SoilKsatDB: global database of soil saturated hydraulic conductivity measurements for geoscience applications

Surya Gupta[1], Tomislav Hengl[2,3], Peter Lehmann[1], Sara Bonetti[4], and Dani Or[1,5]

[1]Soil and Terrestrial Environmental Physics, Department of Environmental Systems Science, ETH, Zürich, Switzerland
[2]OpenGeoHub foundation, Wageningen, the Netherlands
[3]EnvirometriX, Wageningen, the Netherlands
[4]Institute for Sustainable Resources, Bartlett School of Environment, Energy and Resources, University College London, London, UK
[5]Division of Hydrologic Sciences, Desert Research Institute, Reno, NV, USA

Correspondence: Gupta S.
surya.gupta@usys.ethz.ch

Abstract. The saturated soil hydraulic conductivity (Ksat) is a key parameter in many hydrological and climate models. Ksat values are primarily determined from soil basic properties and may vary over several orders of magnitude. Despite the availability of Ksat datasets in the literature, significant efforts are required to combine the data before it can be used for specific applications. In this work, a total of 13,258 Ksat measurements from 1,908 sites were assembled from the published literature and other sources, standardized (i.e., units made identical), and quality-checked in order to obtain a global database of soil saturated hydraulic conductivity (SoilKsatDB). The SoilKsatDB covers most regions across the globe, with the highest number of Ksat measurements from North America, followed by Europe, Asia, South America, Africa, and Australia. In addition to Ksat, other soil variables such as soil texture (11,584 measurements), bulk density (11,262 measurements), soil organic carbon (9,787 measurements), moisture content at field capacity (7,382) and wilting point (7,411) are also included in the dataset. To show an application of SoilKsatDB, we derived Ksat pedotransfer functions (PTFs) for temperate regions and laboratory-based soil properties (sand and clay content, bulk density). Accurate models can be fitted using a Random Forest machine learning algorithm (best concordance correlation coefficient (CCC) equal to 0.74 and 0.72 for measurements from temperate areas and for laboratory measurements, respectively). However, when these Ksat PTFs are applied to soil samples obtained from tropical climates and field measurements, respectively, the model performance is significantly lower (CCC = 0.49 for tropical and CCC = 0.10 for field measurements). These results indicate that there are significant differences between Ksat data collected in temperate and tropical regions and Ksat measured in the laboratory or field. The SoilKsatDB dataset is available at https://doi.org/10.5281/zenodo.3752721 (Gupta et al., 2020) and the code used to extract the data from the literature and the applied random forest machine learning approach are publicly available under an open data license.

1 Introduction

The soil saturated hydraulic conductivity (Ksat) describes the rate of water movement through saturated soils and is defined as the ratio between water flux and hydraulic gradient (Amoozegar and Warrick, 1986). It is a key variable in a number of hydrological, geomorphological, and climatological applications, such as rainfall partitioning into infiltration and runoff (Vereecken et al., 2010), optimal irrigation design (Hu et al., 2015), as well as the prediction of natural hazards including catastrophic floods and landslides (Batjes, 1996; Gliński et al., 2000; Zhang et al., 2018). Accurate measurements of Ksat in the laboratory and field are laborious and time consuming and are often scale dependent (Youngs, 1991). Using infiltrometer measurements in the field enables the measurement of Ksat also in forests and other types of structured soils; however, so far Ksat values have been measured mainly for agricultural soils (Romano and Palladino, 2002).

Efforts to produce reliable and spatially refined datasets of hydraulic properties date back to the 1970's with the proliferation of distributed hydrologic and climatic modeling. These early notable works also provided basic databases (some of which are used in this study) for Australia (McKenzie et al., 2008; Forrest et al., 1985), Belgium (Vereecken et al., 2017; Cornelis et al., 2001), Brazil (Tomasella et al., 2000, 2003; Ottoni et al., 2018), France (Bruand et al., 2004), Germany (Horn et al., 1991; Krahmer et al., 1995), Hungary (Nemes, 2002), the Netherlands (Wösten et al., 2001), Poland (Glinski et al., 1991), and the USA (Rawls et al., 1982). A detailed discussion of the available datasets for Ksat and other hydro-physical properties is provided in Nemes (2011). Collaborative efforts have resulted in the compilation of multiple databases, including the Unsaturated Soil Hydraulic Database (UNSODA) (Nemes et al., 2001), the Grenoble Catalogue of Soils (GRIZZLY) (Haverkamp et al., 1998), and the Mualem catalogue (Mualem, 1976). These databases however, focused on soil types and not on the spatial context of Ksat mapping. In an effort to provide spatial context, Jarvis et al. (2013), and Rahmati et al. (2018) published global databases for soil hydraulic and soil physical properties. Likewise, the European soil data center also started projects such as SPADE (Hiederer et al., 2006) and HYPRES (Wösten et al., 2000), for generating spatially referenced soil databases for several countries. Since HYPRES only includes western European countries, Weynants et al. (2013) gathered data from 18 countries and developed the European HYdropedological Data Inventory (EU-HYDI) database. This dataset is, however, not publicly available and was not included in this compilation. The datasets mentioned above cover almost all climatic zones except tropical regions, where Ksat values can be significantly different due to the strong local weathering processes and different clay mineralogy (Hodnett and Tomasella, 2002). Recently, Ottoni et al. (2018) published a dataset named HYBRAS (Hydrophysical Database for Brazilian Soils) improving the coverage of South American tropical regions. In addition, Rahmati et al. (2018) recently published the Soil Water Infiltration Global database (SWIG) with information on Ksat for the whole globe. In the SWIG database, some Ksat values were extracted from the literature and other Ksat values were deduced from infiltration time series. In contrast to laboratory measurements that determine Ksat as the ratio of flux density to gradient, infiltration-based methods determine Ksat by fitting infiltration dynamics to parametric models of the infiltration process; for a review on analytical models characterizing the infiltration process see Kutílek et al. (1988), Youngs (1991), and Vereecken et al. (2019).

The increasing demand for highly resolved descriptions of surface processes requires commensurate advances in representation of Ksat in modern Earth System Model (ESM) applications. Several existing Ksat datasets miss either coordinates or these are provided with an unknown accuracy thus limiting their applications for spatial modeling. For example, the SWIG dataset misses information on soil depth and assigns entire watersheds to a single coordinate. Similarly, the UNSODA dataset does not provide coordinates and soil texture information for all samples. For a few locations, HYBRAS uses a different coordinate system. Taken together, these limitations imply that, to prepare spatially referenced global Ksat datasets for large scale applications, it requires serious effort to compile, standardize, and quality check all (publicly available) literature.

The objective of the work here is to provide a new global standardized Ksat database (SoilKsatDB) that can be used for geoscience applications. To do so, a total of 13,258 Ksat measurements was compiled, standardized, and cross-checked to produce a harmonized compilation which is analysis-ready (i.e., it can directly be used to test various Machine Learning algorithms for spatial analysis). We compiled data from existing datasets and, to improve the spatial coverage in regions with sparse data, we conducted a literature search to include Ksat measurements in geographic areas that were not yet included in other existing databases. In the manuscript, we first describe the data compilation process and then describe methodological steps used to spatially reference, filter, and standardize the existing datasets. As an illustrative application of the dataset, we derive pedotransfer functions (PTFs) for different climatic regions and measurement methods and discuss their transferability to other regions/measurement methodologies. We fully document all importing, standardization, and binding steps using the R environment for statistical computing (R Core Team, 2013), so that we can collect feedback from other researchers and increase the speed of further updates and improvements. The newly created data set (SoilKsatDB) can be accessed via https://doi.org/10.5281/zenodo.3752721.

2 Methods and materials

2.1 Data sources

To locate and obtain all compatible datasets, a literature search was conducted using different search engines, including Science Direct (https://www.sciencedirect.com/), Google Scholar (https://scholar.google.com/) and Scopus (https://www.scopus.com). We searched soil hydraulic conductivity datasets using *"saturated hydraulic conductivity database"*, *"Ksat"*, and *"hydraulic conductivity curves"* as keywords. The collected datasets are listed in Table 1 together with the number of Ksat observations for each study. They can be classified into three main categories, namely: i) existing datasets (in form of tables) published and archived with a DOI in peer-reviewed publications, ii) legacy datasets in paper/document format (e.g., legacy reports, PhD theses, and scientific studies) and iii) on-line materials.

Existing datasets include published datasets such as HYBRAS (Ottoni et al., 2018), UNSODA (Nemes et al., 2001), SWIG (Rahmati et al., 2018), and the soil hydraulic properties over the Tibetan Plateau (Zhao et al., 2018), from which we extracted the required information as described in Table 2a. The major challenge with making the existing datasets compatible for binding (standardization, removing redundancy) was to obtain the locations for a particular sample as well as the corresponding measurement depths. For instance, the UNSODA database does not provide information on the geographical locations. To fill the gaps and make the data suitable also for spatial analysis, we used Google Earth to find the coordinates based on the given location (generally an address or a location name). Moreover, all datasets were cross-checked to avoid redundancy. For example, the UNSODA data includes the data of Vereecken et al. (2017) and Richard and Lüscher (1983/87) while the SWIG database includes the measurements of Zhao et al. (2018). Hence we removed these from UNSODA and SWIG database and used the original sources.

Table 1. List of articles and digitized Ksat datasets, and number of points (N) per data set used to generate the SoilKsatDB.

Reference	*N*	Reference	*N*	Reference	*N*
Rycroft et al. (1975)	1	Habel (2013)	3	Wang et al. (2008)	19
Waddington and Roulet (1997)	1	Nyman et al. (2011)	3	Deshmukh et al. (2014)	19
Takahashi (1997)	1	Bhattacharyya et al. (2006)	4	Price et al. (2010)	20
Katimon and Hassan (1997)	1	Lopes et al. (2020)	4	Bonsu and Masopeh (1996)	24
El-Shafei et al. (1994)	1	Yasin and Yulnafatmawita (2018)	4	Bambra (2016)	24
Lopez et al. (2015)	1	Daniel et al. (2017)	6	Verburg et al. (2001)	26
Kramarenko et al. (2019)	1	Arend (1941)	7	Southard and Buol (1988)	27
Zakaria (1992)	1	Helbig et al. (2013)	7	Chang (2010)	30
Ramli (1999)	1	Gwenzi et al. (2011)	7	Yao et al. (2013)	33
Singh et al. (2011)	1	Päivänen et al. (1973)	9	Becker et al. (2018)	34
Campbell et al. (1977)	1	Mahapatra and Jha (2019)	9	Baird et al. (2017)	50
Chief et al. (2008)	1	Amer et al. (2009)	9	Keisling (1974)	56
Conedera et al. (2003)	1	Radcliffe et al. (1990)	10	Rahimy (2011)	56
Ebel et al. (2012)	1	Vogeler et al. (2019)	10	Hao et al. (2019)	57
Ferreira et al. (2005)	1	Singh et al. (2006)	10	Kanemasu (1994)	60
Imeson et al. (1992)	1	Kelly et al. (2014)	10	Tete-Mensah (1993)	60
Johansen et al. (2001)	1	Elnaggar (2017)	11	Zhao et al. (2018)	65
Lamara and Derriche (2008)	1	Ganiyu et al. (2018)	12	Hinton (2016)	77
Parks and Cundy (1989)	1	Cisneros et al. (1999)	12	Vieira and Fernandes (2004)	86
Ravi et al. (2017)	1	Niemeyer et al. (2014)	12	Houghton (2011)	88
Smettem and Ross (1992)	1	Sharratt (1990)	14	Tian et al. (2017)	91
Boike et al. (1998)	2	Habecker et al. (1990)	14	Li et al. (2017)	118
Andrade (1971)	2	Nielsen et al. (1973)	14	Forrest et al. (1985)	118
Beyer et al. (2015)	2	Robbins (1977)	15	Richard and Lüscher (1983/87)	121
Blake et al. (2010)	2	Sonneveld et al. (2005)	15	Sanzeni et al. (2013)	127
Bonell and Williams (1986)	2	Quinton et al. (2008)	16	Vereecken et al. (2017)	145
Kutiel et al. (1995)	2	Simmons (2014)	16	Coelho (1974)	176
Martin and Moody (2001)	2	Ouattara (1977)	17	Kool et al. (1986)	240
Mott et al. (1979)	2	Hardie et al. (2011)	17	Nemes et al. (2001)	283
Rab (1996)	2	Baird (1997)	17	Ottoni et al. (2018)	326
Soracco et al. (2010)	2	Kirby et al. (2001)	17	Rahmati et al. (2018)	3637
Varela et al. (2015)	2	Yoon (2009)	17	Grunwald (2020)	6532
Sayok et al. (2007)	3	Jabro (1992)	18		
Abagandura et al. (2017)	3	Greenwood and Buttle (2014)	18		

In the case of legacy datasets (non-digital tabular format, non-peer-reviewed data), we invested a significant effort to digitize, clean and cross-check the data to extract Ksat values. Two datasets were also collected directly from project websites such as the NASA project providing data on hydraulic and thermal conductivity (retrieved from https://daac.ornl.gov/FIFE/guides/Soil_Hydraulic_Conductivity_Data.html and described in Kanemasu (1994)) and the Florida database (http://soils.ifas.ufl.edu) from Grunwald (2020).

