phys., 12, 10423-10440, https://doi.org/10.5194/acp-12-10423-2012, 2012.

Lin, M., Fiore, A. M., Horowitz, L. W., Cooper, O.R., Naik, V., Holloway, J., Johnson, B. J., Middlebrook, A. M., Oltmans, S. J., Pollack, I. B., and et al.: Transport of Asian ozone pollution into surface air over the western United States in spring, J. Geophys. Res., 117, D00V07, 2012.

Lin, M., Horowitz, L. W., Payton, R., Fiore, A. M., and Tonnesen, G.: US surface ozone trends and extremes from 1980 to 2014: quantifying the roles of rising Asian emissions, domestic controls, wildfires, and climate, Atmos. Chem. Phys., 17, 2943-2970, doi:10.5194/acp-17-2943-2017, 2017.

Mallet, V., and Sportisse, B.: Ensemble-based air quality forecasts: A multimodel approach applied to ozone, J. Geophys. Res., 111, D18302, doi:10.1029/2005JD006675, 2006.

Mallet, V., Stoltz, G., and Mauricette, B: Ozone ensemble forecast with machine learning algorithms, J. Geophys. Res., 114, D05307, 2009.

Mathur, R., Xing, J., Gilliam, R., Sarwar, G., Hogrefe, C., Pleim, J., Pouliot, G., Roselle, S., Spero, T. L., Wong, D. C., and Young, J.: Extending the Community Multiscale Air Quality (CMAQ) modelingmodelling system to hemispheric scales: overview of process considerations and initial applications, Atmos. Chem. Phys., 17, 12449-12474, https://doi.org/10.5194/acp-17-12449-2017, 2017.

Mellor, G. L., and Yamada, T.: A hierarchy of turbulence closure models for planetary boundary layers, J. Atmos. Sci., 31, 1791–1806, 1974.

Mellor, G. L., and Yamada, T.: Development of a turbulence closure model for geophysical fluid problems, Rev. Geophys. Space Phys., 20, 851–875, 1982.

Nopmongcol, U., Jung, J., Kumar, N., and Yarwood, G.: Changes in US background ozone due to global anthropogenic emissions from 1970 to 2020, Atmos. Environ., 140, 446–455, 2016.

Nopmongcol, U., Liu, Z., Stoeckenius, T., and Yarwood, G.: ModelingModelling intercontinental transport of ozone in North America with CAMx for the Air Quality Model Evaluation International Initiative (AQMEII) Phase 3, Atmos. Chem. Phys., 17, 9931-9943, https://doi.org/10.5194/acp-17-9931-2017, 2017.

Pleim, J. E.: A combined local and nonlocal closure model for the atmospheric boundary layer. Part I: model description and testing, J. Appl. Meteorology Climatol., 46, 1383–1395, 2007.

Pleim, J., and Ran, L.: Surface Flux ModelingModelling for Air Quality Applications. Atmosphere, 2, 271-302, 2011.

Pouliot, G., Denier van der Gon, H. A. C., Kuenen, J., Zhang, J., Moran, M. D., and Makar, P. A.: Analysis of the emission inventories and model-ready emission datasets of Europe and North America for phase 2 of the AQMEII project, Atmos. Environ.,115, 345–360, 2015.

Ramboll: User's Guide to the Comprehensive Air quality Model with extensions (CAMx). (Available from http://www.camx.com/files/camxusersguide_v6-50.pdf), 2018.

Russell, A., and Dennis, R.: NARSTO critical review of photochemical models and modelingmodelling, Atmos. Environ., 34, 2234–2283, 2000.

Sakulyanontvittaya, T., Duhl, T., Wiedinmyer, C., Helmig, D., Matsunaga, S., Potosnak, M., Milford, J., and Guenther, A.: Monoterpene and Sesquiterpene Emission Estimates for the United States, Environ. Sci. Technol., 42, 1623-1629, doi:10.1021/es702274e, 2008.

Formatted: Font: Not Italic

Formatted: Font: Not Italic

~~Schwede, D., Pouliot, G., and Pierce, T.: Changes to the biogenic emissions inventory system version 3 (BEIS3). 4th Annual CMAS Models-3 Users' Conference, 2005.~~

Simpson, D., Fagerli, H., Jonson, J. E., Tsyro, S., Wind, P., and Tuovinen, J.-P.: Transboundary Acidification, Eutrophication and Ground Level Ozone in Europe, PART I, Unified EMEP Model Description, p. 104, 2003.

Solazzo, E., Bianconi, R., ~~Vautard, R., Appel, K. W., Moran, M.D., Hogrefe, C., Bessagnet, B., Brandt, J., Christensen, J. H., Chemel, C., Coll, I., van der Gon, H. D., Ferreira, J., Forkel, R., Francis, X. V., Grell, G., Grossi, P., Hansen, A. B., Jericevic, A., Kraljevic, L., Miranda, A. I., Nopmongcol, U., Pirovano, G., Prank, M., Riccio, A., Sartelet, K. N., Schaap, M., Silver, J. D., Sokhi, R. S., Vira, J., Werhahn, J., Wolke, R., Yarwood, G.,Zhang, J., Rao, S. T., and Galmarini, S.: Model evaluation and ensemble modelling and for surface-level ozone in Europe and North America, Atmos. Environ., 53, 60–74, 2012.~~

~~Solazzo, E. and Galmarini, S.: Error apportionment for atmospheric chemistry-transport models – a new approach to model evaluation, Atmos. Chem. Phys., 16, 6263-6283, doi:10.5194/acp-16-6263-2016, 2016.~~

~~Solazzo, E., Bianconi, R.,~~ Hogrefe, C., Curci, G., Tuccella, P., Alyuz, U., Balzarini, A., Baró, R., Bellasio, R., Bieserk J., Brandt, J., Christensen, J. H., Colette, A., Francis, X. V., Garcia-Vivanco, M., Jiménez-Guerrero, P., Im, U., Manders, A., Nopmongcol, U., Kitwiroon, N., Pirovano, G., Pozzoli, L., Prank, M., Sokhi, R. S., Unal, A., Yarwood, G., and Galmarini, S.: Evaluation and error apportionment of an ensemble of atmospheric chemistry transport ~~modeling~~modelling systems: multivariable temporal and spatial breakdown, Atmos. Chem. Phys., 17, 3001–3054, doi: 10.5194/acp-17-3001-2017, 2017.

~~van Loon, M., Vautard, R., Schaap, M., Bergström, R., Bessagnet, B., Brandt, J., Builtjes, P. J. H., Christensen, J. H., Cuvelier, C., Graff, A., Jonson, J. E., Krol, M., Langner, J., Roberts, P., Rouil, L., Stern, R., Tarrason, L., Thunis, P., Vignati, E., White, L., and Wind, P.: Evaluation of long-term ozone simulations from seven regional air quality models and their ensemble average, Atmos. Environ., 41,2083–2097, 2007.~~

~~Vautard, R., Van Loon, M., Schaap, M., Bergström, R., Bessagne, B., Brandt, J., et al., Is regional air quality model diversity representative of uncertainty for ozone simulation? Geophys. Res. Lett., 33, L24818, doi:10.1029/2006GL027610, 2006.~~

~~Yarwood, G., Morris, R. E., Yocke, M. A., Hogo, H. and Chico, T.: Development of a Methodology for Source Apportionment of Ozone Concentration Estimates from a Photochemical Grid Model, Presented at the 89th AWMA Annual Meeting, Nashville TN, June 23-28, 1996.~~

~~Yang, Y. J., Wilkinson, J. G., and Russell, A. G.: Fast, direct sensitivity analysis of multidimensional photochemical models, Environ. Sci. Technol., 31, 2859–2868, 1997.~~

Zhang, L., Brook, J. R. and Vet, R.: A revised parameterization for gaseous dry deposition in air-quality models, Atmos. Chem. Phys., 3, 2067–2082, doi:10.5194/acp-3-2067-2003, 2003.

