1 Reply to Referee #1

We are replying quickly to Referee #1's comments as we disagree with the criticism that this manuscript and the dataset lacks sufficient innovation to be acceptable for publication, and believe that it is important to clarify the nature of the manuscript and of the CAMS-GLOB-SOIL v2.2 database. We strongly believe that datasets used by Copernicus, EMEP or indeed any Earth System modellers need to be documented, and we believe ESSD is the appropriate journal for such documentation.

General comments:

Ref#1 This study implemented the YL95 soil NO emissions scheme with various updates from other publications in the EMEP MSC-W chemistry transport model, and generated global soil NO emissions at the spatial resolution of 0.5 degree by 0.5 degree during 2000-2018. Soil NO is a significant contributor to global NO emissions, and generating global NO emissions inventory is important. However, it is very difficult to find innovation in this study, thus I cannot recommend its publication in ESSD.

Reply:

We were disappointed to read the referees comments, but disagree with them, especially those related to 'innovation', and the need to publish manuscripts documenting datasets which are in open-source repositories. This paper was submitted to the journal 'Earth System Science Data', which we believe is a journal intended to describe datasets which are used in Earth System and similar models. The dataset we are describing, CAMS-GLOB-SOIL v2.2, is already available online, with a doi as required by ESSD. The CAMS-GLOB-SOIL data are in use already within both the Copernicus CAMS projects and the Air Convention's EMEP projects. Given this usage, we feel it is very important that the dataset is described in a publication, and that the main purpose of such a publication is to describe the methods and some consequences of the use of the data.

The referee argues that it is difficult to find innovation. As discussed below, we think that there are several innovative features. In any case it is important to note that this work was funded initially by the Copernicus CAMS81 project, and the need was for a practical dataset that improved upon the data initially used *in CAMS, and which could form the basis for future improvements. As with all CAMS81 datasets (e.g. for anthropogenic emissions, biogenic VOC), the soil NO methods build upon previously existing knowledge and data-sets, but modify them in an effort to better suit the needs of the CAMS modelling work. Such needs include:*

- *Datasets that can be kept up-to-date, and updated quickly. (We will soon calculate emissions for 2019 and 2020 with the same system, and this will continue as long as CAMS and/or EMEP have need of new data.)*
- *Datasets that can be extended to past and future scenarios: modelling of scenario years such as 2050 or historical years such as 1900 need to be possible, often with limited resources provided to adapt emissions to such years.*
- *Datasets which use meteorology from ECMWF, to be consistent as far as possible with the CAMS IFS model.*
- *Datasets that can be continuously tested within the CAMS community, and modified as a result of feedback and requests from different modelling and observation groups. This requirement by itself is a major reason why the CAMS-GLOB-SOIL dataset was developed inside the CAMS project, and not just taken "off the shelf" from some external source.*
- *Datasets that provide more detail than 'traditional' soil-NO emissions, especially with the contributions of fertilizer-induced emissions clearly delineated.*

The referee suggests (below) that since Weng et al 2020 have published emissions for 2000-2018, then there is no need for another inventory. Actually the work which lies behind CAMS-GLOB-SOIL was started in 2018, and the basic methods built up well before Weng et al 2020 was published. We believe there is room for more than one soil-NO emission dataset in the atmospheric sciences community. Indeed, given the massive uncertainties associated with soil-NO emissions we believe that alternative datasets are very much needed in order to provide some indication of these uncertainties.

Specific comments:

Ref#1

This study just implemented an old NO emissions scheme with various updates from other publications, but it lacks innovation. Moreover, Weng et al. (2020) has already generated global soil NO emissions at the resolution of 0.5 degree by 0.625 degree for 1980-2017. What is the innovation of this study compared with Weng et al. (2020). A little finer resolution is not sufficient to make it publish.

Reply:

The referee further claims that we have implemented 'an old emission scheme' (YL95) with various datasets. This is misleading, as the same statement could be made about the Hudman et al 2012 inventory, and indeed the HEMCO emissions as used in Weng et al. (2020). These inventories all rely on the ideas introduced in YL95 of biome-based emission factors, which are modified by environmental and nitrogen-input factors. The Hudman type improvements (especially the use of smooth functions to represent the soil-water effects) do look very elegant, and we hope to introduce something like that in future versions of CAMS-GLOB-SOIL. However, though Hudman et al stressed themselves that the soil water itself could not be validated (we mention this on line 100). This type of uncertainty was one of the driving factors to our approach to use 'pragmatic' solutions for CAMS-GLOB-SOIL.

It is anyway more true to say that we start with the Steinkamp and Lawrence 2011 (SL11) emissions, which updated YL95 on the basis of 100s of new measurements. We have then developed simplified methods of dealing with fertilizer- and deposition- induced emissions, soil water effects and pulsing. Compared to most other papers on soil-NO emissions, we stress throughout the difficulties associated with both the input data and the methods used, explaining the pragmatic choices that were made.

