Note: Reviewer's comments are presented in black font; authors' responses are presented in blue plain font; manuscript text quotations are presented in blue bold font.

Anonymous Referee #2

We would like to thank Reviewer #2 for her/his time devoted and the constructive and helpful comments.

The authors present an evaluation of climatological tropopause folds in CAMSRA reanalysis against WOUDC ozonesonde and IAGOS measurements over Europe. A proven labelling algorithm is used to identify tropopause folds in 2003-2018. The overestimation of upper tropospheric O3 in CAMSRA results from the chemical data assimilation. The study presents useful validation and the manuscript is well-written, which merit its publication in ACP. However, there are some improvements that are required both in the writing and discussion. Some of the statements require detailed explanation to support.

We would like to thank Reviewer #2 for the general comments. Our point-by-point responses to the Reviewers comments are presented below.

Major comments:

- Please be more specific about the locations of folding hot spots. Why you are especially interested over Europe? I suggest to add some additional references on the global STT over different hot spots of fold activities.

Similar to our response to a comment raised by Reviewer #1, the present work was performed within the framework of a postdoctoral fellowship of the first author by the Greek State Scholarships Foundation (IKY) (Reinforcement of Postdoctoral Re-searchers - 2nd Cycle" (MIS-5033021)) co-financed by Greece and the European Union (European Social Fund—ESF). The aim of the project is the investigation of stratospheric intrusions and their role on tropospheric ozone levels and air quality over Europe with the synergistic use of CAMS reanalysis and observational data. Thus, and as a prerequisite step in this direction, the evaluation of the CAMS reanalysis O3 during such events is limited over the European region.

In the Revised Manuscript (RM) we have included additional references for the global STT over different hot spots of fold activity. The following are now included in the RM (L40-L44): **"The springtime western United States region is a hot spot of deep folding events with well-known implications for tropospheric ozone and air quality (Langford et al., 2009; Lin et al., 2012, 2015; Knowland et al., 2017). Recently, Luo et al. (2019) explored the seasonal features of tropopause folds over the Tibetan Plateau where folds occur frequently (Tyrlis et al., 2014), while other studies investigated the effect of tropopause folds on lower tropospheric ozone levels and air quality in China (Lu et al., 2019; Zhao et al., 2021b, a)."**

- How did you perform the control simulation of IFS without the use of chemical data assimilation? Since you claimed that the chemical data assimilation is the key factor resulting in most of the biases, you need to be more specific about your experiment design.

The following is now included in the RM as a description for the control run (L120-126): **"As it would have been computationally too expensive to produce a control analysis experiment that was identical to CAMSRA but did not actively assimilate observations of reactive gases, a forecast run was carried out that applied the same settings (model code, resolution, emissions) as used in CAMSRA. The control run was carried out as a sequence of 24 hours. The meteorological initial conditions were taken from CAMSRA, but the initial conditions for the atmospheric composition species, including ozone, from the previous forecast. It thus allows us to detect the impact of the assimilation of e.g. ozone data by comparing its ozone fields with CAMSRA."**

- I have concerns about the great differences between O3 and O3s in the upper troposphere (300-400hPa) as shown in your Figure 4 and 5. Green profiles differ largely away from the red ones at 300 hPa. Will they merge near the tropopause level? If not, I am very worried about the quality of the upper tropospheric O3 as well as the stratospheric ozone tracer in CAMSRA. can you please show a horizontal map of both the climatologies of O3 and O3s at 100 hPa and 850 hPa, respectively? Additionally, there have been various methods when calculating the O3s tracer among different models. Please add more details about the stratospheric ozone tracer (O3s) in CAMSRA. What is the choice of tropopause in defining the tropopause? I'm curious about it because the diagnostics of STT might be sensitive to the choice of tropopause definition.

