Supporting Information for "Higher-resolution tropopause folding accounts for more stratospheric ozone intrusions"

Samuel Bartusek[1,2], Yutian Wu[1], Mingfang Ting[1,2], Cheng Zheng[1], Arlene Fiore[3], Michael Sprenger[4], Johannes Flemming[5]

[1]Lamont-Doherty Earth Observatory, Columbia University, Palisades, NY, USA

[2]Department of Earth and Environmental Sciences, Columbia University, New York, NY, USA

[3]Department of Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA, USA

[4]Institute for Atmosphere and Climate Science, ETH Zürich, Zürich, Switzerland

[5]European Centre for Medium-Range Weather Forecasts (ECMWF), Reading, UK

Contents of this file

1. Text S1

2. Figures S1 to S8

Introduction

This file contains a text section providing details of the modifications we made to the tropopause folding algorithm, and a Supplementary Figures section containing seven figures mentioned in the main article.

Text S1

The folding algorithm we apply to CAMSRA and ERA5 is based on the algorithm originally developed by Sprenger, Maspoli, and Wernli (2003) and further sophisticated by Skerlak, Sprenger, and Wernli (2014) (the labelling portion of the algorithm) and Skerlak, Sprenger, Pfahl, Tyrlis, and Wernli (2015). Its labelling routine produces a label from 1–5 for each grid cell (in 3-D) at each timestep. As alluded to in the main text, these labels geometrically separate grid cells as belonging to either the troposphere or stratosphere, mostly based on their potential vorticity (PV) value but with a few exceptions where PV cannot itself determine which body a certain grid cell belongs to. The labels correspond as follows: 1, troposphere; 2, stratosphere; 3; stratospheric cutoff or diabatically produced PV anomaly; 4, tropospheric cutoff; 5, surface-bound cyclonic PV anomaly. Labels 1, 3, and 5 therefore constitute the troposphere and labels 2 and 4 constitute the stratosphere, where labels 3, 4, and 5 designate the exceptions with PV not indicative of its surrounding body. See Skerlak et al. (2015) for further details.

As mentioned in the main text, it was necessary to make modifications to successfully apply it to ERA5 data. We found that because of ERA5's very high resolution it was susceptible to finding pathways of high-PV air connecting the stratosphere all the way to the surface that are thin enough to be obscured at lower resolution. For such timesteps, the entire stratosphere would constitute a single surface-connected high-PV region, thus receiving label 5 (troposphere), and folding identification would be disallowed anywhere due to filters that help avoid spurious fold identification (see Skerlak et al. (2015) for details of such spurious cases that justify the filters).

The spread of label 5 into the stratosphere was partly attributable to the algorithm's strategy of horizontally propagating labels 5 and 3 into areas of label 2, as long as the area of label 2 is connected to a label 5 grid cell at a higher level. In ERA5 this allowed a single area of label 5 high up in the atmosphere at any location to propagate very extensively horizontally and downward. Our first modification was to deactivate this horizontal propagation behavior, which was introduced for mostly aesthetic reasons in the first place. Specifically, if one compares Figure 1 in Skerlak et al. (2014) against Figure 1 in Skerlak et al. (2015), this behavioral change between the two iterations is responsible for the label 2 "stratospheric funnel" seen in Skerlak et al. (2014) (where label 2 extends through the label 5 blob all the way to the surface) instead being "filled in" with label 5, such that label 5 propagates up to the level of thinnest funnel diameter, as seen in Skerlak et al. (2015).

However, despite this modification, label 5 (or 3) could still sometimes spuriously propagate throughout the stratosphere, invalidating some timesteps. We therefore introduced new conditions to replace appropriate label 5/3 regions that are connected to the stratosphere with label 2, but adopted three conditions to ensure conservativeness.

1. We first impose a condition that such a label 5/3 parcel must be within the upper half of the troposphere (i.e., if a grid cell's pressure distance from the local tropopause is smaller than that from the surface). This condition is very similar to one introduced in Skerlak et al. (2014) wherein label 2 was allowed to propagate horizontally (if contacting label 5) only in the upper half of the troposphere. We lift this lower-troposphere restriction for the approximate Tibetan Plateau region (25°–40°N, 75°–110°E)—its close surface

proximity to the tropopause means that even label 5 regions in the lower troposphere can lead to stratospheric label 5 propagation and missed fold identification (nevertheless, we still find very small to zero frequency differences between CAMSRA and ERA5 in this region (see Figure 2), which by comparisons of cases seems likely to somewhat represent a masking of otherwise increased ERA5 folding frequency, due to persisting spurious label 5 propagation—the frequency differences in this region shown in the main text are therefore likely conservative).

2. Additionally, we only allow relabelling of 5/3 to 2 if the tropopause is greater than 200 hPa from the surface, which for example helps avoid spurious fold identification in winter in Antarctica where very low tropopause heights and high topography with strong surface cooling can create high-PV layers correctly assigned label 5, as discussed in Skerlak et al. (2015).

3. Finally, we modified the algorithm's usage of specific humidity as an indicator of stratospheric air. In the version in Skerlak et al. (2015), as shown in their Figure 4, the threshold $q = 0.1$ g kg^{-1} helps separate low-altitude high-PV airmasses (moist tropospheric air) that merge with a real fold (dry stratospheric air), by determining a level up to which label 5/3 can propagate. Here, we use this threshold in a more restrictive way as a third condition. We disallow any relabelling from 5/3 to 2 for grid cells exceeding it, and we furthermore relabel all cells labelled 2 exceeding it to 5/3—essentially, we use the threshold as a 3-D contour outside of which label 2 is never allowed, as opposed to a vertical level affecting the relabelling of 2 to 5/3, which permitted label 2 to sometimes persist into air moister than the threshold.

