We thank the reviewer for these thoughtful comments on our manuscript. We have provided point-by-point responses to address your concerns as follows:

1. The manuscript defines a successful genesis prediction as one in which the model-predicted genesis time falls within ±24 h of the observed genesis time. However, this criterion does not appear to be applied consistently in Table 2 and the related discussion. For example, Fiona in 14kBestRO is predicted at +36 h, which is outside the stated ±24 h window. Therefore, Fiona should not be described as a successful case unless the definition is changed. The authors may instead present Fiona as a delayed-genesis case with improved pre-genesis structure, or select another case that satisfies the stated success criterion. This revision is important for maintaining consistency between the methodology and the case-study interpretation.

   Thank you for your feedback regarding the consistency of the ± 24-h success criterion. We acknowledge that our interpretation of 'hits' in the original manuscript led to ambiguity, particularly concerning cases like Fiona in the 14kBestRO experiment.

   You are correct because Fiona was predicted at +36 h; it falls outside our stated ± 24-h success window and should not be counted as a hit. We will revise the manuscript to strictly adhere to the ±24-h criterion for all POD calculations. We will update the text to classify Fiona as a 'delay-genesis case' rather than a successful hit, and we will clarify the terminology for non-intensifying cases. This revision aligns our methodology with the interpretation of individual case studies and ensures full consistency throughout Table 2 and subsequent discussion.

2. In addition, the treatment of cases marked by triangles also requires clarification. If these short-lived or non-persistent genesis signals are counted as 0.5 hits, the authors are suggested to clearly distinguish among genesis detected within the ±24 h window, genesis detected outside the verification window, and short-lived genesis-like signals that do not persist or intensify.

   Thank you for the suggestion to improve the clarity of our classification criteria. We agree that distinguishing between these different types of events is necessary to avoid ambiguity.

   In our revised manuscript, we will clarify our methodology to distinguish between the following three categories:

   (1) Successful Hits: Genesis events where the model-predicted time falls within the ±24-h verification window. These are the only events included in the POD calculation.
   (2) Delayed/Early Genesis: Genesis events detected by the model but occurring outside the ±24-h verification window. These are tracked in Table 2 (marked as + or -) but are not counted as hits.

(3) Non-intensifying Disturbances (Triangles): Signals where initial genesis is detected, but the system fails to persist or intensify. These are explicitly distinguished from ‘hits’ and are not included in the POD calculation.

We will remove the ‘0.5 hit’ terminology entirely, as it was inconsistent with this classification framework. By clearly separating these three categories in the revised text and Table 2, we ensure that our quantitative results are consistent with our defined verification criteria.

3. The study includes six tropical cyclones from the 2022 Atlantic hurricane season. The best-performing experiment, 14kBestRO, detects only a limited number of genesis events, and two cases, Ian and Lisa, are not predicted by any experiment. Therefore, broad statements that increased and more homogeneous RO coverage improves tropical cyclogenesis prediction should be tempered. If increasing the sample size is beyond the scope of the present study, the authors may provide uncertainty estimates for the probability of detection or revise the wording to emphasize that the results are suggestive rather than conclusive.

   We thank the reviewer for this thoughtful comment. We agree that the six selected tropical cyclones do not constitute a sufficiently large sample to draw broad conclusions regarding the impact of enhanced RO sampling on tropical cyclogenesis prediction. We intended to demonstrate the potential impact of increased and more homogeneous RO sampling using representative cases from the 2022 Atlantic hurricane season, rather than to establish a statistically robust conclusion.

   To address the reviewer’s concern, we will revise the wording throughout the manuscript, including the Abstract, Results, and Conclusions, to emphasize that the results are specific to the selected cases and are suggestive rather than conclusive. We will also discuss the limited sample size as a limitation of the present study and note that a larger set of tropical cyclogenesis events will be needed to establish a statistically robust conclusion regarding the benefits of enhanced RO sampling.

4. The manuscript states that increased RO data coverage helps improve the pre-genesis environment, with increased specific humidity and relative vorticity. This statement should be qualified. It is unclear whether the assimilation of RO observations always leads to increases in moisture and vorticity. If so, what is the physical explanation for this systematic increase? If not, please revise the statement to more accurately reflect the case-dependent impact of RO assimilation.

