Review of the manuscript "*Analyzing ozone variations and uncertainties at high latitudes during Sudden Stratospheric Warming events using MERRA-2*" by Shima Bahramvash Shams et al.

This study examines the evolution of polar ozone during six recent major sudden stratospheric warming events (SSW) using the MERRA-2 reanalysis. The analysis is preceded by evaluation of the MERRA-2 ozone using independent observations and focusing on the region and periods of interest, which provides a nice complement to previous validation studies that looked mainly at the global picture. The authors find that the impacts of SSWs on ozone were largest in 2009 and 2018 ("elongated vortex" cases) compared to the other events ("displaced vortex") and identify vertical advection as the key mechanism responsible for the formation of these positive ozone anomalies during all the SSWs analyzed in this study.

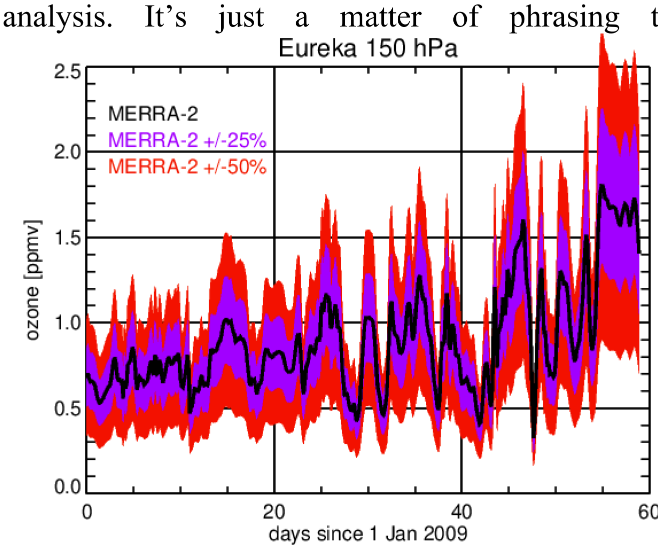
This paper is interesting and certainly suitable for publication in ACP subject to some minor revisions as delineated below. It's nice to see a detailed study of a number of recent SSWs in one place along with an ozone budget analysis for each of them. It certainly adds to our understanding of the role of transport during these events. It's also great to see another paper that demonstrates the value of reanalysis ozone data. It is only recently that our community began to take advantage of the global coverage, high resolution and dynamical consistency that these products offer. The additional validation of MERRA-2 ozone, focused on the region of interest is especially valuable. While I ended up having a rather large number of comments and suggestions, none of them are serious objections, and I believe all of them can be easily addressed. The most important ones concern the need to highlight the novel aspects of this work in the context of other similar studies (some of them not cited) (general comment #1), the choice of these particular six events and a possibility of adding 2021 to the list (general comment #2), and terminology (#3). I hope my comments will be useful.

General comments

1. Can you place your results in the context of previous papers discussing the role of dynamical ozone resupply, e.g., Tegtmeier et al., 2008; Strahan et al. (2016)? It would help if you clearly delineated the novel aspects of your study against the backdrop of the existing literature of the subject.
2. Butler et al. (2017) as well as the SSW Compendium (https://csl.noaa.gov/groups/csl8/sswcompendium/majorevents.html) list several more SSWs than those discussed in the paper: 2007, 2008, and 2010. I assume there's a good reason for the selection discussed in this study to be what it is, but it needs to be explained, especially since the definition of SSW used here is the same as that applied in the Compendium. In addition, ideally, I would like to see 2021 added to the analysis. Seven events are better than six. Note that it would likely be the first paper that talks about ozone during the 2021 SSW.
3. There is a terminological confusion regarding the use of the word "model" where you really mean either data assimilation system, assimilated data, or reanalysis. I tried to catch those instances in my specific comments below. I think it's important to remember that a reanalysis is **not** a model simulation. Rather, it is fundamentally a data-driven product. Calling a reanalysis a "model" is equivalent to calling a retrieved satellite data set "a

radiative transfer model". The role of the general circulation model in a data assimilation system is to propagate information from observations in time and space over a period of several hours. That's all it does. Things do get a bit muddled in places/periods devoid of observations, but your study looks at regions well constrained by data, so that's not an issue here.