As seen in Figure S1 below, our modifications altogether produce a dominantly conservative effect on folding frequencies, for Medium and Deep folding in particular. Our final modified version of the algorithm (specifically, a Fortran code file containing both the 3-D labelling routine and the tropopause fold detection routine based on that label field) is available at [*insert Zenodo link*].

References

Skerlak, B., Sprenger, M., Pfahl, S., Tyrlis, E., & Wernli, H. (2015). Tropopause folds in era-interim: Global climatology and relation to extreme weather events. *Journal of Geophysical Research: Atmospheres*, *120*(10), 4860–4877. doi: https://doi.org/10.1002/2014JD022787

Skerlak, B., Sprenger, M., & Wernli, H. (2014, Jan). A global climatology of stratosphere–troposphere exchange using the era-interim data set from 1979 to 2011. *Atmospheric Chemistry and Physics*, *14*(2), 913–937. doi: https://doi.org/10.5194/acp-14-913-2014

Sprenger, M., Maspoli, M. C., & Wernli, H. (2003). Tropopause folds and cross-tropopause exchange: A global investigation based upon ecmwf analyses for the time period march 2000 to february 2001. *Journal of Geophysical Research: Atmospheres*, *108*(D12). Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2002JD002587\%4010.1002/\%28ISSN\%292169-8996.STACCATO1` doi: https://doi.org/10.1029/2002JD002587

