Reply to Referee 1

We thank Denny Oliveira for his thought- and useful comments that will help to improve our manuscript substantially. Below, we provide a detailed response to the comments and questions of the referee. We upload a version of the revised manuscript which includes all track-changes.

I recommend minor–major revisions before acceptance: the authors should (i) explicitly connect the results to empirical models and storm recovery mechanics (see Major point #1), (ii) discuss possible NO cooling timing on model predictions, and (iii) clarify the use of SABER data..

Major comments

1. **Discussion of possible NO effects on empirical thermospheric neutral mass density models. Although the manuscript clearly shows (and discusses) that externally precipitating electrons fuel NO production and that increased NO can drive infrared cooling and overcooling, a link to empirical models during storm recovery should be made more explicit. The manuscript benchmarks Kompot against the empirical NRLMSIS model and repeatedly notes that Kompot does notinclude externally precipitating electrons, i.e., Kompot (and many empirical/parametric approaches) therefore will miss NO produced by precipitating electrons. This important limitation is explicitly stated. However, the paper does not yet clearly walk the reader through the specific mechanism and timing by which omission of precipitation (and the resulting NO) leads to errors in empirical thermospheric models during the recovery phase (when NO cooling can cause densities to fall below pre-storm levels). The paper mentions that underestimating cooling can overestimate expansion/drag (thus implying impacts on forecasting), but an explicit paragraph that: (a) names typical empirical models (NRLMSIS, etc.), (b) explains how those models are forced/parametrized during storms and recovery, and (c) quantifies (or gives literature evidence for) the size and timing of the bias during recovery would strengthen the manuscript. See lines where the implication is implied but not spelled out.**

   In this case, I recommend the authors add a short subsection in Discussion explicitly entitled something like "Implications for empirical models and storm recovery" that explains why omission of precipitation-driven NO leads to errors specifically during the recovery phase (timing: NO lifetime/diffusion ~1 day is mentioned and important). Also, if possible, provide a short numerical estimate or point to literature values (see below) on how big the cooling bias can be and whether it systematically moves empirical model outputs relative to observations.

There has been previous work done on NO cooling effects on empirical models. For example, Oliveira and Zesta (2019) noted that the lack of NO information in the Jacchia-Bowman 2008 (JB2008) model is most likely a major source for density errors during recovery phase of storms, particularly during extreme events. Licata et al. (2021) also observed the same features with CHAMP and GRACE data, but they noted that the HASDM (High Accuracy Satellite Drag Model) was able to capture cooling effects due to NO (recovery) and CO_2 (pre-storm) phases. Oliveira et al. (2021) also noted with a superposed epoch analysis that HASDM was able to capture NO effects and even an overcooling effect supported by observations (CHAMP and GRACE), but JB2008 failed miserably during the recovery phase of the storm. One more. Zesta and Oliveira (2019) were able to quantify the timing of such cooling effects, noting that the thermosphere heats and cools faster for the more extreme geomagnetic storms. I think the NRL-MSIS results showed by the authors are expected, since the lack of NO effects also have profound impacts on model results during storm recoveries in the case of JB2008. I think this discussion should be added to support the authors' conclusion stating that, e.g., "[...] NO molecules have [not has] protective effect on LEO satellites." (line 367)

We thank the referee for this important suggestion. We agree and based on this, and the referee's next comment, we re-organized our entire discussion section and subdivided it into subsections, with one of these being the suggested "Implications for empirical models and storm recovery" where we discuss this in more detail (lines 400-430 in the manuscript version with track-changes).

One small side note to ""[...] NO molecules have [not has] protective effect on LEO satellites.": Here, "has" is actually correct since it refers to "The related overcooling of the thermosphere" and not to "NO molecules", so we kept it as it is. Besides, however, we really thank the referee for reading everything very carefully!

2. **The Kompot runs are steady /1-D background solutions (daily averaged XUV forcing and homopause boundary from NRLMSIS) and the NO production via the Shematovich model is solved to steady-state. The manuscript acknowledges Kompot does not include precipitation and that the NO/diffusion lifetimes (~1 day) matter. However, I strongly recommend the authors make clearer (in Methods or Discussion) the limits of these steady/1-D assumptions for transient recovery behavior (e.g., how the one-day chemical/diffusion timescale compares to recovery timings). The authors could tie such discussion with the heating and cooling times provided by Zesta and Oliveira (2019). Advise that full 3-D, time-dependent runs would be needed to fully capture spatial and temporal evolution of NO cooling during recovery.**

We agree again with the referee. We also added an additional subsection to the discussion called “Limitations and caveats” where we discuss this point and other limitations pointed out by both referees (lines 357-399 in the manuscript version with track-changes).