We thank the Reviewer for this comment. Indeed, O3 and O3s concentrations differ between 300 and 400 hPa. As initially stated in the manuscript O3s, the stratospheric ozone tracer is identical with ozone in the stratosphere. This was the intension but unfortunately it turns out that this is not the case, due to an issue in the coupling of O3s to O3 in the stratosphere. So CAMSRA O3s in the stratosphere is only the modeled (Cariolle) ozone as data assimilation were not applied for that. As suggested by the Reviewer, Figure R1 presents the CAMSRA 2003-2018 climatology of O3 (left) and O3s (right) concentrations at 100 hPa (top) and 850 hPa (bottom). The abovementioned issue is reflected on the differences between O3 and O3s at 100 hPa, as well as on the vertical profiles of O3 and O3s concentrations over the 2003-2018 period (Figure R2). Tropopause in CAMSRA is calculated based on the temperature lapse rate, switching the chemistry scheme from CB05 (troposphere) to CARIOLLE (stratosphere) accordingly.

It must be noted, that in the present study O3s is used only for qualitative purposes to support the fact that during the selected STT events the induced CAMSRA O3 increase is associated with stratospheric O3 transport. Thus, we do not refer to specific O3s amounts as of transported to the troposphere anywhere in the manuscript. So, the fact that O3s is not generated from the O3 fields after assimilation is not crucial for the purpose of O3s usage here. In the RM manuscript (L129-134) we have replaced the O3s description with the following: **"Apart from O3, a stratospheric ozone tracer (O3s) is also used from CAMSRA providing a diagnostic of O3 STT. In principal, O3s in IFS is defined identically with O3 in the stratosphere, yet, in CAMSRA O3s is equal to the modeled (Cariolle scheme) O3 tracer and not the assimilation-resulted O3. In the troposphere O3s is subject to transport and chemical destruction just like O3. The tropopause in CAMSRA is calculated based on the temperature lapse rate, switching the chemistry scheme from CB05 (troposphere) to CARIOLLE (stratosphere) accordingly. It should be noted, that O3s is used here only as a qualitative diagnostic of ozone STT, to support evidence of stratospheric ozone downward transport during the folding events."**

