Earth Syst. Sci. Data Discuss.,
https://doi.org/10.5194/essd-2020-149-RC2, 2020

Earth System Science Data Discussions

Interactive comment on "SoilKsatDB: global soil saturated hydraulic conductivity measurements for geoscience applications" *by* Surya Gupta et al.

Anonymous Referee #2

Received and published: 13 August 2020

1

* Gupta and co-workers present an interesting dataset about pedo-hydrological properties. They collected data with a focus on saturated hydraulic conductivity from various publications and repositories around the globe. Such a dataset highly deserves publication and has strong potential to contribute to the advance of pedo-hydrological sciences. However, I see quite some room for improvement of the manuscript to really stretch out for this potential and to meet the standards of ESSD.

Some of the co-authors are my "idol-pedologists", who are always inspiring my own research. I feel slightly humble and confused to find this manuscript in such a sloppy and imprecise setting about methods, scale, pedometrics and functional soil descrip-

tion. I have not found any methodological reference about the steps taken to derive, compile and evaluate the data. Instead, the amount of time to digitise and compile the data is emphasised. I am full of confidence that the authors can and will rework their study to a more coherent and scientifically founded state. I hope my comments can constructively guide this process.

2 Major comments:

2.1 Clarification of the conceptual and methodological meta-information:

Throughout the manuscript the authors are not very shy in promoting the central role of Ksat for hydrological applications. Quite to the contrary, there is no word about the conceptual framing and implicit assumptions of Ksat and the respective methods to measure it. Ksat (saturated hydraulic conductivity) is commonly understood as the invert of the Darcy filter resistance (as implicitly argued in most of the manuscript). Ksat is also interpreted as infiltration capacity (as claimed in the abstract). The methods to measure saturated hydraulic conductivity and infiltration capacity differ strongly with respect of their conceptual assumptions. Infiltration capacity is even more under debate, since it has to account for surface conditions too. I clearly see that resolving this debate is not in the scope of this data publication. However, I recommend to be much more clear about your conceptual setting in general and to avoid overrating Ksat measurements.

Moreover, I strongly assume that mixing different measurement techniques will inevitably introduce biases to the data. To my experience, each method has limits which lead to different estimates of Ksat. In addition, the repeatability of "free drain" experiments (i.e. ring infiltrometers and to some degree also Amoozemeter) is very limited. Tension-controlled measurements have a much better performance, which can really

be repeated with similar readings. In the lab, the situation is much more controlled. But the difference between 100 ml and 250 ml ring samples can be substantial. Also here, different techniques and procedures might introduce biases. Such methodological biases cannot be recovered in the final dataset if they are not reported (at least where possible).

2.2 Global coverage and number of samples:

The authors have done a phantasmic job in compiling all the data. However, I am under the impression that there is little thought given to well-known scale issues. I understand that the authors try to leave this to the interpretation by the users of the database by reporting the geographic location. However, the manuscript holds several examples where coverage, data density and similar are referred to countries, continents or studies. I cannot really judge the value of the dataset based on the presented accumulation. Maybe defining a site as some pedological unit would be helpful. Alternatively, at least main textural and climatic classes could guide the overview?

Tab. 1 lists the data sources. Half of the datasets contribute only 10% of the data points. Half of the data stems from one publication about Florida soils. Moreover, it is obvious that there is a substantial amount of data still out there, which has not been published in a way that you could locate it. This gives rise to three questions: How does the skewed distribution of data sources influence the final product? How does the skewed distribution of data points in general imprint on the final product? How could colleagues add their data to the dataset?

I am also under the impression that the mere number of samples does not give me much insight with the necessary meta information about location, site conditions and method. 1000 double ring measurements at one sight might weigh little over 50 precise analyses with tension hood infiltrometers or lab measurements. . .

In addition, the dataset you describe actually contains 152042 entries for soil hydraulic properties. Ksat is only reported in 13267 entries. So why do you emphasise Ksat so much?

2.3 Confidence index:

Using a subjective confidence index about location and overall method appears rather unnecessarily sloppy to me. First of all, I suspect location much less of an issue than the reported values - especially at the scope of the dataset. The authors appear to emphasise otherwise. Second, I see quite easy to implement ways reducing subjectivity: For the location one could instead give some sort of standard deviation (e.g. if you only know the basin than the location is the centroid $\pm$ half the basin's extent). For the actual value, I find it of dramatic importance to report the used method whenever possible. Simply assuming field measurements to be less trustworthy than lab ones has weak reasons. Understandably, the authors do not analyse any coherence with neighbouring measurements or possible biases in different labs. However, this essential meta information needs to be conveyed to allow others to make use of the data. This also holds for the analyses of texture and Corg.

