Response to the Editor Shu-Chih Yang:
Citation: https://doi.org/10.5194/egusphere-2023-2299-EC1

Thank you very much for your own review. In response to your comments, we provide four figures in the supplement as Figs. S1-S4, adding a few paragraphs in the manuscript (Page 10). Please find our point-by-point response to your valuable comments below. Our responses are marked in blue.

Although the authors claim that the manuscript is categorized as a "technical" paper, the associated implementation is successful with **comprehensive** justification. I agree with the reviewer's comments about showing the results related to water vapor since it is vital to know whether implementing the IAU method affects the moisture accuracy of MPAS-A.

- ⇨ We totally agree with you that any new implementation should be thoroughly investigated for correctness. For that, we have cross-checked all mathematical expressions and their corresponding codes multiple times. However, we want to take this opportunity to clarify that ensuring successful development is distinct from providing comprehensive justification for the impact study, which involves various scientific aspects. We also note that the main motivation and purpose of the IAU implementation are to effectively suppress initial noise resulting from dynamic imbalances, as already demonstrated in Figs. 2 and 3.
- ⇨ But we agree that it would be interesting to examine the impact of IAU on moisture as part of the control variables. In response to your comment, we provide two figures below.
- ⇨ First, we compare water vapor mixing ratio (Qv [g/kg]) in 6-h forecasts between CTRL and IAU against ERA-5 analysis globally over a total of 92 6-hourly cycles from April 21 to May 13, 2018 (e.g., after a one-week spin-up period). Although root-mean-square errors are the same as 0.7 in both experiments (not shown), the systematic bias indicates that IAU produces slightly better agreement with ERA-5 analysis than CTRL, especially in the boundary layer (< 2 km).

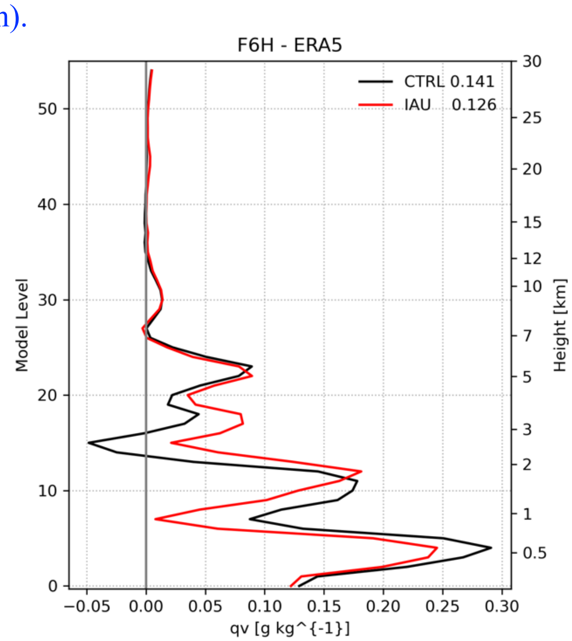

Figure S1. Vertical profile of water vapor mixing ratio (Qv [g/kg]) in 6-h forecasts from CTRL (black) and IAU (red), compared to the ERA-5 analysis, averaged over the cycles from April 21 to May 13, 2018. The mean error below 3 km is indicated next to each experiment name.

⇨ We also present a panel plot below, where we verify the analysis and the first-day forecast against sounding observations between the surface and 250 hPa for the entire month. Compared to the CTRL run, RMS errors are slightly worse at the analysis time (which corresponds to a 3-h forecast in the IAU run, as opposed to the analysis in CTRL) by up to 2% in (a). But in (b), as forecasts start from the analysis, the errors decrease during the first 6 h, and exhibit a statistically significant reduction of approximately 4% at 18 h in the troposphere. Hence, we cannot claim that IAU consistently improves moisture analysis and forecasts, but it is fair to say that it does lead to some improvements.

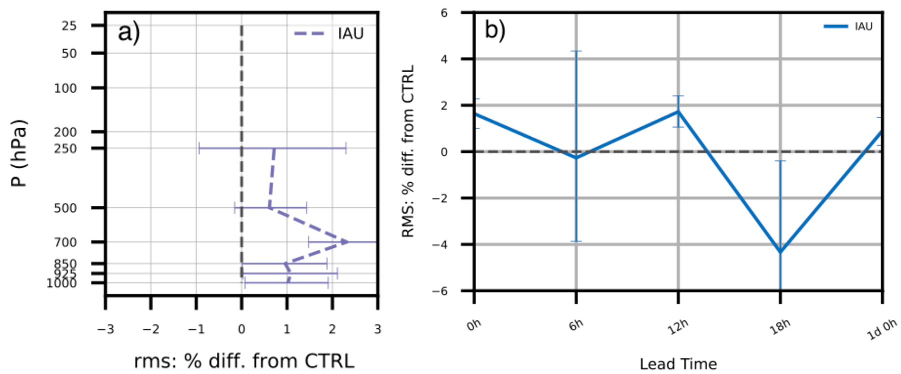

Figure S2. The Qv rms differences from the CTRL in percentages in IAU verified against radiosonde observations in a) the vertical profile at the analysis time and b) the time series of subsequent 24h forecasts of the corresponding metrics between 250 and 1000 hPa.

