Thank you very much for your time and effort regarding our manuscript. We have carefully revised the manuscript according to the reviewers' comments. Detailed responses are in blue, in-line with reviewer input below. In addition, we included Dr. Lin Zhao as a co-author because he made contributions to the intellectual merit of this work in our many previous discussions.

Reviewer 1#:

The authors created high spatial resolution data of organic carbon distribution in the Third pole by compiling all the field data and using machine learning methods. The dataset can be very useful to help the scientific community to understand the carbon cycle. I found the paper is well organized. My major concern is that the authors should clearly explained what are the new findings in comparison with several previous reports on the plateau, e.g., Ding et al. 2016, Wang et al. 2020.

Response: Thanks very much for your review. We carefully read your comments and made substantial revisions according to your comments, we believe the quality of the manuscript has been greatly improved.

To clearly explain what are the new findings in comparison with several previous reports on the plateau, e.g., Ding et al. 2016, Wang et al. 2020. We have added these explanations in the methods section in the revised version as follows:

In this study, we provided the new version of 1-km resolution maps of SOCS across the Third Pole at 0–300cm depth intervals, and largely makes up for the deficiencies of previous studies (Ding et al., 2016; Ding et al., 2019; Wang et al., 2020). On the one hand, our predictions have higher resolution than those studies. Take an example and focus on a 4.5×10^4 *km*^{2} *local area situated in the Budongquan area of Qinghai province, China (Fig. 8). It can be seen from the excerpts of the map that our prediction is much more detailed than previous studies. Thus, our predictions better represented spatial variation of the SOCS across the Third pole region, especially for those regions with large heterogeneity. On the other hand, these reports most focused on the permafrost regions rather than the whole Third Pole (Ding et al., 2016; Wang et al., 2020). To date,* *few studies have investigated the SOC storage and spatial patterns in areas of seasonally frozen ground in the Third Pole region. In this study, we created high spatial resolution data of SOCS distribution in the whole Third Pole by compiling all the field data and using machine learning methods, thus providing more accurate data than previous studies.*

