Major comments:

1. This analysis uses annual mean ozone profiles, which provide no information on ozone's high day-to-day variability in the troposphere. The method is also applied to full profiles that include the troposphere and stratosphere. The result is a very smooth ozone field dominated by the stratosphere. As I describe below, the resulting clusters seem to be insensitive to prominent tropospheric ozone features. For this reason, I think the analysis needs to be applied to the troposphere and the stratosphere separately.

Yes, thank you for these suggestions. For our application, we are not specifically interested in day-to-day variability in the atmosphere, but we agree that including more temporal information in our classification approach would be prudent. We have modified our approach to include seasonal variations by switching from annual mean profiles to seasonal mean profiles. Furthermore, we have also excluded some of the near-surface pressure levels, such that the classification is less affected by near-surface ozone processes. The suggestion to classify the troposphere and stratosphere separately is interesting; unfortunately, there are not enough pressure levels in the stratosphere alone to justify classifying it entirely separately from the troposphere. In any case, we chose this approach because it retains more of the entire profile structure, which we view as an advantage - we want the resulting classes to be influenced by the upper troposphere, the stratosphere, and the interaction between the two.

2. What is the impact (or limitation) of using annual averages? Ozone concentrations vary widely from summer to winter in both the troposphere and the stratosphere. How different are the clusters if the analysis is applied separately to summer and winter months? Another problem with the annual average is that mid-latitudes are heavily influenced by polar air masses in winter (low tropopause), and by tropical air masses in summer (high tropopause). So the annual average is just an unrealistic homogenization of very different air masses, and does not reflect the typical ozone profiles one might find in any given month or season.

Yes, this is an important point. We have switched to classifying seasonal mean profiles, which allows seasonal variation. The analysis in the paper has been updated to reflect this seasonal variation. Thank you for the suggestion.

The analysis makes no use of observations, and with no evaluation against real-world data we are unable to understand the accuracy of the method. Stauffer et al. (2018) clustered ozone profiles at more than 2 dozen ozonesonde stations worldwide. I realize the authors can't use sparse observations as the basis for this global-scale analysis, but they can certainly evaluate the results against observations. The authors should examine the observed profiles above the ozonesonde stations that lie within each of the clustered regions. Do the profiles within each region have similar characteristics? If so, then the method is applicable to the real-world; if not, then the usefulness of the method is questionable. What is the result when the observations are then examined by season? Are the observations within each cluster similar to each other in summer, and also in winter? Or does everything break down (see my comment above about seasonal variability in the mid-latitudes).

In this paper, we are focusing on GMM as a model analysis and comparison tool. We agree that applications to observational data would be interesting, but it is beyond the scope of this short technical note. In any case, it is not necessarily obvious how one would carry out such a comparison, given the sparse observational coverage of atmospheric ozone. Specifically, one might try to fit the GMM using observational profiles only and then use those classes to validate and analyze a model run. This has been done for ocean temperature profiles for a specific region in the European Arctic (Thomas and Müller, 2022, https://doi.org/10.1016/j.ocemod.2022.102092). However, given that observational coverage of atmospheric ozone is not especially uniform, it would be difficult to generate a sufficiently general training dataset for GMM. Any resulting GMM would be biased towards the ozone profiles seen at the observing site. We believe we have shown that GMM can, at the very least, be a useful tool for model ozone analysis and comparison.

Other comments:

1. Line 59: When reviewing clustering techniques as applied to ozone profiles, the authors should include Stauffer et al. (2016, 2018).

Thank you for providing some excellent references. We cited them and extended our ozone clustering techniques applied in other studies. The paragraph now reads:

Clustering techniques have already been used in ozone concentration studies for understanding long-term variability. Boleti et al. (2020) have applied a multidimensional clustering technique to understand the long-term trend of ozone. Diab et al. (2004) used a six-cluster analysis which resulted in distinct clusters of "background" and "polluted" with below and above ozone mixing ratios from over 100 ozonesonde profiles launched from a subtropical Southern Hemisphere Additional Ozonesondes (SHADOZ) (Thompson et al., 2003) site, Irene, South Africa. Jensen et al. (2012) performed a cluster analysis named self-organizing maps (SOM) (Kohonen, 2012) on over 900 tropical ozonesonde profiles. Their findings with four-cluster results were similar to Diab et al. (2004). Both studies showed that the seasonal influences of biomass burning and convection dominate ozone variability. Stauffer et al. (2016) documented the influence of meteorological conditions on the shape of the ozone profile from the troposphere to the lower stratosphere by applying SOM clustering technique to ozonesonde data from specific northern hemisphere midlatitude geographical regions. Later they expanded the study for global ozonesonde sites to show the variation of ozone profiles cluster for various regions and how they vary based on meteorology and chemistry depending on latitudes (Stauffer et al., 2018).

2. Line 61: The perceived methodology and aim of Chang et al. 2017, as stated in the manuscript, is not correct. Chang et al. 2017 are not seeking to cluster similar ozone monitoring sites. Rather they are trying to quantify the regional-scale, long-term trend of ozone while accounting for the spatial distribution of the sites and the correlation between sites. This method accounts for the uneven distribution of sites and prevents any heavily-sampled sub-region from exerting an out-sized influence on the trend

We agree that the citation was not correct. Thank you for pointing that out. We got rid of that part.

