Authors' response to reviewer #2

Please note that in this document, colors' codes are as the referee's comments in black, the authors' responses in blue. Authors' changes in the manuscripts are shown in blue bold. The tables and figures are referred based on their numbers in the revised manuscript with markup (RM).

This study uses observational and CAMS reanalysis and forecast data to investigate O_3 variability and errors in CAMS global systems for several stations over Iran. To this end, the observed and CAMS O_3 time series are decomposed in three spectral components short, medium, and large, with the latter one not examined. Subsequently, an LSTM neural-network is applied to downscale the components and then investigate the CAMS performance and error sources.

Overall, I find this to be an interesting study, yet, with several points that need clarification and improvement, to suggest publication in GMD.

The authors highly acknowledge the referee for spending time and providing valuable and punctilious comments. An investigation of the L term could add more values to the article. But due to limitations in measured data (e.g., time series with long time periods over Iran), the assessment of this term was not possible. Most of the suggestions were applied in the RM; here are a few points to be mentioned.

Main comments

1. Title of the paper: The authors use surface-ozone data from both observations and CAMS (RA and FC). Is this correct? If yes, I don't understand the rationale behind the use of term "tropospheric ozone" in the title and in the abstract. Tropospheric ozone variability is governed by different processes and spatiotemporal scales compared to surface ozone, thus, I believe that the title is somehow misleading.

That is absolutely right. The term "Tropospheric" had been kept fixed since the early version of the paper, which was more concerned about ozone at upper levels. In the RM, the title was modified as:

L1: **Assessment of surface ozone products …**

2. I feel that more interpretation of the results is needed in the manuscript. For example, uncertainties in CAMS emissions inventories and deposition, and the fact that the two products use different emissions inventories should be discussed, especially for the S term.

Very important point. Emissions have large impact on both terms. Uncertainties in the spatial variations of the emissions inventories were discussed.

L361-363: **That could arise from the uncertainties in O_3 precursor emissions affecting modeled local photochemistry and likely S variability. The largest value of the MSE is associated with the O_3^{SD} of the stations located in the city of Tehran. That can be associated with the uncertainties in CAMS emissions inventories, which may have larger impact in cites with high anthropogenic**

emissions sources. The stations in the northern part of the city (e.g., stations 4, 5, 6, 7, 8, and 9) show a larger MSE than the stations in the southern part (e.g., stations 10, 11, 14, 15, 16, 17, and 19). That can be associated with the deficiency of the emissions inventories in capturing the local emissions changes within urban areas.

L370-373: **That can be associated with the uncertainties in the spatial variations of the emissions inventories used in CAMS. Although the CAMS anthropogenic emission inventories account for emissions from different sectors, such as transportation, residential and energy sectors, as well as biogenic fluxes, they have a temporal and spatial allocation with a monthly spatial grid resolution of 0.1°x0.1°.**

L380-418: **The S variability results from the effect of daytime photochemical production, downward transport of O_3 rich from upper levels, combined with O_3 loss by depositions (in the surface layer).**

L491-502: **That could arise from the uncertainty of O_3 precursors in CAMS, as they are not adjusted by data assimilation systems. CAMS-GLOB-BIO (used in CAMSFC, see Table 1) provides monthly average of the global biogenic emissions, which are calculated using the MEGAN (used in CAMSRA, see Table 1), driven by ERA-Interim meteorological fields.**

L511-516: **The impact of these (meso scale) factors cannot be captured in a global emissions inventory with limited resolutions. That induces large model uncertainties, in particular for the S variability, which has large associations with pollutant species. Besides, for some periods the emissions are not available and so prescribed, which means they are either kept fixed since the last year available or extrapolated (projected) with a climatological trend. MACCity emission inventory has not been updated after 2010, and recent years are only based on projections of past trends. CAMS-GLOB-ANT provides the monthly average of the global emissions of 36 compounds over the period 2000-2019.**