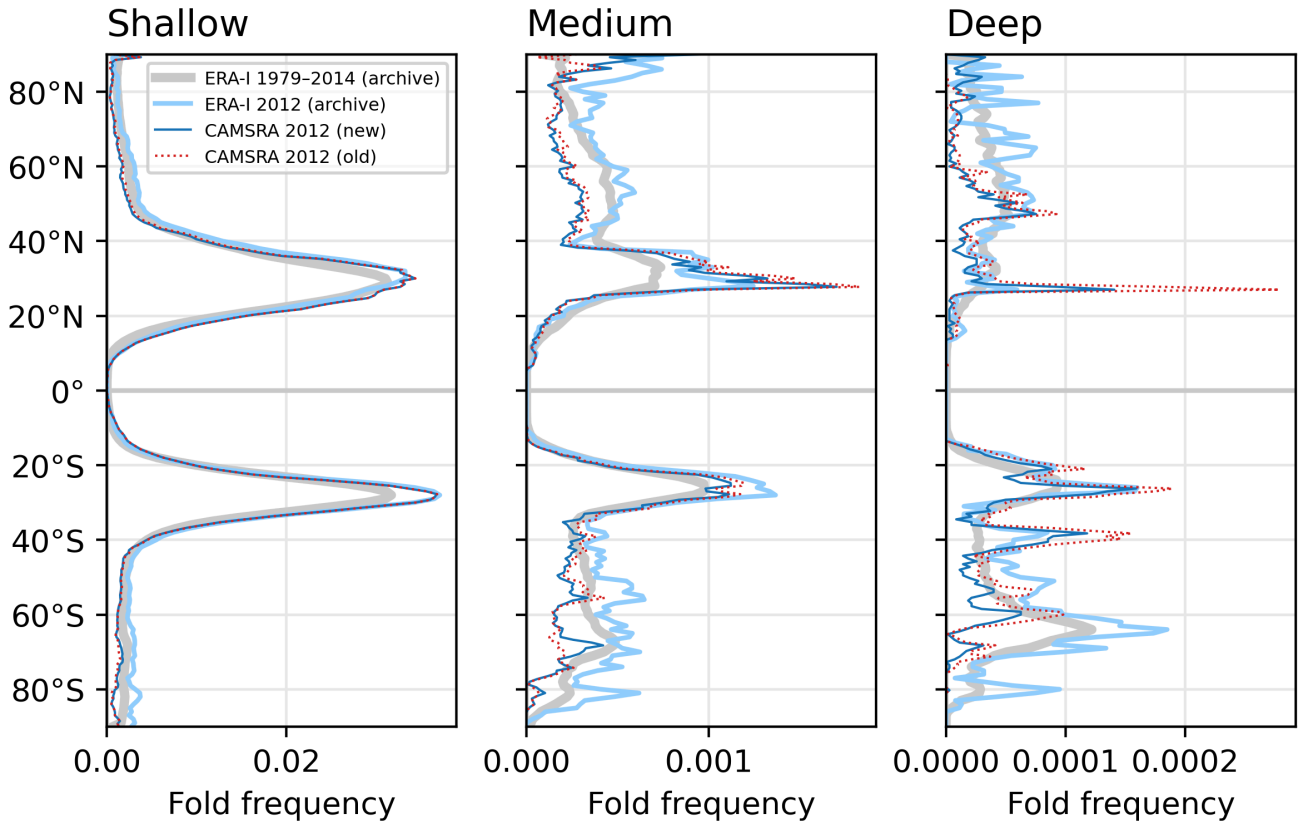


Figure S1. Comparison of 2012 folding to other years, and between folding detection algorithms. ***Left to right***: Zonal-mean fold frequencies are shown for each depth category. The thick light gray line shows the zonal mean tropopause folding frequency over 1979–2014 in ERA-Interim, provided by the ETH Zürich archive (available at http://eraiclim.ethz.ch/). The light blue line isolates the year 2012, showing that 2012 is representative of the underlying average frequency. The dotted red line shows the zonal average frequencies over 2012 in CAMSRA, generated by a newer version of the 3-D labelling algorithm (from Skerlak et al. (2015)). This version introduced more conservativeness in identifying folds than previous versions, likely accounting for most of the difference between it and the ERA-Interim 2012 frequencies (light blue), since the ERA-Interim and CAMSRA meteorologies were produced by the same model (IFS, albeit different model cycles) and at the same resolution. The differences are almost everywhere a reduction in frequency, and are proportionately stronger for Medium and Deep folds. Finally, the dark blue line shows frequencies over 2012 in CAMSRA generated by our modified version of the algorithm, showing that our edits were conservative, reducing frequency nearly everywhere compared with the dotted red line.

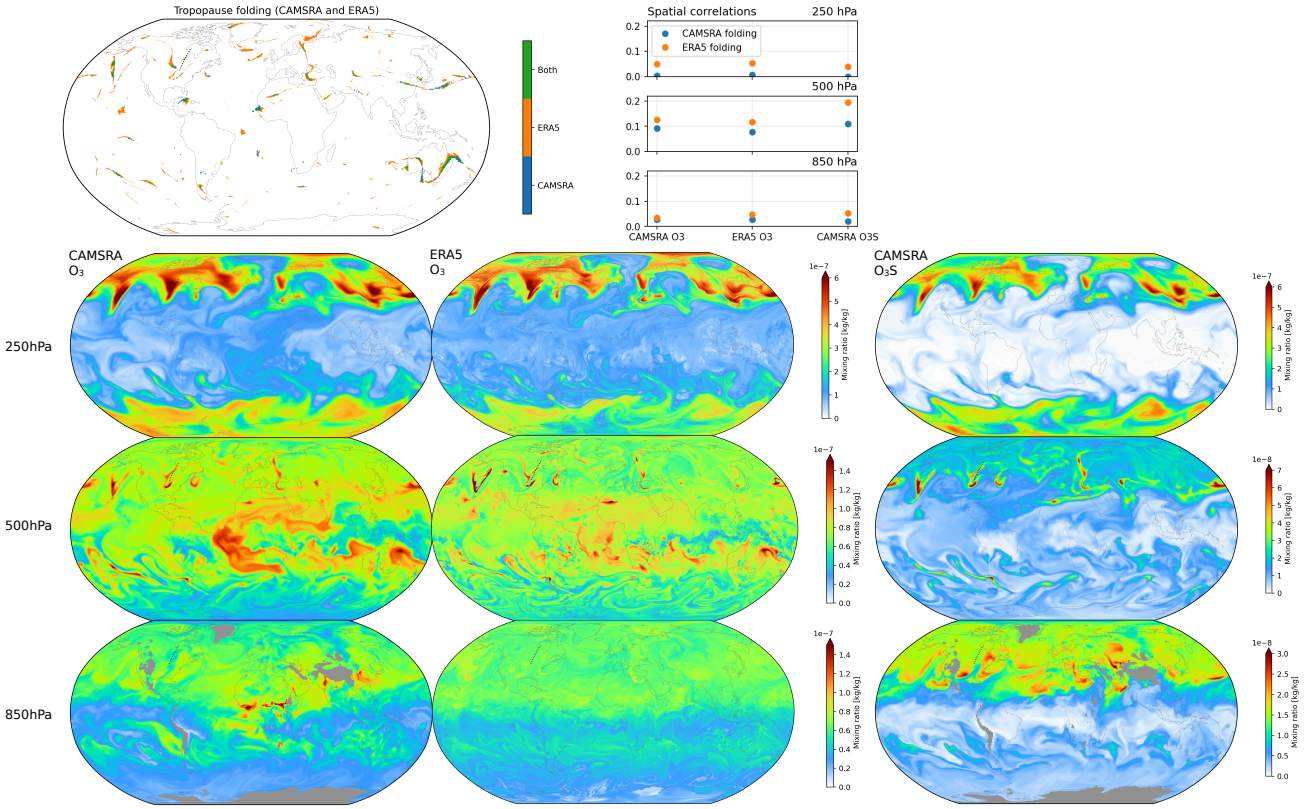


Figure S2. Folding and ozone snapshots at various levels from CAMSRA and ERA5. ***Top left:*** Tropopause folding in both CAMSRA and ERA5 on 1/1/2012 at 1200Z (same timestep as Figure 1a,c–d). ***Top right:*** For the same timestep, the spatial correlations between CAMSRA folding and each ozone snapshot shown below (blue markers), versus the same with ERA5 folding (orange markers). ***Second row:*** For the same timestep, CAMSRA ozone, ERA5 ozone, and CAMSRA stratospheric ozone tracer O_3S (left to right) at the 250 hPa level. ***Third and fourth rows:*** As in second row but for the 500 and 850 hPa levels. (Top panel and 500 hPa O_3 are as in Figure 1c and d but over the whole globe.)