   We thank the reviewer for this insightful comment. We agree that our original wording was too general and misleading. The assimilation of RO is expected to produce more accurate specific humidity and relative vorticity analyses. The analysis increment may be positive or negative. We will revise the statement to more accurately reflect the case-dependent impact of RO assimilation.

Our diagnostics show that differences in analysis across experiments during the data assimilation period are generally modest and vary among assimilation cycles (Figure 10). For example, in the Hurricane Fiona case, enhanced RO sampling produces higher relative vorticity during some assimilation cycles, while the differences in relative humidity are relatively small. However, these analysis differences become more pronounced during the subsequent forecast period, with the experiments using enhanced RO sampling developing a more favorable thermodynamic and dynamic environment characterized by higher moisture and stronger cyclonic vorticity. This suggests that the improved pre-genesis environment results from the nonlinear evolution of the forecast initialized from improved analyses, rather than from a systematic increase in moisture or vorticity directly caused by RO assimilation.

To address the reviewer's concern, we will revise the corresponding statements throughout the manuscript (including the Abstract and Section 4.2) to clarify that the enhanced moisture and cyclonic vorticity are associated with the forecast evolution in the selected cases and should not be interpreted as a universal consequence of assimilating additional RO observations.

5. Fig. 3 appears to illustrate the tracking algorithm's criteria, but its role in the main argument is not sufficiently clear. If the figure is only illustrative, it could be moved to the supplement. Similarly, Fig. 4 shows Gaston in the GFS analysis and in the three experiments, but the manuscript does not clearly explain what this figure is meant to demonstrate. Please clarify the purpose of the two figures.

   We appreciate the suggestion regarding the illustrative nature of these figures. We agree that their presence in the main body was not essential to the primary argument, and we will therefore relocate them to the supplementary material.

6. Fig. 5 shows that the number of assimilated RO profiles in 14kPoorRO is only slightly larger than that in 8kRO in the mid- and lower troposphere, and is sometimes even smaller at certain DA cycles. Within the 500-km radius of the vortex center, the number of RO profiles in 14kPoorRO is also comparable to that in 8kRO. It is therefore unclear why Fig. 7 shows broadly positive impacts, as indicated by negative relative RMSE differences, for 14kPoorRO. Could the authors further elaborate on this?

   We thank the reviewer for this insightful comment. We agree that the total number of assimilated RO profiles in the model domain, particularly within 500 km of the vortex center, is comparable between the 8kRO and 14kPoorRO experiments. However, the forecast impact is influenced not only by the number of assimilated observations but also by their spatial and temporal distribution. Unlike the relatively uniform sampling 8kRO, the 14kPoorRO observations are concentrated within specific longitudinal bands (e.g., Figure 1c), and these bands shift every 6 hours. Consequently, although the total number of assimilated profiles is similar, the atmospheric regions sampled at each data assimilation cycle differ between the two experiments. This evolving sampling pattern provides additional information over different portions of the atmosphere/regions throughout the assimilation period, which likely contributes to the improved forecasts shown in Fig. 7.

7. In addition, Fig. 7 shows that 14kPoorRO performs better than 14kBestRO in several upper-level variables. However, the profile counts shown in Fig. 5 do not clearly demonstrate that 14kPoorRO has more upper-level RO data than 14kBestRO. The manuscript attributes this improvement to increased RO sampling at high latitudes and possible improvements in upper-level jets, but this explanation is not sufficiently supported by the current diagnostics. The authors are encouraged to provide additional evidence to support this interpretation.

   We thank the reviewer for this helpful comment. We agree that the original manuscript did not provide sufficient evidence to support our interpretation regarding the impact of RO sampling on upper-level circulation. To address this concern, we performed additional diagnostics of the wind forecast errors at model level 35 (~246 hPa) averaged over the six selected tropical cyclone cases. The results show that the 14kPoorRO experiment produces larger reductions in both zonal and meridional wind RMSE compared with the 8kRO and 14kBestRO experiments, with particularly pronounced improvements at higher latitudes (Figure R1). These results suggest that the spatial distribution of RO observations in 14kPoorRO contributes to improved representation of the upper-level circulation. We will revise the manuscript to clarify that the improved upper-level forecast skill is associated with differences in RO sampling distribution rather than simply the number of assimilated observations.