4. You use the terms lower and middle stratosphere somewhat loosely. Please define them somewhere in the methods section.
5. You choose to exclude the lower stratosphere from this analysis citing the high uncertainty of MERRA-2 ozone there. While it's completely OK to make that choice I have a problem with the stated motivation. It's certainly true that the uncertainties get larger closer to the tropopause, but I don't think one can say, based on the validation that you did or the results in Wargan et al (2017), that there's no useful information down there, i.e. that the true variability is completely obscured by uncertainty. The ozonesonde comparisons in Fig. 3 suggest difference standard deviations of about 25%. Below I plot the time series of MERRA-2 ozone at Eureka at 150 hPa along with 25% and 50% envelopes. Clearly, the dynamical variability around the 2009 SSW is still very much discernible in the sense that the magnitude of the large jumps exceeds the uncertainty. Additionally, note that Albers et al. (2018) found useful ozone information in MERRA-2 in the lowermost stratosphere and Knowland et al. (2017) as well as Jaeglé et al. (2017) successfully studied stratospheric intrusions into the troposphere using MERRA-2 ozone. Again, I agree that any results in the LS would have larger uncertainty, but I can't agree that they would be worthless, and the present wording in the paper seems to imply that. I'm not suggesting extending the analysis. It's just a matter of phrasing things in a more nuanced way.



6. In several places the language of the paper implies (or even outright) states causal relationships where causality is not immediately demonstrated. I provide some examples in my specific comments. There are sentences like "*the vortex displacement toward the southeast (Europe) prior to the 4 major SSW as seen in 2006, 2008, 2013, and 2019 (hereafter the displaced vortex SSWs) caused an early positive ozone anomaly*", "*an elongated polar vortex is observed before the SSWs which caused a dominant negative ozone anomaly*", "*SSWs and their impact on ozone*" (the title of a section). I get it that Figures 11 and 12 and the discussion in Section 6 do provide evidence for a causal mechanism, but the reader doesn't know that in advance. I suggest rephrasing these causal

statements as something more "relational" (Fig 6 and the wording on P24 L3-5 are very good) and explaining that a dynamical mechanism will be elucidated in Section 6. You can then state your results in causal terms in the conclusions as you do. It's really about streamlining the argument.

Specific comments

P2. L19-22. I believe that this "ozone gets made in the tropics and transported to the high latitudes" description is simplified to the point of being incorrect (albeit quite commonly used). Please, take a look at the discussion around Fig. 5.11 in Bresseur & Solomon 2005 (in my edition it starts on page 286 with "It is sometimes stated that ozone is produced where its mixing ratio maximizes…"). This does not affect anything in your paper of course.

P2. L24-25. Would it make sense to cite the recent review paper on SSWs by Baldwin et al., 2021?

P4. L3-4. "numerical and assimilation models". I struggle with this terminology. I understand that by "numerical models" you mean general circulation models or numerical weather prediction models. I can live with the short-hand version "numerical models". But I think the word model should not be applied to data assimilation systems, which comprise a model component and a statistical analysis scheme. I suggest "numerical models and data assimilation systems". See my general comment #3.

P4. L15-16. At first read this sentence confused me: why focus on zonal structures? SSWs are very non-zonal! But you're actually doing much more than that: the Greenland sector, polar averages, maps. I suggest rephrasing it.

P4. L17. "assimilation model". See above. Why not just say reanalysis?

P5. L7-9. The description of MERRA-2 needs some rewriting. It's not clear to me what it means that there's a variety of models incorporated in MERRA-2. Do you mean GCM, land model, parameterized chemistry, etc.? There aren't multiple "general circulation model**s**" in it as the text implies. There is one. Also, I don't know what you mean by "extended reanalysis". No other reanalysis is incorporated in MERRA-2. If you mean something like OSTIA (the SST data set that provides boundary conditions for the GCM), it's not a reanalysis, at least not in the same sense as MERRA-2.

P5. L16. I wouldn't say that it's been extensively used in trend studies. To my knowledge only Wargan et al. (2018) derived trends from (suitably corrected for discontinuities) MERRA-2 (there was a follow-on paper by Orbe et al. 2020, but MERRA-2 ozone is sort of tangential there). Also, please consider adding a very interesting study by Albers et al. (2018) to these citations.

