

## ***Interactive comment on “A review on the global soil datasets for earth system modeling” by Yongjiu Dai et al.***

### **Anonymous Referee #1**

Received and published: 31 October 2018

#### 1. General comments

This is a timely review of global scale soil data sets that are used to underpin Earth System Models, and the still numerous, associated uncertainties. Such soil data sets have evolved greatly since the coarse 1-degree resolution map generalised by Zobler (1986) resulting in a new generation of digital soil maps, and the underpinning soil point data sets and/or covariates layers. That being said, I have a number of queries and comments. For example, rather little attention is given to difficulties associated with the limited comparability of soil analytical data worldwide and uncertainty propagation. Further, several recent global soil databases of possible interest for ESM modelling have not been considered in the review and discussion.

The manuscript would benefit from a thorough English edit by a native speaker.

C1

#### 2. Specific comments

L15-16: Rephrase this as e.g.: Soil is an important regulator of earth system processes, but remains one of the least well-described data layers in such models.

L17: Function as → provide

L22: Abundant soil observations are not ‘enough’; these should have been analysed according to comparable analytical methods and quality-assessed (which is seldom the case, see Batjes et al. 2017). What about the geographical distribution, or possible clustering, of the available (i.e. shared) soil profile data?

L24: By their nature, pedotransfer functions generally are not portable from one region to the other. Please add some discussion.

L24-25: Speculative as written, provide some arguments for this.

L27-28: What about uncertainty in the co-variates?

L35-36 / 45: You may consider the following reference here: <http://dx.doi.org/10.1002/2015GB005239>.

L43: Remove available

L45: How do you define ‘better’ here? Please clarify.

L47-48: Also other types of soil data, for example soil biology (see ref. line L35-36). See also discussion in <https://doi.org/10.1111/gcb.13896>.

L56: Useful to say that the range of soil data collected during a soil survey, will vary with scale and projected applications of the data (i.e. type of soil survey, routine versus surveys/studies aimed at answering specific user demands).

L72: How would you define reliable soil data? Remove from this sentence.

L76: Rather refer to measurements here.

C2

L87-88: Should add HWSD (FAO/IIASA/ISRIC/ISS-CAS/JRC, 2012) as reference for this type of 'traditional' approach.

L93: usually not ready for ... → ... not appropriately scaled or formatted for ...

L113-114: ... representing main soil types in a landscape unit characterised by soil profiles considered representative for the main component soils of the respective mapping units.

L124: Rephrase this: ... (FAO, 2003b, Zobler 1986) and these products are known to be outdated. The information on the initial SMW and DSMW has since been updated for large sections of the world in the HWSD product (FAO/IIASA/ISRIC/ISS-CAS/JRC, 2012), which has recently been revised in WISE30sec (<http://dx.doi.org/10.1016/j.geoderma.2016.01.034>).

L124-125: Start new paragraph for the regional and national level data.

L132: multiply → multiple

L133: soil properties are observed (e.g. site data) or measured (e.g. pH, sand, silt, clay content)

L138-141: Important to mention here that data served through WoSIS have been standardised, with special attention for the description/comparability of soil analytical methods worldwide. See: <http://dx.doi.org/10.17027/isric-wdcsoils.20180001>. Also an important element for the discussion is that many countries, although having a large collection of soil profile data, are not yet sharing such data. See for example: <https://doi.org/10.1016/j.grj.2017.06.001>

L141: The initial list of attributes corresponds with the GlobalSoilMap specifications, with additional properties added/considered later in WoSIS (see <http://dx.doi.org/10.17027/isric-wdcsoils.20180001>).

L164: The linkage methods assigns a best-estimate for each soil property (and soil in-

C3

terval) under consideration to each component soil unit of a polygon (see e.g. HWSD). [see also 359-360]

L171-173: For a more comprehensive review see also: <http://dx.doi.org/10.1016/j.geoderma.2016.01.034> and <http://dx.doi.org/10.1002/ldr.2656>.

L178: FYI, WISE30sec considers seven layers up to 200 cm depth and 20 soil properties.

L201: Possibly, also mention the GSOC effort of the GSP here, see: <https://doi.org/10.5194/soil-4-173-2018>

L205: ... which is currently the most detailed, though not necessarily most accurate, estimation of ...