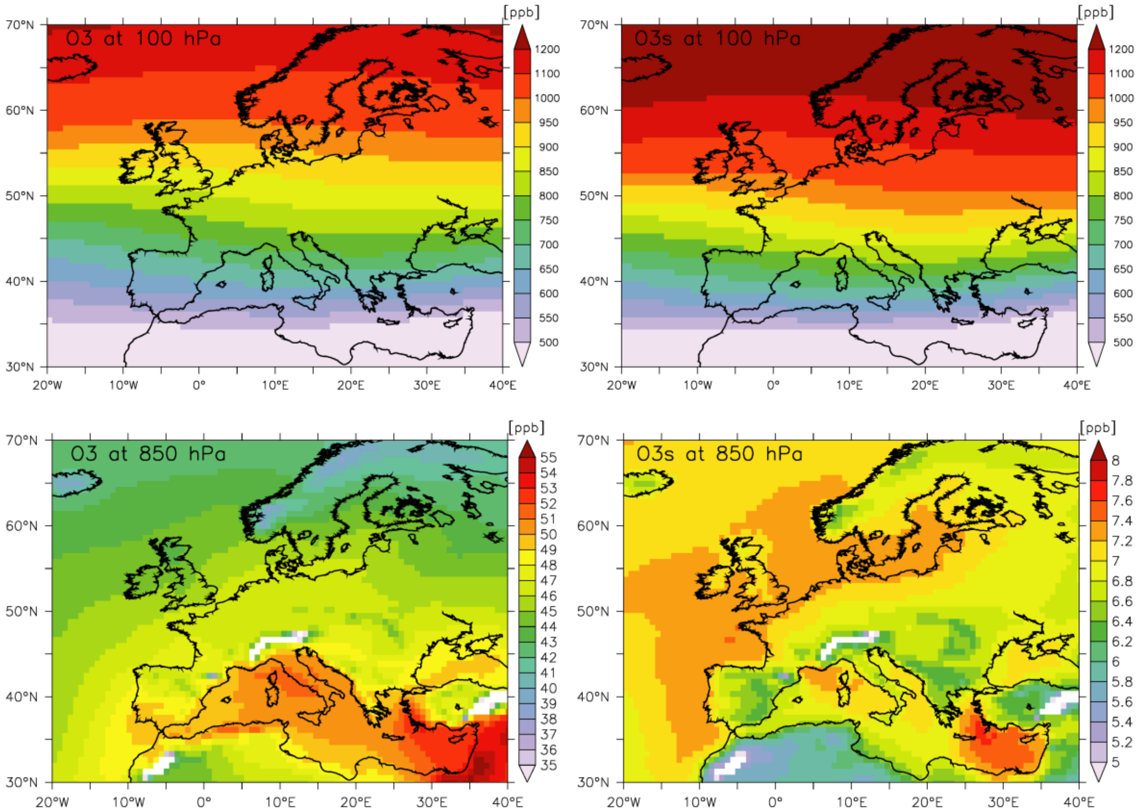

Figure R1. CAMSRA 2003-2018 climatology of O3 (left) and O3s (right) concentrations at 100 hPa (top) and 850 hPa (bottom).

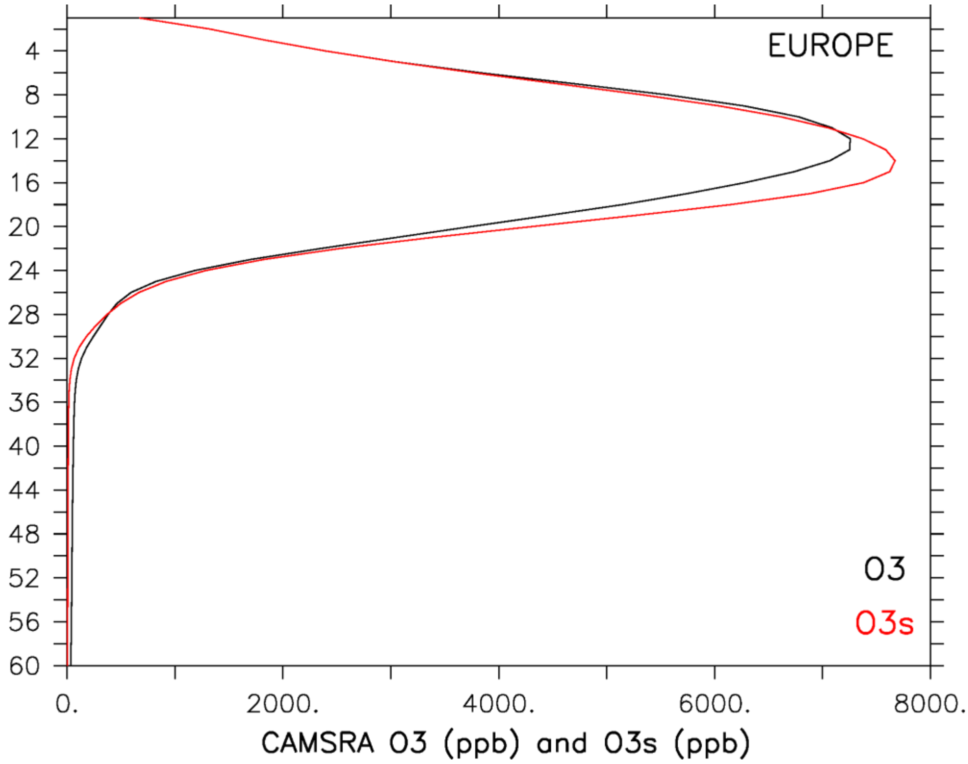

Figure R2. Vertical profiles of CAMSRA 2003-2018 climatology of O3 (black) and O3s (red) concentrations. The vertical axis displays CAMSRA model levels (from surface to the top).

- Please add more details about the two evaluation metrics (FGE vs MNMB) chosen in your study. Why you chose these two statistics and what are differences between the two? What does the score mean, respectively?