3. Line 95-98: I don't understand why the study is limited to 1-1000 hPa. This omits a large section of the globe, i.e., land regions more than 100-200 m above sea level. I realize the method cannot tolerate missing values, but why not conduct the study for profiles in the range of 1-950 hPa; this way, you retain most of the land areas.

Thank you for the suggestion. We are now focusing on 1-850 hPa to retain more land areas.

4. Line 110: This statement is problematic:"The motivation behind withholding the geographical information is that there is no reason for the vertical ozone structure of the profile to be unique to a given region (Maze et al., 2017)." Using a paper that deals with ocean temperature, the authors

seem to suggest that there is no discernable structure in the global ozone distribution and that one region is no different from another. Yet, plenty of observation-based studies identify clear structure in the global ozone distribution that varies with season [Kley et al., 1996; Thouret et al., 1998; Oltmans et al., 1996, 2004; Thompson et al., 2003; Cooper et al., 2007; Gaudel et al., 2020;]. Therefore, certain profile types are more likely to occur in some regions than in other regions. This statement needs to be revised.

We agree that the original statement was unclear. We have revised the text and hope that you find the new statement suitable.

5. Line 173: To say that the tropopause is around 300 hPa is a gross over-simplification. As can be seen in Figure 2, there are plenty of profiles in which the tropopause is around 150 hPa, which is common in the tropics.

We have changed this statement to :

The tropopause height, above which the ozone starts increasing, varies between 300-150 hpa depending on the location of the profiles.

6. Line 175: The statement that ozone increases near the surface is problematic because ozone is plotted in units of mPa. If ozone is plotted in units of ppbv (the typical unit for evaluating air pollution levels in the troposphere), then we would see that the average ozone profile has more ozone in the upper troposphere, especially at mid-and high latitudes (see the ozone profile papers that I cited above). Furthermore, Jaffe and Wigder (2012) is not a sufficient reference because they only discuss ozone at the surface and do not mention the vertical distribution of ozone.

We use ozone partial pressures (mPa) in this study because if we use mixing ratios the ozone profiles would span many orders of magnitude, with surface ozone mixing ratios in the 10s of ppb, and stratospheric ozone mixing ratios reaching a maximum of ppm, and so the profiles would be dominated by the shape in the stratosphere, and it would be very difficult to see what is happening in the troposphere. The choice of mPa shows more clearly the relative structure between profiles in the troposphere. We have amended this paragraph to explicitly highlight the fact that we are discussing ozone concentrations and replaced the reference with that of Monks et al., which provides a more complete review of tropospheric ozone distributions and processes. The paragraph now reads:

Our purpose is to identify coherent patterns within the collection of profiles using unsupervised machine learning. Overall, the profiles reveal relatively high ozone concentrations in the lower and middle stratosphere which peak and then decrease gradually in the upper stratosphere. The tropopause height, above which the ozone concentrations start increasing, varies between 300-150 hPa depending on the location of the profiles. The peak starts decreasing at around 70 hPa and higher altitudes above (Figure 2). In the troposphere, ozone concentrations are fairly constant and then increases towards the surface, in part due the availability of ozone precursors from biomass burning and anthropogenic emissions sources (e.g., Monks et al., 2015).

7. Line 216: Why is the high surface ozone only attributed to biomass burning? This cluster spans the major fossil fuel combustion regions of the northern hemisphere, which are known to drive ozone production across the region.

This has been amended and now refers to ozone precursors from different sources. The sentence now reads:

In the troposphere, ozone concentrations are fairly constant and then increase towards the surface, in part due to the availability of ozone precursors from biomass burning and anthropogenic emissions sources (e.g., Monks et al., 2015).

8. Line 265: The statement that ozone precursor emissions generally increase under SSP5-8.5 isn't really correct, as emissions continue to decrease in developed nations but increase in the developing world. This discussion should also consider the findings of Zanis et al., 2022.

This statement has been expanded upon in the final draft, and we have also included some of the discussion from Zanis et al.. The relevant section now reads:

Here we examine the structure of atmospheric ozone in the 2095-2100 years of the SSP5-8.5 experiment. In this experiment ozone mixing ratios are generally higher throughout much of the troposphere and upper stratosphere. In the troposphere, the drivers of this increase are complex. Under the assumptions of the SSP5-8.5 scenario, global mean emissions of NOx and CO are lower in 2095 than the present day, while global mean emissions of CH4 are higher (Gidden et al., 2019). However, changes to ozone precursor emissions alone do not drive tropospheric ozone changes, which is also affected fby climate change, with increasing tropospheric temperatures changing biogenic volatile organic compounds (BVOC) emissions, the availability of tropospheric water vapor, and stratosphere-to-troposphere transport of ozone, which taken together drive increases to tropospheric ozone concentrations (e.g., Griffiths et al., 2021; Turnock et al., 2021; Zanis et al., 2022). In the stratosphere this increase is simpler to understand. Upper stratospheric ozone

increases under all SSPs as ozone depleting substances decrease, but increases more in scenarios which assume larger increases in greenhouse gas emissions due to the resulting CO2-induced cooling of the stratosphere and the impacts this has on gas phase chemistry (e.g., Haigh and Pyle, 1982; Jonsson et al., 2004).