The authors sincerely thank anonymous referee for the careful reading and constructive suggestions. The comments were very helpful for improving both the clarity of the manuscript and the rigor of the mechanism discussion. Below we provide a point-by-point response. Reviewer comments are reproduced in black, and our responses are given in blue.

General Comments

This work presents an updated version of the orographic gravity wave drag parameterization for the CMA-GFS model. The revised scheme accounts for the impact of non-hydrostatic mountain wave dynamics on surface wave momentum flux and the resulting wave forcing. This development is increasingly relevant as model horizontal resolution continues to improve, making purely hydrostatic assumptions less adequate. The motivation for this work is well founded. The reported improvements are generally positive, although at times relatively modest; nonetheless, they contribute to alleviating persistent model biases. Interestingly, the largest improvements are found in the Southern Hemisphere summer, where however wave drag is weaker due to seasonal conditions, and the authors do not focus much on it.

Reply: The study focuses primarily on the NH because the experiments are conducted in boreal winter, when parameterized OGWD is strongest in this hemisphere. By contrast, the SH is in austral summer, during which stratospheric OGWD is much weaker because of critical-level absorption, leading to comparatively smaller circulation responses. Therefore, the quantitative evaluation of forecast skill is mainly performed for the NH, where the impact of the revised OGWD scheme is physically more pronounced and statistically more robust.

One other general consideration I have is about the evaluation of the updated parameterization against ERA5. While ERA5 provides a high-quality reanalysis dataset, its horizontal resolution (0.25° × 0.25°) is coarser than that of the CMA-GFS simulations used here (0.125° × 0.125°). It would therefore be useful for the authors to clarify why ERA5 is considered a more reliable reference for representing non-hydrostatic effects despite its lower resolution. Presumably, this is due to the assimilative nature of reanalysis products, but this justification should be explicitly stated in the manuscript.

Reply: Thanks. In the revised manuscript, "*The model initial conditions are derived from the 0.25° × 0.25° ECMWF Reanalysis v5 (ERA5) dataset (Hersbach et al., 2020), which are also used as reference for the evaluation of the CMA-GFS forecasts. Although the horizontal resolution of ERA5 is coarser than that of the CMA-GFS v4.0 simulations, ERA5 is adopted here as the verification reference because it is a dynamically consistent reanalysis constrained by a broad range of assimilated observations. In this study, ERA5 is not used to resolve the non-hydrostatic subgrid-scale orographic gravity waves themselves. Instead, it serves as a benchmark for evaluating the large-scale circulation response and medium-range forecast skill associated with the revised OGWD parameterization.*" is added in section 2.b.

Specific Comments

Lines 82-86: is this not simply the inverse of Na/U? - which is a well-known concept (see e.g., Zangl et al, 2003, Guarino et al., 2017, etc.).

Reply: The quantity introduced there is indeed the inverse of the well-known parameter Na/U.

Line 103: "along with many other improvements in the model dynamics and physics." Here a reference, if available, would be needed.

Reply: Thank you, the reference is added (Shen et al., 2023).

Shen, X. S., Su, Y., Zhang, H. L., et al.: New version of the CMA-GFS dynamical core based on the predictor-corrector time integration scheme, Journal of Meteorological Research, 37, 273-285, https://doi.org/10.1007/s13351-023-3002-0, 2023.

Lines 114-116: "partitioning the momentum stress with the Scorer parameter *when model grid mainly locates at the downstream of the subgrid orography*." This is not clear at all, I have read it multiple times and I still don't understand what this means.

Reply: In the revised manuscript, it is rewritten as: "It is noticed that in the existing implementation, nonhydrostatic effects are only partly reflected through a Scorer-parameter-based partitioning of the momentum stress when the model grid point is located downstream of the subgrid orography. However, the launch-level surface WMF itself still follows the original hydrostatic KA95 formulation" .

Lines 144, 146: are (1) and (2) from KA95?

Reply: Yes, equations (1) and (2) are taken from the original KA95 formulation.

Line 147: how is "low level" defined? Its definition is important here.

Reply: The low level is from the surface to $2\sigma_h$ (standard deviation of SSO height).

Line 155: it is not clear what these two constants are, as they have not been introduced before.

Reply: The two constants are empirical constants obtained through extensive experiments.

Line 169: define explicitly tau tilde.

Reply: $\tilde{\tau}$ represents the nonhydrostatic surface wave momentum flux.

Line 171: derivation for equation (5) needs to be expanded; some discussion is provided below but I suggest restructuring this presentation for clarity.

Reply: Thank for the suggestion. Replace the explanatory paragraph after Eq. (5) with: "Eq. (5) is adopted from Xu et al. (2021) for three-dimensional isotropic terrain. Specifically, the nonhydrostatic correction is defined as the ratio of the analytically derived nonhydrostatic surface WMF to its hydrostatic counterpart, which yields an algebraic correction factor depending only on the horizontal Froude number".

Line 210-211: do you mean that you are always selecting the 10th day of each forecast for all 31 simulations? For example:

Run 1 starting on 1 Dec gives output of 10 Dec; Run 2 starting on 2 Dec -> 11 Dec and so forth...

this should be better explained.