Zhang, L., Jacob, D.J., Downey, N.V., Wood, D.A., Blewitt, D., Carouge, C. C., van Donkelaar, A., Jones, D.B., Murray, L.T., and Wang, Y.: Improved estimate of the policy-relevant background ozone in the United States using the GEOS-Chem global model with 1/2º×2/3º horizontal resolution over North America, Atmos. Environ. 45, 6769-6776, 2011.

Zhang, Y., Seigneur, C., Bocquet, M., Mallet, V., and Baklanov, A.: Real-Time Air Quality Forecasting, Part II: State of the Science, Current Research Needs, and Future Prospects, Atmos. Environ., 60, 656–676, 2012.

Formatted: Default Paragraph Font, English (United Kingdom)

List of Tables

List of Figures

~~Figure 10. Difference between WRF/CAMx and WRF/CMAQ (WRF/CAMx minus WRF/CMAQ) in the seasonal averaged DM8A BCT, and DM8A ozone at all receptors.~~

~~Figure 11. Difference in DM8A ozone between rural and urban sites (rural minus urban) from observation, WRF/CMAQ, and WRF/CAMx.~~

Table 1: Model description of the four simulations involved in the study

	WRF/CMAQ	WRF/CAMx	COSMO-CLM/CMAQ	WRF/DEHM
Institute	U.S. EPA	RAMBOLL Environ (U.S.)	Helmholtz-Zentrum Geesthacht (Germany)	Aarhus University (Denmark)
Global Meteorology	NCEP	NCEP	NCEP	ECMWF
Regional ~~Modeling~~Modelling System	WRF3.4/ CMAQ5.0.2	WRF3.4/ CAMx6.2	COSMO-CLM/ CMAQ5.0.1	WRF/ DEHM
Horizontal Resolution	12km	12km	24km	~~50km~~17km
~~Gas Phase Chemistry~~	~~CB05-TUCL~~	~~CB05~~	~~CB05-TUCL~~	~~Brandt et al. (2012)~~
Dry Deposition for Ozone and Tracers	Pleim and Ran (2011)	Zhang et al. (2003)	Pleim and Ran (2011)	Simpson et al. (2003)
~~Wet Deposition for Ozone~~	~~YES~~	~~YES~~	~~YES~~	~~NO~~
Wet Deposition for Tracers	YES	NO	YES	YES
~~Impact of Sub-Grid Clouds on Radiation in RCMs~~	~~NO~~	~~NO~~	~~YES~~	~~NO~~
~~Impact of Sub-Grid Clouds on Ozone Photochemistry in CTMs~~	~~YES~~	~~YES~~	~~YES~~	~~YES~~
Sub-Grid Cloud Mixing in CTMs	YES	NO	YES	YES

Formatted Table

Formatted Table

Formatted Table

Formatted Table

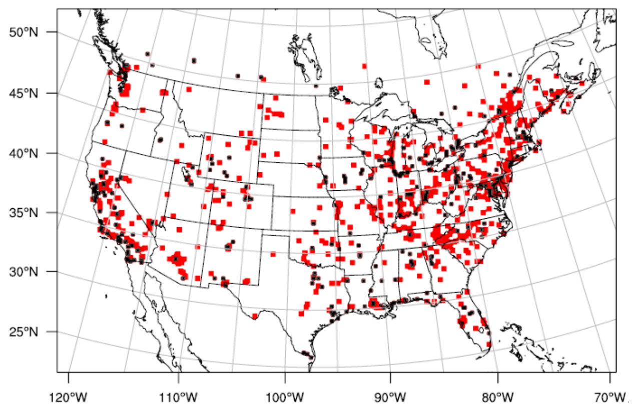
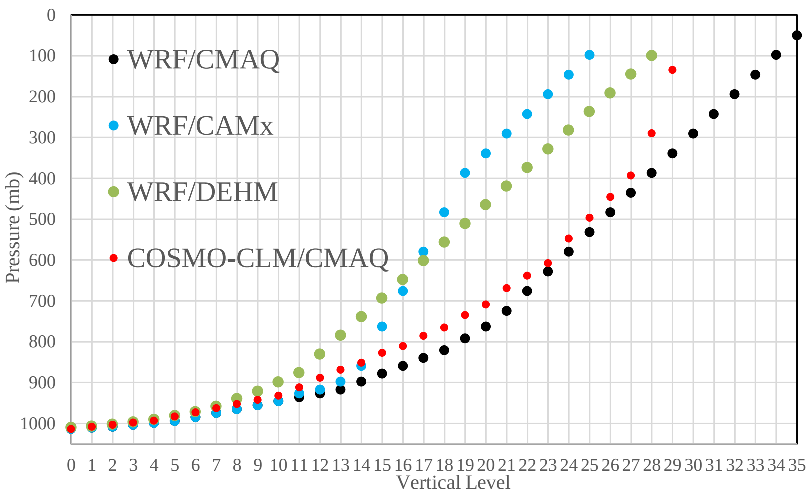


Figure 1: The studied area for this research, and the observation sites for ozone available in the data system of ENSEMBLE, with black dots for urban sites and red dots for rural sites.



Parameterization for Vertical Turbulent Mixing	ACM2 (Pleim 2007)	K-theory (Ramboll, 2018)	ACM2	K-theory (Christensen,1997)

Table 2: Description of the sensitivity simulations

Sensitivity Simulations	Description
WRF/CMAQ_noddry	Same as WRF/CMAQ but no dry deposition for inert tracers
WRF/CMAQ_nodwet	Same as WRF/CMAQ but no wet deposition for inert tracers
WRF/CMAQ_nocldmix	Same as WRF/CMAQ but no sub-grid cloud mixing for inert tracers
WRF/CMAQ_27aL	Same as WRF/CMAQ but using the same vertical grid structure as WRF/CAMx
WRF/DEHM_noddry	Same as WRF/DEHM but no dry deposition for inert tracers

Table 3: Seasonal differences in DM8A BCT (ppb) between simulations after averaging over all land grid cells over the analysis domain. Values in parenthesis are the standard deviations.

	winter	spring	summer	fall
WRF/CMAQ_noddry minus WRF/CMAQ	10.3 (±2.9)	10.3 (±3.2)	12.6 (±4.4)	9.7 (±2.2)
WRF/CMAQ_nocldmix minus WRF/CMAQ	-0.2 (±0.3)	-1.0 (±0.6)	-2.4 (±0.9)	-0.8 (±0.4)
WRF/CMAQ_27aL minus WRF/CMAQ	0.5 (±0.3)	1.5 (±0.6)	1.3 (±0.4)	0.7 (±0.2)
WRF/CAMx minus WRF/CMAQ	4.6 (±3.2)	2.4 (±1.6)	5.2 (±2.6)	1.9 (±1.8)
WRF/CAMx minus WRF/CMAQ_27aL	4.2 (±3.5)	0.9 (±1.7)	3.9 (±2.6)	1.1 (±1.8)
COSMO-CLM/CMAQ minus WRF/CMAQ	1.3 (±1.8)	0.1 (±1.7)	0.1 (±2.1)	1.2 (±1.1)
WRF/DEHM_noddry minus WRF/DEHM	9.0 (±2.3)	15.0 (±4.3)	17.6 (±6.0)	13.2 (±2.8)

Table 4: Seasonal differences in DM8A BC1/BCT (%) between simulations after averaging over all land grid cells over the analysis domain. Values in parenthesis are the standard deviations.

	winter	spring	summer	fall
WRF/CMAQ_noddry minus WRF/CMAQ	-1.4 (±2.0)	-1.1 (±2.3)	0.4 (±2.6)	-1.2 (±2.2)
WRF/CMAQ_nocldmix minus WRF/CMAQ	1.5 (±1.6)	2.4 (±1.3)	4.6 (±2.2)	2.8 (±1.5)
WRF/CMAQ_27aL minus WRF/CMAQ	-8.1 (±3.5)	-8.5 (±2.1)	-7.5 (±3.0)	-9.1 (±2.6)
WRF/CAMx minus WRF/CMAQ	-14.6 (±3.9)	-14.8 (±4.1)	-13.5 (±8.2)	-15.6 (±5.9)
WRF/CAMx minus WRF/CMAQ_27aL	-6.4 (±4.8)	-6.3 (±3.2)	-6.0 (±5.7)	-6.8 (±4.8)
WRF/DEHM_noddry minus WRF/DEHM	0.4 (±1.0)	1.4 (±2.0)	3.7 (±2.2)	1.4 (±1.7)

Table 5: Differences in BC1/BCT (%) averaged over the 24 hours in the diurnal cycle over seven sub-regions in summer between WRF/CMAQ and its sensitivity simulations, and between WRF/CAMx and WRF/CMAQ_27aL. The numbers in parenthesis represents the range of the diurnal differences in BC1/BCT, namely the (maximum – minimum) within the diurnal cycle of the difference in BC1/BCT.