P5. L21. Please, provide a citation for the data collection used in this study. Information on how to cite the MERRA-2 pressure-level output is here:
https://disc.gsfc.nasa.gov/datasets/M2I3NPASM_5.12.4/summary under "data citation". For the

model-level output: https://disc.gsfc.nasa.gov/datasets/M2I3NVASM_5.12.4/summary (I'm providing both in case I misunderstood which one you used; see the next comment).

P5. L21. Why use pressure-level output if much higher vertical resolution model-level output is available? But P8 L16 talks about using the model levels. So which collection is really used?

P5. I think it should be mentioned that MERRA-2 assimilates MLS ozone down to 177 hPa between 2004 and 2015, then switches to version 4.2 retrievals going down to 215 hPa. This version (and vertical range switch) results in some differences between the pre- and post 2015 periods, whereby the latter has likely more accurate ozone, especially in the lower stratosphere.

P5. L22. "Assimilated/reanalysis models". Why not just say "assimilated products" or "reanalyses"? And what do you mean by "variations in models"? Model uncertainties?

P6. L25. Are your results affected by the "ozonesonde problem" identified in Stauffer et al. (2020)?

P8. L21, L24, L26. Model data → reanalysis data, model profile → reanalysis profile, model → reanalysis.

P9 L9-16, See my general comment #2.

P9 L15-16. I don't understand this sentence. Could you rephrase it, please?

P9 L21 The symbol Δ. (the Laplace operator followed by a dot) should be replaced by $\nabla \cdot$ (divergence)

P10 L8. Since you previously said that subscripts denote derivatives you can't use M_y and M_z as these are not the derivatives of M. I can't remember off-hand the notation in Andrews et al. but I suggest $M_{(y)}$ and $M_{(z)}$. I can't exclude the possibility that some authors use M_y and M_z but it strikes me as unnecessary abuse of notation.

P10 L15-16. The GEOS model used in the MERRA-2 data assimilation system does include chemical ozone production (albeit simplified). Do you simply mean to say that P and L is neglected in your analysis (Fig 11 and 12) but that doesn't lead to significant non-closure of the budget as evidenced in Fig. 11? Please, clarify.

P10 L20. "assimilated models" → data assimilation systems

P10 L20-24. I know what you mean but the way it's written this is somewhat contradictory. You say that the ground-based observations are too sparse, and then, in the next sentence you say that it's a "dense network".

P12 L2-3. It's not designed to do that. The MERRA-2 DAS does not include bromine chemistry.

P12 L11. How is significance evaluated? Also, note that lack of statistical significance does not necessarily mean that a result is not useful.

P13 L13. Larger percentage values may also result from low mean ozone concentrations in the denominator.

P13 L21. It's not clear to me what this means. Do you simply mean 75% of ozone molecules?

P14 L10. Why were those values chosen? Because they are representative of the vortex edge? Also, I think "105 K…" should be "10^{-5} K…"; please use proper notation for the exponents. Note that 1 Potential Vorticity Unit = 10^{-6} m^2 s^{-1} K kg^{-1}.

P14 L11. Approximately what altitude does 850 K correspond to?

P16 L4. Causality cannot be inferred from these maps alone. I suggested changing "caused" to "accompanied by" or something like that. You can explain that a causal mechanism will be established later in the paper.

P16 L3-11. Are the "elongated" and "displacement" events related to the more familiar characterization of SSWs as "displacement" and "split" or wave-1 vs. wave-2 events?

P16 L12-20 and Figure 6. This is really nice but I would like to see more detail. What latitude is the PV averaged over? What about the TCO? Is it evaluated over some region, latitude band or location? Can you provide a correlation coefficient and maybe draw a regression line? Let me say preemptively that I realize that the sample is small so technically there may be an issue with low statistical significance, but I wouldn't overestimate the importance of the latter, rather arbitrary notion.

P16 L24. Are these area-weighted averages?

P18 L2-4. I assume the "due to the effects…" bit refers to the 2006 situation. The wording suggests a causal link between the minor warming and the exceptional trajectory of the Greenland vs. zonal TCO in 2006 but no detailed analysis is presented to support that link, at least not until Section 6. Additionally, it looks to me like the increase is similar between the Greenland sector and the zonal mean also in 2013.