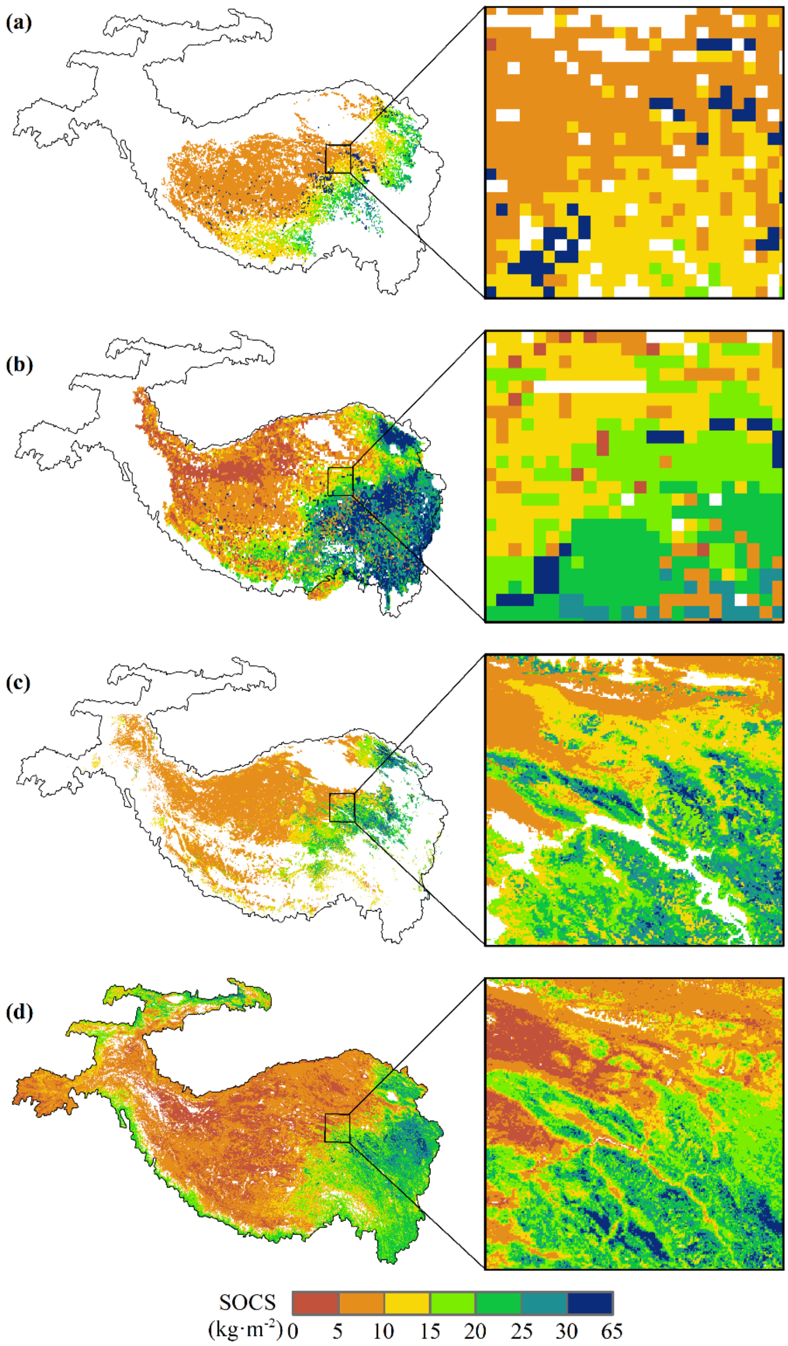

Figure 8. Comparison of spatial details of the predictions with the previous studies: SOCS at 0–300 cm depth in the map excerpt of Budongquan area of Qinghai province, China. (a) Ding et al., 2016; (b) Ding et al., 2019; (c) Wang et al., 2020; (d) This study.

Specific comments:

Figure 1 ,2 and 4 are never mentioned in the manuscript, please add reference mark in right place.

Response: Thanks for your suggestion, we have added references mark in right place.

Many SOC data were collected in this study, however, the data of China's national soil survey were not included in this manuscript, why?

Response: During the collection process of SOC data, we referred to the information of the second China's national soil survey. however, this survey was conducted during 1978 - 1984, and the sampling points lacked accurate location information in Tibetan Plateau due to the limitation of technical means. In addition, most of them lacked data of coarse gravel content and bulk density, which could cause a large uncertainty source in SOC stocks calculation. In contrast, all soil points in this study were sampled after 2000, and all of them contained clear information on gravel content, bulk density and location, so the calculated SOC storage was more accurate.

Please add the accuracy assessment in manuscript between your dataset and other global or regional SOC datasets, such as: SoilGrids and HWSD.

Response: Thanks for your suggestion, we have added the accuracy assessment in the revised version between our dataset and other global SOC datasets (SoilGrids250m and WISE30sec) as follows:

In addition, our predictions were much more accurate than the existing global SOC datasets. Figure 9 shows accuracy assessments of our SOCS prediction, the SoilGrids250m from Hengl et al., (2017) and the WISE30sec SOCS data from Batjes., (2016) at 0-2m depth intervals based on the 213 SOC stocks data from Ding et al., (2016) and Zhao et al., (2018). We found that our prediction had a higher R^2 *value and* *lower RMSE value than SoilGrids250m and WISE30sec. The lowest accuracy was found for the WISE30sec maps, showing the advantage of digital soil mapping based on machine learning over conventional mapping method based on the vegetation/soil units (Liu et al., 2020). The remarkably lower accuracy of SoilGrids250m than our predictions mainly because of serious over-estimation of bulk density, and neglected the influence of coarse gravel content (Hengl et al., 2017). Soil profile data used in SoilGrids250m at Third Pole region are mainly from second China's national soil survey, which lacked accurate information on coarse gravel content and bulk density (Shi et al., 2016). In addition, almost all of these soil profiles are within 1-m depth, which could be a great instability in calculating the deeper SOCS by SoilGrids250m. Moreover, the global model building could be less accurate than the regional model building when focusing on a regional extent (Vitharana et al., 2019; Liu et al., 2020). Thus, our predictions were much more accurate than the existing maps of SOCS.*