3. The role of stratospheric ozone contribution to the surface ozone is only discussed once in the Discussion. Stratospheric ozone can affect surface ozone levels indirectly through vertical downward transport of ozone from the lower stratosphere and/or the upper troposphere in larger time scales (Zanis et al., 2014) or directly through intense stratospheric intrusions (rarer) (Akritidis et al., 2010, Chen et al., 2022). The broader Iran region is a well-known hot spot of stratosphere-to-troposphere transport that might affect day-to-day O3 variability in some cases. The CAMS reanalysis product includes a tracer for stratospheric ozone that might be useful tracer for stratospheric ozone. This is a suggestion that the authors might consider for the present or future studies. At least, the authors should consider for the discussion that by not including such a source (stratospheric ozone) of surface ozone is a potential source of error.

Good point. Yes, Iran is recognized as a hot spot for folding activities and rare deep folds, in particular in spring and summer. The fold frequency distribution over the eastern Mediterranean and the Middle East shows two distinct maxima, one of which is located over Iran. Several studies show that the subsidence of stratospheric ozone, under given weather systems over Iran. This point was further discussed in the RM.

Assessing of the stratospheric ozone tracer is a great idea, and could help in the investigation of the error sources. We consider this point for future study due to lack of time.

L520-531: **Stratospheric ozone can affect surface ozone levels indirectly through vertical downward transport of ozone from the lower stratosphere and/or the upper troposphere in larger time scales (Zanis et al., 2014) or directly through intense stratospheric intrusions (rarer) (Akritidis et al., 2010; Chen et al., 2022). Over Tehran, a major portion of O_3 during spring is transferred from the stratosphere (Aliakbari Bidokhti and Shariepour, 2007). A study by (Shariepour and Aliakbari Bidokhti, 2013) showed that several mid-latitude low pressure weather systems accompanied by tropopause folding affect northern Iran (Caspian Sea), and can cause downward transport of stratospheric ozone rich air towards the surface. During summer, the occurrence of tropopause folding and their intensity over the Eastern Mediterranean and the Middle East regions are majorly controlled by the Asian monsoon. Since the zone of upper level baroclinicity and fold occurrences spread northwestward over this region, it first reaches Iran (Tyrlis et al., 2014).**

L610-612: **In addition, uncertainties in stratospheric ozone and emission inventory might affect this error. Analysis of a tracer for stratospheric ozone can help to identify non-local ozone sources (a recommendation for future work).**

Akritidis, D., Zanis, P., Pytharoulis, I., Mavrakis, A., and Karacostas, Th.: A deep stratospheric intrusion event down to the earth's surface of the megacity of Athens, Meteorol. Atmos. Phys., 109, 9–18, doi:10.1007/s00703-010-0096-6, 2010.

Aliakbari Bidokhti, A.A. and Shariepour, Z.: Analysis of surface ozone variability in the vicinity of synoptic (meteorology) station of Geophysics Institute (Tehran University) for the year 2002, J. Environ. Stud. 33 (42), 63–74, 20.1001.1.10258620.1386.33.42.7.7, 2007.

Chen, Z., Liu, J., Qie, X., Cheng, X., Shen, Y., Yang, M., Jiang, R., and Liu, X.: Transport of substantial stratospheric ozone to the surface by a dying typhoon and shallow convection, Atmos. Chem. Phys., 22, 8221–8240, https://doi.org/10.5194/acp-22-8221-2022, 2022.

Shariepour, Z. and Aliakbari Bidokhti, A.A.: Investigation of Surface Ozone over Tehran for 2008-2011, 39, 191–206, https://doi.org/10.22059/jesphys.2013.35607, 2013.

4. I agree with the main comment raised by Reviewer 1. The manuscript should be more oriented to presenting the main objectives of the study and then the main implications. For sure, this will make it more reader friendly.