BC1/BCT (%)	WB	NB	MT	GP	NE	SE	ATL
WRF/CMAQ_noddry minus WRF/CMAQ	1.6 (1.5)	-0.4 (1.4)	1.8 (0.4)	2.3 (0.7)	2.9 (1.0)	1.5 (0.8)	0.2 (0.01)
WRF/CMAQ_nocldmix minus WRF/CMAQ	2.3 (1.5)	5.6 (1.0)	1.6 (0.7)	5.0 (1.6)	5.4 (1.7)	6.5 (1.1)	13.8 (0.1)
WRF/CMAQ_27aL minus WRF/CMAQ	-8.7 (0.7)	-10.1 (1.7)	-3.8 (0.6)	-5.2 (0.7)	-7.8 (0.5)	-4.6 (0.6)	-5.9 (0.1)
WRF/CAMx minus WRF/CMAQ_27aL	-7.0 (1.2)	-11.8 (2.6)	-2.1 (0.6)	-2.8 (0.9)	-3.8 (1.0)	2.0 (1.9)	16.1 (0.3)

Table 6: Seasonal averaged differences between COSMO-CLM/CMAQ and WRF/CMAQ (COSMO-CLM/CMAQ minus WRF/CMAQ) in BC1 at the surface, in BC2 and BC3 at the surface and at specific elevation in summer at 2pm LST over seven sub-regions.

		WB	NB	MT	GP	NE	SE	ATL
BC1 at the surface		-2.3	-1.5	-0.6	0.8	-0.2	4.0	2.9
BC2	7km	-24.1	-17.6	-25.9	-19.1	-23.6	-15.3	-14.8
	5km	-7.9	-5.5	-15.9	-11.7	-11.7	-12.3	-8.0
	3km	-4.3	-3.5	-8.5	-11.6	-7.1	-15.3	-7.4
	surface	3.2	1.8	-2.6	-4.2	-3.1	-9.1	-10.9
BC3	7km	16.3	14.2	22.5	15.4	25.1	17.2	20.4
	5km	2.6	1.8	7.1	6.1	8.0	12.3	10.1
	3km	-0.5	0.4	4.5	3.6	2.5	1.4	1.6
	surface	0.5	0.1	4.2	2.8	1.0	0.8	-0.5

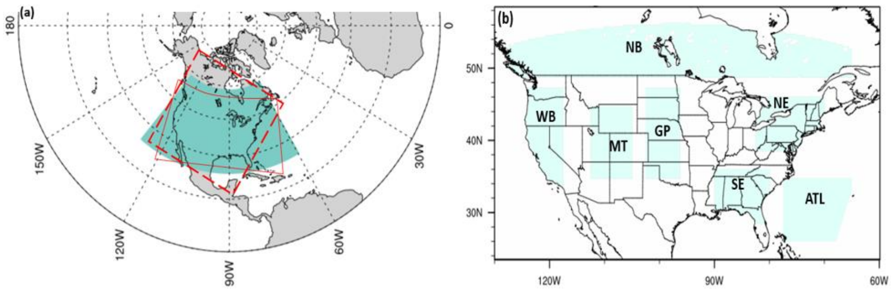


Figure 1. (a) simulation domain for WRF/CMAQ and WRF/CAMx (in red solid line) and the simulation domain for WRF/DEHM (in red dashed line), and the analysis domain in this study (shaded area in green). The simulation domain of COSMO-CLM/CMAQ is of the same size as WRF/CMAQ, but shifts westward by 48km. (b) the sub-regions in the analysis domain.

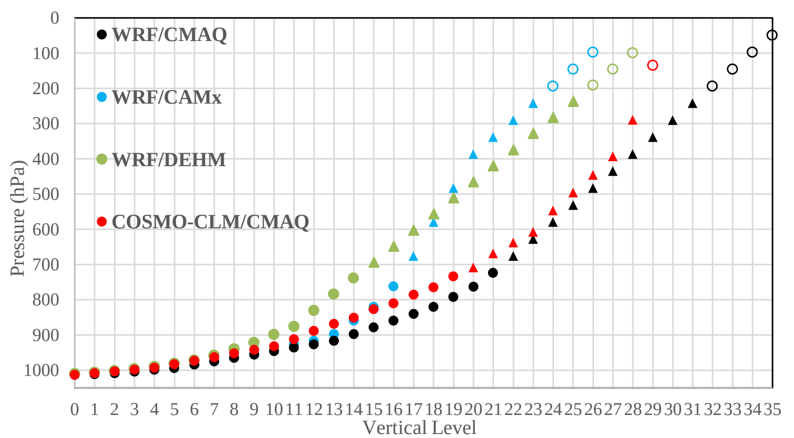
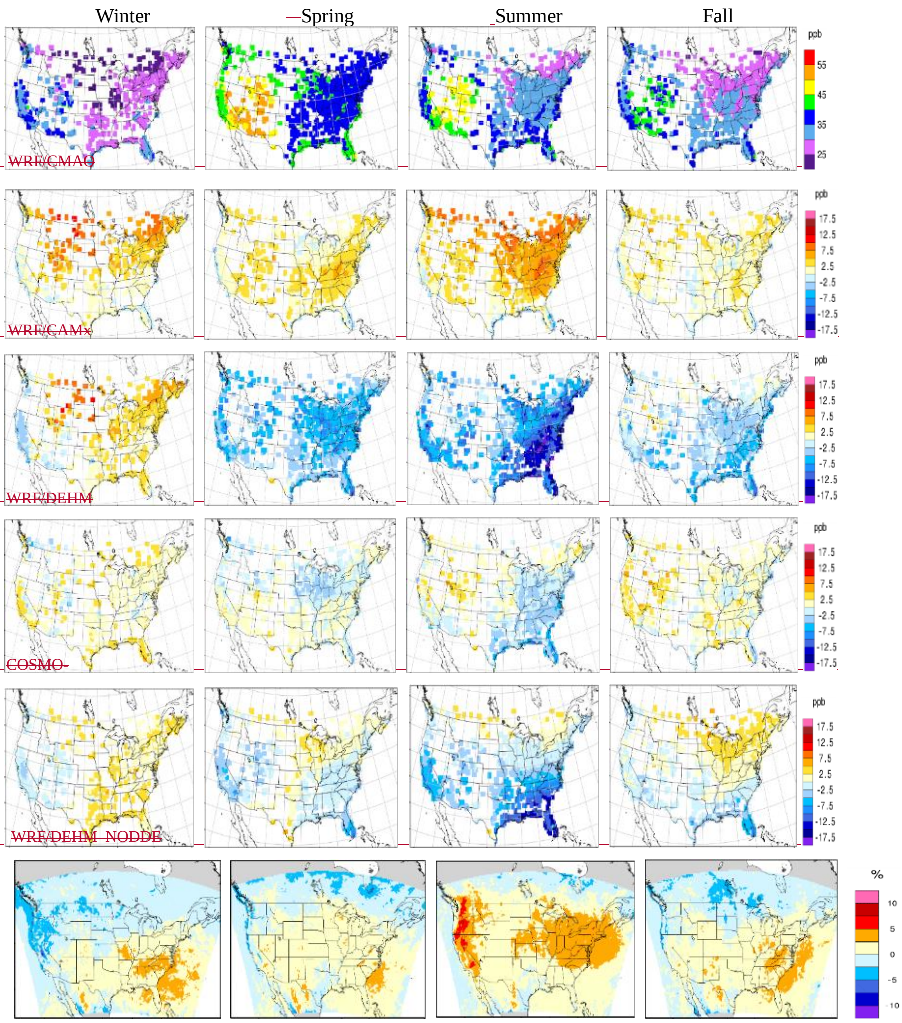


Figure 2~~:~~. Vertical grid ~~resolution~~structures for the chemical transport models used in the four ~~modeling systems.~~simulations, with the filled circles for the vertical levels for BC1, filled triangles for BC2, and open circles for BC3, respectively.