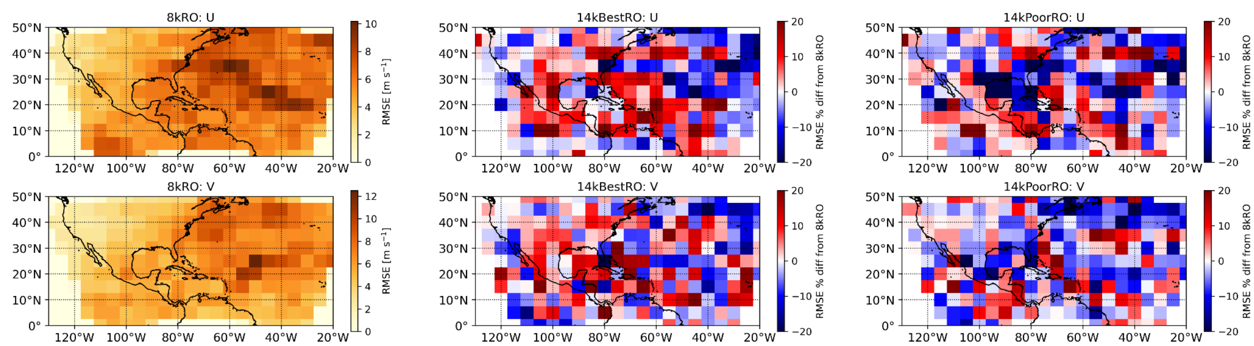

Figure R1. Spatial distribution of root mean square error (RMSE) for the zonal (upper panels) and meridional (lower panels) wind component forecasts. The 8kRO experiment serves as the baseline (left), with the 14kBestRO (middle) and 14kPoorRO (right) experiments expressed as a percentage difference relative to the baseline. Results are averaged across a selection of six tropical cyclone cases throughout the 120-h forecast period.

8. In relation to the Fiona case, Fig. 10 shows the evolution only up to the observed genesis time, whereas the model-predicted genesis in 14kBestRO occurs 36 h later. Therefore, Fig. 10 does not directly diagnose the simulated genesis process in the model. The authors are encouraged to extend the analysis to the model-predicted

genesis time, or at least include additional diagnostics around that time, to better support the interpretation of why genesis occurs in the 14kBestRO experiment.

We thank the reviewer for this valuable suggestion. To better diagnose the simulated genesis process, we extended the analysis to the end of the forecast period and added three new diagnostics: (1) an extended version of Fig. 10 (Figure R2), (2) maps of 700-hPa relative humidity and 500-hPa potential vorticity at the end of the forecast (Figure R3), and (3) time series of 850-hPa relative vorticity and 700-hPa specific humidity averaged within a 500-km radius of the vortex center (Figure R4).

The additional diagnostics indicate that the successful prediction of genesis in the 14kBestRO experiment is associated with a more persistent evolution of the cyclonic circulation together with a more favorable thermodynamic environment. Compared with the other experiments, 14kBestRO maintains stronger cyclonic vorticity through much of the middle troposphere during the forecast and consistently higher 700-hPa specific humidity before the predicted genesis time (t = 72 h; Figures R2d-e). At the end of the forecast, the 500-hPa potential vorticity field also shows a better-organized disturbance in 14kBestRO than in 14kPoorRO (Figures R3c-d).

Although the 14kPoorRO experiment develops stronger low-level relative vorticity and higher upper-level relative humidity after approximately forecast hour 60, these favorable characteristics occur relatively late in the forecast and are not accompanied by the earlier sustained evolution of moisture and cyclonic circulation found in 14kBestRO (Figure R2e and Figure R4). Consequently, these late enhancements are insufficient to compensate for the less favorable evolution during the earlier pre-genesis stage. The additional diagnostics suggest that the successful prediction of genesis in 14kBestRO is associated with the improved temporal evolution of the combined thermodynamic and dynamical environment.

We will revise the discussion in Section 4 accordingly to emphasize the forecast evolution leading to model-predicted genesis, rather than relying solely on comparisons at the observed genesis time.

