

Geosci. Model Dev. Discuss., referee comment RC2
<https://doi.org/10.5194/gmd-2021-166-RC2>, 2021
© Author(s) 2021. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on gmd-2021-166

Anonymous Referee #2

Referee comment on "Evaluation of global EMEP MSC-W (rv4.34) WRF (v3.9.1.1) model surface concentrations and wet deposition of reactive N and S with measurements" by Yao Ge et al., Geosci. Model Dev. Discuss., <https://doi.org/10.5194/gmd-2021-166-RC2>, 2021

Review of Ge et al. gmd-2021-166

Overarching comment:

The manuscript provides a comprehensive evaluation of reactive nitrogen concentrations and wet deposition simulated by the WRF/EMEP-MSC-W modeling system against a suite of surface measurement networks over Europe, Asia, and North America. The goal of the manuscript is to establish the credibility of this modeling system in simulating the atmospheric transport and fate of reactive nitrogen on global scales for studying processing affecting concentrations and quantifying the impacts of policy scenarios. The results suggest that the performance of the modeling system is on par with the performance of other contemporary global-scale modeling systems. While I do not expect this general conclusion to change, there are several aspects of the manuscript that should be strengthened prior to its final publication as detailed in my main and specific comments below. The manuscript is well written and structured clearly, the presentation quality of all figures and tables is good, and references to other published studies on similar topics are provided where appropriate. I enjoyed reading Section 4 that provides a nice discussion of the inherent limitations of using surface measurements as criterion for establishing model credibility.

Main comments:

- It is not clear to me whether the authors used the available observations at their native

temporal resolution (e.g. hourly, daily, weekly) and matched these native-resolution observations to the corresponding model values (for non-missing hours / days / weeks) to compute observed and modeled annual averages for the subsequent analyses, or whether they relied on pre-computed observed annual averages provided by some or all of the networks and then matched them to modeled annual averages (computed from all 8760 hours in a given year). Section 2.2 and the discussion in Section 4 suggest the latter approach was taken, but given that this approach can lead to temporal mismatching as discussed in Section 4, I don't see why such a choice would have been made. At least for the North American and European networks I am familiar with (EMEP, NAPS, NTN, AMoN, and EPA Air Data), native resolution data is readily available (e.g. https://aqs.epa.gov/aqsweb/airdata/download_files.html, <http://nadp.slh.wisc.edu/data/NTN/ntnAllsites.aspx>, <https://data-donnees.ec.gc.ca/data/air/monitor/national-air-pollution-surveillance-naps-program/Data-Donnees/2010/ContinuousData-DonneesContinu/HourlyData-DonneesHoraires/?lang=en>) and should be used in the analysis to eliminate this factor of uncertainty.

- Given that one of the key uses of ACTMs such as WRF/EMEP-MS-CW is to quantify the effects of emission scenarios as stated by the authors in both the abstract and summary, an analysis of how well the model captured historic changes in concentrations and deposition (due to changes and variations in both emissions and meteorology) can help establish its credibility for that purpose. The two simulations performed for 2010 and 2015 with the ECLIPSEe emission inventories seem to offer such an opportunity. I strongly encourage the authors to consider adding a section that quantifies emission changes between 2010 and 2015 and then also compares observed and modeled concentration and deposition changes between the two years. While for some pollutants and regions the changes may be dominated by interannual meteorological variability, for others the concentration changes may show a clear link to emission changes.

Specific comments:

Abstract, line 27: change "measurement" to "measurements"

Abstract, line 33: suggest changing "most" and "least" to "largest" and "smallest"

Section 1, line 49: Do we know which components (SIA, OM) drive this increase in PM_{2.5} burden reported by Shaddick et al. (2020)?

Section 1, lines 70 – 75: The work by Tan et al. (<https://acp.copernicus.org/articles/18/6847/2018/>) could also be referenced here

Section 2.1, line 87: consider changing "with implementation" to "used for applications"

Section 2.1, lines 104 – 112: Please provide additional information on the WRF configuration options and input datasets used in these simulations (e.g. land use / land cover database, microphysics scheme, cumulus parameterization scheme, PBL scheme, radiation scheme, land-surface model, etc.). Please also discuss whether the land use / land cover information used in the WRF simulations is consistent with the information used in the EMEP MSC-W biogenic emission and deposition calculations.

Section 2.1, lines 133 - 149: It appears that no day-of-week or hour-of-day temporal profiles were applied to either the ECLIPSEe or HTAP emissions. If this is correct, what was the rationale for not applying such profiles?

Section 2.2, lines 173 – 195: please see my first main comment regarding the use observations with their native temporal resolution.

Section 2.2, lines 181 – 185: What was the rationale for not including wet deposition measurements from CAPMoN (<https://www.canada.ca/en/environment-climate-change/services/air-pollution/monitoring-networks-data/canadian-air-precipitation.html>) to strengthen the analysis in Section 3.3?

Section 2.2, lines 186 – 187: Please make sure to clearly distinguish between “R” and “RT” in the supplement and their first uses in the text, tables, and figures.

Section 3.1.1., line 216: consider changing “total gridded differences” to “the total number of grid cells”

Section 3.1.2, lines 292 – 295: I suggest extracting and comparing the HTAP and ECLIPSEe emissions for the grid cells corresponding to these locations (maybe in addition to some surrounding grid cells) to gain further insights into the relative importance of potential localized emission differences (i.e. localized model uncertainty) vs. localized phenomena affecting the measurements.

Section 3.1.2, line 328: I suggest avoiding the term “bias” since the analysis presented here does not constitute a rigorous verification of emission inventories.

Section 3.2.1, lines 374 – 375: “there is no significant difference between model and measurements in most Southeast Asia countries” – This seems like an overstatement given that FAC2 is only about 0.5 for EANET NH3.

Section 3.2.1, lines 394 – 397: This again seems like an overstatement and selective reading of evaluation metrics. The FAC2 values for many of the EANET pollutant measurements are quite low, and many of these FAC2 values as well as NME and NMB values are worse for EANET than NNDMN.

Section 3.2.3, lines 471 – 472: What was the rationale for not including AIRBASE over Europe? If it was included, the number and type of sites might be more comparable to the NO₂ and SO₂ measurement sites over North America.

Section 3.2.3, lines 474 – 476: It is likely the location of the sites which are often source-oriented. SO₂ emissions from the largest sector (power generation) are directly measured and included in the EPA emission inventory that was used in HTAP. Does the ECLIPSEe inventory not use this available information about measured SO₂ emissions from the U.S. power sector?

Section 3.2.3, lines 480 – 481: As for SO₂, this is likely due to the local-scale nature of many NO₂ measurement sites.

Section 3.2.3, line 482: I suggest replacing “secondary” with “aerosol” since there are other secondary pollutants (e.g. HNO₃) in Table 4 besides sulfate, nitrate, and ammonium.

Section 3.2.4, lines 505 – 506: As noted in my first main comment, higher temporal resolution data is also available over the U.S. and Canada and should be used here as well.

Section 3.2.4, Figure 11: Please add results for available North American networks (e.g. AMoN) here and in Figure S6.

Section 4, lines 666 – 676: Please see my first main comment about using observational data at their native temporal resolution for the analysis presented in this manuscript.

Section 4, lines 704 – 705: “unmatched temporal coverage” should not be an issue if model-observation matching is performed at the native temporal resolution of the observations as discussed in my main comment.

Section 5, lines 763 – 764: Please see my second main comment about adding an analysis of observed and modeled 2010 vs. 2015 changes in concentrations to evaluate the

modeling system's ability to capture such changes which are at least partly driven by changes in emissions.