The main objective of the study is to show the capability of the reanalysis products in capturing the surface ozone variability over Iran. Many scientists in Iran use reanalysis datasets (instead of observational data) and provide advices for policy applications. This issue can extend to other chemical species. These points were considered in the RM as:

L15-16: **Results show the benefit of the LSTM method compared to using the raw CAMS products for providing O_3 over Iran.**

L20-22: **This study demonstrates that coarse-scale global model data such as CAMS needs to be downscaled for regulatory purposes or policy applications at local scales. Our method can be useful not only for the evaluation but also for the prediction of other chemical species, such as aerosols.**

L616-620: **To date, most of the studies of ozone and other pollutants in Iran rely on reanalysis products, without using decompositions or downscaling procedures. Our findings show that the CAMSRA and CAMSFC datasets have some deficiencies in simulating ozone, in particular over the cities with high emissions of ozone precursors. Downscaling improves these products and makes them suitable for the study of ozone in major metropolitan areas. The method used in this study is not only applicable for the evaluation of the global models but also for prediction purposes.**

Comments

- P1, L21: A recent review of air pollution impacts on health is worth cited here by Pozzer et al. (2023).

  Yes, that is right. It was added.

  L25: **Malley et al., 2015; Pozzer et al., 2023).**
  Pozzer, A., Anenberg, S. C., Dey, S., Haines, A., Lelieveld, J., and Chowdhury, S.: Mortality attributable to ambient air pollution: A review of global estimates, GeoHealth, 7, https://doi.org/10.1029/2022GH000711, 2023.

- P5, L135-136: "whereas long-term and seasonal variation is mainly related to solar radiation." What about long-range transport and stratosphere-to-troposphere transport (Monks et al., 2000).

  That was a missing point. It was added in the RM as

  L180-181: **whereas long-term and seasonal variations are mainly related to solar radiation, long-range transport and transport from the stratosphere to the troposphere (Monks, 2000).**

  Monks, P. S.: A review of the observations and origins of the spring ozone maximum, Atmos. Environ., 34, 3545–3561, https://doi.org/10.1016/S1352-2310(00)00129-1, 2000.

- P6, L166-167: I appreciate the fact that the authors performed a hyperparameter tunning instead of using default values.

That was because of the outputs were more sensitive to epoch, we only tuned this parameter. In the RM, most of the hyperparameters were tuned, and the results were improved. The hyperparameters to be optimized and their value ranges are shown in Table A2.

L340-348: **We tuned hyperparameters, which allow the learning algorithm to run until the error from the model, i.e., the loss function, has been sufficiently minimized. As there are no given values to set these numbers, the optimum values were obtained by multiple trial-and-error tests (see Table A5).**

Table A2. The hyperparameter settings of the LSTM model

hyperparameter	**values**
Train portion	65 %
Test portion	35 %
Epoch	1…30
Batch size	**[12, 24, 48, 72, 96, 120]**
Optimizer	ADAM
Units (hidden layer)	**2…10**
Dropout rate	**0.001**
Learning rate	**0.001…0.1**
Loss function	MSE

Table A5. The optimum units, dropout, learning rate and batch size to perform the LSTM model. Here, the T and F refer to True and False.

models	**CAMSRA**		**CAMSFC**	
Stations' number	**S**	**M**	**S**	**M**
1	**10, T, 0.04, 24**	**10, T, 0.04, 24**	**2, T, 0.09, 48**	**4, F, 0.04, 48**
2	**4, F, 0.04, 48**	**4, F, 0.04, 48**	**10, T, 0.04, 24**	**10, T, 0.04, 24**
3	**10, T, 0.04, 24**	**4, F, 0.04, 48**	**2, T, 0.09, 48**	**4, F, 0.04, 48**
…	**…**	**…**	**…**	**…**

- P6, L175-176: Regarding the training of the LSTM model:

  Are the data shuffled prior to the training process?

  To preserve the logical sequence of the data, the data was not shuffled. But since the shuffling happens on the batches axis and not on the time axis, this parameter is set to True in the RM.

Did you apply "early stopping" during training to help avoid overfitting?

Good point for the choice of epochs number. This parameter was applied in the RM.

- P9, L250-251: "That might reflect that the more predictors, the better the model would not be." Or that the M term is of less complexity and easier to be modeled?

Yes, that is right.

L351-352: That might reflect that the M component is easier to be modeled due to less complexity.