3. **SABER NO flux maps are used; the manuscript states that event-1 shows increases consistent with overcooling while event-2 does not. This is good. However, consider adding a brief note on the limits of SABER sampling (anti-sun viewing, gaps, hemispheric coverage) and how that affects the interpretation of polar NO enhancements vs. global effects — the paper already points this out (good), but a sentence tying that observation limitation into inference about recovery would help.**

This is a very good point! We slightly expanded the discussion in the revised manuscript and made this clearer (lines 95-98 397 in the manuscript version with track-changes). In addition, we also added a schematic figure on the observational technique through which the data is actually obtained by TIMED/SABER (new Figure 3 incl. caption).

SABER generally views towards the anti-Sun side of the spacecraft. This prevents solar infrared radiation from overlaying the desired thermospheric signal, but it also results in an asymmetric global coverage over any 60-day period (Russell et al., 1999), and to the visible polar gaps that can be seen in the data for both events.

4. **I recommend the authors also cite Knipp et al. (2017) to support the claim of electron precipitation in producing storm-time NO molecules. The authors also mention that NO molecules are more numerously produced when the CME-driven storms are preceded by interplanetary shocks.**

Thanks for pointing us to this reference! We added this reference to the manuscript (lines 40, 230, 244, 401 in the manuscript version with track-changes) and briefly describe its findings to support this claim (lines 245-246 in the manuscript version with track-changes).

Minor comments

Caption of figure 2: repectively ◊ respectively.

Caption of Table 1. “TIMMED/SEE” ◊ TIMED/SEE.

Line 378. “author’s” ◊ authors.

Spell out DMSP the first time it is mentioned. The same for LST.

Thanks, we took care of the minor comments in the revised manuscript.

References

Russell, J. M. I., Mlynczak, M. G., Gordley, L. L., Tansock, J., and Esplin, R.: An Overview of the SABER Experiment and Preliminary Calibration Results, Space Dynamics Laboratory Publications, 114, https://doi.org/https://doi.org/10.1117/12.366382, 1999.

Reply to Referee 2

We thank the referee for her/his valuable comments and the positive evaluation, which improves our manuscript further! Below, we provide a detailed response to the comments and questions of the referee. We upload a version of the revised manuscript which includes all track-changes.

L 37: "It was found that the thermospheric NO concentration correlates strongly with space weather events and solar activity": Reference?

References are, e.g., Barth et al. (2004), Mlynczak et al. (2015), and Knipp et al. (2017), which we added to the revised version of the manuscript (line 38 in the manuscript version with track changes. Please be aware that latexdiff, i.e., track changes via LaTeX, erroneously does not make a linebreak here. It is, however, correctly formatted in the manuscript version without track-changes.)

It would be good to mention what "suprathermal" oxygen atoms are.

"Suprathermal atoms" are atoms that have kinetic energies/temperatures above the mean temperature of the bulk atmosphere. We define them, more specifically, as atoms that have kinetic energies corresponding to temperatures >=4000 K (see, e.g., Shematovich et al. 2011). We add this definition to the manuscript (line 37 in the manuscript version with track-changes).

Figure 1: Why is the shading wider for CHAMP early on and GRACE has narrower shading?

While details are a bit more nuanced, the size of the shaded area is in general a function of the orbit's eccentricity. The plot indicates the same trend for both satellites, i.e., the difference between apogee and perigee decreases over time, which means their orbits become more circular as the orbital energy gets lost. The sudden changes in CHAMP's early years are due to orbital manoeuvres which are commonly performed near apogee/perigee, as this is most energy efficient. The shaded areas therefore indicate that CHAMP's orbit had a larger eccentricity early-on compared to GRACE. We added a brief explanation about the shaded areas to the figure caption.