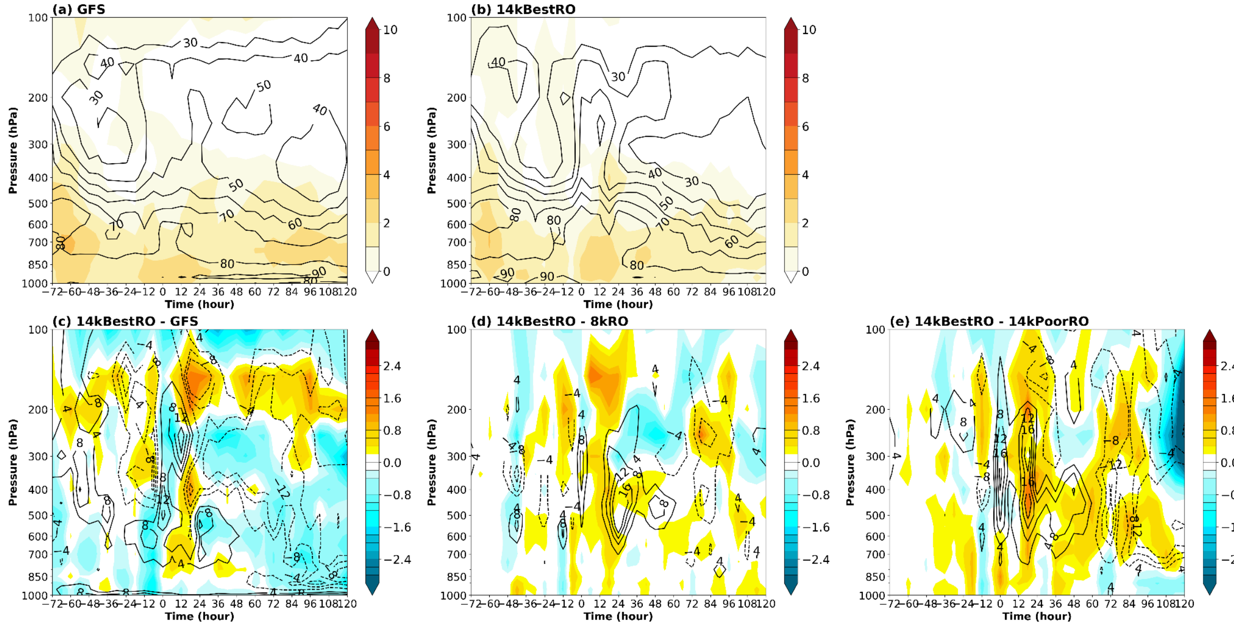

Figure R2. Same as Fig. 10 but extending to the end of the forecast period.