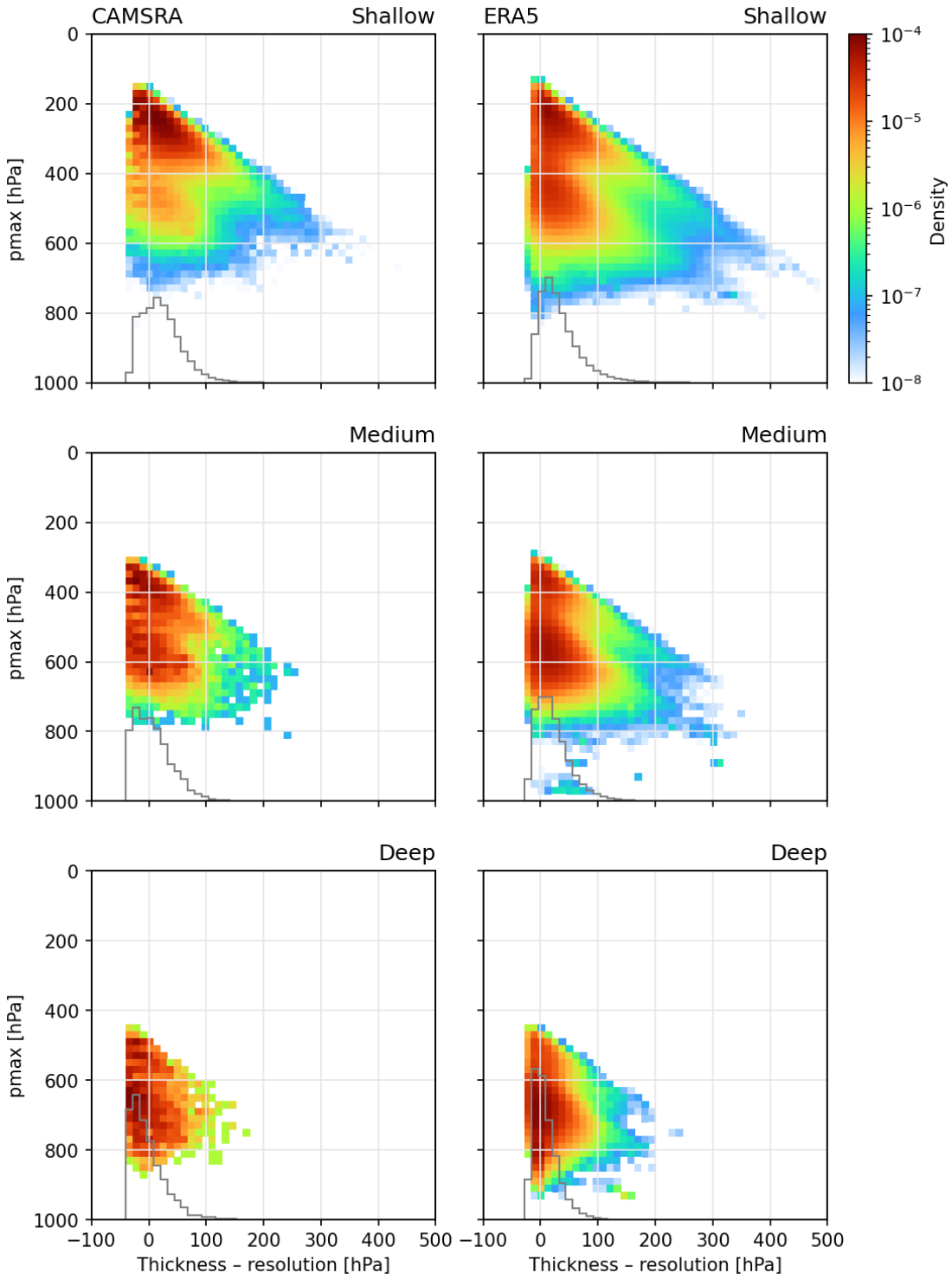
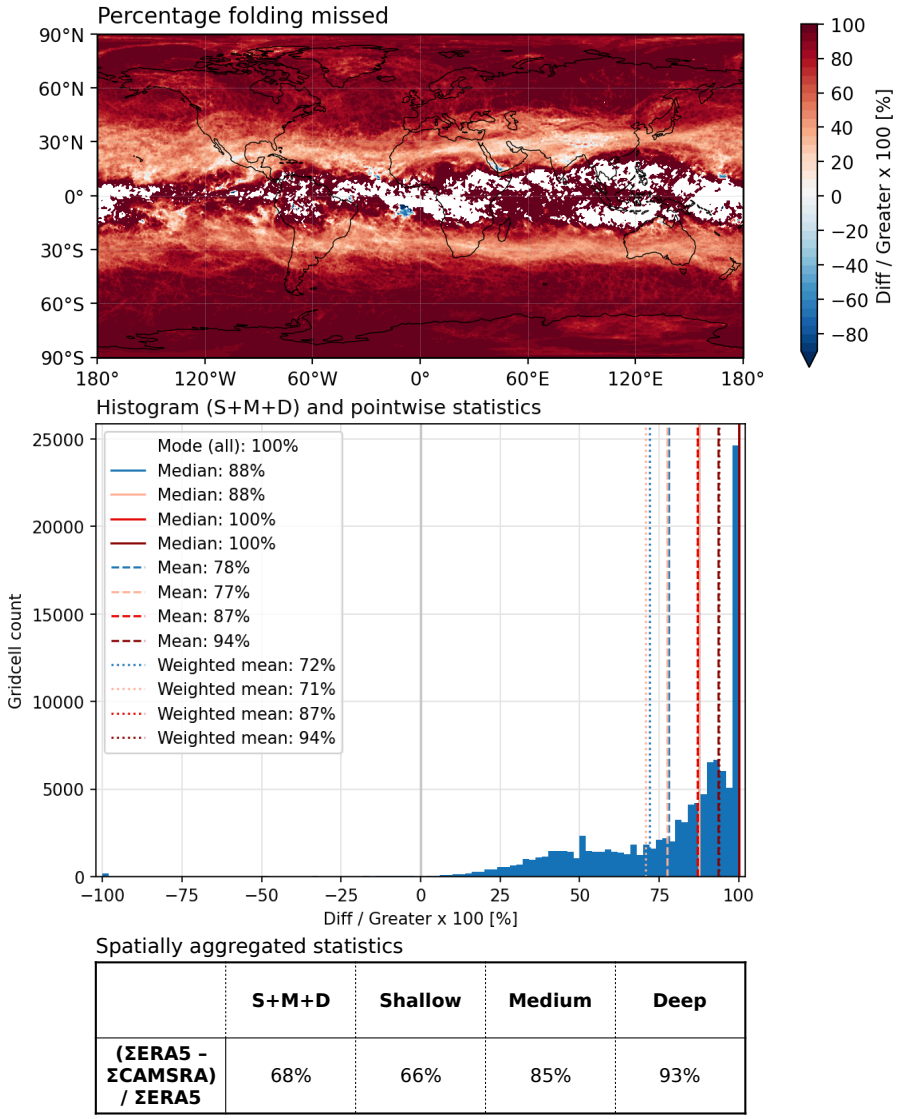