There are many biomes and climatic regions, such as desert dunes, peatlands and frozen soils, for which very few Ksat measurements were publicly available. We have intensively searched for additional data for these areas and found 39 studies (each contains less than 5 Ksat measurements) to cover these regions. We thus digitized Ksat values from these studies (shown either in bar charts or line plots), georeferenced the maps where necessary, and then converted the data into tabular form. In some cases, we also contacted colleagues that worked in these regions to retrieve additional data.

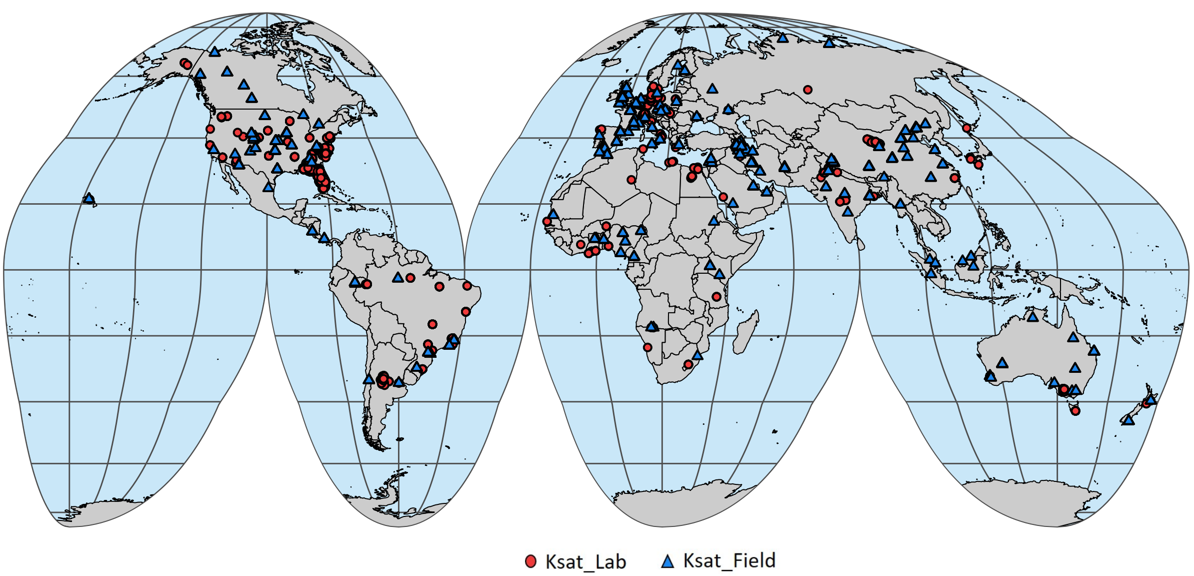

Figure 1. Spatial distribution of Ksat measurements based on (red) laboratory and (blue) field measurements, respectively, in the SoilKsatDB. A total of 1,908 locations are shown on the map.

2.2 Georeferencing Ksat values and definition of spatial accuracy

Georeferencing of Ksat measurements is important for using the data for local, regional, or global hydrological and land surface models. Although many studies provided information on the geographical location of the measurements, studies conducted particularly in the 70's and 80's only provided the name of the locations and approximate distance from a reference location. A limited accuracy of the position value may affect the application of the Ksat value in a spatially distributed model. For example, in case of a location with contrasting hydraulic properties, it must be known to which subregion the measured value can be assigned and the user must know if the given location is accurate enough. For that purpose we assigned an accuracy value (*'accuracy classes'*) to each measurement as explained next. We assigned each Ksat value to one of seven *'accuracy classes'* ranging from highest (0–100 m) to lowest (more than 10,000 m or non available information (NA)) accuracy. For example, Forrest et al. (1985), Zhao et al. (2018), and Ottoni et al. (2018) provided exact coordinates of the locations, thus we assigned a location accuracy of 0–100 m (i.e., highly accurate; see Table 3 for more details). For other references, we digitized provided maps or sketches with locations of the points. We first georeferenced these maps using ESRI ArcGIS software (v10.3) and then digitized the coordinates from georeferenced images. Some of the documents we digitized (e.g. Nemes et al. (2001)) provided the names of specific locations, and hence we used Google Earth to obtain the coordinates. We estimate that the spatial location accuracy of these points is roughly between 0 and 5 km. Similarly, spatial maps in jpg format (e.g. Becker et al. (2018)) were geo-referenced with 100–500 m location accuracy. In contrast, few studies (e.g. Yoon (2009)) provided the exact location of the sampling with assumed location accuracy of 10–20 m. In the SWIG database, the information related to location (coordinates for each point) was missing, so we went through each publication referenced in Rahmati et al. (2018) and added coordinates.

2.3 Standardization

The database was cleaned to remove 716 unrealistic low Ksat values as outliers deduced from infiltration time series in SWIG database. Moreover, in the SWIG database, soil depth information was not available, so we assumed that infiltration experiments were conducted in the topsoil and assigned a depth of 0–20 cm. Furthermore, we computed sand (particles > 50 m), silt (2-50 m), and clay fraction (< 2 m) for the UNSODA database based on the available particle-size data, assuming a log-normal distribution, as described in Nemes et al. (2001).

After data extraction from literature, geo-referencing, and standardization (conversion of all values to the same units), all information was collected in tabulated form in the new database SoilKsatDB (https://doi.org/10.5281/zenodo.3752721). The database consists of 23 columns (various sample properties) and 13,259 rows (a header and 13,258 samples). An excerpt of the database with all 23 columns is shown in Table 2b.

2.4 Statistical modeling of Ksat

To show a possible application of the database, we computed various PTFs. The PTF models were fitted using a random forest (RF) machine learning algorithm (Breiman, 2001) in the R environment for statistical computing (R Core Team, 2013). We fitted the RF model for log-transformed (log_{10}) Ksat values as a function of primary soil properties. In this application, PTFs for Ksat were built based on bulk density and sand and clay content. The observed correlation between these primary soil properties and Ksat motivates their use as key variables for the estimation of PTFs. Organic carbon (OC) was not used to build the PTFs because (i) this information was missing for 15% of measurements and (ii) the correlation between OC and Ksat was poor (i.e. 0.005, Pearson's correlation coefficient). We derived two PTFs for Ksat:

1. *PTF for temperate regions*: the map of Ksat locations were overlaid on the Köppen-Geiger climate zone map (Rubel and Kottek, 2010; Hamel et al., 2017) and then divided the measurements based on climatic regions (temperate, tropical, boreal, and arid) to account for differences in climate and related weathering processes (Hodnett and Tomasella, 2002). A

Table 2a. Description and units of some key variables listed in the database. The list can be found in the readme-file (table 1) and dataset *'sol_ksat.pnts_horizons'* on Zenodo (https://doi.org/10.5281/zenodo.3752721). We used the same codes as in the National Cooperative Soil Survey (NCSS) Soil Characterization Database (National Cooperative Soil Survey, 2016).

Headers	Description	Units
ID	Unique ID	—
site_key	Data set identifier	—
longitude_decimal_degrees	Ranges up to +180 degrees down to -180 degrees	Decimal degree
latitude_decimal_degrees	Ranges up to +90 degrees down to -90 degrees	Decimal degree
location_accuracy_min	Minimum value of location accuracy	m
location_accuracy_max	Maximum value of location accuracy	m
hzn_top	Top of soil sample	cm
hzn_bot	Bottom of soil sample	cm
hzn_desgn	Designation of soil horizon	—
db	Bulk density	g cm^{-3}
w3cld	Soil water content at 33 kPa (field capacity)	vol %
w15l2	Soil water content at 1500 kPa (wilting point)	vol %
tex_psda	Soil texture classes based on USDA	—
clay_tot_psa	Mass of soil particles, < 0.002 mm	%
silt_tot_psa	Mass of soil particles, > 0.002 and < 0.05 mm	%
sand_tot_psa	Mass of soil particle, > 0.05 and < 2 mm	%
oc_v	Soil organic carbon content	%
ph_h2o_v	Soil acidity	—
Ksat_lab	Soil saturated hydraulic conductivity from lab	cm day^{-1}
Ksat_field	Soil saturated hydraulic conductivity from field	cm day^{-1}
source_db	Source of the data	—
location_id	Combination of latitude and longitude	—
hzn_depth	Mean depth of soil horizon	—

total of 8,296 Ksat values for measurements in temperate-climate that contain information on sand, clay, and bulk density, were used to develop the PTF. The data set was randomly divided into a training (6,637 samples, 80%) and testing dataset (1,659 samples, 20%). PTFs for temperate regions include all depths of Ksat measurements (40% Ksat values from top soil 0-20 cm). In a validation step, we applied the PTF determined for temperate regions to tropical regions. Compared to temperate regions, the tropics are affected by different soil formation processes resulting in different clay mineral types. With such validation, we intend to discuss the transferability of PTFs across different regions.

Table 2b. Example of Ksat database structure with key variables (from left to right: unique ID, reference, longitude and latitude (decimal degree), minimum and maximum accuracy (m), top and bottom of soil sample (cm), horizon designation, bulk density (g cm^{-3}), moisture content at field capacity and wilting point (%), soil textural class, clay, silt and sand content (%), soil organic carbon content (%), soil acidity, saturated hydraulic conductivity measured in lab or field (cm day^{-1}), source of the data, location id and mean soil depth). NA is 'no value'. Column names are explained in Table 2a. The dataset file *'sol_ksat.pnts_horizons'* can be found on Zenodo (https://doi.org/10.5281/zenodo.3752721).

ID	site_key	longitude_ decimal_ degrees	latitude_ decimal_ degrees	location_ accuracy_ min	location_ accuracy_ max	hzn_ top	hzn_ bot	hzn_ desgn	db	w3cld	w1512	tex_ psda	clay_ tot_ psa	silt_ tot_ psa	sand_ tot_ psa	oc_ v	ph_ h20_ v	ksat_ lab	ksat_ field	source_ db	location_ id	hzn_ depth
1	Becker_2018	-110.13	31.73	0	100	0	15	NA	NA	NA	NA	Sandy loam	NA	NA	NA	NA	NA	NA	26.40	ETH_literature	ID_-110.13_31.73	7.5
2	Becker_2018	-110.09	31.72	0	100	0	15	NA	NA	NA	NA	Sandy loam	NA	NA	NA	NA	NA	NA	27.84	ETH_literature	ID_-110.09_31.72	7.5
3	Becker_2018	-110.09	31.69	0	100	0	15	NA	NA	NA	NA	Sandy loam	NA	NA	NA	NA	NA	NA	21.60	ETH_literature	ID_-110.09_31.69	7.5
4	Becker_2018	-110.05	31.74	0	100	0	15	NA	NA	NA	NA	Loam	NA	NA	NA	NA	NA	NA	23.76	ETH_literature	ID_-110.05_31.74	7.5
5	Becker_2018	-110.04	31.72	0	100	0	15	NA	NA	NA	NA	Sandy loam	NA	NA	NA	NA	NA	NA	39.12	ETH_literature	ID_-110.04_31.72	7.5
6	Becker_2018	-110.04	31.69	0	100	0	15	NA	NA	NA	NA	Sand	NA	NA	NA	NA	NA	NA	102.96	ETH_literature	ID_-110.04_31.69	7.5
7	Becker_2018	-110.02	31.67	0	100	0	15	NA	NA	NA	NA	Sand	NA	NA	NA	NA	NA	NA	111.36	ETH_literature	ID_-110.02_31.67	7.5
8	Becker_2018	-110.01	31.72	0	100	0	15	NA	NA	NA	NA	Sand	NA	NA	NA	NA	NA	NA	63.36	ETH_literature	ID_-110.01_31.72	7.5
9	Becker_2018	-110.00	31.72	0	100	0	15	NA	NA	NA	NA	Sand	NA	NA	NA	NA	NA	NA	133.44	ETH_literature	ID_-110.00_31.72	7.5
10	Becker_2018	-109.99	31.71	0	100	0	15	NA	NA	NA	NA	Sandy loam	NA	NA	NA	NA	NA	NA	9.84	ETH_literature	ID_-109.99_31.71	7.5

Table 3. Number of samples (N) assigned to each spatial accuracy class. NA are samples without information on spatial accuracy.