⇨ As we believe that it would be best to stay focused on surface pressure to represent the entire column, we have decided to include them in the Supplemental material, with the comment in the text (L282) as "The moisture verification for the 6-h forecast is provided in the supplement.". Thank you for your understanding.

In addition, it is worth addressing the issue of the degradation of the northern polar region with IAU. According to Fig. 7b, the rms difference is about 20% in the north polar region and 10% near 25-degree S during the first-day forecast! If the authors attributed the difference to their own error in GFS analysis, I suggest showing the same figure with other analyses as the reference, like the EC analysis.

⇨ Thank you for your careful review on Figure 7. We admit that it is challenging to explain the forecast degradation in a specific region, especially given that we did not find anything particularly suspicious in Fig. 3. We fully understand the editor's concern, though, and we would like to examine the issue through the observation-space diagnostics, rather than the verification against any particular analysis. This work was meant for the analysis cycling (with our own data assimilation), so it is legitimate to check the performance with respect to observations. In Figure S3, where we verified our background forecasts for surface pressure against measurements, we noticed one red dot near the North Pole (to the north of Greenland) in the bottom panel. And there are more red dots to the north of Russia, all of which seemed to contribute to the larger errors shown in Fig. 7b. Also, you correctly captured the red area near 25°S, which seems associated with the red dots in the southern part of the tropics, as depicted in Fig. S3b. But as you can see through the global distribution, we do not have enough observations near the Poles and over the ocean to sufficiently constrain the model state. The month-long error statistics reveal almost no colors (e.g., little deviations) over the CONUS domain with a dense observing network, indicating that IAU itself does not degrade forecast errors when the model states are well constrained through data assimilation. It remains unclear whether the impact of IAU has intertwined with model errors in certain data-sparse regions. Note that model errors are not accounted for in this hybrid 3D-EnVar framework.