Figure S3. Fold thickness vs. model level resolution. ***Top row:*** For all Shallow folded columns during 2012 in CAMSRA (left) and ERA5 (right), a bivariate density histogram is shown. x-axis: the difference between the fold's vertical thickness ($pmax$–[$pmin$+dp]; see Figure 4) and the model level thickness at the column's lowest tropopause crossing. y-axis: the pressure at the lowest tropopuase crossing ($pmax$). ***Middle and bottom rows:*** Same as top row but for Medium and Deep folding occurrences.



Spatially aggregated statistics

	S+M+D	Shallow	Medium	Deep
(ΣERA5 – ΣCAMSRA) / ΣERA5	68%	66%	85%	93%

Figure S4. Statistics of folding missed by CAMSRA (or ERA5) over 2012. ***Top:*** As in Figure 2i but showing the whole globe instead of a zonal mean. At each gridcell, the percentage of folding in whichever dataset is higher frequency that is missed by whichever dataset is lower frequency (signed positive if ERA5 is higher frequency). In other words, the difference in folding frequency (ERA5 minus CAMSRA), expressed as a percentage of the greater of the two. Areas with no color indicate no folding in either dataset. ***Middle:*** Histogram of all gridcell values in the map above. Statistics are shown in the legend: the modes, medians, means, and area-weighted means for each of the folding depth categories are shown, with accompanying vertical lines (S+M+D, Shallow, Medium, and Deep in blue, pink, red, and dark red). For example, across all gridcells, it is most common for 100% of ERA5 folding at a given location during 2012 to be missed by CAMSRA (i.e., at a given gridcell, CAMSRA folding frequency is zero while ERA5 frequency is non-zero). On average across all locations, nearly 100% of ERA5 Deep folding is missing in CAMSRA (ranging from 94% to 100% by averaging type). ***Bottom:*** Statistics that ignore the location and timing of folding: the total number of CAMSRA folding occurrences over all locations and times in 2012 is compared to that of ERA5, for each depth category. For example, the total number of Deep folded columns is 93% lower in CAMSRA than in ERA5. Note that these sums are taken on the same grid (that of CAMSRA).

