

Title: Leveraging JEDI for Atmospheric Composition: A unified framework for evaluating observations and model forecasts

Authors: Wei et al.

General comments

This manuscript describes leveraging JEDI components, along with MetPlus, to evaluate model forecasts against observations for the atmospheric composition applications, mostly benefiting from JEDI's modular and model-agnostic design. The authors demonstrated this capability through various cases comparing model equivalents from wrf-chem and MERRA-2, with widely-used observations in this community such as trace gases, AOD, surface PM2.5 in the ground-based or spaceborne platforms. As pointed out by the authors, the manuscript's focus is more on a technical demonstration of this capability rather than scientific findings. This is appropriate and falls within the GMD scope. To possibly help its scientific aspect, however, I have the following general comments.

1. Some of this described capability has been explored and tested in multiple JEDI-related scientific publications. I would encourage the authors to do more relevant literature review and to cite these papers where appropriate. It should help complement the current technical focus on this manuscript by supplementing additional scientific validations from others.
2. Description and interpretation of some figures and tables is insufficient or irrelevant (please see specific comments below). In several instances, the manuscript simply says JEDI has this capability and then refers to figures or tables, with no further explanatory description or summarization about them. While this manuscript is more technical-focused, an accurate and even short description of included figures or tables would be needed for a scientific publication to enhance readers' understanding.

As JEDI is undergoing rapid development and will be widely used for future research and operations in our community, this manuscript would be a great contribution. Overall, this manuscript is very nicely written and organized. I believe the above comments are easy to address. Therefore, my recommendation to the editor is minor revision (see specific comments as follows).

Specific comments/Technical corrections

- Line 100: Typo: "Model forecasts states" → "Model forecast states"

- Section 3.7: Please clarify VIIRS AOD used in this manuscript is NOAA product or NASA product. Because their algorithms are slightly different over the land.
- Figure 3: Please add some interpretation for Fig. 3 and be more relevant and accurate. Otherwise, it is not clear why it is included here. Especially, when compared with different obs, the model forecast opposite biases, e.g., slightly positive bias wrt TROPOMI and large negative bias wrt TEMPO. How will this “*offers insight into systematic over- or underestimation patterns in time that can inform on atmospheric model process and emission inventory improvements*” ? Are comments about whether to show diurnal cycles in TEMPO and TROPOMI around Line 315 reflected in Fig. 3?
- Table 1: Similar comments as Figure 3.
- Line 360: Is VIIRS AOD NOAA/NESDIS product or the NASA product, because they apply different algorithms
- Figure 7, 8 and Table 2: Similar comments as Figure 3. It’s not clear why they are suggestive of “*This highlights for example the utility of AERONET as a benchmark for both model evaluation and observation operator performance.*”.
- Near Line 410: $H(x)$ is called “observation equivalents”, but “model equivalents” elsewhere. Please be consistent and use “model equivalents” throughout the manuscript.