- P26, Table 2: How should someone interpret the fact that for station 23 only Q (specific humidity? Not included in Table A3) is important. Moreover, there are stations that the O_3^{RA} (or O_3^{FC} in Table A4) is not important; how should someone also interpret this?

The Q was a typo; it should be named SH. It was corrected in Table 2 and Table A4. Station 23, i.e., Rasht (known as the city of rain), is located between the mountains (Alborz) and coast (Caspian Sea). Its local environment is rainy with a humid subtropical climate. So, the humidity has a large effect on the ozone level at this site, similar to the Western Mediterranean regions. The variability of observed O_3 at those stations are more associated with another parameters such as NO or NO_2 (than O_3^{RA} or O_3^{FC}) as shown in Fig. 4. These points were added as:

L328-333: **There are a few stations where O_3^{RA} (O_3^{FC}) is not selected as an important variable, which is related to the small (weak) associations between O_3^{RA} (O_3^{FC}) and O_3^{OBS}. For instance, SH is selected as the main factor effecting the M term at station 23, i.e., Rasht. This station is located between the mountains (Alborz) and coast (Caspian Sea), with a local environment of rainy with a humid subtropical climate. That is similar to the Western Mediterranean regions, where a lack of strong synoptic advection, combined with the orographic characteristics and the land-sea breezes, favors episodes of high ozone levels over this region (Millan et al., 1999; Velchev et al., 2011; Wentworth et. al., 2015).**

Millán, M. M., Mantilla, E., Salvador, R., Carratalá, R., Sanz, M. J., Alonso, L., Gangioti, G., and Navazo, M.: Ozone cycles in the Western Mediterranean basin: interpretation of monitoring data in complex coastal terrain, J. Appl. Meteorol., 39, 487–508, 1999.

Velchev, K., Cavalli, F., Hjorth, J., Marmer, E., Vignati, E., Dentener, F., and Raes, F.: Ozone over the Western Mediterranean Sea – results from two years of shipborne measurements, Atmos. Chem. Phys., 11, 675–688, doi:10.5194/acp-11-675-2011, 2011.

Wentworth, G. R., Murphy, J. G., Sills, D. M. L.: Impact of lake breezes on ozone and nitrogen oxides in the greater toronto area, Atmospheric Environment, 109, 52-60, doi: 10.1016/j.atmosenv.2015.03.002, 2015.

- P29, Table A3: The sea surface temperature is listed here as a meteorological variable for CAMSFC. As the SST fields are only over sea, for which coordinates are the SST data extracted for each station?

  The SST was treated as other variables. We noticed that its value are fixed to 273.16 K over the land. It was removed from the list of meteorological variables in RM.

- A small paragraph on what might drive the differences between O_3^{FC} and O_3^{RA} in the discussion is needed.

  The main difference between O_3^{RA} and O_3^{FC} was their model resolutions. We expected that the model with a finer resolution (i.e., O_3^{FC}) would provide better results. Our study shows more or less similar results for both models (at least for the S and M variability). Their difference might be more apparent in the bias term. Another main differences of two products, are their emissions inventories. A couple of species, i.e., NO_x and CO, of the CAMS-GLOB-ANT and MACCity were assessed to show their differences. These points were added as:

  L462-470: **Although both datasets share many of the same parameters in common, there are several differences that distinguish O_3^{RA} from O_3^{FC}. O_3^{FC} is produced by a model with finer horizontal and vertical resolutions. Different anthropogenic and biogenic emissions have been used in both models (see Table 1). CAMS-GLOB-ANT (used in CAMSFC) provides up-to-date emissions of air pollutants and greenhouse gases, at the spatial and temporal resolution required by the model (0.1°x0.1°). CAMSRA uses MACCity emission inventory with a resolution of 0.5°x0.5°. Figure _ shows a comparison of CAMS-GLOB-ANT and MACCity for a couple of ozone precursors, i.e., NO_x and CO. Compared to CAMS-GLOB-ANT, MACCity, based on scenario, provides a higher emissions for both species. CAMS-GLOB-ANT shows more details of the emissions variability due to finer spatial resolution. In both inventories, the highest emissions area are located over Tehran.**