Figures 2 and 4: The differences between CHAMP and GRACE responses are visible clearly. A brief note on the altitude dependence of thermospheric response in the figure caption would strengthen significance.

Good point, thanks! CHAMP's larger decline in altitude is expected, since the absolute increase in the thermospheric density between the onset of such an event and the maximum thermospheric density during the event will typically be larger at the lower compared to the higher orbit (i.e., CHAMP's compared to GRACE's orbit). This can clearly be seen in both figures. We highlighted this in both figure captions.

Kompot is not magnetically coupled and does not incorporate particle precipitation and geomagnetic energy input. Since the results rely heavily on Kompot as the background atmosphere, the authors should comment more explicitly on: how missing storm-time heating (Joule heating, electron precipitation, convection) affects interpretation; whether Kompot's baseline densities during events may be biased high or low; and how these limitations influence NO production estimates.

As the referee points out correctly, Kompot does not include any magnetospheric energy input. These can increase thermospheric temperature and drive global expansion of the atmosphere (e.g., Wang et al. 2020). At a given altitude, therefore, neutral densities typically increase during the main phase of a storm. Since these processes are not included in Kompot, we may expect that our model is biased towards lower temperatures and baseline densities. Since the peak of NO cooling slightly shifts upwards during the main phase of a storm due to the expansion of the atmosphere, we can further expect that the peak of NO production and cooling could slightly shift upwards in altitude (e.g., Li et al. 2019, Luo et al. 2024), but the column-integrated NO production should only be weakly affected as this is primarily controlled by the column density and not the specific altitude of production. NO production is further dependent on temperature (by order of unity; see, e.g., Shematovich et al. 2024), which can lead to a slight increase in NO production. However, the energy distribution and energy flux of the precipitating electrons, as also illustrated by our study, will be the dominant driver affecting column-integrated NO production. Uncertainties in the assumed precipitating energy flux and distribution might therefore be larger than the errors introduced by neglecting the storm-time background response in Kompot.

As already stated in our reply to the first referee, we re-organized our discussion section into various subsections and we added this specific caveat to the new subsection "Limitations and caveats" (lines 357-399 in the manuscript version with track-changes).

A brief discussion comparing Kompot outputs to empirical models (e.g., NRLMSIS or JB2008) during storm conditions would strengthen confidence.

This is a good point, thanks. As stated above, Kompot does not include explicit magnetospheric storm forcing; its thermospheric background therefore remains close to quiet conditions. However, the lower boundary of our model at 80 km is initialized by using temperature and neutral densities from NRLMSIS (Picone et al., 2002; Emmert et al., 2021) for the respective days of the events and for a solar zenith angle of 66°, as this best reproduces a global average of the Earth's atmosphere (Johnstone et al., 2018). NRLMSIS indeed includes empirical storm-time variability through the *Ap* and *ap* indices, but its geomagnetic activity dependence is known to be weak at 80 km (Emmert et al., 2022). The lower boundary of our background atmosphere model therefore is realistic for average daily climatic conditions but only presents a weak storm-time adjustment.

The dominant storm-time effects in the thermosphere occur above ~100 km and storm-induced temperature and density enhancements are therefore not inherently included in our Kompot simulations as they are empirically captured by NRLMSIS or JB2008. As already discussed above, this is likely to mainly affect the altitude of the NO production and cooling peaks.

We also added this discussion to the newly written "Limits and caveats" section.

Figures 6 and 7: It is also not stated anywhere if it's a global average response for the Kompot simulations.

Since Kompot is a 1D model, we simulate a global average response of the atmosphere. Kompot was benchmarked against NRLMSIS in Johnstone et al. (2018) where it was shown that the global average atmosphere can be reproduced best by assuming a solar zenith angle of 66°. We highlighted this in the text more explicitly (lines 160-161 in the manuscript version with track-changes) and added it to the figure captions as well (figures 7 and 8 in the revised manuscript).

I agree with Reviewer #1 on explaining more about the SABER data usage.

As already stated in our reply to referee #1, we extended the discussion on the SABER data usage (lines 97-100 in the manuscript version with track-changes) and added an additional figure (new Figure 3) that schematically describes the measurement technique used by SABER.

L 178: "hotter" can be replaced by higher

L 338: "this averaging procedure is not an easy procedure" can be replaced by "this averaging procedure is not easy"

Some references are duplicated, e.g., Barth et al. 1999a/1999b appears to be identical.

