Reply to Referee (ESSDD)

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

Anonymous Referee #2

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 description. 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.

RE: We thank the Reviewer for the critical assessment of our manuscript and for the numerous comments and suggestions. In the revised version, we have clarified the methodology and the database description and included additional analyses. We have improved in the dataset and removed the typos as you mentioned in your feedback. We have provided answers to your questions as listed below (in red).

Major comments:

Q: 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).

RE: We agree that it is essential to provide information related to methods. Now we have included the information of Ksat method (and soil texture, bulk density, and organic carbon methods) in the CSV file "sol_ksat.pnts_metadata.csv" (see version 0.3, https://doi.org/10.5281/zenodo.3752721) to recover the methodological bias in the final dataset. We also added a new table 4 listing all the methods used for the measurements and gave references to the methods.

Q: 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?

RE: We thank the Reviewer for appreciating our effort. With regards to specific points raised:

a) We have now included the information related to climate zones based on Köppen-Geiger climate zones map (Rubel and Kottek, 2010, Hamel et al., 2017) and pedological units based on openlandmap.org in the CSV file "sol_ksat.pnts_cl_pedo.csv" (see version0.3, https://doi.org/10.5281/zenodo.3752721)

b) The Reviewer is right, in the manuscript we mentioned that 50% of the Ksat data is from Florida and we agree that this would impact statistically the final product. We would like to give this liberty to the users to use this dataset as per their requirement.

C) Our database and code is publicly available (https://gitlab.com/openlandmap/compiled-ess-point-data-sets/-/tree/master/themes/sol/SoilHydroDB) and users can contribute new data by either opening a new issue or directly by adding code and doing a pull request (https://docs.gitlab.com/ee/user/project/merge_requests/creating_merge_requests.html).

Q: 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±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.

RE: Thank you for this suggestion. We do not use a confidence index anymore and just list the location accuracy (as shown in Table 3). We have also emphasized in the text that the actual measurement errors are usually unknown and digitized legacy soil data from scientific reports and similar should be used with caution (P17L9-10).

Q: 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.

RE: Thank you for your feedback. Our main objective was to show pedotransfer functions as a way to use the dataset (although we understand that there is a plethora of additional applications for the dataset). To better convey this, we modified the text and better explained the purpose for this application (P7L3-7and P8L1-3). Moreover, we removed the section on multilinear polynomial regression, focused on PTFs derived from random forest (as state of the art approach), and better described the importance of different variables in the result section. We have now showed how we derived the PTF (https://github.com/ETHZ-repositories/Ksat_database_2020/blob/master/Ksat_data_PTF_supplimentry_code.pdf)

Minor comments:

Q: 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.

RE: We modified the text in the abstract (P1L1).

Q: 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.

RE: We modified the text in the abstract (P1L2).

Q: P1L4: There is substantial literature about the scope- and scale-dependency of transferring measured ksat values to model applications. Using many data points 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.

RE: In the modified manuscript, we refer to the scale dependency (P15L14-15 and P16L1). We modified the text in the manuscript (P1L3)

Q: 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?

RE: In a new figure (Figure 3d), we list the number of samples per soil textural class. In this work, standardization refers to make units of datasets identical (this has been clarified in the manuscript - We modified the text in the abstract (P1L5).

Q: 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...

RE: Thank you for this question. Data density was provided to give an overview to the users about the compilation of data from different continents. In the revised version, we also provide information on distribution of samples across different climatic regions (P11L21-22). We have made some modifications to the text (P1L6-7). We also agree that it might be important to relate this data with soil maps/classes. Therefore, we overlaid the Ksat values on the openlandmap.org layer and extracted the values of soil classes (please see sol_ksat.pnts_cl_pedo.csv (see version 0.3, https://doi.org/10.5281/zenodo.3752721))

Q: 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...

RE: We agree with the Reviewer. We have now provided the method for these properties as much as we could extract from the respective papers. Please have a look at the CSV file "sol_ksat.pnts_metadata.csv" (see version0.3 'https://doi.org/10.5281/zenodo.3752721). We have also modified the text (P12L1-3)

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

RE: Dataset covers all climatic regions (this is quantified in the revised manuscript). Here, we extracted the Ksat values belonging to the temperate climate region by overlaying the climate zone map. Further, the PTF was derived using these points. Then, PTF was tested for Ksat values belonging to the tropical climate region.

Q: 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?

RE: Sorry for the confusion. In the manuscript, our goal was to address two different aspects.

1. Firstly, we overlaid the 13,267 points on the climate zone map (now explained in the method section) and extracted only those points where information on sand, clay, and bulk density was available. Then, we extracted only points in the temperate climate zone ksat values and fitted the model to 80% of these measurements using the Random forest approach. The fitted model was tested on the remaining temperate data points (20 %) and on tropical Ksat values. In this case, we mixed both lab and field measurements.
2. In the second case, we separated 13267 points based on lab and field methods (9162 and 4133, respectively). For lab data, we fitted the model based on 80% of the lab-based ksat values and

tested it on the remaining (20%) lab-based data values and on all the field-based Ksat values. In this application, we did not differentiate between different climatic regions.

We have now clarified this in the methods (P7L3-7and P8L1-3).

Q: 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.?

RE: Thank you for your suggestions. We will consider these options in the future.

Q: 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?

RE: The publically available maps show the **predictions** of Ksat. However, the underlying measured data are not publicly available. We tried initially to ask this data from the authors, but due to government restrictions, they could not share.

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

RE: We have modified the text (P2L25-27). Please see also Table 4.