Winter
Spring
Summer
Fall
WRF/CMAQ
WRF/CAMx
WRF/DEHM
COSMO
WRF/DEHM_NODDE
ppb
55
45
35
25
17.5
12.5
7.5
2.5
-2.5
-7.5
-12.5
-17.5
%
10
5
0
-5
-10

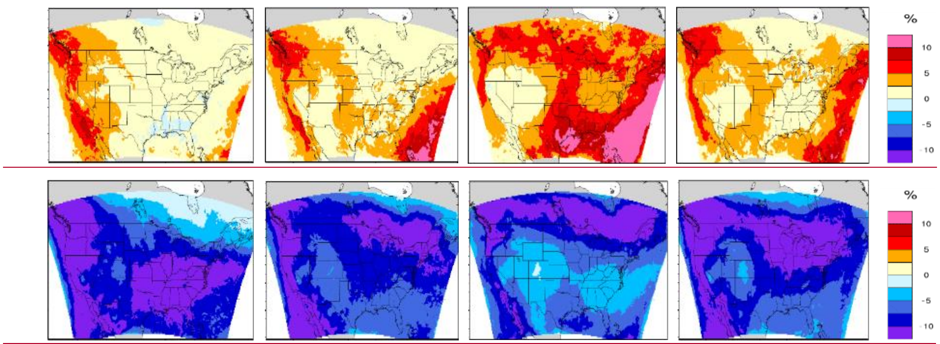
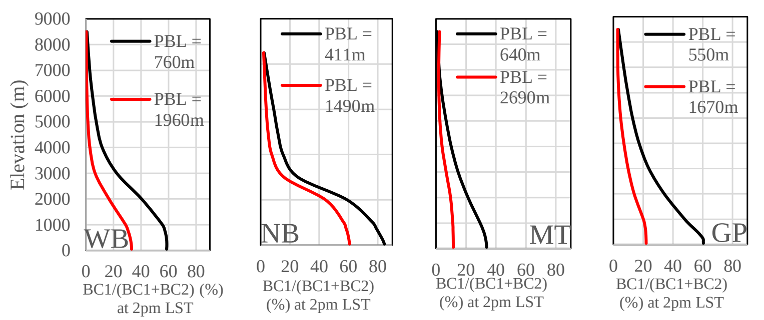


Figure 3~~: Seasonal average of~~. Differences in the seasonal averaged DM8A ~~BCT from WRF/CMAQ (first row), and the difference in DM8A~~ BC1/BCT between WRF/~~CAMx~~CMAQ_noddry and WRF/CMAQ~~,~~ (first row), between WRF/~~DEHM~~CMAQ_nocldmix and WRF/CMAQ~~,~~ (second row), and between ~~COSMO-CLM~~WRF/CMAQ_27aL and WRF/CMAQ (~~from the second to the fourth row) at all receptors. Results~~ third row). All results are shown ~~are~~ as ~~other model~~sensitivity simulation minus WRF/CMAQ. The ~~last row is the difference~~areas in ~~DM8A BCT~~ white or grey are the grid cells where are out of the simulation domain of WRF/CMAQ.



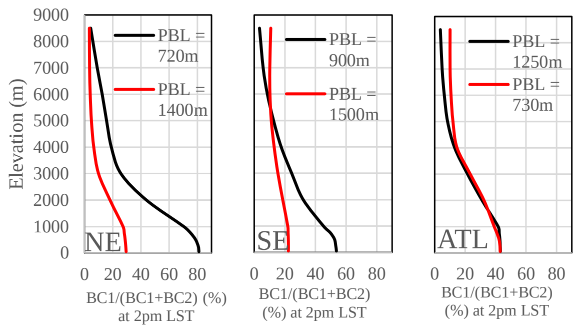


Figure 4. Seasonal averaged vertical profiles of BC1/(BC1+BC2) (in percentage) for WRF/CMAQ in winter (in black) and in summer (in red) over the sub-regions at their local standard time of 2pm, respectively. The numbers in the legend show the seasonal averaged maximum PBL height during the daytime over each sub-region.

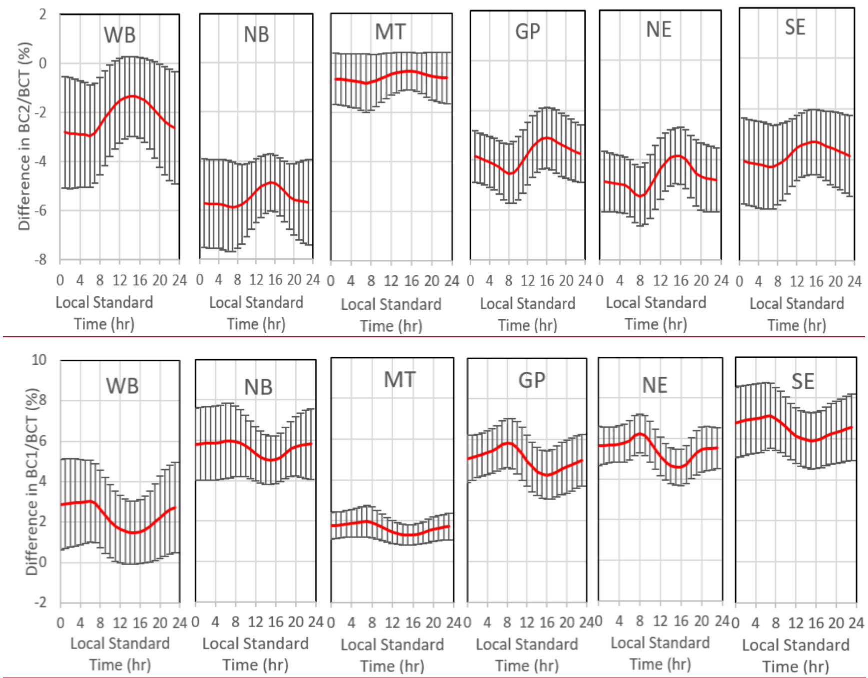


Figure 5. Diurnal cycles for the differences in BC2/BCT (%) (top row) and in BC1/BCT (%) (bottom row) between WRF/~~DEHM_NODDEP~~CMAQ_nocldmix and WRF/CMAQ~~_NODDEP~~ (WRF/~~DEHM_NODDEP~~CMAQ_nocldmix minus WRF/CMAQ~~_NODDEP).~~) during summer over sub-regions. For each sub-region, the regional average is in red line with the standard deviation in black bars.