Thanks for these further comments. We implement all these changes in the updated manuscript.

References:

Barth, C. A., Baker, D. N., and Bailey, S. M.: Seasonal variation of auroral electron precipitation, Geophysical Research Letters, 31, L04809, https://doi.org/10.1029/2003GL018892, 2004.

Emmert, J. T., Drob, D. P., Picone, J. M., Siskind, D. E., Jones Jr., M., Mlynczak, M. G., Bernath, P. F., Chu, X., Doornbos, E., Funke, B., Goncharenko, L. P., Hervig, M. E., Schwartz, M. J., Sheese, P. E., Vargas, F., Williams, B. P., and Yuan, T.: NRLMSIS 2.0: A Whole-Atmosphere Empirical Model of Temperature and Neutral Species Densities, Earth and Space Science, 8, e2020EA001 321, https://doi.org/https://doi.org/10.1029/2020EA001321, e2020EA001321 2020EA001321, 2021.

Emmert, J. T., Jones Jr, M., Siskind, D. E., Drob, D. P., Picone, J. M., Stevens, M. H., Bailey, S. M., Bender, S., Bernath, P. F., Funke, B., Hervig, M. E., and Pérot, K.: NRLMSIS 2.1: An Empirical Model of Nitric Oxide Incorporated Into MSIS, Journal of Geophysical Research: Space Physics, 127, e2022JA030 896, https://doi.org/https://doi.org/10.1029/2022JA030896, e2022JA030896 2022JA030896, 2022.

Johnstone, C. P., Güdel, M., Lammer, H., and Kislyakova, K. G.: Upper atmospheres of terrestrial planets: Carbon dioxide cooling and the Earth's thermospheric evolution, Astronomy and Astrophysics, 617, A107, https://doi.org/10.1051/0004-6361/201832776, 2018.

Knipp, D. J., Pette, D. V., Kilcommons, L. M., Isaacs, T. L., Cruz, A. A., Mlynczak, M. G., Hunt, L. A., and Lin, C. Y.: Thermospheric nitric oxide response to shock-led storms, Space Weather, 15, 325–342, https://doi.org/10.1002/2016SW001567, 2017.

Li, Z., Knipp, D., andWang,W.: Understanding the Behaviors of Thermospheric Nitric Oxide Cooling During the 15 May 2005 Geomagnetic Storm, Journal of Geophysical Research: Space Physics, 124, 2113–2126, https://doi.org/https://doi.org/10.1029/2018JA026247, 2019.

Liu, H., Gao, H., Li, Z., Xu, J., Bai, W., Sun, L., and Li, Z.: Response of NO 5.3 μm Emission to the Geomagnetic Storm on 24 April 2023, Remote Sensing, 16, https://doi.org/10.3390/rs16193683, 2024.

Mlynczak, M. G., Hunt, L. A., Marshall, B. T., Russell, J. M., Mertens, C. J., Thompson, R. E., and Gordley, L. L.: A combined solar and geomagnetic index for thermospheric climate, Geophysical Research Letters, 42, 3677–3682, https://doi.org/10.1002/2015GL064038, 2015.

Picone, J. M., Hedin, A. E., Drob, D. P., and Aikin, A. C.: NRLMSISE-00 empirical model of the atmosphere: Statistical comparisons and scientific issues, Journal of Geophysical Research (Space Physics), 107, 1468, https://doi.org/10.1029/2002JA009430, 2002.

Shematovich, V., Bisikalo, D., Tsurikov, G., and Zhilkin, A.: Non-Thermal Processes of Nitric Oxide Formation during Precipitation of Auroral Electrons into the Upper Atmospheres of Terrestrial Planets, Astronomy Reports, 68, 843–864, https://doi.org/10.1134/S1063772924700744, 2024.

Wang, X., Miao, J., Aa, E., Ren, T., Wang, Y., Liu, J., and Liu, S.: Statistical Analysis of Joule Heating and Thermosphere Response During Geomagnetic Storms of Different Magnitudes, Journal of Geophysical Research: Space Physics, 125, e2020JA027 966, https://doi.org/https://doi.org/10.1029/2020JA027966, e2020JA027966 2020JA027966, 2020.