Q: 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.

RE: We have modified the sentence (P2L28-29) but we are not sure if we understood the reviewer correctly. In advanced Earth System models, the spatial resolution (~1km) also for very large regions is comparable for many RS-based products.

Q: P2L26ff.: I agree. In my opinion, this is a discussion topic on 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.

RE: We agree. We have now removed the sentence.

Q: 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.

RE: Thanks for this suggestion. Now, we have modified the text. Please see the subsection "statistical modelling of Ksat".

Q: 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.

RE: The Reviewer is right. However, to facilitate the user, we have standardized the geographic system.

Q: 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...

RE: We agree with the Reviewer. We removed this sentence from the manuscript.

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

RE: The README file has been corrected.

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

RE: The table provides a glimpse of the CSV file and its inclusion was recommended by the editor.

Q: 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.

RE: We have modified this part of the manuscript. We have provided the Ksat method for each study and separated it from location accuracy (please have a look at table 3).

Q: 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 colleagues 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"?

RE: In the SWIG database, 1845 Ksat measurements were extracted from the literature, and Ksat for other samples were computed using the infiltration database, fitting infiltration data series to Ksat. Some Ksat values computed using infiltration database were less than 10^{-14} m/day, which seems unreasonable, so these values were not included in the database. We have modified in the text (P6L14-16).

Cross-checking: Here cross-checking means that we crosschecked all the datasets to avoid the mistakes considering the same dataset two times. For example, SWIG database included the database from Zhao et al. (2018) in the Tibetan plateau. We removed Zhao et al (2018) from SWIG and presented the data of Zhao as separated database.

Q: 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?

RE: We thank the Reviewer for noticing this - These 32 values in the soil texture class are errors. It means that the total of sand, silt, and clay % is more or less than 102 or 98%. However, after reanalyzing the data, we found that 75 values have the same problem. Hence, we provided soil texture class as “Error”. We have modified in the text (P12L8-9).

Q: 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.

RE: The SWIG dataset holds 3637 samples. No we have added the Ksat methods. Please have a look at "sol_ksat.pnts_metadata.csv" (see version 0.3, https://doi.org/10.5281/zenodo.3752721) file and table 4)

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

RE: Now we have provided the method information for each sample (Please see "sol_ksat.pnts_metadata.csv" (see version 0.3, https://doi.org/10.5281/zenodo.3752721) file and table 4)

Q: P9L6: See above about the index.

RE: As stated above, we don't use the confidence index anymore.

Q: 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...

RE: We agree with the Reviewer that it might not be the correct way to provide a subjective confidence degree based on the measurement method. Hence, in the revised version, we removed the confidence degree based on the measurement method and only provided the positional accuracy based on the location (P6L20-24).

Q: 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).

RE: We agree with the Reviewer and modified in the text (P6L20-24).

Q: 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.

RE: We have modified the text accordingly to make it more clear to the users (P7L3-7and P8L1-3).

Q: 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).

RE: It is because; there are 4 studies in the dataset which have both field and lab measurements. We mentioned this in the metadata CSV file "sol_ksat.pnts_metadata.csv" (see version 0.3, https://doi.org/10.5281/zenodo.3752721).

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

RE: One site is equal to one location id (Combination of latitude and longitude).

Q: 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…

RE: We thank the Reviewer for pointing this out. Now we have given the Ksat points distribution based on continents and climate region (P11L21-23).

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

RE: We thank the Reviewer for pointing this out. In the revised version, we rechecked the numbers and fixed typos. It is important to show the mean values of soil properties under various soil texture classes for the users.

Q: P10L24: What are statistical properties?

RE: We have modified the subsection from "Statistical properties" to "Statistical properties of SoilKsatDB".

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

RE: We modified the captions to highlight that the proportions are not correct. However, we prefer to keep the figure because it is illustrative to show for how many samples the different soil properties are measured.

Q: 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?

RE: Thank you for your feedback. It is illustrative to show that databases with many field data and from different regions show the highest spread of data. Now we have also added the violin plot for soil texture classes (see figure 3).

Q: 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.

RE: Thanks for noticing this. We have now made the color coding static for figure 4 and revised the captions.

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

RE: We incorporated figure 7 as new panel in Figure 3 to present the statistics of measured value in one concise Figure. In this section, we just report the key differences and discuss the origin of the differences later on.

References

Hamel, P., Falinski, K., Sharp, R., Auerbach, D. A., Sánchez-Canales, M., & Dennedy-Frank, P. J. (2017). Sediment delivery modeling in practice: Comparing the effects of watershed characteristics and data resolution across hydroclimatic regions. Science of the Total Environment, 580, 1381-1388.

Rubel, F., & Kottek, M. (2010). Observed and projected climate shifts 1901–2100 depicted by world maps of the Köppen-Geiger climate classification. Meteorologische Zeitschrift, 19(2), 135-141.

Kutílek, M., & Krejca, M. (1987). Three-parameter infiltration equation of Philip type. Vodohosp. ˇCas, 35, 52-61.

Haverkamp, R., Ross, P. J., Smettem, K. R. J., & Parlange, J. Y. (1994). Three-dimensional analysis of infiltration from the disc infiltrometer: 2. Physically based infiltration equation. Water Resources Research, 30(11), 2931-2935.

Rahmati, M., Weihermüller, L., Vanderborght, J., Pachepsky, Y. A., Mao, L., Sadeghi, S. H., ... & Schütt, B. (2018). Development and analysis of the Soil Water Infiltration Global database. Earth system science data, 10, 1237-1263.