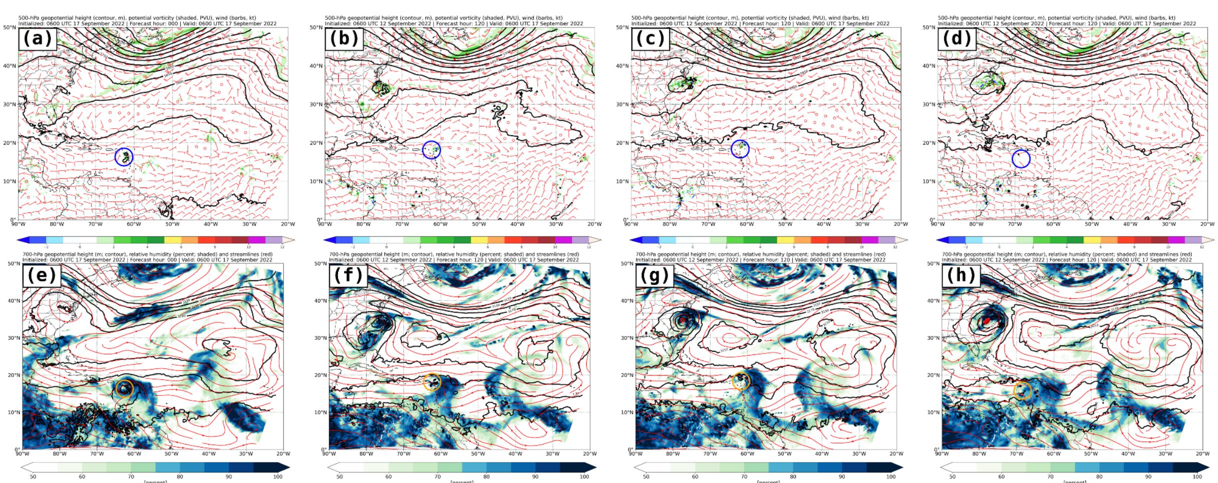

Figure R3. (a-d) 500-hPa geopotential height (black; m), potential vorticity (shaded; PVU), and wind barbs (red), and (e-h) 700-hPa geopotential height (black; m), relative humidity (shaded; percent), and streamlines (red) from (a, e) the GFS analysis, (b, f) 8kRO, (c, g) 14kBestRO, and (d, h) 14kPoorRO experiments at 0600 UTC 17 September 2022. The model forecast was initialized at 0600 UTC 12 September 2022. The observed TC location is at (16.3˚N, 62.6˚W), marked by a blue or yellow circle. The predicted 700-850hPa circulation center locations from the DA experiments are marked by blue or yellow circles.

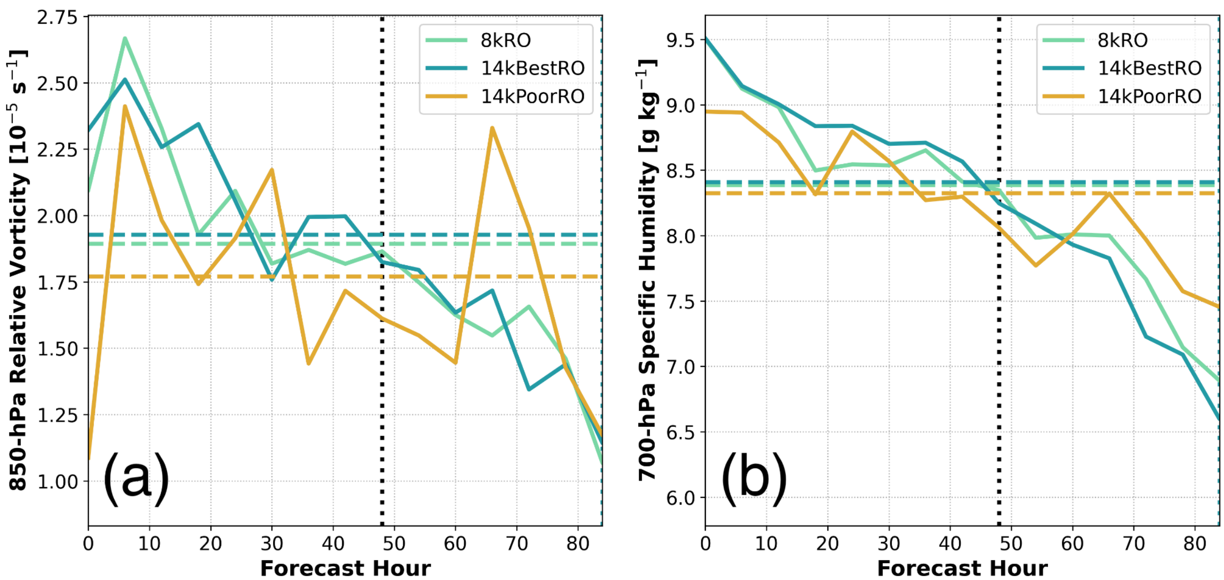

Figure R4. Time series of (a) 850-hPa relative vorticity (10^{-5} s^{-1}) and (b) 700-hPa specific humidity (g kg^{-1}) from forecast hours 0 to 84 for the 8kRO (light green), 14kBestRO (blue), and 14kPoorRO (gold) experiments. All parameters are averaged within a 500-km radius of the vortex center. Horizontal dashed lines denote the time-mean value for each experiment over the forecast period. Vertical dotted lines indicate the genesis times: observed (black) and 14kBestRO predicted (blue).

9. Some cases, especially Ian, Julia, and Lisa, occur near the Caribbean Sea and Central America. Terrain interaction, coastline proximity, sea-surface temperature representation, and model resolution may affect low-level circulation, moisture convergence, and cyclogenesis. The 15-km WRF domain may also limit the ability to represent convective organization during genesis. Could the authors discuss how these geographical and model-configuration factors may influence the failed or delayed genesis forecasts?