Since O3 concentrations are in general increased with altitude exhibiting higher values near the tropopause, we prefer to show normalized metrics for easier and more straightforward interpretation. The fractional gross error (FGE) is a normalized version of the mean error metric, while the modified normalized mean bias (MNMB) is a normalized version of the mean bias. Both metrics are normalized by the mean of the observed and CAMSRA values and have the advantage to behave symmetrically with respect to under- and overestimation, being less sensitive to outliers in the distribution. They are dimensionless and relative, making them suitable for comparison at different parts of the troposphere. The FGE is a measure of the overall error of CAMSRA, while MNMB is a measure of the overall bias. The MNMB metric is useful in the direction of showing if and how much overall CAMSRA under- or overestimates the observed O3 concentrations, while the FGE is used to depict the overall CAMSRA error regardless under- or overprediction. Both metrics are widely used in atmospheric composition evaluation studies related to CAMS (e.g. Katragkou et al., 2015; Akritidis et al., 2018; Inness et al., 2019; Wagner et al., 2021). In the RM we have modified the respective paragraph as follows (L210-215): **"For a quantitative comparison between CAMSRA and observations, we present in Figures 6 the vertical profiles of fractional gross error (FGE) and modified normalized mean bias (MNMB) of CAMSRA O3 for the WOUDC ozonesonde sites and IAGOS airports. The FGE is a normalized version of the mean error, while the MNMB is a normalized version of the mean bias. Both metrics are normalized by the mean of the observed and model (here CAMSRA) value, being dimensionless and relative, thus suitable to use at different heights in the troposphere. FGE and MNMB are insensitive to outliers in the distribution, and range between 0 to 2 and -2 to 2, respectively, behaving symmetrically with respect to under- and overestimation."**

- Section 3.1 and 3.2, what leads to the different biases across stations?

Similar to our response in a comment raised by Reviewer #1: Regarding the climatological comparison between CAMSRA and observations, all stations exhibit an overestimation of CAMSRA O3 in the upper troposphere which is the main feature. Attributing the different overestimations among the examined sites in specific reasons is difficult, but potential factors can indeed be discussed. To our knowledge no site-specific data quality issues are reported, but both ecc and Brewer Mast ozonesonde measurements introduce uncertainties which might affect in different ways the comparison among the examined sites. In addition, the proximity of the selected grid points to the respective ozonesonde site, and the CAMSRA 3-D spatial representation of the IAGOS take-off/landing are also another factor to consider. Accordingly, we have included the following sentence in the RM (L184-187): **"The differences seen in the comparison between the observed and CAMSRA O3 concentrations among the examined sites are subject to the uncertainties introduced by the ozonesonde instrument measurements, as well as the proximity of the selected grid points to the respective ozonesonde sites, and the CAMSRA 3-D spatiotemporal representation of the IAGOS take-off/landing routes."**

As for the differences of CAMSRA-obs comparison during the selected STT events, in addition to the aforementioned factors and even more importantly is the vertical location and geometrical characteristics of these events. More specifically, the pressure level in which a part of the fold appears over the examined site, directly affects ozone concentrations, and shapes its vertical profile. Thus, for observational sites with not so extended number of stt events the individual dynamics and ozone increases in different levels of the troposphere can form somehow unique structures of CAMSRA and observed O3 deviations. In contrast, sites where more stt events were selected (e.g. Frankfurt and Munich) exhibit more smoothed profiles for both CAMSRA and observations.

In more detail, Figure R3 presents a schematic representation of fold detection, also depicting the multiple crossings of the stratosphere in vertical profiles. As mentioned in the manuscript the pressure difference between the middle and upper stratosphere crossing $\Delta p = p_m - p_u$ indicates the vertical extent of the fold, while the vertical area between the pm and pl represents the area of the fold that directly affects the ozone profile over the underlying site. Such crossings during the stt events detected at the ozonesonde sites (not shown for IAGOS as several folds might be present during a selected