Formatted: Line spacing: single

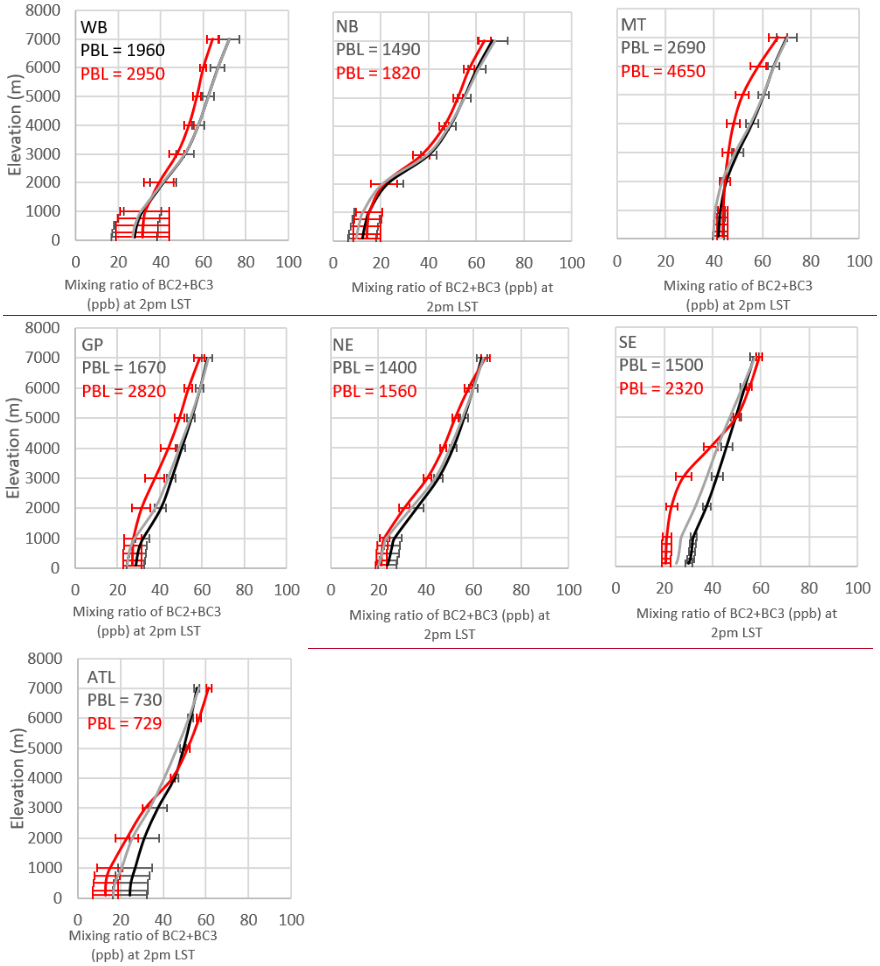


Figure 6a. Seasonal averaged vertical profiles of (BC2+BC3) (in ppb) for WRF/CMAQ(black), WRF/CMAQ_nocldmix (grey) and COSMO-CLM/CMAQ (red) in summer over the sub-regions at their local standard time of 2pm, respectively. The bars represent the standard deviations over the sub-region (The standard deviations of WRF/CMAQ_nocldmix is not shown, as the values are almost the same as those of WRF/CMAQ). The numbers in legend are the seasonal averaged maximum PBL height during the daytime over each sub-region for WRF/CMAQ (in black) and for COSMO-CLM/CMAQ (in red).

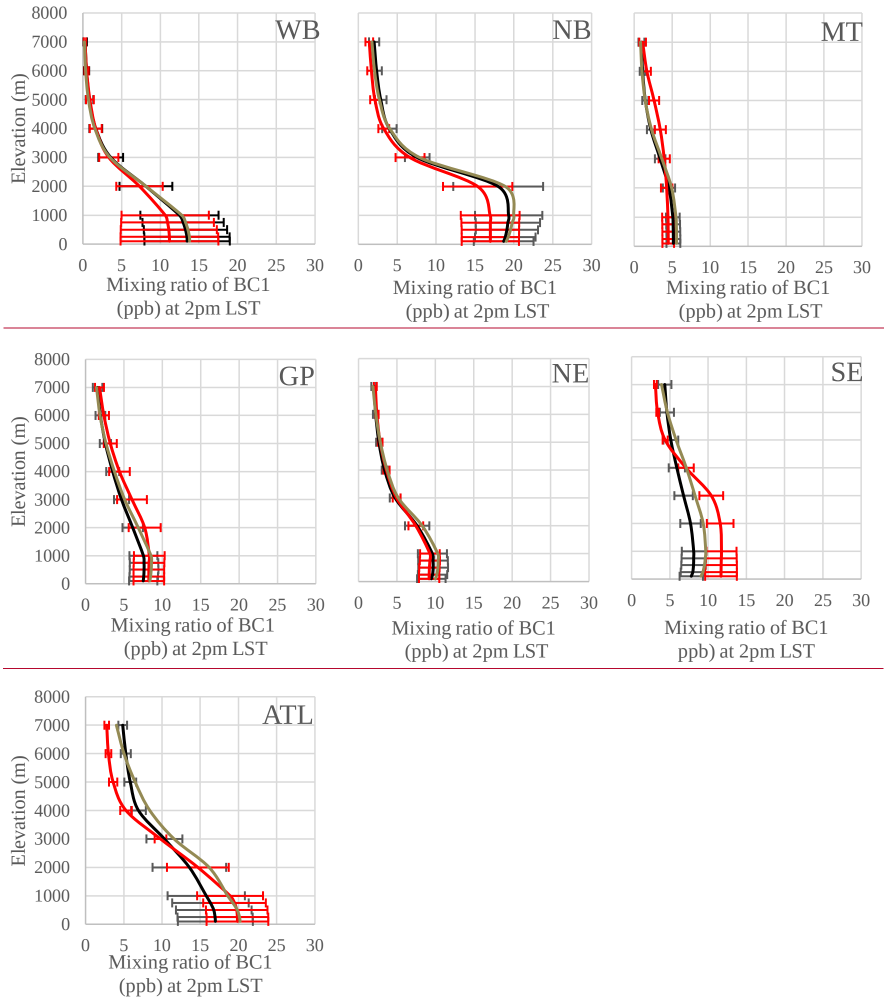
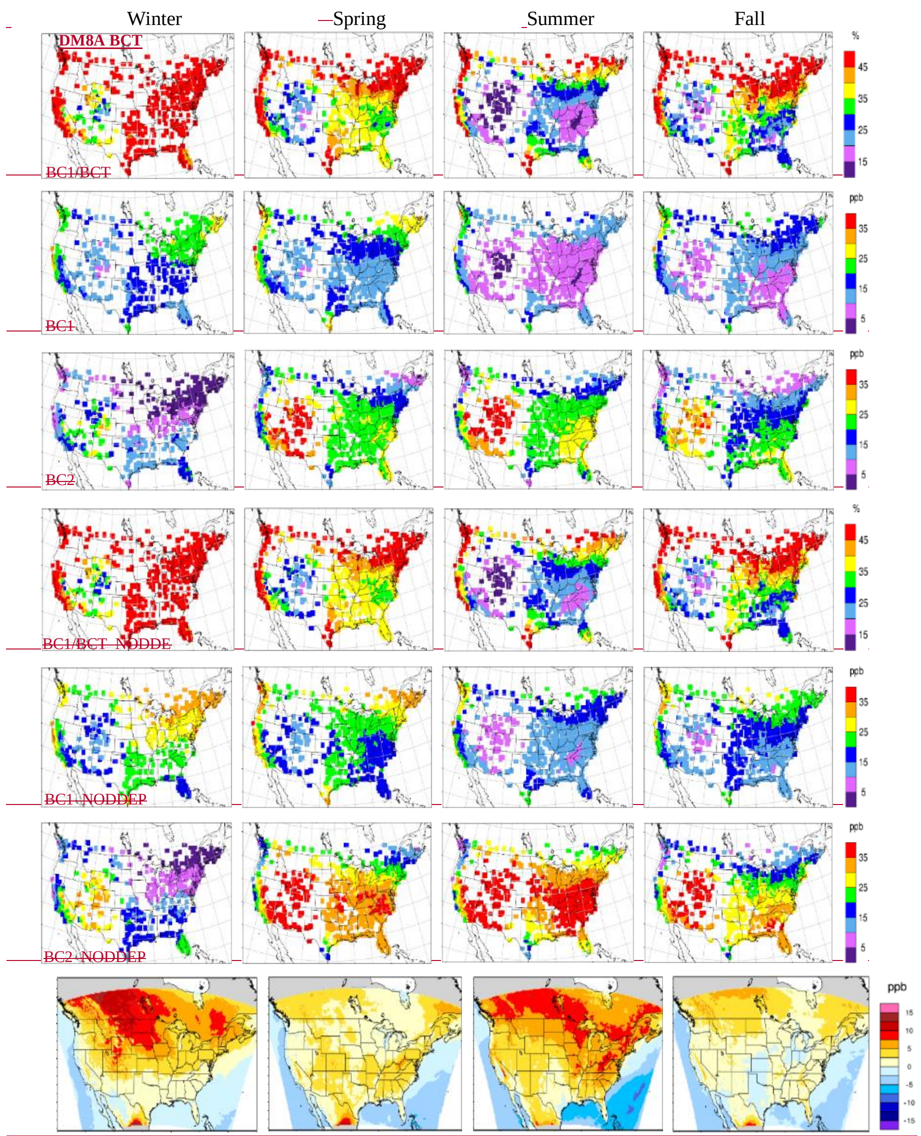


