Reviewer 2#:

This is a very useful dataset that can be used and cited in the future. But the manuscript is written as a scientific research paper rather than a data description paper. I suggest the authors to restructure the paper to describe what the data actually contain. It is unclear to me whether the dataset contains the climate data, which are collected from some other source. So these things need to be more clearly described. Please also clarify what are modeled and what are in situ data. The title should also reflect this.

For the same reason, the introduction, results, and discussion sections are unnecessarily long and not very useful in actually understanding what the data are.

Response:

Thanks very much for your review. We carefully read your comments and restructure the paper according to your comments, we believe the quality of the manuscript has been greatly improved. Detailed responses are in blue, in-line with reviewer input below.

First, this dataset is the soil organic carbon distribution data at different soil depths (0–30 cm, 0–50 cm, 0–100 cm, 0–200 cm, and 0–300 cm) in the frozen ground area of the Third Pole region, and the dataset does not contain the climate data (precipitation and temperature), which are the environmental factors for constructing machine learning models in this study. Similarly, our dataset also does not contain the situ data, which are the input data for constructing machine learning models in this study. Therefore, the original title “Soil organic carbon distribution for 0-3 m soils at 1 km^2 scale of the frozen ground in the Third Pole Regions” reflects the information of the data subject (soil organic carbon), depth intervals (0-3 m), spatial resolution (1 km^2) and study area (Third Pole Regions).

Second, we have added a detailed description of the dataset in the **Data availability** section (section 5) to clarify the dataset information in this study. We have added these explanations in the methods section in the revised version as follows:

The datasets of SOC stocks distribution in GeoTiff format are available at https://doi.org/10.5281/zenodo.4293454 (Wang et al., 2020). The file name is "TP-SOC-d.tif", where d represents soil depth, for example, "TP-SOC-30.tif" represents the spatial distribution of SOC stocks in the Third Pole regions of the upper 30 cm depth interval.

Thirdly, we have truncated some content in the introduction, results and discussion sections, which are not very useful in actually understanding what the data are. In addition, we have added some subtitles and a more detailed description in the collecting and processing of soil sample data to improve the quality of the **Materials and Methods** section.

Finally, we have added the accuracy assessment in the revised version between our dataset and other global SOC datasets (SoilGrids250m and WISE30sec) in the discussion sections to further improve this article.

Minor comments:

#Please describe how and why the sampling area (soil pits) were selected.

Response:

Because of harsh natural conditions and the inaccessibility of traffic in the Third pole regions, the soil samplings from earlier studies were conducted along the major roads, which unable to represent the environmental characters in large areas of the Third pole. Therefore, we conducted a large-scale field sampling from 2009 to 2013, covering all

major climatic regions and vegetation types across the plateau, including the large unpopulated area with harsh natural conditions.

The setting of sampling points for soil survey is widely representative, with at least one sampling point for each soil/geomorphic type. In areas with strong environmental heterogeneity, the layout of soil pits based on different parts of medium topography or micro-topography, especially in mountainous area with complex terrain. Moreover, the layout of sampling points avoids special sections such as roads, railways, engineering facilities, etc., which are greatly affected by human interference activities. Therefore, combined with the available published data and field investigated data, the 458 soil pits (depth of 0–1 m) and 114 soil cores (depth of 0–3 m) in this study can represent the vegetation types and characters in large areas of the Third pole.

Therefore, we have added the **Table 2** in **Data Processing** (section 2.2.1 in Manuscript) to display the number of soil sample points of different vegetation types in Third pole region.

Table 2 Number of soil sample points of different vegetation types in Third pole region

Vegetation types	Forest	Shrub	Grassland	Desert	Cropland
Number	10	22	371	49	6

#- How did the authors deal with gravel and what was the gravel %?

Response:

The weak chemical and biological processes on the third pole resulted in the widespread existence of gravel in soil. Missing the gravel data will affect the estimation of bulk density, which leads to the erroneous estimation of SOC storage. In our study, all soil samples for carbon analysis were air-dried, handpicked to remove plant detritus, and then sieved through a 2mm mesh to calculate the volume percentage of the gravel. In addition, the rest of SOC data, which obtained from Yang et al. (2010), Song et al.

(2016), Xu et al. (2019) and Ding et al. (2016), was treated with the similar method as our experimental procedure.

According to the statistics of 200 soil samples obtained from the field (Table 5), it was found that the gravel content of grassland soil increased with the depth of 2-m profiles, from 4.65% in 0-30cm depth interval to 28.17% in 100-200cm depth interval. The volume percentage of soil gravel in desert areas is overall higher than that in grassland areas, especially in the top 50cm, but no significant change with soil depth increased in the 2-m profiles. It should be noted that the soil sample points in forest, scrub and cropland areas are derived from the literatures, and the volume percentage of gravel is unclear.

Table 5 Statistical description of the gravel content at different depths based on the 200 soil profiles.

Vegetation types	Gravel content (%)			
	0–30 cm	30–50 cm	50–100 cm	100–200 cm
Grassland	4.65 ± 1.98	11.91 ± 3.59	24.89 ± 4.04	28.17 ± 4.53
Desert	18.56 ± 4.15	21.6 ± 4.76	20.84 ± 4.36	21.07 ± 4.69