Comment on essd-2020-280
Anonymous Referee #2

General comments:

This manuscript describes a unique data set of anthropogenic NOx point emissions derived from nearly two years of TROPOMI data, explains the methods used to create it, and discusses its limitations and uncertainties.

The data set presented is limited in several ways: it is incomplete because of the method used, it is only partly quantitative as it suffers from a known but poorly constrained low bias in the current TROPOMI lv2 data, and it is potentially biased because the assumption has to be made, that emissions are constant over the time period used as input. Therefore, this data set should not be used as emission input for CTM runs or for the verification of reported emissions.

In spite of these limitations, this is a very valuable data set as it provides observational evidence for the position and approximate strength of a large number of NOx point emissions sources worldwide. It is a unique data set and fully independent of other data sets of NOx point sources and therefore very well suited for semi-quantitative verification of other inventories, in particular if up-to-date information is needed.

The data set has the potential to be improved by future updates when more TROPOMI data is available for better statistics, and when improved TROPOMI lv2 data versions have been released for reduction of uncertainties.

The manuscript describes the generation of the data set in a clear and sufficiently detailed way, highlighting the differences to the scientific study describing the original method. The
uncertainties and limitations are discussed in detail and give users a good idea of how to use the data, and which caveats to consider.

**Detailed comments:**

The authors introduce a number of ad hoc abbreviations such as PP for power plants, which in my opinion make the text less accessible and do not really help with length. I would suggest using them only in the figures.

Page 2, line 7 – 9: I am confused by this description of AMF and Averaging Kernel. To me, this should be Jacobians or box AMFs instead of Averaging Kernels as the latter are the box AMFs divided by the total AMF derived for the a priori profile. Please reconsider.

Page 2, line 31: “flux increases discontinuously” – I find this formulation a bit awkward as a discrete data set is discontinuous by definition

Page 3, line 18: TROPOMI spatial resolution has already been introduced

Page 3, line 27: Mention native resolution of model data

Page 4: Mention altitude of O3 extraction

Page 5: More details on the gridding would be useful: I assume it is a 2d linear interpolation? Is there a risk of low bias using this approach? Is this approach reducing the spatial resolution of the original data? Mention that is done by orbit (this is explained later but I asked myself here if this is done daily, monthly or on orbits)

Page 7: Is there a risk that the requirement for a high dynamic range excludes some point sources with little variation in wind direction?

Page 9: The upscaling to NOx would benefit from a bit more discussion – NO2 columns are inserted where NO2 concentrations should be used, and in particular when [O3] is taken at varying altitudes, this raises some questions about the validity of the approach. For high altitude regions, the factor between NO2 column and NO2 concentration will also change – does this have an effect?
Page 10, line 22: Is this really the best reference here – Eskes et al. introduced the column AK, not the AMF.

Page 11: I would suggest to already here identify x and y as latitude and longitude

Page 12, line 15: neglects of the => neglects the

Page 14, line 27: above > 30% => above 30%

Page 15: Maybe in India, there is a lack of distributed NOx emissions such as from traffic, helping to reduce the background levels

Page 17, line 7: Ukraina => Ukraine

Page 22, line 9: I am not convinced that larger deviations of wind direction will be readily visible in the divergence as this is applied to the averaged flux where wind direction errors could cancel

Page 25, acknowledgements – isn’t there a standard statement required when using EU Copernicus data?