Figure 6b. Same as 6a, but for BC1 (in ppb).


Winter
Spring
Summer
Fall
DM8A BCT
BC1/BCT
BC1
BC2
BC1/BCT_NODDE
BC1_NODDEP
BC2_NODDEP
%
45
35
25
15
ppb
35
25
15
5
ppb
15
10
5
0
-5
-10
-15

DM8A O_3

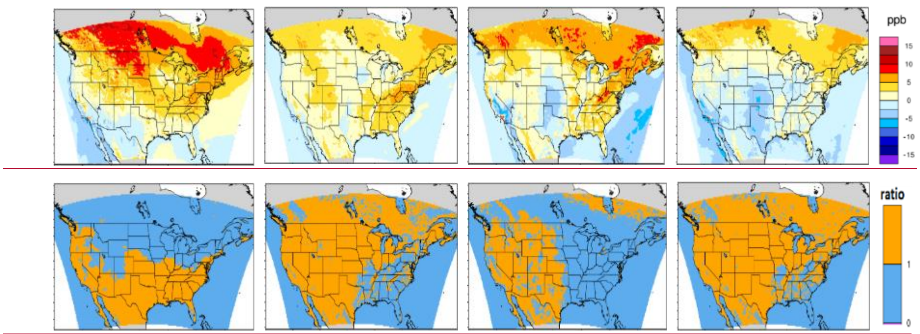
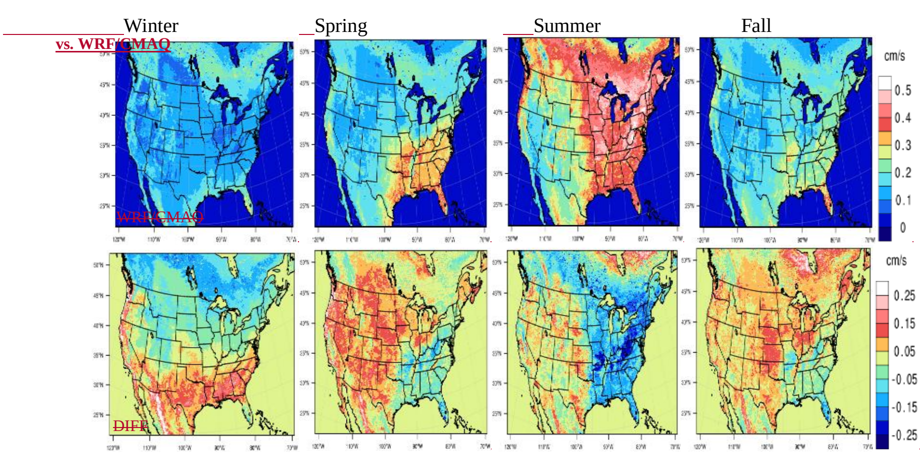


Figure 4: Seasonal averages from WRF/CMAQ of (DM8A BC1)/(7. Differences in the seasonal averaged DM8A BCT) (first row),) and DM8A BC1O3 (second row), and DM8A BC2 () between WRF/CAMx and WRF/CMAQ (WRF/CAMx minus WRF/CMAQ) in the analysis domain. The third row) at all receptors. The last three rows is the seasonal averaged ratio in dry deposition velocity of (WRF/CAMx over WRF/CMAQ) for ozone. The areas in white or grey are the same as the first three rows but for WRF/CMAQ_NODDEPgrid cells where are out of the simulation domain.



Formatted: Justified

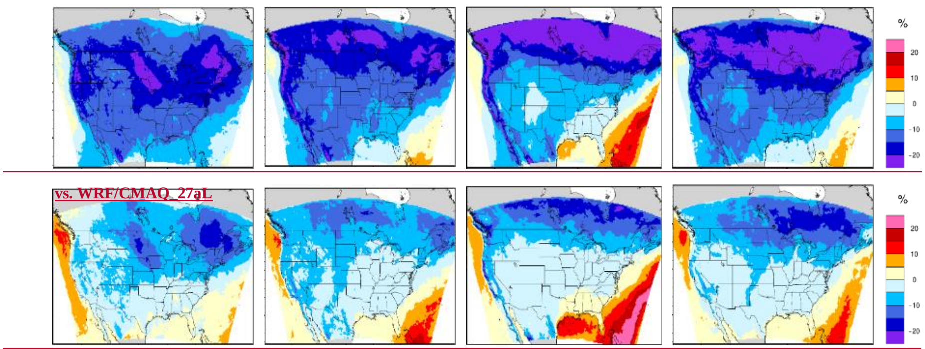


Figure ~~5: Seasonal average of~~8. Differences in the ~~dry deposition velocity of ozone from WRF/CMAQ (first row), and the difference~~seasonal averaged DM8A BC1/BCT between WRF/CAMx and WRF/CMAQ (WRF/CAMx minus WRF/CMAQ) (first row), and between WRF/CAMx and WRF/CMAQ_27aL (WRF/CAMx minus WRF/CMAQ_27aL) (second row~~).~~) in the analysis domain.