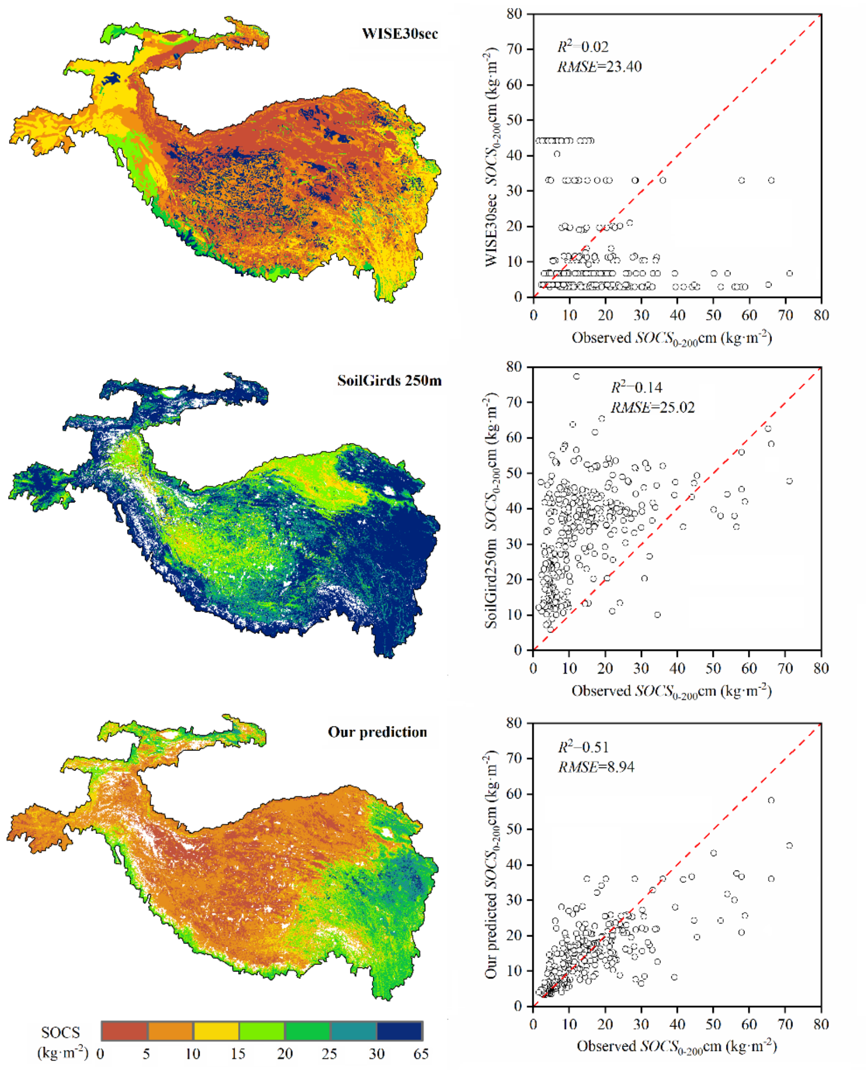

Figure 9. Comparison of the SOCS prediction with the WISE30sec from Batjes., (2016) and the SoilGrids250m from Hengl et al., (2017) at 0–200 cm depth intervals based on the 213 SOCS data from Ding et al., (2016) and Zhao et al., (2018).

Batjes, N.H.: Harmonized soil property values for broad-scale modelling (WISE30sec) with estimates of global soil carbon stocks. Geoderma 269, 61–68, https://doi.org/10.1016/j.geoderma.2016.01.034, 2016.

Hengl T, Mendes de Jesus J, Heuvelink GBM, Ruiperez Gonzalez M, Kilibarda M, Blagotić A, et al.: SoilGrids250m: Global gridded soil information based on machine learning. PLoS ONE 12(2): e0169748. https://doi.org/10.1371/journal.pone.0169748, 2017.

Liu, F., Zhang, G.-L., Song, X., Li, D., Zhao, Y., Yang, J., Wu, H., and Yang, F.: High-resolution and three-dimensional mapping of soil texture of China, Geoderma, 361, https://doi.org/10.1016/j.geoderma.2019.114061, 2020.