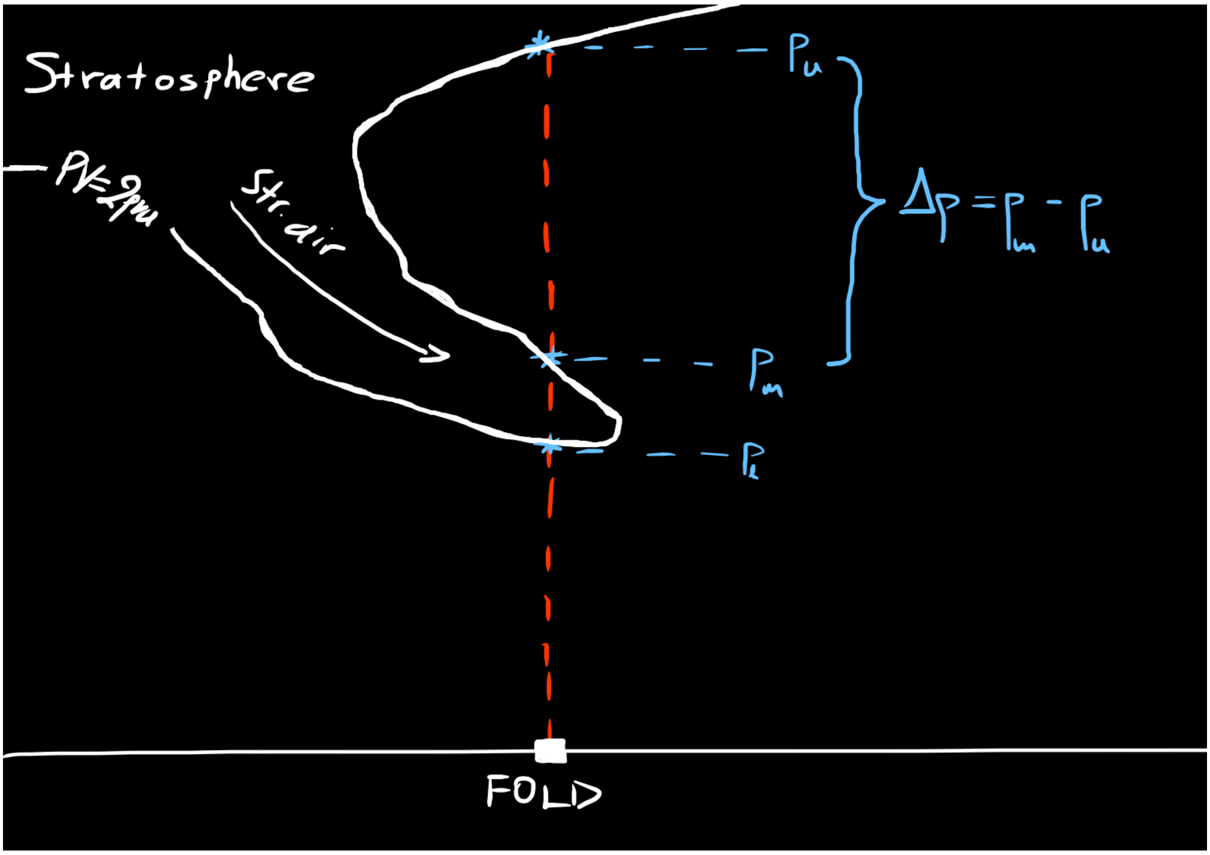

Figure R3. Schematic representation of fold detection.