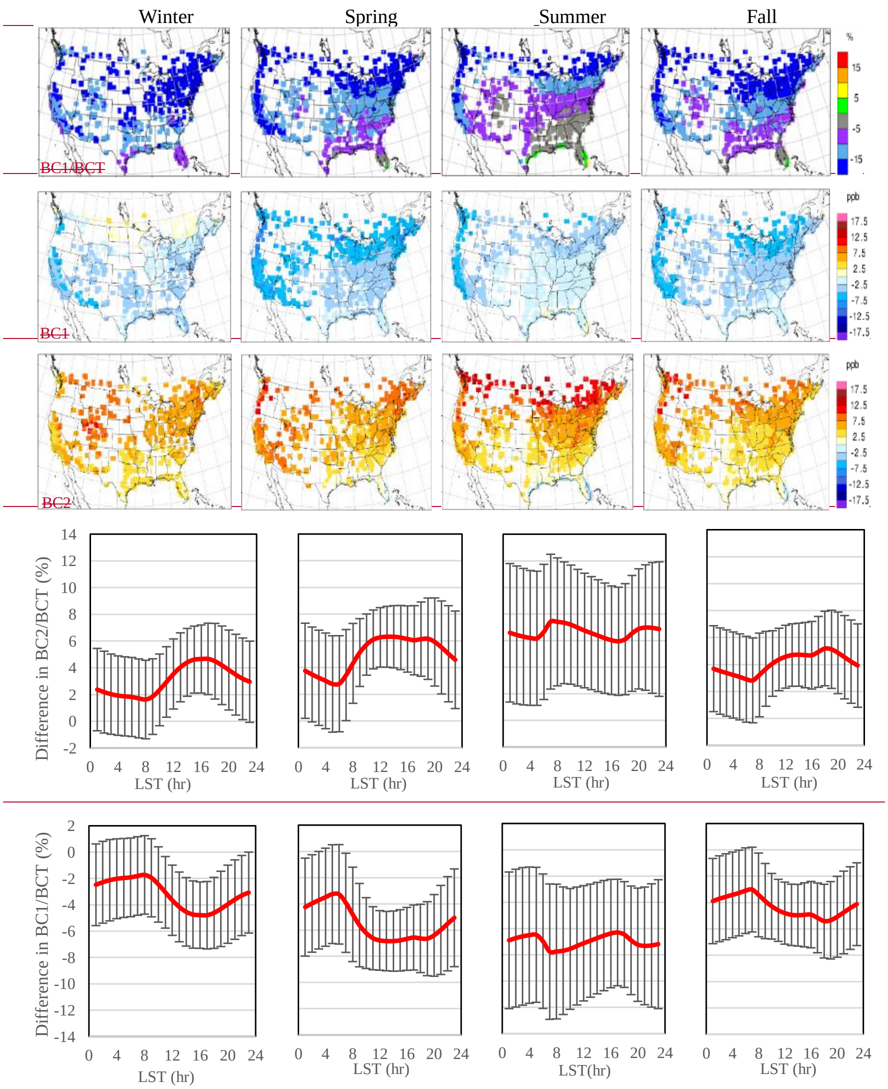


Figure ~~6: Difference~~9a. Differences in the seasonal averaged diurnal cycles of BC2/BCT (%) (top row) and in BC1/BCT (%) (bottom row) between WRF/CAMx and WRF/CMAQ 27aL (WRF/CAMx minus WRF/CMAQ~~) in the seasonal averages of (DM8A~~

BC1)/(DM8A BCT) (first row), DM8A BC1 (second row), and DM8A BC2 (third row) at all receptors 27aL) over the WB sub-region. For each season and hour, the regional average is shown in red with the standard deviation shown in black bars.

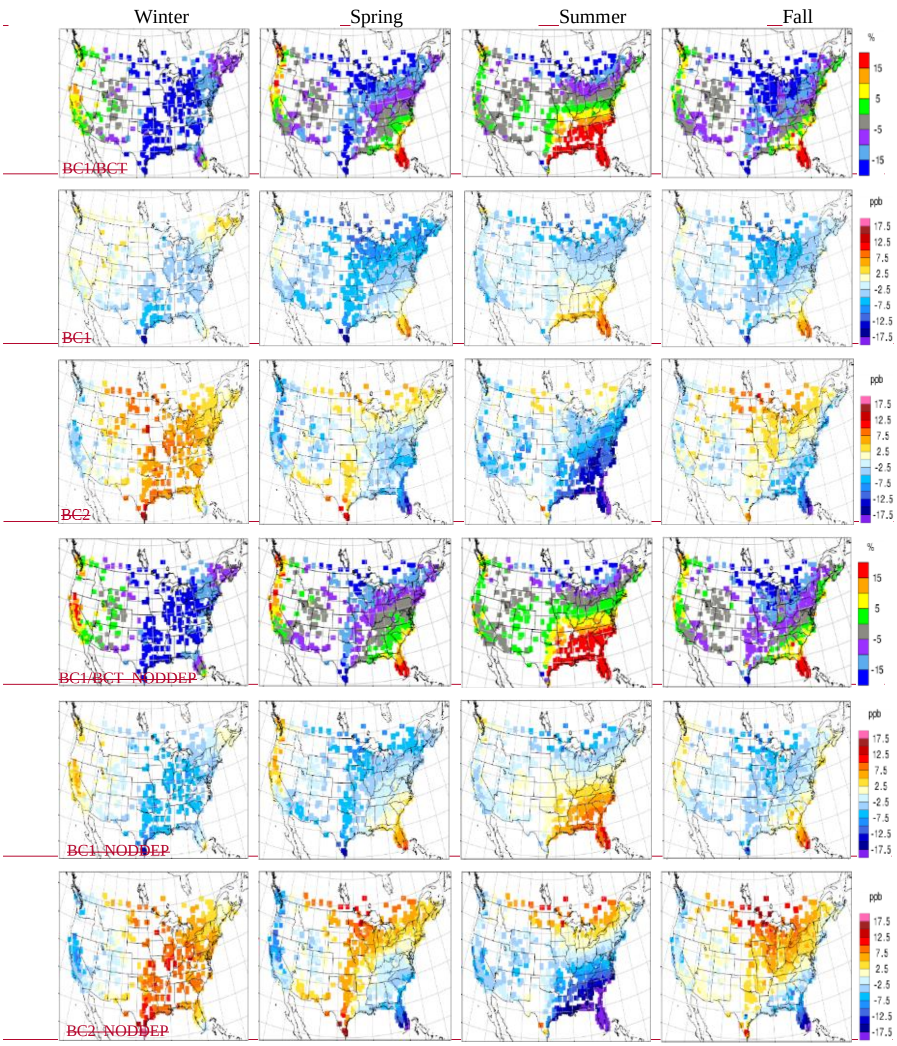


Figure 7: Difference between WRF/DEHM and WRF/CMAQ (WRF/DEHM minus WRF/CMAQ) in the seasonal averages of (DM8A BC1)/(DM8A BCT) (first row), DM8A BC1 (second row), and DM8A BC2 (third row) at all receptors. The last three rows are the same as first three rows but for the difference between WRF/DEHM_NODDEP and WRF/CMAQ_NODDEP (WRF/DEHM_NODDEP minus WRF/CMAQ_NODDEP).

Formatted: Page break before

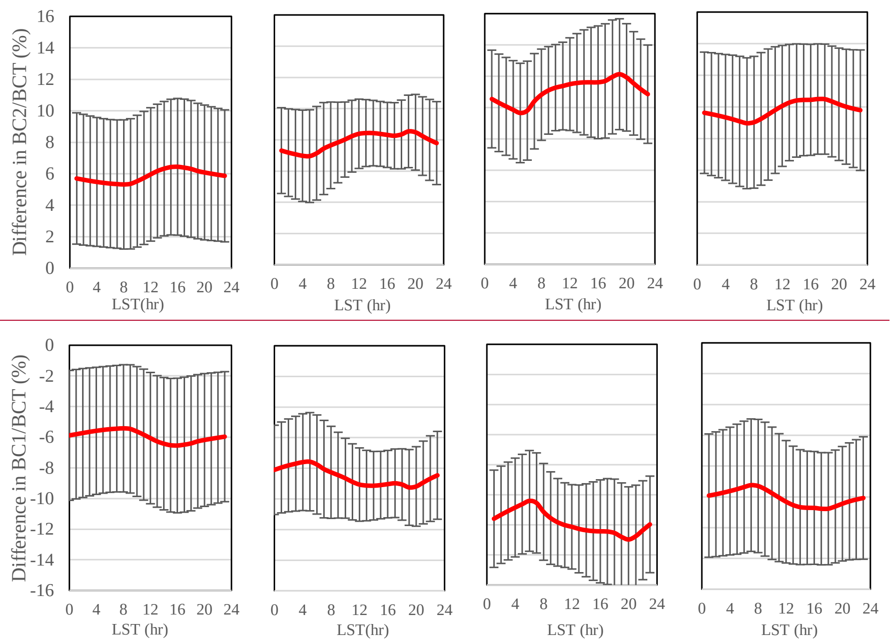


Figure 9b. Same as Figure 9a, but over the NB sub-region.