Shi Jianping, Song Ge.: Soil Type Database of China: A nationwide soil dataset based on the Second National Soil Survey (in Chinese). China Scientific Data, (2):1-12, http://dx.doi.org/10.11922/sciencedb.180.88, 2016.

Vitharana, U., Mishra, U., and Mapa, R. B.: National soil organic carbon estimates can improve global estimates, Geoderma, 337, 55-64, https://doi.org/10.1016/j.geoderma.2018.09.005, 2019.

#L 62-64, Permafrost degradation will not only cause serious geological disasters and affect engineering construction in cold areas…" needs citations here.

Response: we added the reference of Cheng et al., 2007; Cheng et al., 2019; Ding et al., 2021.

Cheng, G., Wu, T.: Responses of permafrost to climate change and their environmental significance, Qinghai-Tibet Plateau, Journal of Geophysical Research Earth Surface, 112, F02S03, https://doi.org/10.1029/2006JF000631, 2007.

Cheng, G., Zhao L., Li R., Wu, X., Sheng., Y, Hu G., Zou D, Jin, H, Li, X., and Wu., Q.: Characteristic, changes and impacts of permafrost on Qinghai-Tibet Plateau (in Chinese), Chin Sci Bull, 64: 2783–2795, https://doi.org/10.1360/TB-2019-0191, 2019.

Ding, Y., Mu, C., Wu, T., Hu, G., Zou, D., Wang, D., Li, W., and Wu, X.: Increasing cryospheric hazards in a warming climate, Earth-Science Reviews, 213, https://doi.org/10.1016/j.earscirev.2020.103500, 2021.

#L.87-90. "Furthermore, the large-scale maps of vegetation and soil types…" needs citations

Response: we added the reference of (Mishra et al., 2013; Mu et al., 2020).

Mishra, U., Jastrow, J. D., Matamala, R., Hugelius, G., Koven, C. D., Harden, J. W., Ping, C.

L., Michaelson, G. J., Fan, Z., and Miller, R. M.: Empirical estimates to reduce modeling uncertainties of soil organic carbon in permafrost regions: a review of recent progress and remaining challenges, Environmental Research Letters, 8, 1402-1416, https://.org/10.1088/1748-9326/8/3/035020, 2013.

Mu, C. C., Abbott, B. W., Norris, A. J., Mu, M., Fan, C. Y., Chen, X., Jia, L., Yang, R. M., Zhang, T. J., Wang, K., Peng, X. Q., Wu, Q. B., Guggenberger, G., and Wu, X. D.: The status and stability of permafrost carbon on the Tibetan Plateau, Earth-Science Reviews, 211, 21, https://doi.org/10.1016/j.earscirev.2020.103433, 2020.

#L.33, L.216, and L.367. what is SOCSs?

Response: Changed.

#L.302 & L.302. Change “in the area of” into “in the areas of”

Response: Changed.

#L.223. Change “Fig. A1” to “Fig. S1” Also in L.323.

Response: Changed.

#L.248. “To test the predictive effects of the two machine learning methods…”, two or three?

Response: Changed.

#L.289-291. “The estimated SOC storage at a depth interval of 0–300 cm in forest, shrub, cropland, grassland, and desert areas was 3.30 Pg, 0.85 Pg, 31.67 Pg, 9.77 Pg, and 0.59 Pg, thus accounting for 7.15%, 1.84%, 68.58%, 21.57%, and 1.28% of the total, respectively”. Mismatching in different vegetation types and SOC storage values.

Response: Changed.

#L.296. “lithosols” Inconsistent font.

Response: Changed.

#L.302. Change “QTP” into “Third pole region”

Response: Changed.

#L.313-314. Hence, most terrestrial SOCS studies have focused on the shallow soil layer within 100 cm …especially that of permafrost zones (Ding et al., 2016; Mu et al., 2015; Wang et al., 2020; Zhao et al., 2018). Wrong citations. The soil depth is deeper down to 100cm in those articles

Response: We have deleted this sentence.

#L.475. Change “In” into “in”.

Response: Changed.

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