2.4 Pedo transfer functions:

I recommend to drop the topic of PTFs. The way it is introduced in the manuscript and the methods applied open hundreds of questions which I do not consider in the scope of the data publication. The current form does not adhere to the state of science in this field.

3 Minor comments:

P1L2: Isn't the infiltration capacity controlling this partioning and it is due to the commonly used models that ksat is considered a key parameter? I suggest to avoid overly strong claims but to emphasise on the value of the data in its own realm.

P1L2f.: Again, this is the concept but the physical processes are taking place in the soil pores. As some of the co-authors pioneer research in this domain, I can surely assume that we do not disagree about this. Hence, I think it is important to be precise about the conceptual underpinnings of the data.

P1L4: There is substantial literature about the scope- and scale-dependency of transferring measured ksat values to model applications. Using many datapoints obtained from a rather difficult to control measurement procedure (i.e. ring infiltrometers, and amoozemeters) might end up in more blur due to the method than insight about infiltration capacity. In the same lines of thought, lab measurements of ksat in differently sized ring samples and under different methods are prone to generate unknown biases on the recorded values for different soil situations. Moreover, it is well known that different landscape settings (e.g. forest vs. agricultural lands) have substantial impact. Hence, I am a little reluctant to follow your argumentation and to be impressed by the mere number of records here.

P1L6: "global database": How does your study relate to other globally available soil data products? How many classes are covered with how many samples? In which respect has standardisation been applied?

P1L7: "data density": Again, how does your data density relate to globally available soil maps/classes? I do not understand why the ranking of a country and continents shall be of importance. Most cover a broad range of climates and landscapes which might not be unique. . .

P1L8: "other soil variables": Again, I cannot judge from the numbers given if and

to what degree the samples are comparable. E.g. soil texture can be measured by quite a spectrum of methods with known biases. The retention properties are not fully covered by these more agronomically motivated references. . .

P1L11 "temperate climatic regions": Does this mean that your dataset mainly covers this climatic region? If so, maybe the title should include this.

P1L12 "random forest": This statement appears rather generic to me. Given some data, a random forest is known to produce very good fits. Moreover, I do not understand the reference to temperate and lab based measurements. You mean that one subset refers to the climate region and the other subset to all climate regions but excluding field measurements? This is difficult to get and set into perspective. How can I differentiate between methodological and conceptual effects here? I mean, could it be that PTFs based on the given variables have been developed in and for lab samples in temperate regions and thus apply well for these but that for field measurements and other climatic regions, the PTFs miss an important predictor?

P1L18 "data license": I am not a fan of Zenodo to publish such valuable data. Why don't you use a more geoscience specific, long-term available repository like Pangeae or GFZ-dataservice etc.?

P2L16f.: I do not understand this. https://esdac.jrc.ec.europa.eu/content/3d-soil-hydraulic-database-europe-1-km-and-250-m-resolution I assume that this is the respective data product and it is public. Do you mean the raw data behind the product? Since one of the co-authors is also author of the data product, why is it omitted?

P2L21f.: Please specify the spectrum of methods for Ksat derivation.

P2L23ff: ESMs operate at scales where even topography is highly aggregated. RS products are very quick in claiming surface properties which only show weak coherence with soil water dynamics. The scale of RS products varies greatly but is well below the

scale of ESMs. Honestly, I do not get your point here. It appears to me that you follow a quite classic but maybe not very contemporary conceptual model of soils as static filters which can be easily predicted once the filter resistance (or Ksat as the invert) is defined. This approach has its merits and does not counteract the value of your dataset. However, I would suggest to precisely clarify this conceptual setting and to refer these assumptions to the set of methods to derive the values of Ksat in the database.

P2L26ff.: I agree. In my opinion, this is a discussion topic how to define a standard for pedohydrological data to ease data processing. I came across several rather generic formulations so far which I strongly suggest to revise and recompile in a discussion section - or simply omit.

P2L32: I think this is an important point. However, in the current form I find little to learn about data standardisation and collection. Please reconsider where and how to make the point of data coherency.

P3L9: If I am not mistaken the only methodological citation goes to machine learning, which you do not at all tackle in the manuscript. Please strongly rework the manuscript to refer to the state of pedological and hydrological sciences.

P3L27f.: Sorry, but coordinate conversion is not an issue any longer as long the geographic system / EPSG code is given. You can directly use https://proj.org with the software of your choice. . . or https://espg.io online.