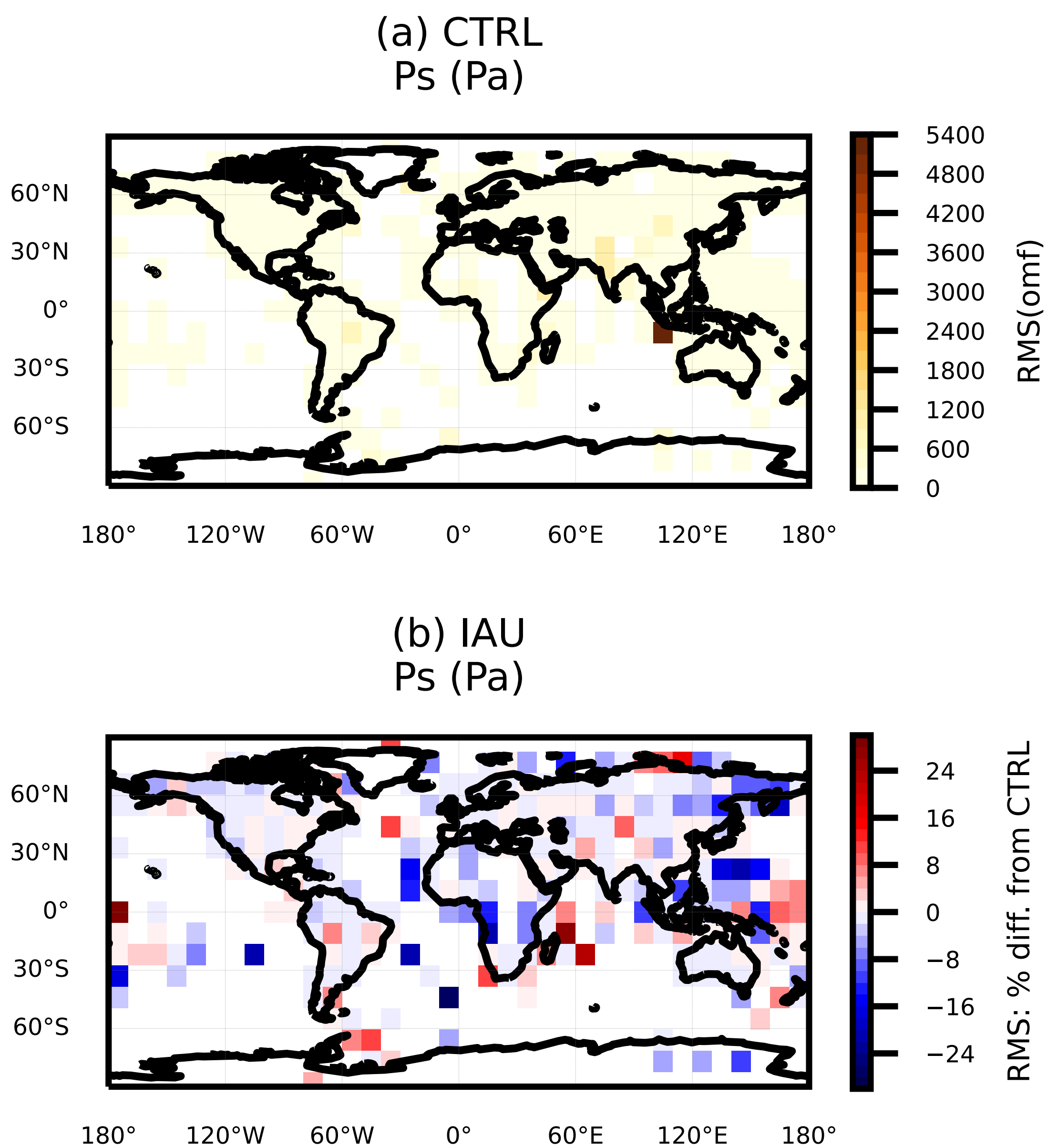

Figure S3. The global distribution of (a) the RMS errors of the CTRL run in 6-h background forecasts for surface pressure ([Pa]), verified against surface observations and (b) the RMS differences from the CTRL run in IAU.

- It is worth noting that *<u>we implemented IAU on the model's unstructured mesh in a generic way</u>*, not specifically tied to geographic regions. For clarity, following the statement "Compared to the GFS analysis, MPAS forecasts in CTRL exhibit the largest (or the fastest) error growth in the Southern Hemisphere. Forecasts in the IAU run, on the other hand, tend to reduce errors in the tropics while increasing errors near the North Pole region.", we added a paragraph in L289-295, stating "This aligns with the findings of Ha et al. (2017), where forecast errors were significantly reduced over the tropics in a variable-resolution mesh, including both resolution-transition and high-resolution parts. Because IAU is implemented on the model's unstructured mesh (which is in a random order), it is not associated with particular geographic locations or mesh configurations. Given its time filtering feature, IAU might be more effective in simulating low-frequency modes dominant over the tropics. It is also noted that the impact of IAU may be nonlinearly intertwined with model errors in data-sparse regions, such as the Poles. However, model errors are not accounted for in the hybrid 3DEnVar system used in this study. Additional area-specific features in the verification are provided in the supplemental material.".

- Fig. S4 further supports our point that the IAU significantly improved forecast errors in most regions, except for the North Pole area and the Southern Hemisphere Ocean, in a statistically significant manner. As illustrated in Fig. 4 in the manuscript, the IAU run assimilated slightly more observations (by 1-2%) throughout the month-long cycling, which is a good indicator that the DA cycling system work more effectively. Overall, it is our belief that our IAU implementation was successfully completed, with positive impacts on the analysis and short-term forecasts.

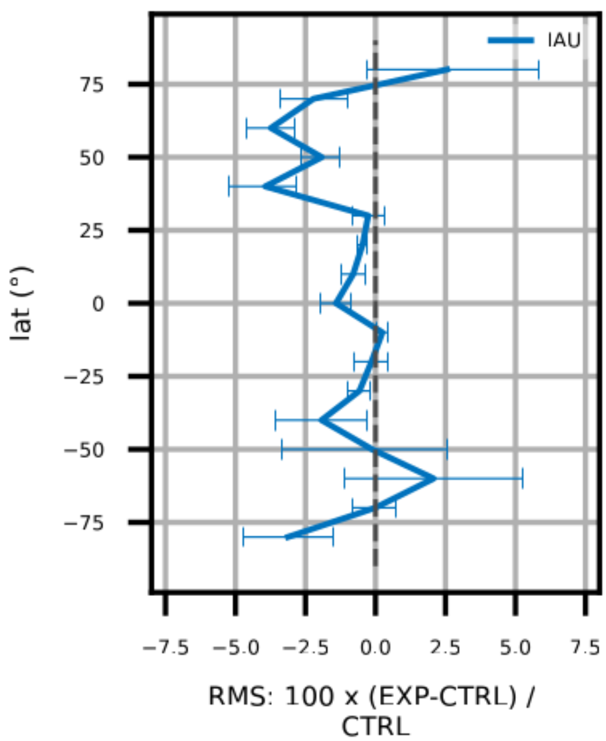

Figure S4. The rms percentage differences from the CTRL in IAU for 6-h background forecasts in surface pressure across latitudes. The error bars denote the standard deviation, corresponding to a 95% confidence level.