   Thank you for this insightful question. The influence of regional geography and model configuration on tropical cyclogenesis predictability in the Caribbean Sea is complex. Regarding model configuration, while our 15-km WRF simulations utilize convective parameterization, this resolution likely limits the representation of critical convective organization. We note that high-resolution (3-km) sensitivity tests conducted for Hurricane Ian, which allow for explicit convection, did not yield genesis, suggesting that resolution alone may not resolve these forecasting challenges. This indicates a need for further investigation into how grid spacing interacts with other physical processes.

   Geographically, the complex terrain of Hispaniola and coastal proximity likely contribute to the disruption of convective intensification. Furthermore, synoptic-scale interactions play a significant role; as noted by Shieh and Colucci (2010), low-level divergent winds across the Caribbean can inhibit convective development. We acknowledge that these multifaceted geographical and model-related factors contribute to the failed or delayed genesis forecasts. We will update the manuscript to reflect these limitations and recognize these areas as key priorities for future research into Caribbean tropical cyclogenesis predictability.

Reference:

Shieh, O. H., and Colucci, S. J.: Local minimum of tropical cyclogenesis in the eastern Caribbean. Bull. Amer. Meteor. Soc., 91, 185-196, https://journals.ametsoc.org/view/journals/bams/91/2/2009bams2822_1.xml, 2010

10. The differences among the experiments in Figs. 11 and 16 are not easy to identify with the current color scales. It is encouraged to adjust the colorbar ranges or use more appropriate color intervals to enhance the contrast among experiments. This would make the differences in relative vorticity more visually distinguishable and help readers better assess the physical interpretation discussed in the text.

    We appreciate this constructive suggestion. By adjusting the colorbar scales, the contrasts across the different panels are now clearly distinguishable. Furthermore, we have resolved a previous omission by including an analysis of the dry air intrusion noted at 600-800 hPa between t = -12 h and t = 0 h within the 14kPoorRO experiment (Figure R5d). This particular intrusion of dry air coincides with a transient reduction in relative vorticity, a feature that we will integrate into our discussion to provide a more comprehensive characterization of the pre-genesis environment.