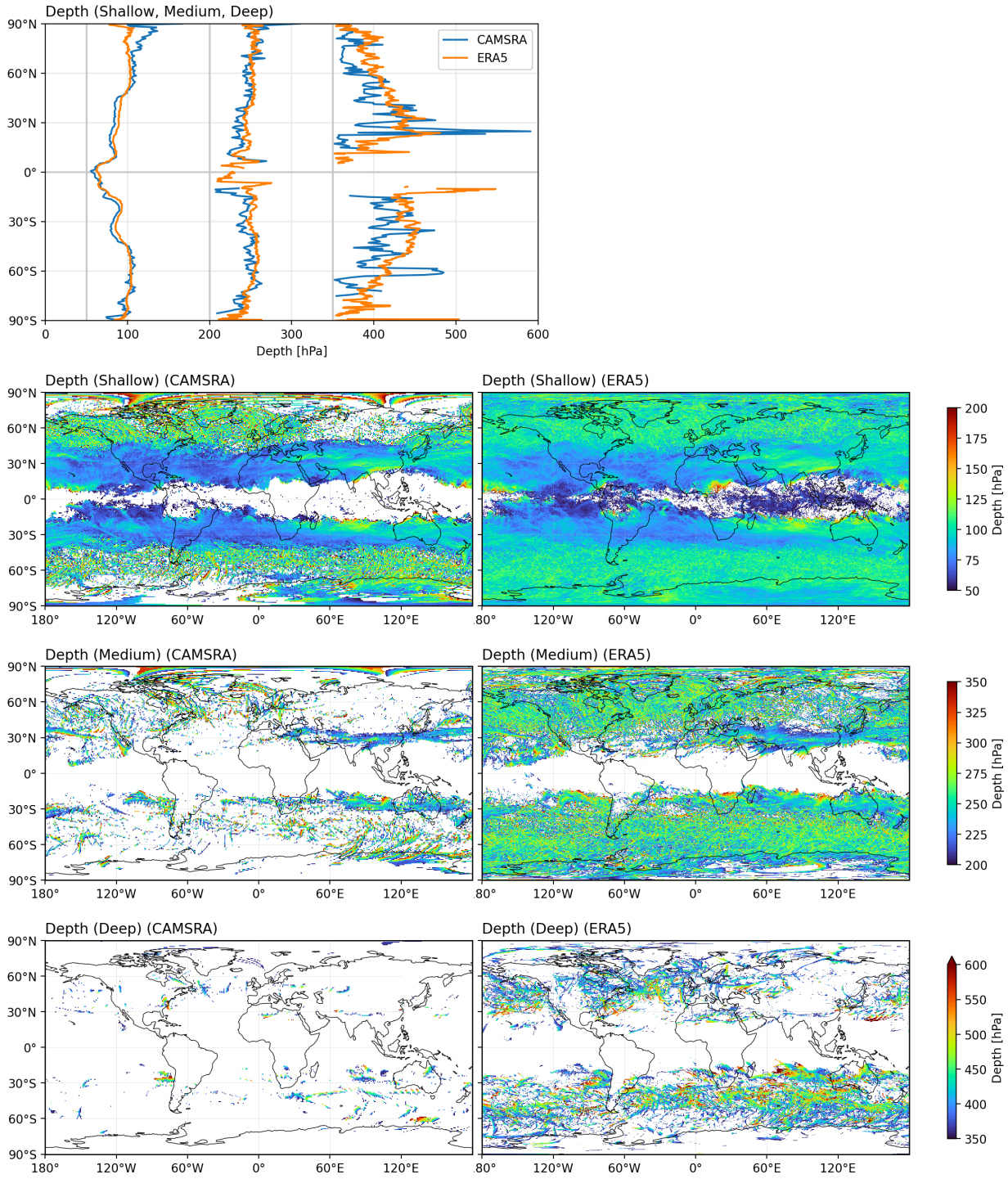


Figure S5. Depth of folding in spatial detail. Depth of folding (dp, see Figure 1a) as in Figure 4c–d's histograms, but shown as zonal means and full maps. ***Top:*** Zonal means of dp in CAMSRA and ERA5 for Shallow, Medium, and Deep folds. x-axis is continuous; gray lines indicate the three depth ranges. ***Second row:*** CAMSRA (left) and ERA5 (right) average folding depths for Shallow folds. White indicates no folding over the whole year. This spatial distribution of depth of Shallow folding tightly mirrors the folding frequency ratio (ERA5/CAMSRA) shown in Figure 1d. ***Third and fourth rows:*** As in second row but for Medium and Deep folding.

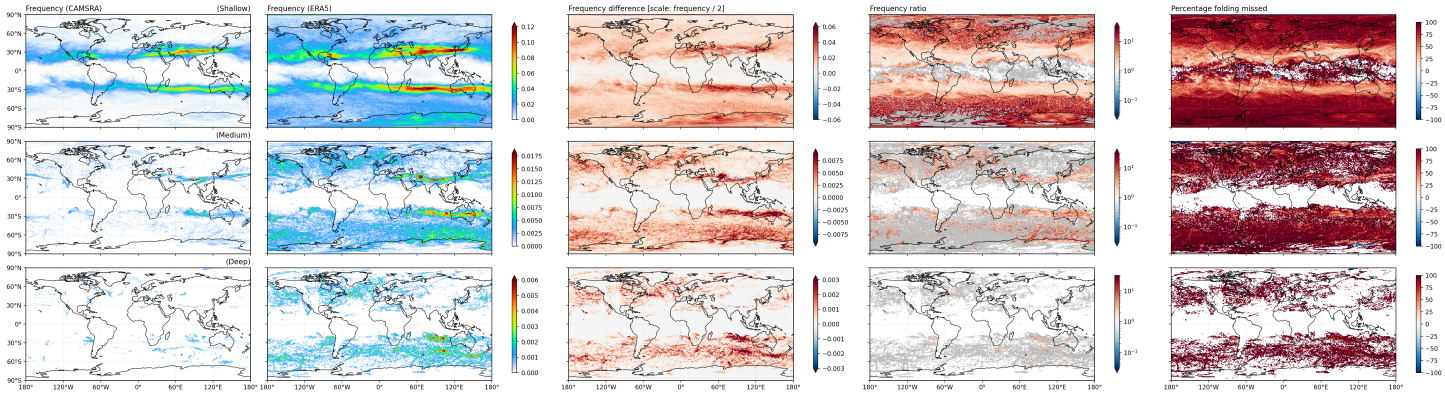


Figure S6. Folding frequency comparison by depth of folding. ***Top:*** (from left to right) Frequency of Shallow folding in CAMSRA and ERA5, their difference (ERA5–CAMSRA), their frequency ratio (ERA5/CAMSRA, with gray indicating zero denominator and non-zero numerator), and the percentage of folding in the higher-frequency of the two datasets that is missed by the lower-frequency of the two, as described in Figure S4's caption. ***Middle and bottom:*** As in top row but for Medium and Deep folding. Colorbar scales for the left three columns change across rows; those for the right two columns do not.