Minimum location error	Maximum location error	N
0 m	100 m	9801
100 m	250 m	972
250 m	500 m	623
500 m	1000 m	499
1000 m	5000 m	959
5000 m	10000 m	263
10000 m	NA	141
Total		**13,258**

2. *PTF for laboratory-based Ksat values*: in a second application, the dataset (total 13,258) was divided into laboratory and field based Ksat values. A total of 9,155 Ksat measurements belongs to laboratory and 4,131 Ksat measurements belong to field. The laboratory dataset (8,491 Ksat measurements with information on soil texture and bulk density information) was used for training (6,793) and testing (1,698) following the same method as used for the PTF for the temperate climate (i.e., 80% for training and 20% for testing). Lab-based PTFs include all depths of Ksat measurements (30% Ksat values from top soil 0-20 cm). Similar to the application of PTF from temperate region for the tropics, we apply the PTF deduced from laboratory data for prediction of Ksat measured in the field. We expect differences because field measurements scan larger soil volumes that may contain soil structural pores.

The *'ranger'* package version 0.12.1 (Wright and Ziegler, 2015) was used to build the PTFs. The PTFs developed for temperate regions and for laboratory data were then applied to test their ability to predict the result for the measurements in

Table 4. Instruments and methods used to estimate Ksat. A key reference with further details is given for all methods. The field methods are subdivided in different subcategories (in some cases, *'ponding'* in the filed methods or *'permeameter'* in the lab methods were listed in original studies without specification). Information on the applied method for each sample is provided in file *'sol_ksat.pnts_metadata_cl_pedo.csv'* (https://doi.org/10.5281/zenodo.3752721).

Lab Ksat methods	*N*	Field Ksat methods		*N*
Constant head method (Klute and Dirksen, 1986)	8055		Single (Bagarello and Sgroi, 2004)	460
Falling head method (Klute, 1965)	765	Ring (N = 1,282)	Double (Bodhinayake et al., 2004)	625
Triaxial cell (ASTM D 5084) (Purdy and Suryasasmita, 2006)	99		BEST (Lassabatere et al., 2006)	197
Pressure plate (Sharratt, 1990)	14		Mini Disc (Naik et al., 2019)	773
Oedometer test (ASTM D2435-96) (Sutejo et al., 2019)	12		Tension (Reynolds et al., 2000)	720
Permeameter	10	Infiltrometer (N = 2,207)	Disc (Soracco et al., 2010)	551
Oedometer test (UNI CEN ISO/TS 17892-5) (Terzaghi, 2004)	9		Guelph (Gupta et al., 1993)	87
			Hood (Schwärzel and Punzel, 2007)	41
			Micro (Sepehrnia et al., 2016)	35
			Guelph (Reynolds and Elrick, 1985)	153
			Aardvark (Hinton, 2016)	142
		Permeameter (N = 349)	Disc (Mohanty et al., 1994)	27
			Constant head (Amoozegar, 1989)	20
			Philip–Dunne (Muñoz-Carpena et al., 2002)	7
			Piezometer slug test (Baird et al., 2017)	72
			Tensiometers Nielsen et al. (1973)	70
			Rainfall simulator (Gupta et al., 1993)	55
		Others (N = 226)	Steady infiltration (Scotter et al., 1982)	16
			Ponding	8
			Auger method (Mohsenipour and Shahid, 2016)	5
Unknown	191		Unknown	67
Total	**9,155**			**4,131**

tropical climate (1,111 Ksat measurements) and for field measurements (1,998 Ksat measurements with information on soil texture and bulk density information), respectively. The code for generating and testing the PTFs is provided in the supplementary information. ANOVA (Analysis of Variance) with post-hoc Tukey's HSD (Honestly Significance Difference) test (Hilton and Armstrong, 2006; Abdi and Williams, 2010) was used to test the significant difference in Ksat between texture classes. ANOVA test indicates that at least one group differs from the other groups but does not explain the patterns of differences between means. Then, the Tukey HSD test was used to compute the significant difference between two means using a statistical distribution as shown in the results in Appendix A. This analysis was important to understand and check whether Ksat values between soil texture classes are significantly different or not.

The relative importance of the covariates for modeling Ksat was assessed by the node impurity, which, for RF regression problems, is computed as the decrease of residual sum of squares (RSS) when a particular covariate splits the data at the nodes of a tree (Hastie et al., 2009, sections 10.13.1, 15.3.2). The variable that provides maximum decline in RSS (and consequently

Table 5. Mean values of soil hydro-physical properties for each soil textural class. The number of samples (N) is given in parenthesis under each soil variable for each soil texture classes. N values marked with * correspond to undefined soil texture class. BD = bulk density (g/cm^3), OC = soil organic carbon content (%), FC = moisture content at field capacity (% vol), WP = moisture content at wilting point (% vol), $Ksat_l$ and $Ksat_f$ is laboratory and field Ksat (cm/day), respectively. For Ksat the geometric mean is reported (due to the sensitivity to a few extreme values for the arithmetic mean). For all other properties the arithmetic mean is provided.

Texture Classes	Clay (N)	Silt (N)	Sand (N)	BD (N)	OC (N)	FC (N)	WP (N)	$Ksat_l$ (N)	$Ksat_f$ (N)
Clay	56.3	23.6	20.0	1.27	2.00	43.2	30.0	8.22	110.07
	(830)	(830)	(830)	(639)	(448)	(447)	(449)	(499)	(331)
Silty Clay	45.2	45.1	9.6	1.18	3.83	49.9	30.2	3.63	196.65
	(181)	(181)	(181)	(175)	(116)	(46)	(46)	(85)	(96)
Sandy Clay	39.3	8.1	52.5	1.52	0.23	34.7	23.4	14.16	—–
	(176)	(176)	(176)	(172)	(140)	(158)	(158)	(172)	(4)
Clay Loam	31.4	38.6	29.9	1.27	2.49	37.2	22.1	13.34	60.56
	(544)	(544)	(544)	(382)	(360)	(76)	(76)	(127)	(417)
Silty Clay Loam	33.1	57.1	9.7	1.24	2.67	46.2	23.9	1.57	48.45
	(335)	(335)	(335)	(283)	(227)	(57)	(56)	(113)	(222)
Sandy Clay Loam	26.3	12.1	61.6	1.53	1.26	28.7	17.1	19.43	14.23
	(1148)	(1148)	(1148)	(966)	(950)	(805)	(759)	(876)	(272)
Silt	7.7	84.6	7.6	1.16	1.65	51.4	7.5	13.27	—-
	(25)	(25)	(25)	(19)	(11)	(12)	(11)	(25)	
Silt Loam	15.2	66.8	17.9	1.34	3.65	35.2	15.6	5.87	44.63
	(810)	(810)	(810)	(618)	(498)	(148)	(138)	(447)	(364)
Loam	19.0	39.1	41.7	1.29	2.16	32.07	14.2	45.62	34.21
	(685)	(685)	(685)	(593)	(561)	(94)	(97)	(219)	(464)
Sandy Loam	13.5	16.8	69.7	1.49	1.33	24.2	11.0	39.71	74.57
	(1601)	(1601)	(1601)	(1492)	(1337)	(806)	(792)	(1078)	(523)
Loamy Sand	7.3	8.5	84.0	1.55	1.13	17.3	6.5	95.37	132.33
	(736)	(736)	(736)	(711)	(674)	(582)	(586)	(637)	(99)
Sand	2.2	3.1	94.6	1.51	0.62	8.2	2.5	488.46	209.55
	(4513)	(4513)	(4513)	(4437)	(4179)	(4063)	(4062)	(4409)	(106)
Total	**11,584**	**11,584**	**11,584**	**10,487**	**9,501**	**7,294**	**7,229**	**8,687**	**2,900**
	(17*)	(0*)	(38*)	(775*)	(286*)	(88*)	(182*)	(468*)	(1,231*)

increase in node purity) is considered as the most important variable; the variable with the second largest RSS decrease is considered the second most important variable, and so on. Furthermore, the accuracy of the predictions was evaluated using bias, root mean square error (RMSE, in log-transformed Ksat measurement), and concordance correlation coefficient (CCC) (Lawrence and Lin, 1989).

Bias and RMSE are defined as:

$$bias = \sum_{i=1}^{n} \frac{(\hat{y}_i - y_i)}{n} \quad (1)$$

$$RMSE = \sqrt{\sum_{i=1}^{n} \frac{(\hat{y}_i - y_i)^2}{n}} \quad (2)$$

where y and $\widehat{y}$ are observed and predicted Ksat values, respectively, and n is the total number of cross-validation points.

The CCC is a measure of the agreement between observed and predicted Ksat value and is computed as

$$CCC = \frac{2 \cdot \rho \cdot \sigma_{\hat{y}} \cdot \sigma_y}{\sigma_{\hat{y}}^2 + \sigma_y^2 + (\mu_{\hat{y}} - \mu_y)^2} \quad (3)$$

where $\mu_{\hat{y}}$ and μ_y are predicted and observed means, $\sigma_{\hat{y}}$ and σ_y are predicted and observed variances and ρ is the Pearson correlation coefficient between predicted and observed values. CCC is equal to 1 for a perfect model.

3 Results

3.1 Data coverage of SoilKsatDB

Based on the literature search and data compilation, we have assembled a total of 13,258 values of Ksat from 1,908 locations (each location has a unique `location_id`) across the globe. Moreover, the database contains a total of 13,286 Ksat values because a few studies have reported both field and lab measurements for the same location. Figure 1 shows the global distribution of the sites used in this study. Most data originate from North America, followed by Europe, Asia, South America, Africa, and Australia. With respect to climatic regions, 10,093 Ksat measurements were taken in temperate regions (8,296 contained texture and bulk density information and were used to build PTF) and 1,443, 1,106, 580, and 36 in tropical, arid, boreal, and polar regions, respectively, as shown in Figure 2b. The points are often spatially clustered with the biggest cluster of points (1,103 locations with 6,532 Ksat measurements) in Florida (Grunwald, 2020). The Ksat database includes 4,131 values from field measurement and 9,155 values from laboratory measurements. In particular, different types of infiltrometers (e.g., Mini disc infiltrometer, Tension infiltrometer) and permeameters (e.g., Guelf permeameter, Aardwark permeameter) were used for the field measurements, whereas constant or falling head methods were mainly used in laboratory analyses (Table 2.2).

Out of the 13,258 Ksat measurements, 11,584 had information on soil texture, 11,262 on bulk density, 9,787 on organic carbon, 7,382 on field capacity, and 7,411 on wilting point, while for 8,994 measurements information for all soil basic properties (bulk density, soil texture, and organic carbon) was available (Figure 2a).

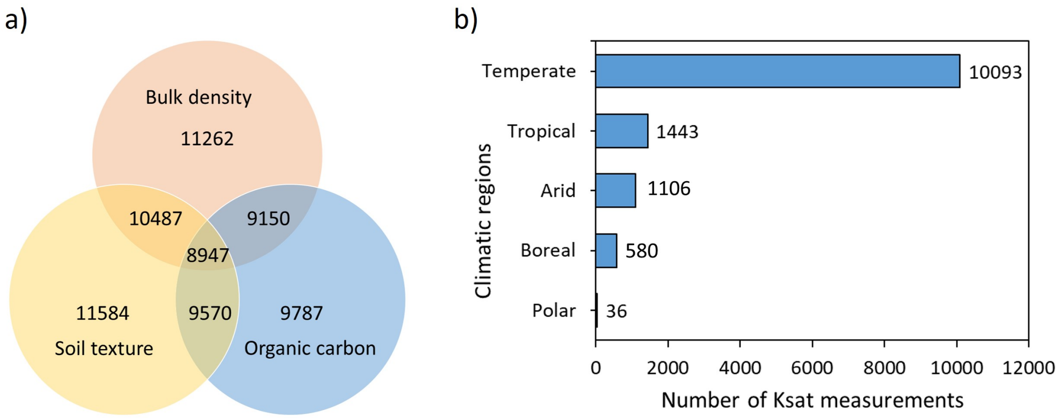

Figure 2. Description of Ksat measurements. (a) Venn diagram illustrating the number of Ksat measurements in the SoilKsatDB for which bulk density, soil texture, and soil organic carbon data were also available. Note that the size of the intersecting areas does not represent the correct fractions. (b) Distribution of Ksat measurements among climatic regions.