1. *The CAMS-GLOB-SOIL database is (as far as we know) the only dataset in which the contributions from background biome, fertilizer, deposition and pulsing are provided separately:*

$$F_{soil} = F_{biome} + F_{Fert} + F_{Ndep} + F_{pulse} \quad (1)$$

Or

$$F_{soil} = F_{nonFert} + F_{Fert} \quad (2)$$

This feature is both innovative, and very important! As we make clear in Sect. 6 there are substantial risks of double-counting when soil NO emission databases are used together with other emission inventories. Sect. 6 explains the problems, and suggests solutions, for dealing with this double-counting risk.

The system is proving extremely important in both CAMS and EMEP model runs, where users have to sometimes include and sometimes exclude the fertilizer component. This separation also provides a useful system whereby the different components can be modified by users. This allows for example users to specify different assumptions about the percentages of N released from fertilizers or N-deposition.

2. *As part of this study we have also documented the usage of soil-NO emissions in EMEP, ECLIPSE, EDGAR and CAMS inventories of 'anthropogenic' emissions. Actually before this comparison was done, many colleagues were not aware of the extent to which soil-NO emissions had been included in these other inventories. We hope that the current paper will give guidance to many atmospheric modellers, even if they use other soil NO emission datasets.*

3. *The CAMS-GLOB-SOIL data use different meteorological and land-cover inputs to the Weng et al data. As the CAMS data are intended to be used with the CAMS Integrated Forecasting System coupled model (C-IFS), it is important that the meteorology is as consistent as possible with C-IFS. (In future the land-cover may also be changed in CAMS-GLOB-SOIL, C-IFS, and indeed other CAMS components such as forest-fire or biogenic VOC calculations to improve harmonisation. Discussions have started on such issues, but this is likely a multi-year process. The fact that such discussion exists among CAMS partners is however a good reason for CAMS to have control of its own soil emission methods.)*

4. *Sect.7 illustrates the results of applying CAMS-GLOB-SOIL at both global and regional scale. The literature has surprisingly few examples which show the impact of such emissions on atmospheric pollutants, either in magnitude or in comparison with observations. Although we restricted the number of results we used since this is a data-description and not a modelling*

paper, we believe the snapshot we present of the impacts of soil NO can be useful for other scientists. (More detailed comparison with measurements, including satellite data, will be an important component of future work, and guide model improvements.)

Ref#1

Although Sect. 3.3.2 explain a bit, I still cannot understand why use air temperature rather than soil temperature. Do you mean soil temperature has very large uncertainty or it requires more coding work to implement?

Reply:

Sect. 3.3.2 provides two explanations. The first reason, denoted (a) on L196, reflects the practical consideration that the soil temperature was not available from the EMEP model's data files. Given the very limited funding provided by CAMS-81 for this work, and the tight deadlines, we built a system around the EMEP model files which were readily available. (The EMEP model data does come from ECMWF, but is the subset needed for our chemical transport model.) The second reason given, denoted (b) on L187, is that soil temperatures are indeed difficult to predict for large grid cells. We provide references to support that assertion.

We can note that in the first version (v1.1) of CAMS-GLOB-SOIL we made use of relationships between air and soil temperature taken from YL95 and from the MEGAN code (Guenther et al., 2006). However, closer examination of these relationships showed some problems (as mentioned in Sect. 4.7). This is clearly an issue that needs more attention, and as noted in Sect. 4.7 and the conclusions we will pursue better methods in future.

We will add more text on these points in the revised manuscript. (Ref #2 also had some comments on these temperature issues; we will address those in a separate reply.)

Ref#1

The introduction lacks the review of current soil NOx emission algorithm.

Reply:

We do not really understand this comment. Sections 1 (introduction) and 2 (Nitrogen Oxide emissions: background) provide more than 3 pages of background *on soil emission algorithms, and refer the reader to much more extensive reviews that exist. We present the basic ideas of YL95, SL11 and Hudman et al (and hence HEMCO and Weng et al), though of course we do not cover all details. Again, we believe ESSD is a journal where the focus should be the description of the dataset being presented, rather than a journal where long reviews of other works should be presented.*

Nevertheless, if the referee missed something specific we will be happy to add comments.

Ref#1

Line 3: delete degrees

Reply:

Thank you. It will be deleted.

Ref#1

There are many grammar errors, and the writing should be polished.

Reply:

We will endeavour to improve the English.

Concluding remarks

As noted above, we disagree with the criticism of Ref#1 that this manuscript and the dataset lacks sufficient innovation to be acceptable for publication. We hope we have clarified the nature of the manuscript and of the CAMS-GLOB-SOIL v2.2 database. We are happy to add additional comments into the revised manuscript to make our points clearer. We strongly believe that datasets used by Copernicus, EMEP or indeed any Earth System modellers need to be documented, and we believe ESSD is the appropriate journal for such documentation.