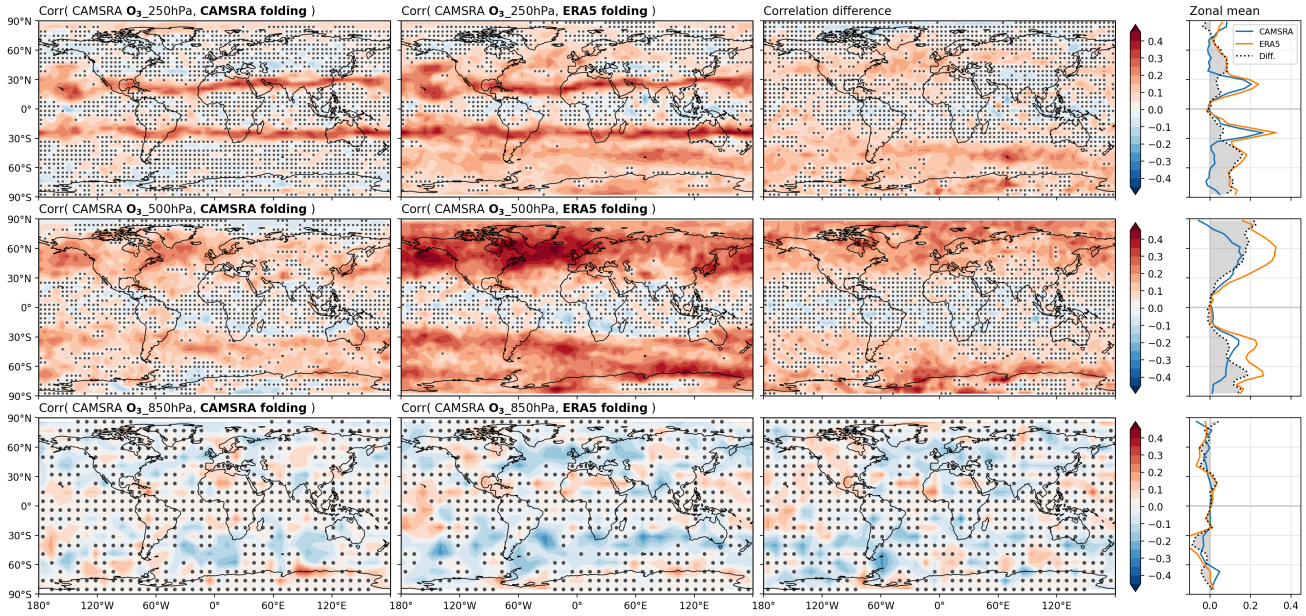


Figure S7. Correlations between CAMSRA tropospheric ozone and folding. As in Figure 3 but using total ozone O_3 instead of the stratospheric ozone tracer O_3S. The same conclusions are supported except at 850 hPa, where many other sources for tropospheric ozone besides the stratosphere are important. The correspondence of these maps at 250 and 500 hPa with those in Figure 3 indicates that O_3S is tightly related to total ozone at those levels.