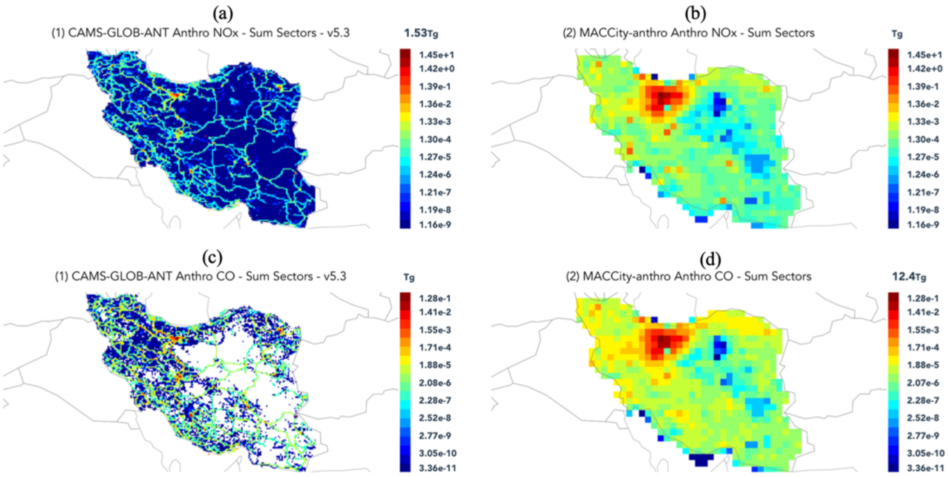

Figure A6. The annual average of surface emissions of the (a)-(b) NO_x and (c)-(d) CO in the CAMS-GLOB-ANT and MACCity emission inventories.

- P12, L364-365: “The most relevant peroxides were found by screening several meteorological variables and chemical species.” I don’t understand this sentence. Please explain or rephrase.

  Sure. It was modified as:

  L580-581: **The potential predictors (inputs) were identified from chemical and meteorological variables at each station.**

Minor comments

- L9: Atmospheric -> Atmosphere. This is by CAMS definition.

  That was corrected.

- L11: datasets -> time series

  That was modified.

- L20: please delete “, or tropospheric ozone at ground level,”

  That was deleted.

- L30: level is -> levels are

  That was modified.

- L36: provide: provides

  That was corrected.

- L37: satellite-> satellites, and also remove “computer”

  These were modified.

- L38: technique-> techniques

That was corrected.

- L90-91: 50 chemical species and seven different aerosols. It provides outputs for several meteorological variables as well. -> 50 chemical and seven aerosol species, providing also several meteorological parameters.

  It was modified as

  L121-122: **The forecast consists of more than 50 chemical and seven different aerosols, providing also several meteorological parameters.**

- L101: . They-> and

  It was modified as

  L142: **accessible to the public and were obtained from the Iranian …**

- L127: Both reanalysis and forecast datasets were..

  It was modified as

  L169: **Both reanalysis and forecast datasets were …**

- L196: Add equation number.

  It was added.

- L224: was extracted-> were extracted

  That was corrected.

- L247 explained variability-> explained variance . Please apply this were applicable.

  A good point, that was modified as

  L349: **indicates the amount of explained variance by the LSTM model**

L350: **that more than 50 % of the O_3 variance is explained by the LSTM**

L426: **S in comparison to that of M is attributed to the larger variance of S (Fig. A2c)**

L433-435: **That reflects the CAMS products, which explain more of the S variance than that of the M term. In other words, most of the variance of the M term in the LSTM is explained by the lagged O_3 (not by the CAMS products)**

L439: **This is similar to the MSE of the LSTM, which is related to the higher variance of S than M.**

- L260: tiny->small

  Its sentences was deleted.

- L260-261: so both models show similar performance. -> with both models exhibiting similar performance.

  That was deleted.

- L307: SDS?

  A typo; that was corrected as

  L475: **by the SD procedures**