Winter Spring Summer Fall

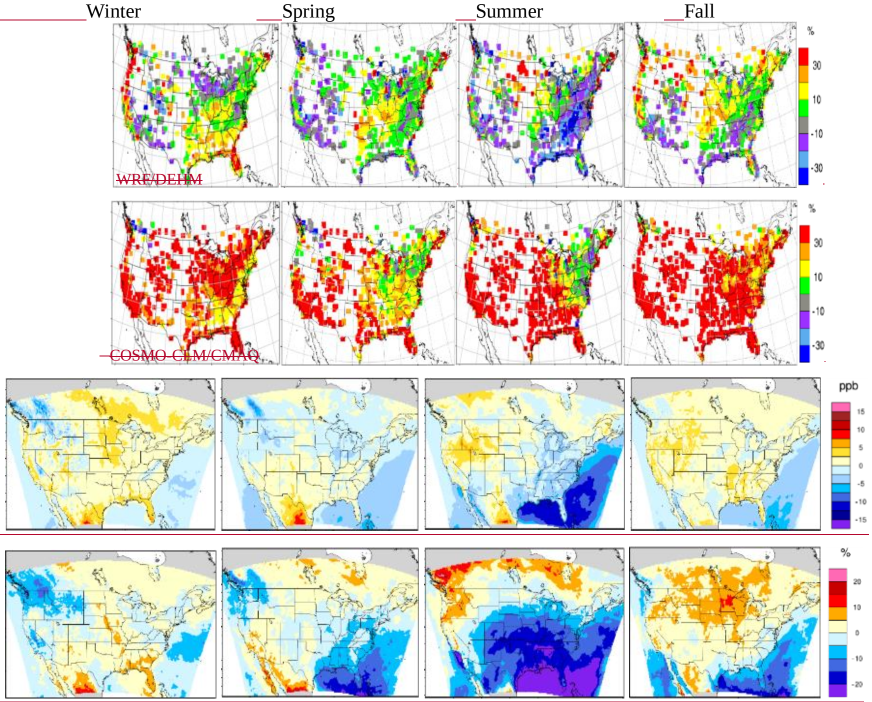
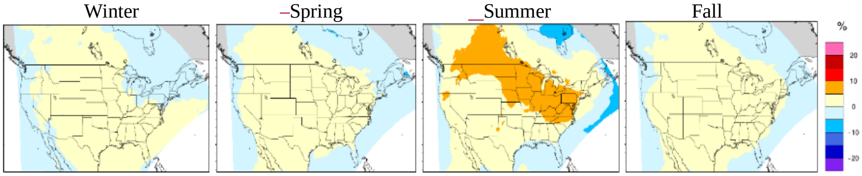


Figure ~~8: Seasonal average of the difference~~10. Differences in ~~PBL between WRF/DEHM and WRF/CMAQ at all receptors~~the seasonal averaged ~~by hourly data during daytime~~DM8A_BCT (first row~~). The difference in PBL is (WRF/DEHM minus WRF/CMAQ) and normalized by WRF/CMAQ. The~~) and DM8A_BC2/BCT (second row ~~is the same as first row but~~) between COSMO-CLM/CMAQ and WRF/CMAQ (COSMO-CLM/CMAQ minus WRF/CMAQ) in the analysis domain. The areas in white or grey are the grid cells that are out of the simulation domain of either of the two models.

Winter Spring Summer Fall



Formatted: Justified, Page break before

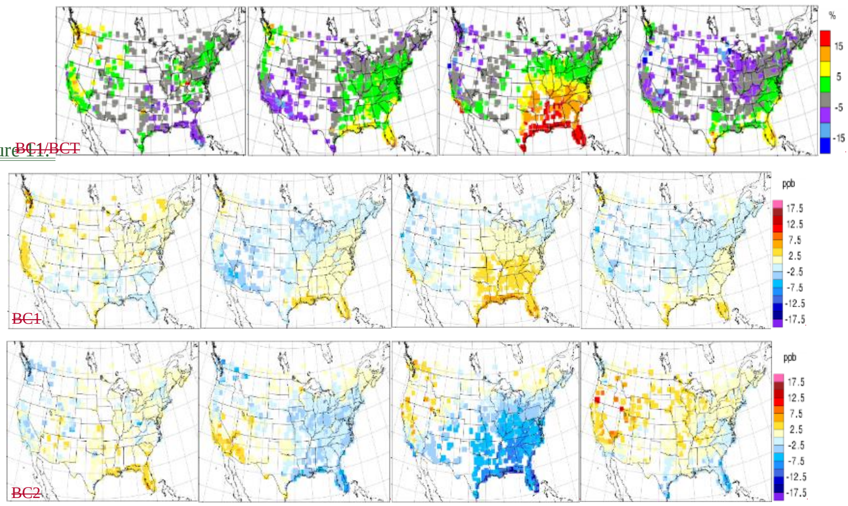


Figure 9: DifferenceDifferences in the seasonal averaged DM8A BC1/BCT between COSMO-CLM/CMAQWRF/DEHM_noddry and WRF/CMAQ (COSMO-CLM/CMAQDEHM (WRF/DEHM_noddry minus WRF/CMAQ) in the seasonal averagesDEHM). The areas in white or grey are the grid cells that are out of (DM8A BC1)/(DM8A BCT) (first row), DM8A BC1 (second row), and DM8A BC2 (third row) at all receptors.

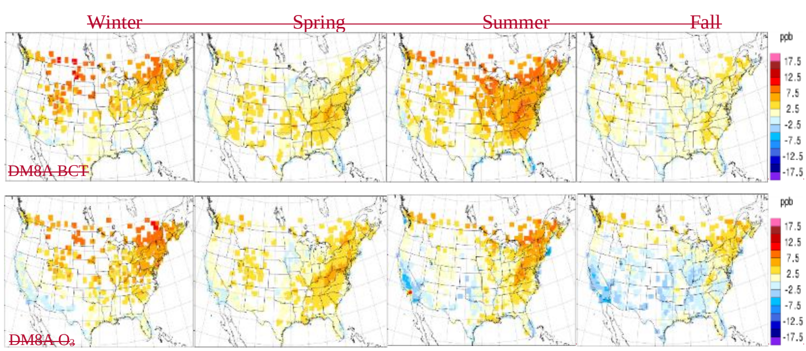


Figure 10: Difference between WRF/CAMx and WRF/CMAQ (the simulation domain of WRF/CAMx minus WRF/CMAQ) in the seasonal averaged DM8A BCT, and DM8A ozone at all receptorsDEHM.

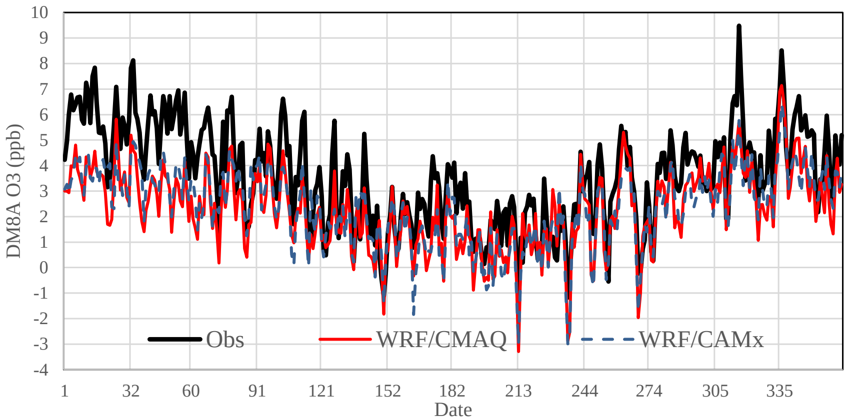


Figure 11: Difference in DM8A ozone between rural and urban sites (rural minus urban) from observation, WRF/CMAQ, and WRF/CAMx.

Formatted: Normal