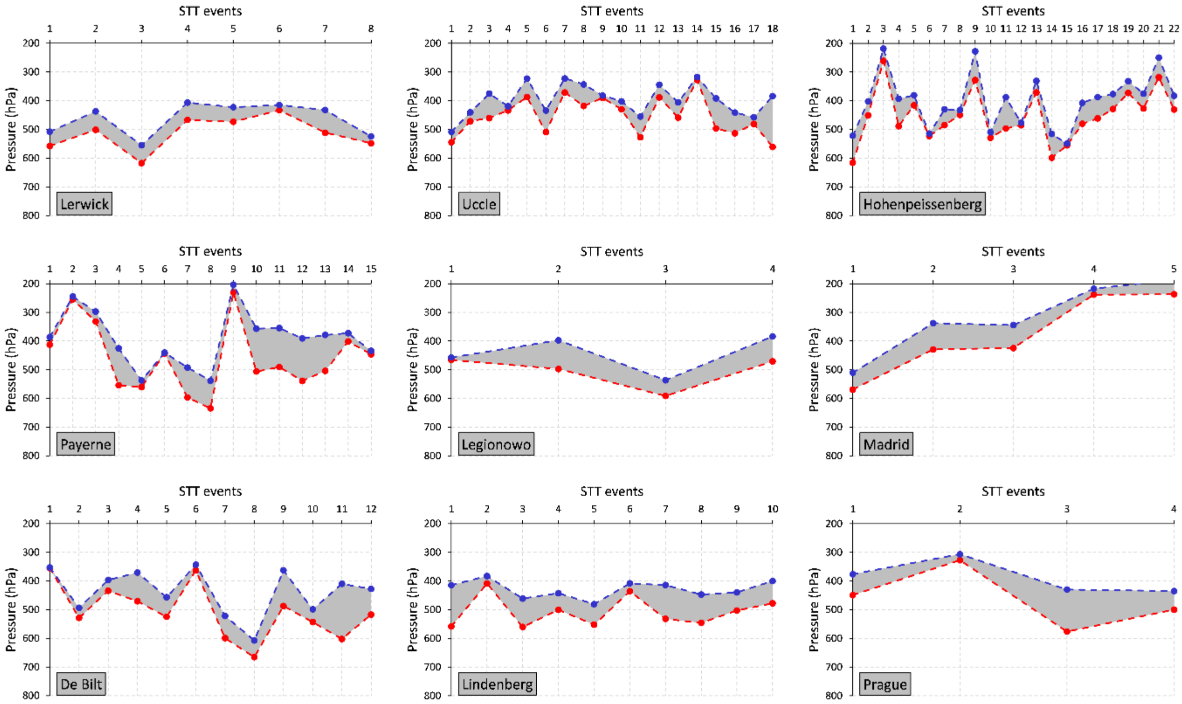

Figure R4. Vertical location of middle (blue) and lower (red) crossings of the stratosphere during STT events at the WOUDC ozonesonde sites.

event following the take-off/landing route) are depicted in Figure R4. The following sentence is now included in the RM (L205-207): **"In addition, the individual dynamics and the different vertical location and geometrical characteristics of the selected STT events, especially for observational sites with not so extended number of events, may form somehow unique structures of CAMSRA and observed O3 deviations".**

Minor comments:

- Lines 30-35: please be more specific about the hot spots of fold activities. Please add some references for transport of VSLS to the lower stratosphere

The hot spots of fold activity are presented in the next paragraph (see L38-50 in the RM). Regarding the transport of VSLS to the lower stratosphere the following phrase was included in the RM (L27-29): **"The latter constitutes an important pathway through which very short lived substances (VSLS), emitted at the surface, can be transported to the lower stratosphere influencing ozone (Levine et al., 2007; Aschmann et al., 2009; Liang et al., 2014)."**

- Line 39, a comma is missing before "resulting"

Done.

Line 68: can you add a few sentences about the differences between the ecc and the Brewer Mast ozonesondes?

The following is now included in the RM (L75-78): **"Both ozonesonde types are based on the same measurement principle of ozone electrochemical detection in potassium iodine. The major differences between ecc and Brewer Mast ozonesondes are that the latter uses only one reaction chamber, and a silver anode instead of a platinum anode, requiring an external electrical potential in contrast to the ecc (Beekmann et al., 1994)."**. Moreover, the Komhyr (1969) and Brewer and Milford (1960) references are now included in the RM.

Line 90: can you be more specific about "CY42R1" and "4D-VAR"?