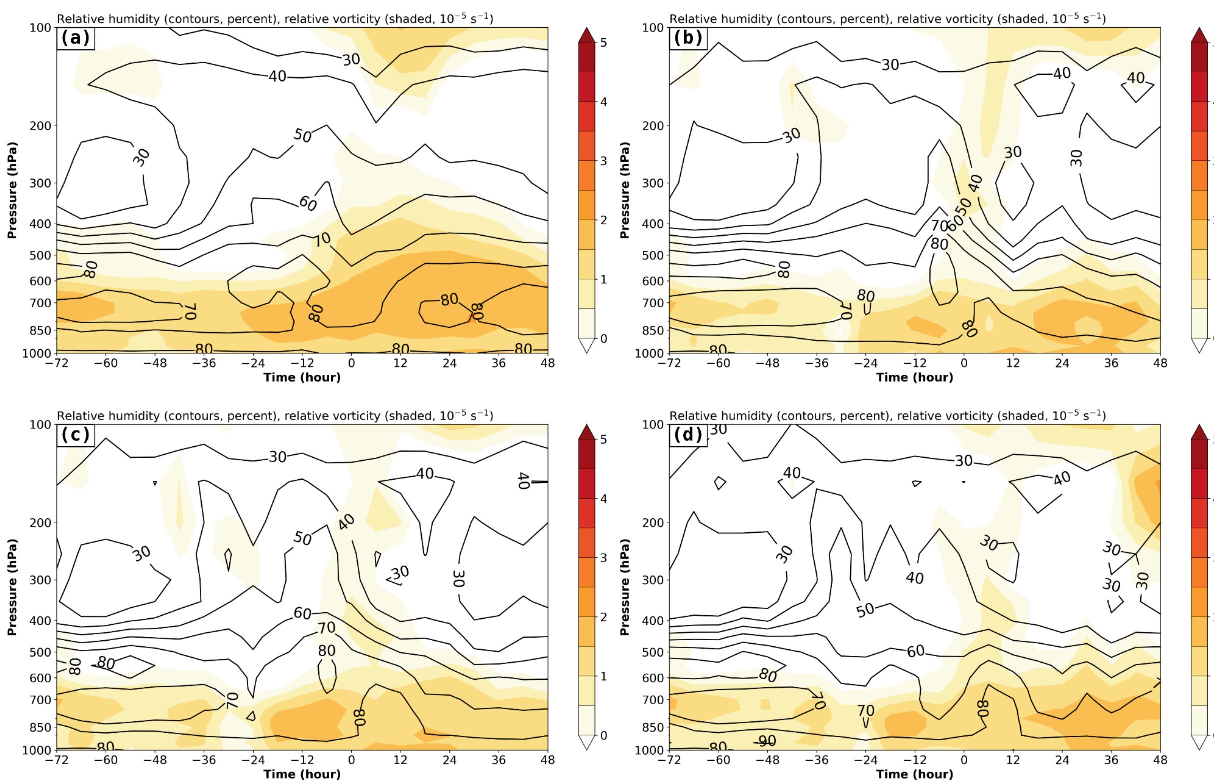

Figure R5: Same as Fig. 11, but with new colorbar ranges.

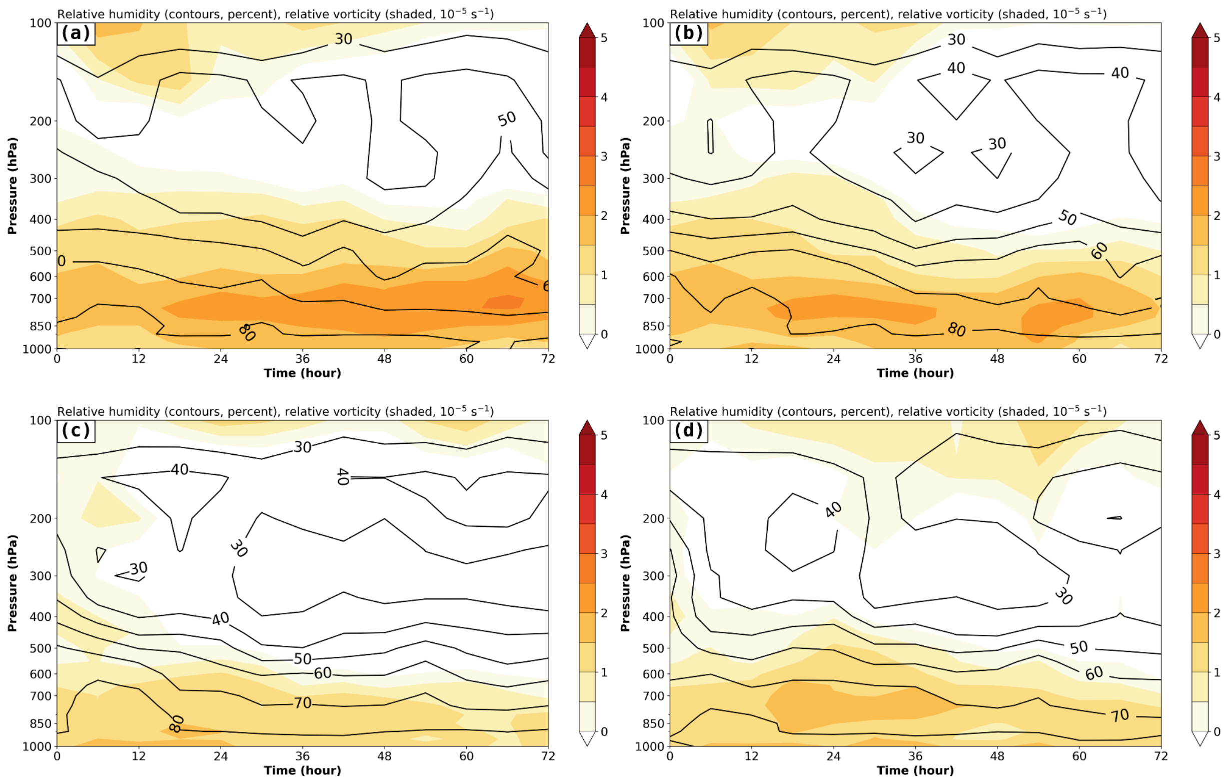

Figure R6: Same as Fig. 16, but with new colorbar ranges.