The methods used to compute these soil properties (as much as we could extract from the literature and existing databases) are listed in the CSV file `sol_ksat.pnts_metadata_cl_pedo.csv` available at https://doi.org/10.5281/zenodo.3752721. Note that in addition to 11,584 soil texture values, 75 measurements have soil texture information with total (sand+silt+clay) less than 98% or greater than 102%. We did not use these values in the PTF development but included it in the database as *"Error"* class in the soil texture column.

3.2 Statistical characteristics of SoilKsatDB

The distribution of measurements based on soil texture classes is shown on the USDA soil texture triangle in Figure 3a. The database covers all textural classes, with a high clustering in sandy soils due to the numerous samples from Florida (Grunwald, 2020), while only few measurements belong to the silt textural class. The increase in Ksat values in clayey and loamy soils for field methods (compared to laboratory methods) is likely due to the effect of soil structure. ANOVA with post-hoc Tukey's HSD test showed that the mean values for all broad soil texture classes are significantly different from each other, except for clayey soils field Ksat values and sandy soils field Ksat values (see Table A1). The violin distribution plot in Figure 3c shows the range of Ksat values for the different databases. Most of the datasets report Ksat values between 10^{-2} and $10^{2.5}$ cm/day, with a wider range of Ksat values observed in measurements from theses and reports (including studies with extreme values from sandy desert soils and low conductive clay soils) and from the SWIG database (databases 9 and 6 in Figure 3c, respectively). Likewise, Figure 3d shows the violin distribution of Ksat based on soil texture classes. The arithmetic mean of Ksat was highest for the sand and loamy sand soils (i.e., 2.68 and 1.99, respectively in $\log_{10}$ cm/day), while the lowest mean values were found for silt and silty loam (i.e., 1.12 and 1.15, respectively in $\log_{10}$ cm/day). Table A2 shows that the Ksat values in sand and loamy

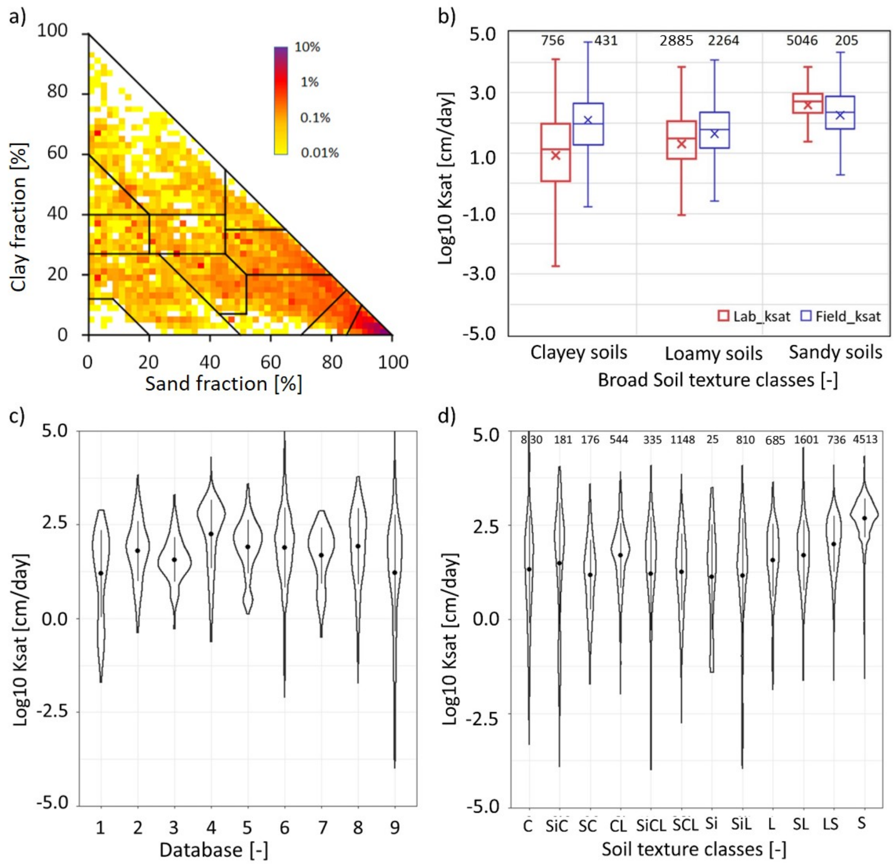

Figure 3. Characterization of collected Ksat values: (a) distribution of soil samples on the USDA soil texture triangle, (b) distribution of Ksat values using broad soil texture classes (sandy soils: sand and loamy sand; loamy soils: sandy loam, loam, silt loam, silt, clay loam and sandy clay loam; clayey soils: sandy clay, silty clay and clay) based on laboratory and field measurements (the number of measurements is shown on the top of the figure). Figure (c) shows the range of Ksat values spanned by each data source. The database numbers 1–9 refer to different sources and databases: 1 = Australia (Forrest et al., 1985), 2 = Belgium (Vereecken et al., 2017), 3 = China (Tian et al., 2017; Li et al., 2017), 4 = Florida (Grunwald, 2020), 5 = HYBRAS (Ottoni et al., 2018), 6 = SWIG (Rahmati et al., 2018), 7 = Tibetan Plateau (Zhao et al., 2018), 8 = UNSODA (Nemes et al., 2001), 9 = all other databases in Table 1. (d) Distribution of Ksat based on soil textural classes with the number of measurements shown on the top of the figure. In the violin diagrams (c and d) the dot represents the mean value, and the line represents the standard deviation for each data set.

sand soil texture classes are significantly different from all other soil texture classes. However, silt, silty clay, and silty clay loam classes are not significantly different from clay, sandy clay, and sandy clay loam Ksat values.

Average values of Ksat and other hydro-physical properties are shown in Table 5. Higher average organic carbon and bulk density values were observed in clayey and loamy soils compared to sandy soils. Ksat values obtained from field measurements were on average higher than those obtained from laboratory Ksat values. Particularly, for the clay texture class much lower

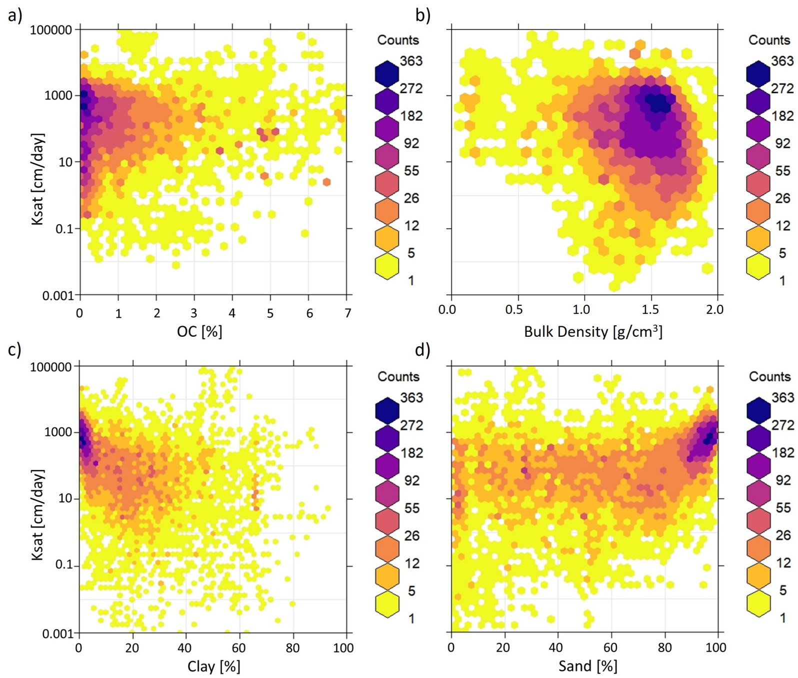

Figure 4. Partial correlation between Ksat and a) soil organic carbon OC (%), b) bulk density (g/cm^3), c) clay (%) and d) sand content (%) as heat maps. Ksat decreases with increasing clay content and bulk density, and increases with sand content. The color of each hexagonal cell shows the number of the measurements in each cell.

Ksat values were observed for laboratory (mean Ksat ≈ 8 cm/day) compared to field (mean Ksat ≈ 110 cm/day) measurements (Table 5). Figure 3b further illustrates the higher range of Ksat values obtained for finer texture soils (clay and loam) compared to coarser soils (sand).

3.3 Ksat PTFs derivation

As a test application of SoilKsatDB, two PTFs were derived for Ksat (i.e., for measurements taken in temperate regions and based on laboratory measurements) using basic soil properties as covariates. General trends between Ksat and soil properties are shown in partial correlation plots in Figure 4. The figure indicates that Ksat decreases with clay content and bulk density, and increases with sand content.

Figure A1 shows the list of relative importance of the covariates to build PTFs for the measurements from temperate regions and laboratory-based measurements. Clay content was found to be the most important variable followed by sand and bulk density for the temperate climate PTF. On the other hand, sand content was the most important variable followed by clay and bulk density for the laboratory-based Ksat PTF. CCC, bias, and RMSE were respectively equal to 0.74, -0.006, and 0.64, for the temperate-based PTF, and 0.72, -0.02, and 0.66 for laboratory-based PTF.

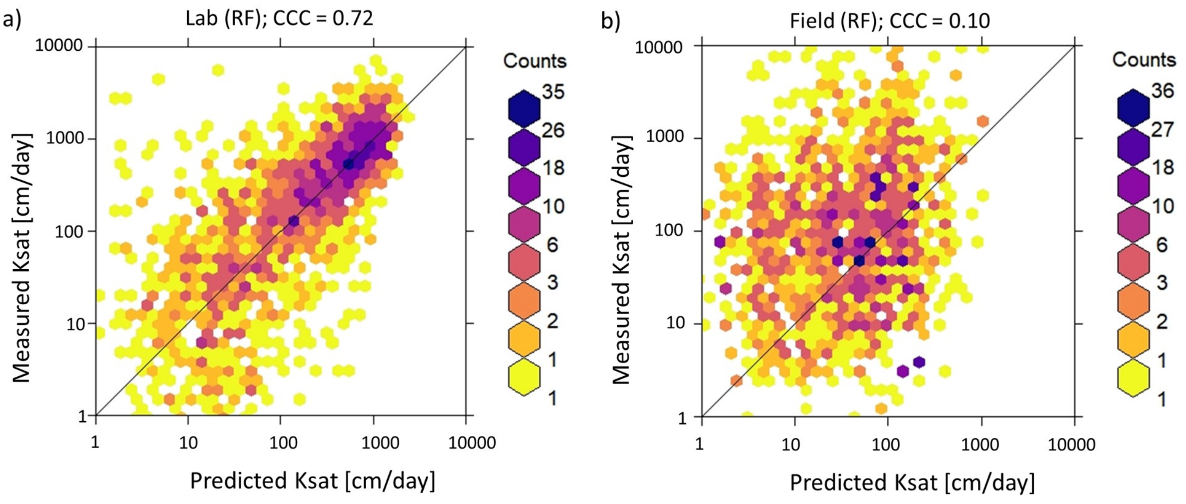

Figure 5. The correlation between observed and predicted Ksat values obtained from random forest (RF) models. The RF-based Pedotransfer function (PTF) model was fitted using data for laboratory measurements of Ksat and tested on both laboratory (a) and field (b) measurements. Results showed reasonable agreement (CCC = 0.72) using RF algorithms for laboratory measurements, but low CCC (0.10) for field measurements. PTFs developed based on laboratory measurements do not provide accurate estimates of Ksat measured in the field.

As we will discuss in more detail in the next section, PTF models derived for temperate and laboratory-based Ksat values underestimated Ksat for tropical- and field-based Ksat values, respectively (see Figure 6b and Figure 5b). CCC, bias, and RMSE values were respectively equal to 0.49, -0.2, and 0.94 for tropical Ksat values, and to 0.10, -0.22, and 1.2 for field measured Ksat values.