Reply: Yes. In the revised manuscript, it is changed to "Figure 1 shows the zonal-mean zonal wind composited from the day 10 forecast output of the 31 simulations initialized at 00 UTC on 1-31 December 2023. Accordingly, the valid times of these day-10 forecasts span 10 December 2023 to 10 January 2024" .

Lines 216-217: "These two jets are separated around 70 hPa and 50°N, with the stratospheric jet being stronger than its tropospheric counterpart." We can't really appreciate which one is stronger, given that the vertical axis stops at 1 hPa, above which the jet intensifies - in fact, in the figure as currently presented the tropospheric subtropical jet appears stronger than the stratospheric one. There are also two independent features, so I would remove this sentence entirely.

Reply: Thank for the suggestion. It is deleted.

A general comment about lines 212 to 223 is that this section can be significantly shortened or removed, it simply describes well-known atmospheric circulation patterns. Better to go directly into comparison with model outputs (Fig.1b,c,d)

Reply: Thank for the suggestion. It is revised.

Line 238: “Antarctic region” comprises 90-60S. I would say over the polar cap.

Reply: Thank for the suggestion. It is revised.

Line 245: “the easterly biases in the mid-upper stratosphere” specify you refer here to NH.

Reply: Thanks. It is revised to “the NH easterly biases in the mid-upper stratosphere”.

Figure 1b,d and its description:

What about the Northern Hemisphere troposphere? At high latitudes (50–90°N), the dipole of anomalies appears to be a mirror image of the stratospheric dipole above. In my view, this strongly suggests that during Northern Hemisphere winter, north of ~60°N, gravity waves are breaking more frequently within the mid–upper troposphere (roughly between 500 and 50 hPa). This would locally decelerate the background flow, thereby reducing wave breaking in the overlying stratosphere, where the flow remains stronger than in the control simulation. Conversely, the opposite mechanism would apply to the positive anomalies south of 60°N. That said, the origin of the longitudinal (east–west) dipole in the anomalies remains unclear and deserves further discussion. To fully explain this behavior, it would be important to examine the lower troposphere. One would expect that non-hydrostatic effects-by increasing wave dispersion and reducing wave amplitude—limit wave breaking at lower levels, thereby shifting the altitude of wave breaking upward. In that case, one could expect to see a vertical structure characterized by alternating positive–negative–positive anomalies from the surface to the stratosphere. If this is not the case (as I am speculating), then the anomalies shown in Figure 2d require a more thorough explanation.

Reply: Fig. 6 has been extended downward to 850 hPa. The vertical structure indeed characterized by alternating positive-negative-positive anomalies from the surface to the stratosphere.

Line 247: 42% is a remarkable improvement. Similarly to the Northern Hemisphere, it would be interesting to see what happens at lower levels.

Reply: Fig. 1 has been extended downward to 850 hPa. The tropospheric westerly biases have also been reduced at lower levels.

Whole section b and line 424 specifically “leading to greater upward motion poleward of”: I am somewhat confused by this section. The downward control states that the residual mean meridional circulation at a given level is determined by the integrated wave drag above that level. This circulation is characterized by upwelling at the tropics and downwelling at the poles. A weakening OGWD in NHE, at the poles, should thus weak this circulation and decrease downwelling at those latitudes, not increase upwelling as stated here. Perhaps results should be framed so that they are presented as: weaker polar downwelling (so less adiabatic warming) -not greater upward motion- which thus makes NHE colder than CTRL?

Reply: The previous version invoked the steady-state downward-control framework too strongly for a 10-day forecast context, and this interpretation was not sufficiently justified. Accordingly, in the revised manuscript we have removed the entire subsection “downward control mechanism”, including Eq. (6) and the associated discussion. Instead, the revised manuscript now adopts a more cautious interpretation focused on diagnostics that are more directly supported by the medium-range forecast experiments, namely the changes in parameterized OGWD, resolved EP flux, and refractive properties associated with Rossby-wave propagation. In this way, the mechanistic discussion is now better aligned with the forecast timescale considered in this study.

Line 428: Are these temperature changes an improvement compared to ERA5?

Reply: Yes.

Technical Corrections

Line 47: “when breaking transferring”

Reply: Thank for the suggestion. It is revised to “As these waves break, they transfer momentum from the surface to higher levels”.

Lines 70-73: I suggests removing sentence between brackets.

Reply: Thank for the suggestion. It is deleted.

Line 113: “descripted” -> described

Reply: Thanks. It is revised.

Line 233: “As in the NH troposphere,” I am not sure the use of “as” grammarly makes sense here.

Reply: Thank for the suggestion. It is corrected to “Similarly, the zonal-mean zonal winds in the NH troposphere”.

Line 389: “by the way”, too colloquial

Reply: Thank for the suggestion. It is deleted.

Line 501: “in total 31 forecasts of 10-day forecasts”

Reply: Thanks. It is corrected to: “In each numerical experiment, there are in total 31 independent 10-day forecasts which are initiated at 00 UTC on each day of December 2023”.