Further details on CAMS reanalysis are provided at https://confluence.ecmwf.int/display/CKB/CAMS%3A+Reanalysis+data+documentation. Although documentation for CY42R1 is not available, details on earlier and later IFS cycles for reference can be found at https://www.ecmwf.int/en/publications/ifs-documentation. In the RM we have included further information on the 4D-VAR system and the meteorological observations assimilated in IFS: (L102-104) **"In more detail, it is based on the minimization of a penalty function that takes the**

deviations of the model's background fields from the observations to provide the optimal forecast during 12-hour assimilation windows (from 09 UTC to 21 UTC and 21 UTC to 09 UTC) by modifying accordingly the initial conditions." and (L107-109) **"In addition, meteorological observations, including satellite, PILOT, in situ, radiosonde, dropsonde, and aircraft measurements are also incorporated in IFS."**.

- Line 92: NO_2

Done.

- Line 99: what is the vertical resolution in the upper troposphere?

There are 13 levels between approximately 400 and 100 hPa. The following is now included in the RM (L115): **"(13 levels between approximately 400 and 100 hPa)"**.

- Line 130: the rest of

Done.

- Line 142: I suggest adding the longitude and latitude information of each station to Figure 3.

Done.

- Line 166: Figure

Done.

References

Akritidis, D., Katragkou, E., Zanis, P., Pytharoulis, I., Melas, D., Flemming, J., Inness, A., Clark, H., Plu, M., and Eskes, H.: A deep stratosphere-to-troposphere ozone transport event over Europe simulated in CAMS global and regional forecast systems: analysis and evaluation, Atmos. Chem. Phys., 18, 15515–15534, https://doi.org/10.5194/acp-18-15515-2018, 2018

Aschmann, J., Sinnhuber, B.-M., Atlas, E. L., and Schauffler, S. M.: Modeling the transport of very short-lived substances into the tropical upper troposphere and lower stratosphere, Atmos. Chem. Phys., 9, 9237–9247, https://doi.org/10.5194/acp-9-9237-2009, 2009

Beekmann, M., Ancellet, G., Megie, G., Smit, H., and Kley, D.: Intercomparison campaign of vertical ozone profiles including electrochem-290ical sondes of ECC and Brewer-Mast type and a ground based UV-differential absorption lidar, Journal of atmospheric chemistry, 19,259–288, https://doi.org/10.1007/BF00694614, 1994.

Brewer, A. W. and Milford, J. R.: The Oxford-Kew Ozone Sonde, Proceedings of the Royal Society of London. Series A, Mathematical andPhysical Sciences, 256, 470–495, http://www.jstor.org/stable/2413928, 1960.

Inness, A., Ades, M., Agustí-Panareda, A., Barré, J., Benedictow, A., Blechschmidt, A.-M., Dominguez, J. J., Engelen, R., Eskes, H., Flemming, J., Huijnen, V., Jones, L., Kipling, Z., Massart, S., Parrington, M., Peuch, V.-H., Razinger, M., Remy, S., Schulz, M., and Suttie, M.:

The CAMS reanalysis of atmospheric composition, Atmos. Chem. Phys., 19, 3515–3556, https://doi.org/10.5194/acp-19-3515-2019, 2019

Katragkou, E., Zanis, P., Tsikerdekis, A., Kapsomenakis, J., Melas, D., Eskes, H., Flemming, J., Huijnen, V., Inness, A., Schultz, M. G., Stein, O., and Zerefos, C. S.: Evaluation of near-surface ozone over Europe from the MACC reanalysis, Geosci. Model Dev., 8, 2299–2314, https://doi.org/10.5194/gmd-8-2299-2015, 2015

Komhyr, W.: Electrochemical concentration cells for gas analysis, 25, 203–210, 1969.

Langford, A. O., Aikin, K. C., Eubank, C. S., and Williams, E. J.: Stratospheric contribution to high surface ozone in Colorado duringspringtime, Geophysical Research Letters, 36, https://doi.org/10.1029/2009GL038367, 2009

Levine, J. G., Braesicke, P., Harris, N. R. P., Savage, N. H., and Pyle, J. A. (2007), Pathways and timescales for troposphere-to-stratosphere transport via the tropical tropopause layer and their relevance for very short lived substances, J. Geophys. Res., 112, D04308, doi:10.1029/2005JD006940.