4 Discussion

4.1 Laboratory vs field estimated Ksat: effect of soil structure

The Ksat values were, on average, higher for the field measurements compared to laboratory measurements for most soil texture classes (Table 5 and Figures 3b and 5). The difference in laboratory and field based Ksat values and larger range of Ksat values for fine textured soil is probably related to the effect of biologically-induced soil structure that might be neglected in laboratory measurements. The omission of soil structures in many laboratory samples limits the possibility to properly reproduce field observations that are likely to be more affected by the presence of biopores (Fatichi et al., 2020) and other soil structural characteristics, such a cracks. In other words, variability in the Ksat values depends on the consideration (and existence) of soil structure by the measurement methods. Soil structural pores change the pore size distribution and subsequently affect Ksat values (Tuller and Or, 2002). Such an effect is more likely to be neglected in laboratory measurements rather than in field studies due to the small size of most laboratory samples. Presence or absence of large structural pores depends on the scale

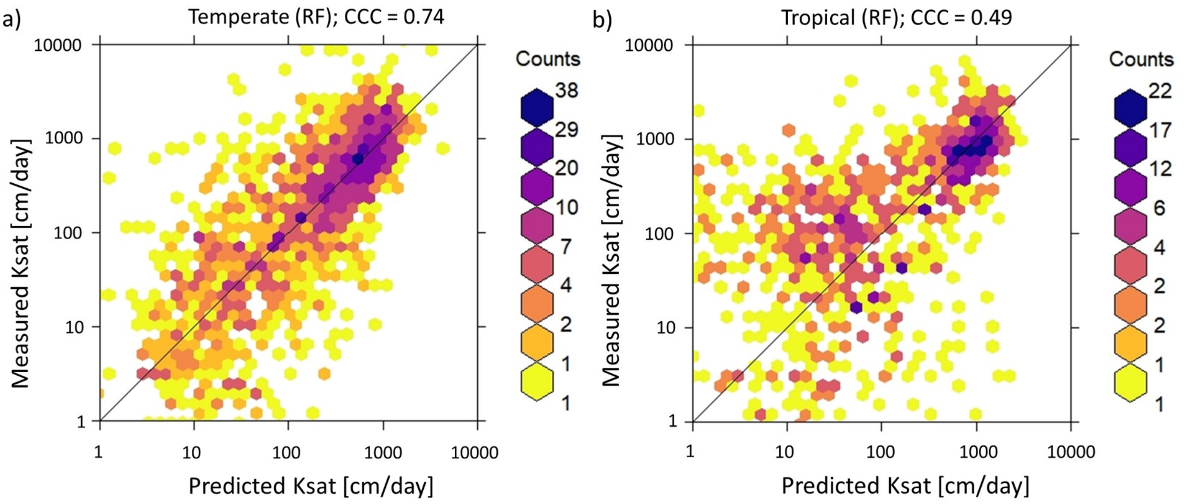

Figure 6. Correlation between observed and predicted Ksat values obtained from the random forest (RF) model. The RF-based Pedotransfer function (PTF) model was obtained by fitting 6,637 training points measured in a temperate-climate and tested on (a) data from temperate climate (1,659 measurements) and (b) data from tropical areas (1,111 measurements). PTFs showed good performance (CCC = 0.74) for the measurements taken in the temperate climate (including both laboratory and field measurements), but lower CCC values were obtained for tropical soil measurements (0.49 for RF). PTFs determined for temperate regions cannot be easily transferred to tropical regions because of the different soil forming processes.

of measurements (that is usually larger in the field). Mohanty et al. (1994), for example, compared three field methods and one laboratory method and found that the sample size affects the measurement of Ksat due to the presence and absence of open-ended pores. Similarly, Ghanbarian et al. (2017) showed that the sample dimensions (e.g., internal diameter and height) also impact Ksat. The authors further developed a sample dimension-dependent PTF, which performed better than other PTFs available in the literature. Likewise, Braud et al. (2017) used three field methods for Ksat measurements and found significant variation between these measurement methods. Davis et al. (1996) also highlighted the necessity to choose the most appropriate scale of measurement for a particular soil sample when undertaking conductivity measurements. They tested small cores (73 mm wide and 63 mm high) and large cores (223 mm wide and 300 mm high) using the constant head method in the laboratory and found a difference of 1 to 3 orders of magnitude.

4.2 Temperate vs tropical soils: effect of soil formation processes

PTFs obtained for temperate soils performed poorly for tropical soils (Figure 6), with Ksat being underestimated by the temperate-based PTFs. This result is in agreement with Tomasella et al. (2000) who derived PTFs using data from tropical Brazilian soils, which did not properly capture observations in temperate soils. We argue that the significant differences for tropical and temperate soils are due to the differences in the soil-forming processes that also define the clay type and mineralogy.

In fact, Oxisols (highly weathered clay soils as a result of high rainfall and temperature in tropical regions) are characterised by inactive (non-swelling) clay minerals. In contrast to tropical soils, active (smectite) and moderately active clay minerals (illite) are the dominant clay minerals in temperate regions. These swelling clay minerals retain water within internal structures with very low hydraulic conductivity. Therefore, such a difference in clay mineralogy is likely responsible for the underestimation of Ksat in tropical soils from PTFs based on measurements in temperate areas. In addition, soil structure formation processes may be different in tropical and temperate regions (perennial activities of vegetation in the tropics) which would also lead to differences between measured Ksat values for the two climatic regions.

4.3 Limitations of SoilKsatDB

We put an effort to combine laboratory and field data from across the globe. However, we acknowledge that there are still gaps in some regions, such as Russia and higher northern latitudes in general, which may result in uncertainties in Ksat estimates in such regions. The SoilKsatDB could also be of limited use for fine-resolution applications because many data points were characterized by limited spatial accuracy and missing soil depth information. Specifically, the spatial accuracy of many points is between tens of meters and several kilometers (see the methodology sections regarding the extraction of the spatial locations using Google Earth). Many of the records in the SoilKsatDB come from legacy scientific reports and the original authors could not be traced and contacted, hence we advise to use this data with caution. In addition, in the SWIG database, the soil depth and measurement method information were not provided, and often one location was used to represent an entire watershed. We tried to revisit each publication and extract the most accurate coordinates of assumed sampling locations. In addition, we assumed that most of the samples were obtained from field measurements as authors used different infiltrometers to compute Ksat, so there might be a few points in our SoilKsatDB that belong to laboratory measurements and that we have incorrectly assigned to field measurements. Moreover, the field measurements in the database are a mix of many Ksat measurement methods.

For each measurement, a location accuracy (0-100 m = highly accurate, >10,000 m = least accurate) was assigned based on the sampling location accuracy. The location accuracy can be used as a weight or probability argument in Machine Learning for Ksat mapping. We are aware that this was a rather subjective decision; a more objective way to assign weights would be to use the actual spatial positioning errors. Because these were not available for most of the datasets, we have opted for the definition of a location accuracy estimated from the available documentation.

4.4 Further developments

The advancement in remote sensing technology opens the doors to link the hydraulic properties with global environmental data. Satellite-based maps of environmental characteristics such as local information on vegetation, climate, and topography for specific areas, which are often ignored by basic PTFs, can be incorporated. For example, Sharma et al. (2006) developed PTFs using environmental variables such as topography and vegetation and concluded that these attributes, at fine spatial scales, were useful to capture the observed variations within the soil mapping units. Likewise, Szabó et al. (2019) used a random forest machine learning algorithm for mapping soil hydraulic properties and incorporated local environmental information such as vegetation, climate, and topography.

5 Data availability

All collected data and related soil characteristics are provided online for reference and are available at https://doi.org/10.5281/zenodo.3752721 (Gupta et al., 2020).

6 Summary and conclusions

We compiled a comprehensive global dataset of Ksat measurements ($N = 13,258$) by importing, quality controlling, and standardizing tabular data from existing soil profile databases and legacy reports, as well as scientific literature. The SoilKsatDB covers a broad range of soil types and climatic regions and hence is useful in global models. A larger variation in Ksat values was observed for fine-textured soils compared to coarse-textured soils, indicating the effect of soil structure on Ksat. Moreover, Ksat values obtained from field measurements were generally higher than those from laboratory measurements, likely due to the impact of soil structural pores in field measurements.

The new database was used to develop PTFs using RF algorithms for Ksat values obtained for temperate climates and for laboratory measurements. PTFs developed for a certain climatic region (temperate) or measurement method (laboratory) could not be satisfactorily applied to estimate Ksat for other regions (tropical) or measurement method (field) due to the role of different soil forming processes (inactive clay minerals in tropical soils and impact of biopores in field measurements).

There are still some gaps in the geographical representation of the data, especially in Russia and the higher northern latitudes, that could induce uncertainty in global modeling. Therefore, the data set can be further improved by covering the missing areas thus allowing better accuracy in modeling applications.

The SoilKsatDB was developed in R software and is available via https://doi.org/10.5281/zenodo.3752721. We have made code and data publicly available to enable further developments and improvements.

Appendix A

Table A1. Listed p-values under 95 percent confidence interval for each class to show the significant difference in Ksat between texture classes (see Figure 3b). The significant difference was computed using ANOVA (Analysis of Variance) with post-hoc Tukey's HSD (Honestly Significance Difference). The values highlighted in yellow show the significant difference in Ksat between two soil texture classes.

Broad soil texture classes	Lab Loamy soils	Lab Sandysoils	Lab Clayey soils	Field Loamy soils	Field Sandy soils	Field Clayey soils
Lab Loamy soils	1	<0.05	<0.05	<0.05	<0.05	<0.05
Lab Sandy soils	<0.05	1	<0.05	<0.05	<0.05	<0.05
Lab Clayey soils	<0.05	<0.05	1	<0.05	<0.05	<0.05
Field Loamy soils	<0.05	<0.05	<0.05	1	<0.05	<0.05
Field Sandy soils	<0.05	<0.05	<0.05	<0.05	1	0.1
Field Clayey soils	<0.05	<0.05	<0.05	<0.05	0.1	1

Table A2. Listed p-values under 95 percent confidence interval for each class to show the significant difference in Ksat between texture classes for Figure 3d. The significant difference was computed using ANOVA (Analysis of Variance) with post-hoc Tukey's HSD (Honestly Significance Difference) test. The values highlighted show the significant difference in Ksat between two soil texture classes.

Texture Classes	C	SiC	SC	CL	SiCL	SCL	Si	SiL	L	SL	LS	S
C	1	0.4	<0.05	<0.05	0.1	0.06	0.4	<0.05	<0.05	<0.05	<0.05	<0.05
SiC	0.4	1	<0.05	0.1	0.06	0.06	0.2	<0.05	0.5	0.09	<0.05	<0.05
SC	<0.05	<0.05	1	<0.05	0.8	0.3	0.8	0.8	<0.05	<0.05	<0.05	<0.05
CL	<0.05	0.1	<0.05	1	<0.05	<0.05	0.051	<0.05	<0.05	0.9	<0.05	<0.05
SiCL	0.1	0.06	0.8	<0.05	1	0.6	0.7	0.6	<0.05	<0.05	<0.05	<0.05
SCL	0.06	0.06	0.3	<0.05	0.6	1	0.6	0.1	<0.05	<0.05	<0.05	<0.05
Si	0.4	0.2	0.8	0.051	0.8	0.6	1	0.9	0.1	0.051	<0.05	<0.05
SiL	<0.05	<0.05	0.8	<0.05	0.6	0.1	0.9	1	<0.05	<0.05	<0.05	<0.05
L	<0.05	0.5	<0.05	<0.05	<0.05	<0.05	0.1	<0.05	1	<0.05	<0.05	<0.05
SL	<0.05	0.09	<0.05	0.9	<0.05	<0.05	0.051	<0.05	<0.05	1	<0.05	<0.05
LS	<0.05	<0.05	<0.05	<0.05	<0.05	<0.05	<0.05	<0.05	<0.05	<0.05	1	<0.05
S	<0.05	<0.05	<0.05	<0.05	<0.05	<0.05	<0.05	<0.05	<0.05	<0.05	<0.05	1

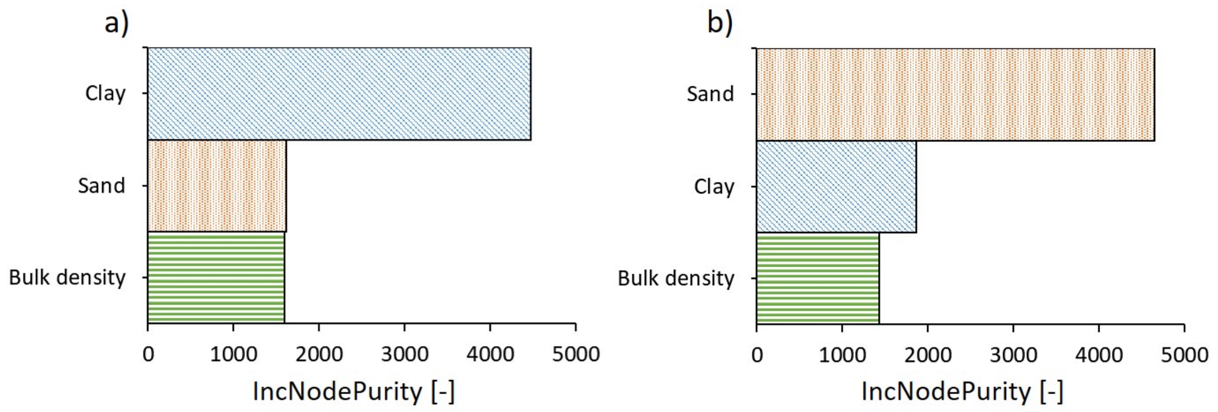

Figure A1. Importance of the variables for developing the PTFs for Ksat using random forest algorithm. The x-axis displays the average increase in node purity (the larger the value, the more important is a covariate). (a) Clay content was the most important variable followed by sand and bulk density for the Random Forest model built on data from temperate regions. (b) Sand content was the most important variable followed by clay and bulk density for the Random Forest model based on laboratory measurements.