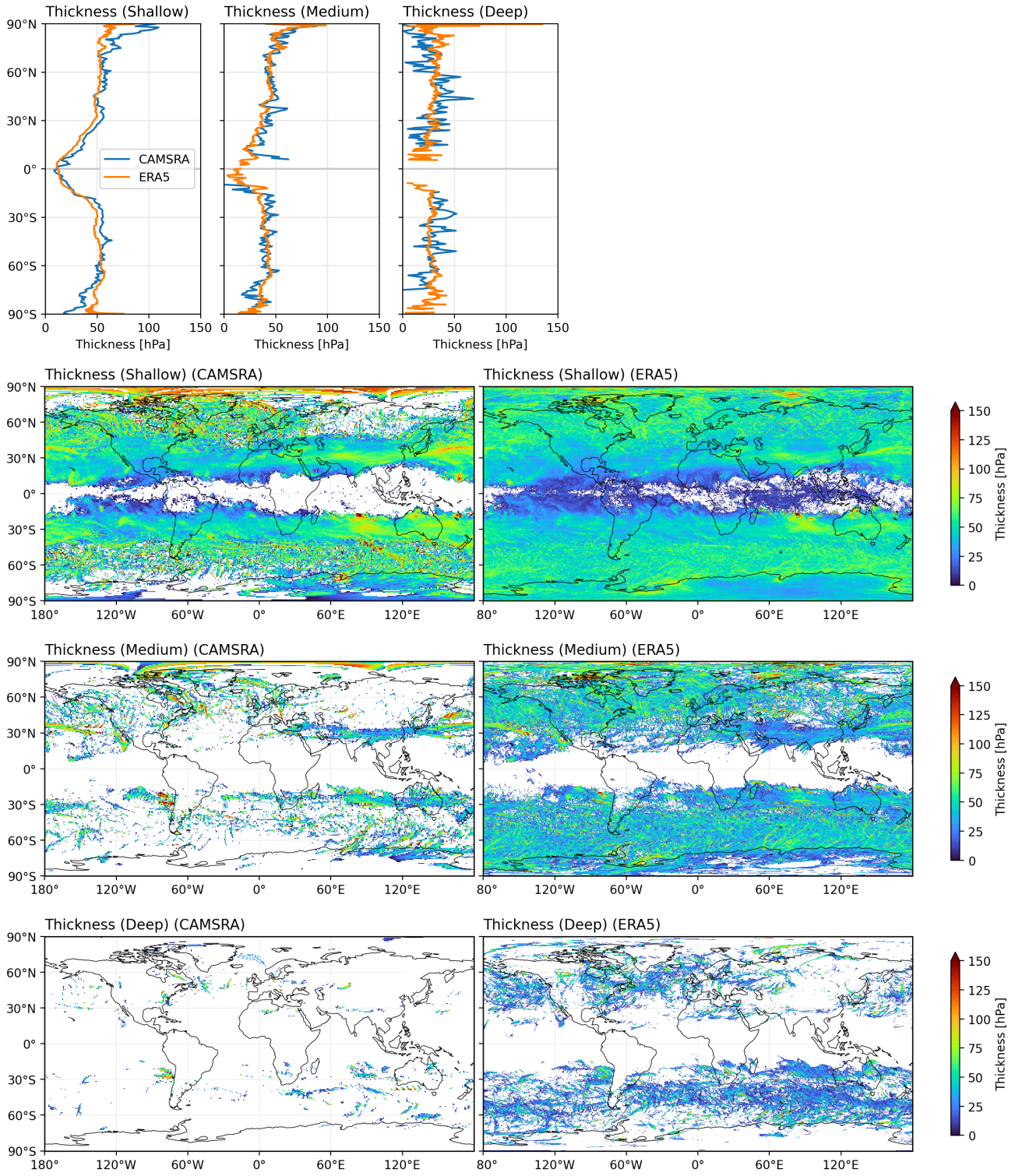


Figure S8. Thickness of folding in spatial detail. As in Figure S5 but for folding thickness (calculated as $pmax - (dp + pmin)$)

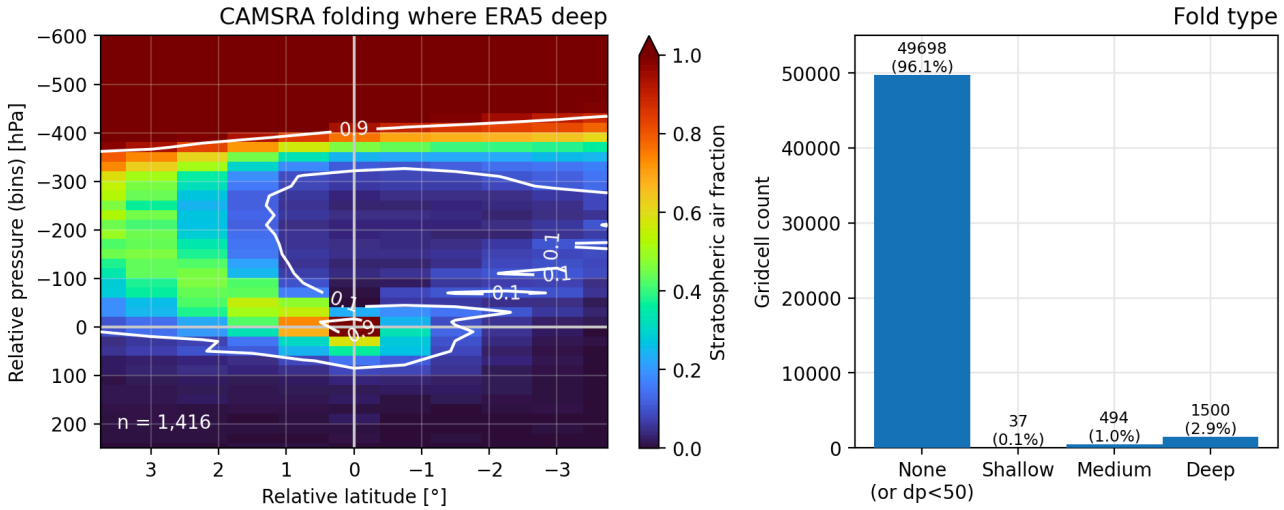


Figure S9. Fold cross-sections and folding types in CAMSRA where ERA5 identifies Deep folding. ***Left:*** A composite (as in Figure 4a–b) across 1,416 latitudinal cross-sections through column ranges in which any type of folding (Shallow, Medium, or Deep) is identified in CAMSRA at the same location and time that Deep folding is identified in ERA5. The 0.9 contour falls around 400 hPa above the top of the intrusion, which is closer than in ERA5 for all ERA5 Deep folding cases (Figure 4b). However, the 0.5 contour (not explicitly shown) is slightly over 350 hPa above, implying that for most of the cases in which CAMSRA does identify folding of any type, that folding is Deep. ***Right:*** Histogram of folding (or non-folding) types identified in all CAMSRA columns corresponding to Deep folding instances in ERA5. When folding is identified in CAMSRA it is most often Deep (1,500 columns) rather than Medium or Shallow (494 or 37 columns), in agreement with the cross-section composite (left [wherein all 2,031 of these columns belong to 1,416 contiguous latitudinal ranges]). However, across all 51,729 such columns, nearly all (96.1%) identify no folding at all in CAMSRA. Together these findings imply that CAMSRA is failing to resolve the tip of folds rather than resolving a fold at the wrong depth.