Liang, Q., Atlas, E., Blake, D., Dorf, M., Pfeilsticker, K., and Schauffler, S.: Convective transport of very short lived bromocarbons to the stratosphere, Atmos. Chem. Phys., 14, 5781–5792, https://doi.org/10.5194/acp-14-5781-2014, 2014

Lin, M., Fiore, A. M., Cooper, O. R., Horowitz, L. W., Langford, A. O., Levy, H., Johnson, B. J., Naik, V., Oltmans, S. J., and Senff,350C. J.: Springtime high surface ozone events over the western United States: Quantifying the role of stratospheric intrusions, Journal ofGeophysical Research: Atmospheres, 117, 2012

Lin, M., Fiore, A. M., Horowitz, L. W., Langford, A. O., Oltmans, S. J., Tarasick, D., and Rieder, H. E.: Climate variability modulateswestern US ozone air quality in spring via deep stratospheric intrusions, Nature communications, 6, 2015

Knowland, E., Doherty, R., and Hodges, K.: The effects of springtime mid-latitude storms on trace gas composition determined from theMACC reanalysis, Atmospheric Chemistry and Physics, 15, 3605–3628, https://doi.org/10.5194/acp-15-3605-2015

Luo, J., Liang, W., Xu, P., Xue, H., Zhang, M., Shang, L., and Tian, H.: Seasonal features and a case study of tropopause folds over theTibetan Plateau, Advances in Meteorology, 2019, https://doi.org/https://doi.org/10.1155/2019/4375123, 2019

Lu, X., Zhang, L., Chen, Y., Zhou, M., Zheng, B., Li, K., Liu, Y., Lin, J., Fu, T.-M., and Zhang, Q.: Exploring 2016–2017 surface355ozone pollution over China: source contributions and meteorological influences, Atmospheric Chemistry and Physics, 19, 8339–8361,https://doi.org/10.5194/acp-19-8339-2019, 2019

Tyrlis, E., Škerlak, B., Sprenger, M., Wernli, H., Zittis, G., and Lelieveld, J.: On the linkage between the Asian summer monsoon andtropopause fold activity over the eastern Mediterranean and the Middle East, Journal of Geophysical Research: Atmospheres, 119, 3202–3221, 2014

A. Wagner, Y. Bennouna, A.-M. Blechschmidt, G. Brasseur, S. Chabrillat, Y. Christophe, Q. Errera, H. Eskes, J. Flemming, K. M. Hansen, A. Inness, J. Kapsomenakis, B. Langerock, A. Richter, N. Sudarchikova, V. Thouret, C. Zerefos; Comprehensive evaluation of the Copernicus Atmosphere Monitoring Service (CAMS) reanalysis against independent observations: Reactive gases. Elementa: Science of the Anthropocene 21 January 2021; 9 (1): 00171. doi: https://doi.org/10.1525/elementa.2020.00171

Zhao, K., Huang, J., Wu, Y., Yuan, Z., Wang, Y., Li, Y., Ma, X., Liu, X., Ma, W., Wang, Y., and Zhang, X.: Impact of Stratospheric Intrusionson Ozone Enhancement in the Lower Troposphere and Implication to Air Quality in Hong Kong and Other South China Regions, Journal430of Geophysical Research: Atmospheres, 126, e2020JD033 955, https://doi.org/https://doi.org/10.1029/2020JD033955, e2020JD0339552020JD033955, 2021a

Zhao, K., Hu, C., Yuan, Z., Xu, D., Zhang, S., Luo, H., Wang, J., and Jiang, R.: A modeling study of the impact of stratospheric in-trusion on ozone enhancement in the lower troposphere over the Hong Kong regions, China, Atmospheric Research, 247, 105 158, https://doi.org/https://doi.org/10.1016/j.atmosres.2020.105158, 2021b