Acknowledgements. The SoilKsatDB is a compilation of numerous existing datasets from which the most significant are: SWIG dataset (Rahmati et al., 2018), UNSODA (Leij et al., 1996; Nemes et al., 2001), and HYBRAS (Ottoni et al., 2018). The study was supported by ETH Zurich (Grant ETH-18 18-1). OpenGeoHub maintains a global repository of Earth System Science datasets at www.openlandmap.org. We thank Zhongwang Wei for helping in collecting the datasets and for insightful discussions. We acknowledge Samuel Bickel (ETH Zurich) for the help with High Performance Computing. We would also like to thank two anonymous reviewers, Dr. Attila Nemes and Dr. Sibylle K. Hassler (topical editor) for their constructive feedback to improve the manuscript.

References

Abagandura, G. O., Nasr, G. E.-D. M., and Moumen, N. M.: Influence of tillage practices on soil physical properties and growth and yield of maize in jabal al akhdar, Libya, Open Journal of Soil Science, 7, 118–132, 2017.

Abdi, H. and Williams, L. J.: Tukey's honestly significant difference (HSD) test, Encyclopedia of research design, 3, 583–585, 2010.

Amer, A.-M. M., Logsdon, S. D., and Davis, D.: Prediction of hydraulic conductivity as related to pore size distribution in unsaturated soils, Soil science, 174, 508–515, 2009.

Amoozegar, A.: A compact constant-head permeameter for measuring saturated hydraulic conductivity of the vadose zone, Soil Science Society of America Journal, 53, 1356–1361, 1989.

Amoozegar, A. and Warrick, A.: Hydraulic conductivity of saturated soils: field methods, Methods of Soil Analysis: Part 1 Physical and Mineralogical Methods, 5, 735–770, 1986.

Andrade, R. B.: The influence of bulk density on the hydraulic conductivity and water content-matric suction relation of two soils, 1971.

Arend, J. L.: Infiltration rates of forest soils in the Missouri Ozarks as affected by woods burning and litter removal, J. For., 39, 726–728, 1941.

Bagarello, V. and Sgroi, A.: Using the single-ring infiltrometer method to detect temporal changes in surface soil field-saturated hydraulic conductivity, Soil and Tillage research, 76, 13–24, 2004.

Baird, A. J.: Field estimation of macropore functioning and surface hydraulic conductivity in a fen peat, Hydrological Processes, 11, 287–295, 1997.

Baird, A. J., Low, R., Young, D., Swindles, G. T., Lopez, O. R., and Page, S.: High permeability explains the vulnerability of the carbon store in drained tropical peatlands, Geophysical Research Letters, 44, 1333–1339, 2017.

Bambra, A.: Soil loss estimation in experimental orchard at Nauni in Solan district of Himachal Pradesh, Ph.D. thesis, Dr. Yashwant Singh Parmar, University of horticulture and forestry, 2016.

Batjes, N. H.: Total carbon and nitrogen in the soils of the world, European journal of soil science, 47, 151–163, 1996.

Becker, R., Gebremichael, M., and Märker, M.: Impact of soil surface and subsurface properties on soil saturated hydraulic conductivity in the semi-arid Walnut Gulch Experimental Watershed, Arizona, USA, Geoderma, 322, 112–120, 2018.

Beyer, M., Gaj, M., Hamutoko, J. T., Koeniger, P., Wanke, H., and Himmelsbach, T.: Estimation of groundwater recharge via deuterium labelling in the semi-arid Cuvelai-Etosha Basin, Namibia, Isotopes in environmental and health studies, 51, 533–552, 2015.

Bhattacharyya, R., Prakash, V., Kundu, S., and Gupta, H.: Effect of tillage and crop rotations on pore size distribution and soil hydraulic conductivity in sandy clay loam soil of the Indian Himalayas, Soil and Tillage Research, 86, 129–140, 2006.

Blake, W. H., Theocharopoulos, S. P., Skoulikidis, N., Clark, P., Tountas, P., Hartley, R., and Amaxidis, Y.: Wildfire impacts on hillslope sediment and phosphorus yields, Journal of Soils and Sediments, 10, 671–682, 2010.

Bodhinayake, W., Si, B. C., and Noborio, K.: Determination of hydraulic properties in sloping landscapes from tension and double-ring infiltrometers, Vadose Zone Journal, 3, 964–970, 2004.

Boike, J., Roth, K., and Overduin, P. P.: Thermal and hydrologic dynamics of the active layer at a continuous permafrost site (Taymyr Peninsula, Siberia), Water Resources Research, 34, 355–363, 1998.

Bonell, M. and Williams, J.: The two parameters of the Philip infiltration equation: their properties and spatial and temporal heterogeneity in a red earth of tropical semi-arid Queensland, Journal of Hydrology, 87, 9–31, 1986.

Bonsu, M. and Masopeh, B.: Saturated hydraulic conductivity values of some forest soils of Ghana determined by a simple method, Ghana Journal of Agricultural Science, 29, 75–80, 1996.

Braud, I., Desprats, J.-F., Ayral, P.-A., Bouvier, C., and Vandervaere, J.-P.: Mapping topsoil field-saturated hydraulic conductivity from point measurements using different methods, Journal of Hydrology and Hydromechanics, 65, 264–275, 2017.

Breiman, L.: Random forests, Machine learning, 45, 5–32, 2001.

Bruand, A., Duval, O., and Cousin, I.: Estimation des propriétés de rétention en eau des sols à partir de la base de données SOLHYDRO: Une première proposition combianant le type d'horizon, sa texture et sa densité apparente., 2004.

Campbell, R. E., Baker, J., Ffolliott, P. F., Larson, F. R., and Avery, C. C.: Wildfire effects on a ponderosa pine ecosystem: an Arizona case study, USDA For. Serv. Res. Pap. RM-191. Fort Collins, CO: US Department of Agriculture, Forest Service, Rocky Mountain Forest and Range Experimental Station. 12 p., 191, 1977.

Chang, Y.-J.: Predictions of saturated hydraulic conductivity dynamics in a midwestern agricultural watershed, Iowa, 2010.

Chief, K., Ferré, T., and Nijssen, B.: Correlation between air permeability and saturated hydraulic conductivity: Unburned and burned soils, Soil Science Society of America Journal, 72, 1501–1509, 2008.

Cisneros, J., Cantero, J., and Cantero, A.: Vegetation, soil hydrophysical properties, and grazing relationships in saline-sodic soils of Central Argentina, Canadian Journal of Soil Science, 79, 399–409, 1999.

Coelho, M. A.: Spatial variability of water related soil physical properties., 1974.

Conedera, M., Peter, L., Marxer, P., Forster, F., Rickenmann, D., and Re, L.: Consequences of forest fires on the hydrogeological response of mountain catchments: a case study of the Riale Buffaga, Ticino, Switzerland, Earth Surface Processes and Landforms: The Journal of the British Geomorphological Research Group, 28, 117–129, 2003.

Cornelis, W. M., Ronsyn, J., Van Meirvenne, M., and Hartmann, R.: Evaluation of pedotransfer functions for predicting the soil moisture retention curve, Soil Science Society of America Journal, 65, 638–648, 2001.

Daniel, S., Gabiri, G., Kirimi, F., Glasner, B., Näschen, K., Leemhuis, C., Steinbach, S., and Mtei, K.: Spatial distribution of soil hydrological properties in the Kilombero floodplain, Tanzania, Hydrology, 4, 57, 2017.

Davis, S. H., Vertessy, R. A., Dunkerley, D. L., Mein, R. G., et al.: The influence of scale on the measurement of saturated hydraulic conductivity in forest soils, in: National Conference Publication-Institution of Engineers Australia NCP, vol. 1, pp. 103–108, Institution of Engineers, Australia, 1996.

Deshmukh, H., Chandran, P., Pal, D., Ray, S., Bhattacharyya, T., and Potdar, S.: A pragmatic method to estimate plant available water capacity (PAWC) of rainfed cracking clay soils (Vertisols) of Maharashtra, Central India, Clay Res, 33, 1–14, 2014.

Ebel, B. A., Moody, J. A., and Martin, D. A.: Hydrologic conditions controlling runoff generation immediately after wildfire, Water Resources Research, 48, 2012.

El-Shafei, Y., Al-Darby, A., Shalaby, A., and Al-Omran, A.: Impact of a highly swelling gel-forming conditioner (acryhope) upon water movement in uniform sandy soils, Arid Land Research and Management, 8, 33–50, 1994.

Elnaggar, A.: Spatial Variability of Soil Physiochemical Properties in Bahariya Oasis, Egypt, Egyptian J. of Soil Sci. (EJSS), 57, 313–328, https://doi.org/10.21608/EJSS.2017.4438, 2017.

Fatichi, S., Or, D., Walko, R., Vereecken, H., Young, M. H., Ghezzehei, T. A., Hengl, T., Kollet, S., Agam, N., and Avissar, R.: Soil structure is an important omission in Earth System Models, Nature Communications, 11, 2020.

Ferreira, A., Coelho, C., Boulet, A., and Lopes, F.: Temporal patterns of solute loss following wildfires in Central Portugal, International Journal of Wildland Fire, 14, 401–412, 2005.

Forrest, J., Beatty, H., Hignett, C., Pickering, J., and Williams, R.: Survey of the physical properties of wheatland soils in eastern Australia, 1985.

Ganiyu, S., Rabiu, J., and Olatoye, R.: Predicting hydraulic conductivity around septic tank systems using soil physico-chemical properties and determination of principal soil factors by multivariate analysis, Journal of King Saud University-Science, 2018.

Ghanbarian, B., Taslimitehrani, V., and Pachepsky, Y. A.: Accuracy of sample dimension-dependent pedotransfer functions in estimation of soil saturated hydraulic conductivity, Catena, 149, 374–380, 2017.

Glinski, J., Ostrowski, J., Stepniewska, Z., and Stepniewski, W.: Soil sample bank representing mineral soils of Poland, Problemy Agrofizyki (Poland), 1991.

Gliński, J., Stępniewski, W., Stępniewska, Z., Włodarczyk, T., Brzezińska, M., et al.: Characteristics of aeration properties of selected soil profiles from central Europe., International agrophysics, 14, 17–31, 2000.

Greenwood, W. and Buttle, J.: Effects of reforestation on near-surface saturated hydraulic conductivity in a managed forest landscape, southern Ontario, Canada, Ecohydrology, 7, 45–55, 2014.

Grunwald, S.: Florida soil characterization data, Soil and water science department, IFAS-Instituite of food and agriculture science, University of Florida, http://soils.ifas.ufl.edu, 2020.