P18 L15. Is the zonal average taken between 60N and 80N as before? Are these area-weighted averages? Are the anomalies calculated with respect to the climatology?

P18 L25-27. "…because of". Again, causality is stated but not demonstrated until later.

P19 L15. I don't think that it's the vertical advection that leads to the poleward transport of PV. Isn't it rather due to planetary wave breaking and resulting mixing of low-PV low latitude air into the high latitudes?

P20 L20 The notation M_y/dy is incorrect. If anything it should be $\partial M_{(y)}/\partial y$. Same for the vertical derivative. More troubling is the use of "y" (which was never defined, I believe; the equations so far were in terms of φ). These two quantities should be the terms of the del operator acting on $\mathbf{M}$

in the spherical coordinates, i.e. one would expect something like $\frac{1}{a\cos\varphi}\frac{\partial}{\partial\varphi}(M_{(y)}\cos\varphi)$. Please make sure that the calculation is done correctly and that the notation is also correct.

P21 L8. Is this connection with "elongated vortex" something that you could explain, or at least provide a viable hypothesis for?

P22 L5. "*errors in the vertical derivatives over high latitude can be large as cos(φ) gets small)*" Why? I would think that it's the horizontal derivative that would be sensitive to that, but it's very likely that I'm missing something obvious.

P22 L24. Again, it's not clear to me how significance is calculated.

P24 L23-25. As this (true) statement is not substantiated by the results of this study, it would be good to cite something here to support it.

P25 L8. Even adding just the 2021 SSW would help!

Technical corrections

P3. L1 Please double check the grammar. "SSWs" is plural

P3. L18-19. Grammar. "one of the strongest" → "some of the strongest"

P10 L5. The "a" (earth's radius) should be italicized.

P10 L10-11. The equation numbering changed from 1, 2,.. to 2.4, 2.5. Please, fix it.

P9 L22. "in this case ozone mixing ratio ***tendency***"

Kris Wargan

References

Albers, J. R., Perlwitz, J., Butler, A. H., Birner, T., Kiladis, G. N., Lawrence, Z. D., ... Dias, J. (2018). Mechanisms governing interannual variability of stratosphere-to-troposphere ozone transport. *Journal of Geophysical Research: Atmospheres*, *123*, 234–260. https://doi.org/10.1002/2017JD026890

Baldwin, M. P., Ayarzagüena, B., Birner, T., Butchart, N., Butler, A. H., Charlton-Perez, A. J., et al. (2021). Sudden stratospheric warmings. *Reviews of Geophysics*, *59*, e2020RG000708. https://doi.org/10. 1029/2020RG000708

Brasseur, G. P*., and* S. Solomon *(*2005*):* Aeronomy of the Middle Atmosphere*, 3rd ed., Springer,* New York*.*

Jaeglé, L., Wood, R., & Wargan, K. (2017). Multiyear composite view of ozone enhancements and stratosphere-to-troposphere transport in dry intrusions of northern hemisphere extratropical cyclones. Journal of Geophysical Research: Atmospheres, 122. https://doi.org/10.1002/2017JD027656

Orbe, C., K. Wargan, S. Pawson, and L.D. Oman, 2020: Mechanisms linked to recent ozone decreases in the Northern Hemisphere lower stratosphere. *J. Geophys. Res. Atmos.*, **125**, no. 9, e2019JD031631, doi:10.1029/2019JD031631.

Stauffer, R. M., Thompson, A. M., Kollonige, D. E., Witte, J. C., Tarasick, D. W., Davies, J., et al. (2020). A post-2013 dropoff in total ozone at a third of global ozonesonde stations: Electrochemical concentration cell instrument artifacts? *Geophysical Research Letters*, 47, e2019GL086791. https://doi.org/10.1029/2019GL086791

Strahan, S.E., A.R. Douglass, and S.D. Steenrod, Chemical and dynamical impacts of stratospheric sudden warmings on Arctic ozone variability, *J. Geophys. Res. Atmos*., *121*, 11,836–11,851, doi:10.1002/2016JD025128, 2016.

Tegtmeier, S., M. Rex, I. Wohltmann, and K. Krüger, Relative importance of dynamical and chemical contributions to Arctic wintertime ozone, *Geophys. Res. Lett.*, *35*, L17801, doi:10.1029/2008GL034250, 2008.