L214: See also: Tifafi M, Guenet B and Hatté CCG 2018. Large differences in global and regional total soil carbon stock estimates based on Soil-Grids, HWSD and NCSCD: Intercomparison and evaluation based on field data from USA, England, Wales and France. *Global Biogeochemical Cycles*, 42-56. <http://dx.doi.org/10.1002/2017GB005678>

Note: This paper is erroneously referred to as Marwa et al. 2018 in manuscript. This should be: Tifafi et al. 2018.

L214: Check if this is for 0-100 cm; likely these estimates are for 0-200 cm (see also recent sources mentioned above).

L224: Large sections of HWSDv1.2 still draw on the now outdated DSMW.

L295-296: See earlier comments.

L299: WISE30sec presents estimations of uncertainty, unlike the HWSD and GSDE.

L300: Needs some discussion and references to publications on the subject.

C4

L320: Larger number of soil properties for GSDE, but what about the accuracy of the predictions? (not given as indicated earlier).

L303: Rephrase. . . . SoilGrids products currently consider the list of attributes as defined by the GlobalSoilMap consortium.

L323: Most PTFs are not portable (i.e. locally or regionally validated).

L331-332: add database (word is missing in sentence)

L360-361: . . . component soil unit in most cases, and thus a one-to-many relationship exists between the SMU and the profile attributes of the respective soil units. . .

L397-398: Possibly, rephrase this sentence.

L410: remove high from sentence

L441-442: Provide some justification (a sentence or two) for this statement.

L451-452: Speculative as written. Please provide some evidence for this.

L460: and quantifying uncertainty in the predictions

L461: 'need to gain popularity in . . .'. Basically, the "proof of the pudding is in the eating".

L462: What I miss in this paper, is a discussion of the inherent uncertainty attached to using soil profile data coming from various sources. Often, little consideration is given to differences in analytical methods used for analysing e.g. soil organic carbon content worldwide (see Shangguang et al 2014, who consider this as 'a major imitation to their approach'). For a discussion of issues see e.g.: <http://dx.doi.org/10.17027/isric-wdcsoils.20180001>

L463-464: More soil profiles is not necessarily the solution. More quality-assessed data, analysed according to comparable analytical methods, are needed to support such efforts. Reference should be made to 'new' types of data

C5

as derived from proximal sensing (e.g. <http://dx.doi.org/10.5194/soil-2017-36>), and associated limitations. Reference, in this respect, could also be made to the GLOSOLAN effort, initiated by the GSP (<http://www.fao.org/global-soil-partnership/resources/events/detail/en/c/1037455/>) and work of GSP Pillar 5 towards harmonisation (<http://www.fao.org/3/a-bs756e.pdf>). Also, importantly, the geographical distribution and possible clustering of the shared soil profiles.

L471-475: True, but how many of these profiles are actually being shared for the greater benefit of the international community? See paper by Arrouays et al. 2017 for a discussion.

L479: Some reference to the ongoing work of the Global Soil Partnership, Pillars 4 and 5, is needed here.

L948: Table 2 is not complete; 'recent' datasets not yet considered in the review should be added here ( <http://dx.doi.org/10.1002/ldr.2656> ; <http://dx.doi.org/10.1016/j.geoderma.2016.01.034> ). Idem for Table 3.

L952: Table 3. Change title to "Derived soil properties considered in three global soil datasets". Essentially, this is a simple enumeration of derived soil properties. However, the fact that many different analytical methods have been used to derive a given soil property (e.g. soil organic carbon Walkley & Black method or LECO total analyses) or which CEC (e.g. measured at 'field pH' or in a buffer-solution at 'pH7' or 'pH8') has been considered is not mentioned here (in a footer perhaps). In their study, Shangguan et al. (2014) rightly indicate that this has not been the case and indicate that they see this an important limitation. However, there are still no straightforward mechanisms for harmonising the data (cf. GSP Pillar 5 and GLOSOLAN activities, as mentioned above).

Potaasium → Potassium

---

Interactive comment on SOIL Discuss., <https://doi.org/10.5194/soil-2018-32>, 2018.

C6