Gupta, R., Rudra, R., Dickinson, W., Patni, N., and Wall, G.: Comparison of saturated hydraulic conductivity measured by various field methods, Transactions of the ASAE, 36, 51–55, 1993.

Gupta, S., Hengl, T., Lehmann, P., Bonetti, S., and Or, D.: SoilKsatDB: a global compilation of soil saturated hydraulic conductivity measurements, Zenodo, https://doi.org/10.5281/zenodo.3752721, 2020.

Gwenzi, W., Hinz, C., Holmes, K., Phillips, I. R., and Mullins, I. J.: Field-scale spatial variability of saturated hydraulic conductivity on a recently constructed artificial ecosystem, Geoderma, 166, 43–56, 2011.

Habecker, M., McSweeney, K., and Madison, F.: Identification and genesis of fragipans in Ochrepts of north central Wisconsin, Soil Science Society of America Journal, 54, 139–146, 1990.

Habel, A. Y.: The role of climate on the aggregate stability and soil erodibility of selected El-Jabal Al-Akhdar soils-Libya, Alexandria Journal of Agricultural Research, 58, 261–271, 2013.

Hamel, P., Falinski, K., Sharp, R., Auerbach, D. A., Sánchez-Canales, M., and Dennedy-Frank, P. J.: 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, 2017.

Hao, M., Zhang, J., Meng, M., Chen, H. Y., Guo, X., Liu, S., and Ye, L.: Impacts of changes in vegetation on saturated hydraulic conductivity of soil in subtropical forests, Scientific reports, 9, 8372, 2019.

Hardie, M. A., Cotching, W. E., Doyle, R. B., Holz, G., Lisson, S., and Mattern, K.: Effect of antecedent soil moisture on preferential flow in a texture-contrast soil, Journal of Hydrology, 398, 191–201, 2011.

Hastie, T., Tibshirani, R., and Friedman, J.: The Elements of Statistical Learning; Data Mining, Inference and Prediction, Springer, New York, 2 edn., 2009.

Haverkamp, R., Zammit, C., Bouraoui, F., Rajkai, K., Arrúe, J., and Heckmann, N.: GRIZZLY: Grenoble catalogue of soils: Survey of soil field data and description of particle-size, soil water retention and hydraulic conductivity functions, Lab. d'Etude des Transferts en Hydrol. et Environ., Grenoble, France, 1998.

Helbig, M., Boike, J., Langer, M., Schreiber, P., Runkle, B. R., and Kutzbach, L.: Spatial and seasonal variability of polygonal tundra water balance: Lena River Delta, northern Siberia (Russia), Hydrogeology Journal, 21, 133–147, 2013.

Hiederer, R., Jones, R. J., and Daroussin, J.: Soil Profile Analytical Database for Europe (SPADE): reconstruction and validation of the measured data (SPADE/M), Geografisk Tidsskrift-Danish Journal of Geography, 106, 71–85, 2006.

Hilton, A. and Armstrong, R. A.: Statnote 6: post-hoc ANOVA tests, Microbiologist, 2006, 34–36, 2006.

Hinton, H.: Land Management Controls on Hydraulic Conductivity of an Urban Farm in Atlanta, GA, 2016.

Hodnett, M. and Tomasella, J.: Marked differences between van Genuchten soil water-retention parameters for temperate and tropical soils: a new water-retention pedo-transfer functions developed for tropical soils, Geoderma, 108, 155–180, 2002.

Horn, A., Stumpfe, A., Kues, J., Zinner, H.-J., and Fleige, H.: Die Labordatenbank des Niedersächsischen Bodeninformationssystems (NIBIS)-. Teil: Fachinformationssystem Bodenkunde, Geologisches Jahrbuch. Reihe A, Allgemeine und regionale Geologie BR Deutschland und Nachbargebiete, Tektonik, Stratigraphie, Paläontologie, pp. 59–97, 1991.

Houghton, T. B.: Hydrogeologic characterization of an alpine glacial till, Snowy Range, Wyoming, Ph.D. thesis, Colorado State University. Libraries, 2011.

Hu, W., She, D., Shao, M., Chun, K. P., and Si, B.: Effects of initial soil water content and saturated hydraulic conductivity variability on small watershed runoff simulation using LISEM, Hydrological Sciences Journal, 60, 1137–1154, 2015.

Imeson, A., Verstraten, J., Van Mulligen, E., and Sevink, J.: The effects of fire and water repellency on infiltration and runoff under Mediterranean type forest, Catena, 19, 345–361, 1992.

Jabro, J.: Estimation of saturated hydraulic conductivity of soils from particle size distribution and bulk density data, Transactions of the ASAE, 35, 557–560, 1992.

Jarvis, N., Koestel, J., Messing, I., Moeys, J., and Lindahl, A.: Influence of soil, land use and climatic factors on the hydraulic conductivity of soil, Hydrology and Earth System Sciences, 17, 5185–5195, 2013.

Johansen, M. P., Hakonson, T. E., and Breshears, D. D.: Post-fire runoff and erosion from rainfall simulation: contrasting forests with shrublands and grasslands, Hydrological processes, 15, 2953–2965, 2001.

Kanemasu, E.: Soil Hydraulic Conductivity Data (FIFE), ORNL Distributed Active Archive Center, https://doi.org/10.3334/ORNLDAAC/107, 1994.

Katimon, A. and Hassan, A. M. M.: Field hydraulic conductivity of some Malaysian peat, Malaysian Journal of Civil Engineering, 10, 1997.

Keisling, T. C.: Precision with which selected physical properties of similar soils can be estimated, Ph.D. thesis, Oklahoma State University, 1974.

Kelly, T. J., Baird, A. J., Roucoux, K. H., Baker, T. R., Honorio Coronado, E. N., Ríos, M., and Lawson, I. T.: The high hydraulic conductivity of three wooded tropical peat swamps in northeast Peru: measurements and implications for hydrological function, Hydrological Processes, 28, 3373–3387, 2014.

Kirby, J., Kingham, R., and Cortes, M.: Texture, density and hydraulic conductivity of some soils in San Luis province, Argentina, Ciencia del suelo, 19, 20–28, 2001.

Klute, A.: Laboratory measurement of hydraulic conductivity of saturated soil, Methods of Soil Analysis: Part 1 Physical and Mineralogical Properties, Including Statistics of Measurement and Sampling, 9, 210–221, 1965.

Klute, A. and Dirksen, C.: Hydraulic conductivity and diffusivity: Laboratory methods, Methods of Soil Analysis: Part 1 Physical and Mineralogical Methods, 5, 687–734, 1986.

Kool, J., Albrecht, K. A., Parker, J., Baker, J., et al.: Physical and chemical characterization of the Groseclose soil mapping unit, 1986.

Krahmer, U., Hennings, V., Müller, U., and Schrey, H.-P.: Ermittlung bodenphysikalischer Kennwerte in Abhängigkeit von Bodenart, lagerungsdichte und Humusgehalt, Zeitschrift für Pflanzenernährung und Bodenkunde, 158, 323–331, 1995.

Kramarenko, V., Brakorenko, N., and Molokov, V.: Hydraulic conductivity of peat in Western Siberia, in: E3S Web of Conferences, vol. 98, p. 11003, EDP Sciences, 2019.

Kutiel, P., Lavee, H., Segev, M., and Benyamini, Y.: The effect of fire-induced surface heterogeneity on rainfall-runoff-erosion relationships in an eastern Mediterranean ecosystem, Israel, Catena, 25, 77–87, 1995.

Kutílek, M., Krejča, M., Haverkamp, R., Rendon, L., and Parlange, J.-Y.: On extrapolation of algebraic infiltration equations, Soil Technology, 1, 47–61, 1988.

Lamara, M. and Derriche, Z.: Prediction of unsaturated hydraulic properties of dune sand on drying and wetting paths, Electron. J. Geotech. Eng, 13, 1–19, 2008.

Lassabatere, L., Angulo-Jaramillo, R., Soria Ugalde, J., Cuenca, R., Braud, I., and Haverkamp, R.: Beerkan estimation of soil transfer parameters through infiltration experiments—BEST, Soil Science Society of America Journal, 70, 521–532, 2006.

Lawrence, I. and Lin, K.: A concordance correlation coefficient to evaluate reproducibility, Biometrics, pp. 255–268, 1989.

Leij, F., Alves, W., Van Genuchten, M. T., and Williams, J.: The UNSODA Unsaturated Soil Hydraulic Database; User's Manual, Version 1.0, Rep. EPA/600/R-96, 95, 103, 1996.

Li, X., Liu, S., Xiao, Q., Ma, M., Jin, R., Che, T., Wang, W., Hu, X., Xu, Z., Wen, J., et al.: A multiscale dataset for understanding complex eco-hydrological processes in a heterogeneous oasis system, Scientific data, 4, 170 083, 2017.

Lopes, V. S., Cardoso, I. M., Fernandes, O. R., Rocha, G. C., Simas, F. N. B., de Melo Moura, W., Santana, F. C., Veloso, G. V., and da Luz, J. M. R.: The establishment of a secondary forest in a degraded pasture to improve hydraulic properties of the soil, Soil and Tillage Research, 198, 104 538, 2020.

Lopez, O., Jadoon, K., and Missimer, T.: Method of relating grain size distribution to hydraulic conductivity in dune sands to assist in assessing managed aquifer recharge projects: Wadi Khulays dune field, western Saudi Arabia, Water, 7, 6411–6426, 2015.

Mahapatra, S. and Jha, M. K.: On the estimation of hydraulic conductivity of layered vadose zones with limited data availability, Journal of Earth System Science, 128, 75, 2019.

Martin, D. A. and Moody, J. A.: Comparison of soil infiltration rates in burned and unburned mountainous watersheds, Hydrological Processes, 15, 2893–2903, 2001.

McKenzie, N., Jacquier, D., and Gregory, L.: Online soil information systems–recent Australian experience, in: Digital soil mapping with limited data, pp. 283–290, Springer, 2008.

Mohanty, B., Kanwar, R. S., and Everts, C.: Comparison of saturated hydraulic conductivity measurement methods for a glacial-till soil, Soil Science Society of America Journal, 58, 672–677, 1994.

Mohsenipour, M. and Shahid, S.: Estimation OF saturated hydraulic conductivity: A Review, Malasia: Academia Edu. Recuperado de http://bit. ly/2WShxfW, 2016.

Mott, J., Bridge, B., and Arndt, W.: Soil seals in tropical tall grass pastures of northern Australia, Soil Research, 17, 483–494, 1979.

Mualem, Y.: Catalogue of the hydraulic properties of unsaturated soils, Technion Israel Institute of Technology, Technion Research & Development, 1976.

Muñoz-Carpena, R., Regalado, C. M., Álvarez-Benedi, J., and Bartoli, F.: Field evaluation of the new Philip-Dunne permeameter for measuring saturated hydraulic conductivity, Soil Science, 167, 9–24, 2002.

Naik, A. P., Ghosh, B., and Pekkat, S.: Estimating soil hydraulic properties using mini disk infiltrometer, ISH Journal of Hydraulic Engineering, 25, 62–70, 2019.

National Cooperative Soil Survey: National cooperative soil survey characterization database, United States Department of Agriculture, Natural Resoucres Conservation, Lincoln, NE, 2016.

Nemes, A.: Unsaturated soil hydraulic database of Hungary: HUNSODA, Agrokémia és Talajtan, 51, 17–26, 2002.

Nemes, A.: Databases of soil physical and hydraulic properties, Encyclopedia of agrophysics, pp. 194–199, 2011.

Nemes, A. d., Schaap, M., Leij, F., and Wösten, J.: Description of the unsaturated soil hydraulic database UNSODA version 2.0, Journal of Hydrology, 251, 151–162, 2001.

Nielsen, D., Biggar, J., and Erh, K.: “Spatial variability of field-measured soil water properties. Hilgardia, 42 (7), 215–259., 1973.

Niemeyer, R., Fremier, A. K., Heinse, R., Chávez, W., and DeClerck, F. A.: Woody vegetation increases saturated hydraulic conductivity in dry tropical Nicaragua, Vadose Zone Journal, 13, 2014.

Nyman, P., Sheridan, G. J., Smith, H. G., and Lane, P. N.: Evidence of debris flow occurrence after wildfire in upland catchments of south-east Australia, Geomorphology, 125, 383–401, 2011.

