The authors present an ambitious study to analyze the information content of surface ozone over Iran (particularly Tehran) in two CAMS-based data products: the CAMS reanalysis, and the CAMS NRT analysis. For this the authors make use of in-situ observations of ozone, along with other chemical and meteorological quantities at various stations across Iran, but particularly in the cities Tehran and Tabriz.

We value their efforts in studying the usefulness of CAMS data for their applications, and their investigation in mathematical methods to analyze, and differentiate errors of various kinds. Reporting on this progress, and its application to locations across the world that may often be overlooked, is much appreciated. Nevertheless, I believe there are several serious flaws which makes it hard to extract useful information from this study, which unfortunately makes me advise to reject this manuscript for publication in ACP, before these aspects are addressed.

My main criticism concerns the plain application of mathematical methods to differentiate and analyze the representativity of various time scales in the data products: Prior to applying these methods, in my view the authors should better justify the selection of input data for this, both on observation and modeling side:

- from the observation side, I notice that the majority of observation stations are very close to one another, i.e. inside the city of Tehran. This makes that only a single grid-box in the global (re-)analysis data is used (and one at Tabriz station). Therefore the statistics is difficult to interpret for the whole of Iran, as suggested in the title of the manuscript. A closer assessment, and possibly corresponding filtering, of the observations that can be used to better represent ozone variability on the scales of the global model, e.g. by using a site classification method (urban / rural / etc) could be beneficial. E.g. the authors should better describe the ‘Geophysics’ station (line 82). A
map with distribution of stations across Iran (and/or Tehran) would be beneficial.

- Such assessment of the observation network is the more important because the global (re-)analysis datasets definitely lack ability to describe (ozone) pollution levels on a city-scale, as also remarked by the authors. Therefore this analysis should be done with care, and should be discussed. Also it is not 100% clear which model level is selected, and what is the rationale for selecting a model level above the surface for comparison against surface observations. Such an approach is useful for observations that are situated on (isolated) hill / mountain tops, but otherwise this can add to large discrepancies. It might be best to simply select the surface layer for this assessment.

- The authors should better explain the status of the two CAMS datasets. Different to what the authors describe, I think these can best be described as a Reanalysis, and a NRT Analysis dataset, with various similarities (based on very similar input datasets, and data assimilation systems), but with their differences (in terms of model resolution, chemistry, input emissions). The statement on line 71 (“CAMS provides global forecasts, called analysis”) is confusing in this respect.

- Also the selection of additional quantities that explain the variability in O3, as discussed in Sec. 3.3 (and eqn 6) appears ad-hoc, and it is not clear where these quantities are taken from (also from CAMS model data, or observed?). E.g. why is radiation (or solar zenith angle / cloud cover) not included as a proxy to explain variability in O3?

Furthermore, the authors present statistics that is difficult to interpret. E.g. their Figure 1 the diurnal cycle plots are illegible, while the Figure 1(d) better should include longer time series. The units in expressions 1-4 are not specified, nor the meaning of variable ‘p’. The naming convention dfo, dfa, fdr (line 90) is needlessly complex for the reader. (why ‘df’ in the first place?). The Table 1 shows numbers for the regression coefficients that cannot be inter compared as such, as they apparently depend on the mean variability, while it is unclear what is the meaning of the second number in the column. Also it is unclear why the authors select station ‘8’ for this analysis. To my judgement the study would benefit from presentation of plain intercomparison of model-observation time series, and corresponding statistics, to justify and analyze these more advanced mathematical methods.