P3L30: How certain are you about this re-attribution of locations? Could this somehow be given as metadata in the database?

P3L33f.: I thank you very much for doing this work and providing the data. However, I do not expect digitising to be an issue worth debating here. There are many ways including automated processing. Definitely MS Word is not a necessary step but your choice of processing. . .

Table 2a: The README in the dataset gives slightly different entries. Please make coherent.

Table 2b: I do not see why table 2b is given. All information is or can be provided in table 2a already.

Table 3: As stated above, I suggest to fully rework the matter of confidence measures. Your proposed subjective index can only obscure the data – Especially since you combine spatial precision with lab/field method assumptions.

P7 Sec. 2.3 "Standardization and quality assignment": I do not see if or how this has been performed. Despite agreeing to your judgement about very small Ksat values, I would be interested why the colleages did not perform such "cleaning" in the original data. How can they possibly measure 10e-14m/day? I suspect some strange averaging with small numbers behind this. What do you mean with "cross-checking"?

Table 4 bottom row: I do not understand 32*. You report 11635 Samples for texture but 32 without texture class? Once you know the composition, the texture class is defined.? How many of the Ksat_lab samples have been measured in the field, too? I think this table is not very helpful. Maybe once the main topics and questions are clarified, a couple of easy plots would be more helpful to understand the dataset?

P9L1 "SWIG": Am I right assuming that this dataset holds 65 samples? If this is roughly 1% of the total number, I am not quite sure why this is highlighted here. Again, I would strongly recommend to include such specific metadata in the final table/database – especially because I suspect many other samples to suffer from similar issues.

P9L3: I think this is a very important subsection. Since it popped up ealier already, I recommend to restructure the section to give a specific place for this debate. Moreover, I find it very difficult to speak about accuracy here. Can this be really judged posteriorly? The overall method here could use so many means from pedometrics but largely avoids to get precise.

P9L4f.: Why? Are the methods mostly unknown? I suspect this to be of dramatic importance to report the used method whenever possible.

P9L6: See above about the index.

P9L9: I strongly disagree. Why should a sample carried to the lab have a better depth precision than an experiment in the field? The procedure to measure the depth is one of the most simple ones in pedology. The issue might be about the actual measurement though. E.g. if I use an Amoozemeter, I can precisely position the water supply probe but the recorded value might not reflect Ksat in the sense of hydrological models. . .

P9L10f.: This points right into the essence of whether Ksat reflects infiltration capacity (as claimed in the abstract) or if it is the invert of the Darcy filter resistance (as implicitly argued thoughout the manuscript). I recommend to be much more clear about your conceptual setting again. With respect to the air entry and/or full saturation (which I see as two distinct issues) there is clear reference in the respective measurement procedures. Hence I would not agree that lab and field mostly differ in this respect but in the definition of the sample boundaries. In the lab, the sample is (more or less) well contained in a ring (with all known issues about it). In the field, the lateral component of capillary water movement is mostly unknown. In addition, there is little control about the vertical extent of the sampled location and conductive macropores and/or less permeable cross-sections. . . (to name one example)

P9L12: Why should spatial accuracy (I suspect something like numbers of digits) be a quality attribute? Sec. 2.4: I can not at all follow your method here. What kind of PTF, what predictors, what training sets etc. pp. As stated above, I suggest to remove the PTFs.

P10L15 "13,267 values": Please clarify this number (which I see is the count in the file). In Table 4 you report 11,727 from field and lab (13,294 including those without texture classes).

P10L15 "sites": What is counted as one site?

P10L17: I find this list very difficult. You mix countries and continents. What is the information in it? Maybe it would be better to define the distribution of sites? Next line you refer to the state of Florida with half the samples. . .

P10L21: Sorry, but the numbers in table 4 are slightly different. . . Moreover, I do not gain any insight from them.

P10L24: What are statistical properties?

Fig 2: I find this plot not only superfluous but reporting incorrect proportions. Please drop.

Fig 3b: I do not understand this. A) Table 1 gives far more than 9 databases. B) Why should I look at a distribution of Ksat per database (holding an unspecified ensemble of sites) instead of any other site attribute?

Fig 4: Please keep the colour coding static! Maybe convert the counts to percentages of the data? How about plotting all plots in one line with the respective marginal distributions? This is one of the most insightful plots and deserved far more description in the caption and text.

P12L6f.: This does not surprise me. However, you address this topic later. Why do you refer to it here?

Rest of the document omitted because I expect it to be substantially reworked during major revisions.

Interactive comment on Earth Syst. Sci. Data Discuss., https://doi.org/10.5194/essd-2020-149, 2020.