Ottoni, M. V., Ottoni Filho, T. B., Schaap, M. G., Lopes-Assad, M. L. R., and Rotunno Filho, O. C.: Hydrophysical database for Brazilian soils (HYBRAS) and pedotransfer functions for water retention, Vadose Zone Journal, 17, 2018.

Ouattara, M.: Variation of saturated hydraulic conductivity with depth for selected profiles of Tillman-Hollister soil, Ph.D. thesis, Oklahoma State University, 1977.

Päivänen, J. et al.: Hydraulic conductivity and water retention in peat soils., Suomen metsätieteellinen seura, 1973.

Parks, D. S. and Cundy, T. W.: Soil hydraulic characteristics of a small southwest Oregon watershed following high-intensity wildfires, in: In: Berg, Neil H. tech. coord. Proceedings of the Symposium on Fire and Watershed Management: October 26-28, 1988, Sacramento, California. Gen. Tech. Rep. PSW-109. Berkeley, Calif.: US Department of Agriculture, Forest Service, Pacific Southwest Forest and Range Experiment Station: 63-67, vol. 109, 1989.

Price, K., Jackson, C. R., and Parker, A. J.: Variation of surficial soil hydraulic properties across land uses in the southern Blue Ridge Mountains, North Carolina, USA, Journal of Hydrology, 383, 256–268, 2010.

Purdy, S. and Suryasasmita, V.: Comparison of hydraulic conductivity test methods for landfill clay liners, in: Advances in Unsaturated Soil, Seepage, and Environmental Geotechnics, pp. 364–372, 2006.

Quinton, W. L., Hayashi, M., and Carey, S. K.: Peat hydraulic conductivity in cold regions and its relation to pore size and geometry, Hydrological Processes: An International Journal, 22, 2829–2837, 2008.

R Core Team: R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, http://www.R-project.org/, 2013.

Rab, M.: Soil physical and hydrological properties following logging and slash burning in the Eucalyptus regnans forest of southeastern Australia, Forest Ecology and Management, 84, 159–176, 1996.

Radcliffe, D., West, L., Ware, G., and Bruce, R.: Infiltration in adjacent Cecil and Pacolet soils, Soil Science Society of America Journal, 54, 1739–1743, 1990.

Rahimy, P.: Effects of Soil Depth and Saturated Hydraulic Conductivity Spatial Variation on Runoff Simulation by the Limburg Soil Erosion Model, LISEM: A Case Study in Faucon Catchment, France, University of Twente Faculty of Geo-Information and Earth Observation (ITC), 2011.

Rahmati, M., Weihermüller, L., Vanderborght, J., Pachepsky, Y. A., Mao, L., Sadeghi, S. H., Moosavi, N., Kheirfam, H., Montzka, C., Van Looy, K., et al.: Development and analysis of the Soil Water Infiltration Global database, 2018.

Ramli, M.: Management of Groundwater Resources from Peat in Sarawak, 1999.

Ravi, S., Wang, L., Kaseke, K. F., Buynevich, I. V., and Marais, E.: Ecohydrological interactions within “fairy circles” in the Namib Desert: Revisiting the self-organization hypothesis, Journal of Geophysical Research: Biogeosciences, 122, 405–414, 2017.

Rawls, W. J., Brakensiek, D. L., and Saxtonn, K.: Estimation of soil water properties, Transactions of the ASAE, 25, 1316–1320, 1982.

Reynolds, W. and Elrick, D.: In situ measurement of field-saturated hydraulic conductivity, sorptivity, and the α-parameter using the Guelph permeameter, Soil science, 140, 292–302, 1985.

Reynolds, W., Bowman, B., Brunke, R., Drury, C., and Tan, C.: Comparison of tension infiltrometer, pressure infiltrometer, and soil core estimates of saturated hydraulic conductivity, Soil Science Society of America Journal, 64, 478–484, 2000.

Richard, F. and Lüscher, P.: Physikalische Eigenschaften von Böden der Schweiz. Lokalformen. Eidg. Anstalt für das forstliche Versuchswesen. Sonderserie., 1983/87.

Robbins, C. W.: Hydraulic conductivity and moisture retention characteristics of southern Idaho’s silt loam soils, 1977.

Romano, N. and Palladino, M.: Prediction of soil water retention using soil physical data and terrain attributes, Journal of Hydrology, 265, 56–75, 2002.

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

Rycroft, D., Williams, D., and Ingram, H.: The transmission of water through peat: I. Review, The Journal of Ecology, pp. 535–556, 1975.

Sanzeni, A., Colleselli, F., and Grazioli, D.: Specific surface and hydraulic conductivity of fine-grained soils, Journal of Geotechnical and Geoenvironmental Engineering, 139, 1828–1832, 2013.

Sayok, A., Ayob, K., Melling, L., Goh, K., Uyo, L., and Hatano, R.: Hydraulic conductivity and moisture characteristics of tropical peatland-preliminary investigation, Malaysian Society of Soil Science (MSSS), 2007.

Schwärzel, K. and Punzel, J.: Hood infiltrometer—a new type of tension infiltrometer, Soil Science Society of America Journal, 71, 1438–1447, 2007.

Scotter, D., Clothier, B., and Harper, E.: Measuring saturated hydraulic conductivity and sorptivity using twin rings, Soil Research, 20, 295–304, 1982.

Sepehrnia, N., Hajabbasi, M. A., Afyuni, M., and Lichner, L.: Extent and persistence of water repellency in two Iranian soils, Biologia, 71, 1137–1143, 2016.

Sharma, S. K., Mohanty, B. P., and Zhu, J.: Including topography and vegetation attributes for developing pedotransfer functions, Soil Science Society of America Journal, 70, 1430–1440, 2006.

Sharratt, B. S.: Water retention, bulk density, particle size, and thermal and hydraulic conductivity of arable soils in interior Alaska, 1990.

Simmons, L. A.: Soil hydraulic and physical properties as affected by logging management, Ph.D. thesis, University of Missouri–Columbia, 2014.

Singh, I., Awasthi, O., Sharma, B., More, T., Meena, S., et al.: Soil properties, root growth, water-use efficiency in brinjal (Solanum melongena) production and economics as affected by soil water conservation practices, Indian Journal of Agricultural Sciences, 81, 760, 2011.

Singh, R., Van Dam, J., and Feddes, R. A.: Water productivity analysis of irrigated crops in Sirsa district, India, Agricultural Water Management, 82, 253–278, 2006.

Smettem, K. and Ross, P.: Measurement and prediction of water movement in a field soil: The matrix-macropore dichotomy, Hydrological processes, 6, 1–10, 1992.

Sonneveld, M., Everson, T., and Veldkamp, A.: Multi-scale analysis of soil erosion dynamics in Kwazulu-Natal, South Africa, Land Degradation & Development, 16, 287–301, 2005.

Soracco, C. G., Lozano, L. A., Sarli, G. O., Gelati, P. R., and Filgueira, R. R.: Anisotropy of saturated hydraulic conductivity in a soil under conservation and no-till treatments, Soil and Tillage Research, 109, 18–22, 2010.

Southard, R. and Buol, S.: Subsoil saturated hydraulic conductivity in relation to soil properties in the North Carolina Coastal Plain, Soil Science Society of America Journal, 52, 1091–1094, 1988.

Sutejo, Y., Saggaff, A., Rahayu, W., et al.: Hydraulic conductivity and compressibility characteristics of fibrous peat, in: IOP Conference Series: Materials Science and Engineering, vol. 620, p. 012053, IOP Publishing, 2019.

Szabó, B., Szatmári, G., Takács, K., Laborczi, A., Makó, A., Rajkai, K., and Pásztor, L.: Mapping soil hydraulic properties using random-forest-based pedotransfer functions and geostatistics, Hydrology and Earth System Sciences, 23, 2615–2635, 2019.

Takahashi, H.: Studies on microclimate and hydrology of peat swamp forest in Central Kalimantan, Indonesia, in: Biodiversity and Sustainability of Tropical peatlands, Samara Publishing Limited, 1997.

Terzaghi, K.: Geotechnical investigation and testing-Laboratory testing of soil-Part 5: Incremental loading oedometer test 2, W3C XML, 1, 2006, 2004.

Tete-Mensah, I.: Evaluation of Some Physical and Chemical Properties of Soils Under two Agroforestrv Practices, Ph.D. thesis, University of Ghana, 1993.

Tian, J., Zhang, B., He, C., and Yang, L.: Variability in soil hydraulic conductivity and soil hydrological response under different land covers in the mountainous area of the Heihe River Watershed, Northwest China, Land degradation & development, 28, 1437–1449, 2017.

Tomasella, J., Hodnett, M. G., and Rossato, L.: Pedotransfer functions for the estimation of soil water retention in Brazilian soils, 2000.

Tomasella, J., Pachepsky, Y., Crestana, S., and Rawls, W.: Comparison of two techniques to develop pedotransfer functions for water retention, Soil Science Society of America Journal, 67, 1085–1092, 2003.

Tuller, M. and Or, D.: Unsaturated Hydraulic Conductivity of Structured Porous MediaA Review of Liquid Configuration–Based Models, Vadose Zone Journal, 1, 14–37, 2002.

Varela, M., Benito, E., and Keizer, J.: Influence of wildfire severity on soil physical degradation in two pine forest stands of NW Spain, Catena, 133, 342–348, 2015.

Verburg, K., Bridge, B. J., Bristow, K. L., and Keating, B. A.: Properties of selected soils in the Gooburrum–Moore Park area of Bundaberg, CSIRO Land and Water Technical Report, 9, 77, 2001.

Vereecken, H., Weynants, M., Javaux, M., Pachepsky, Y., Schaap, M., Genuchten, M. T., et al.: Using pedotransfer functions to estimate the van Genuchten–Mualem soil hydraulic properties: A review, Vadose Zone Journal, 9, 795–820, 2010.

Vereecken, H., Van Looy, K., Weynants, M., and Javaux, M.: Soil retention and conductivity curve data base sDB, link to MATLAB files, 2017.

Vereecken, H., Weihermüller, L., Assouline, S., Šimnek, J., Verhoef, A., Herbst, M., Archer, N., Mohanty, B., Montzka, C., Vanderborght, J., et al.: Infiltration from the pedon to global grid scales: An overview and outlook for land surface modeling, Vadose Zone Journal, 18, 1–53, 2019.

Vieira, B. C. and Fernandes, N. F.: Landslides in Rio de Janeiro: the role played by variations in soil hydraulic conductivity, Hydrological Processes, 18, 791–805, 2004.

Vogeler, I., Carrick, S., Cichota, R., and Lilburne, L.: Estimation of soil subsurface hydraulic conductivity based on inverse modelling and soil morphology, Journal of Hydrology, 574, 373–382, 2019.

Waddington, J. and Roulet, N.: Groundwater flow and dissolved carbon movement in a boreal peatland, Journal of Hydrology, 191, 122–138, 1997.

Wang, T., Zlotnik, V. A., Wedin, D., and Wally, K. D.: Spatial trends in saturated hydraulic conductivity of vegetated dunes in the Nebraska Sand Hills: Effects of depth and topography, Journal of Hydrology, 349, 88–97, 2008.

Weynants, M., Montanarella, L., Toth, G., Arnoldussen, A., Anaya Romero, M., Bilas, G., Borresen, T., Cornelis, W., Daroussin, J., Gonçalves, M. D. C., et al.: European HYdropedological Data Inventory (EU-HYDI), EUR Scientific and Technical Research Series, 2013.

Wösten, J., Pachepsky, Y. A., and Rawls, W.: Pedotransfer functions: bridging the gap between available basic soil data and missing soil hydraulic characteristics, Journal of hydrology, 251, 123–150, 2001.

Wösten, J. et al.: The HYPRES database of hydraulic properties of European soils., Advances in GeoEcology, pp. 135–143, 2000.

Wright, M. N. and Ziegler, A.: Ranger: a fast implementation of random forests for high dimensional data in C++ and R, arXiv preprint arXiv:1508.04409, 2015.

Yao, S., Zhang, T., Zhao, C., and Liu, X.: Saturated hydraulic conductivity of soils in the Horqin Sand Land of Inner Mongolia, northern China, Environmental monitoring and assessment, 185, 6013–6021, 2013.

Yasin, S. and Yulnafatmawita, Y.: Effects of Slope Position on Soil Physico-chemical Characteristics Under Oil Palm Plantation in Wet Tropical Area, West Sumatra Indonesia, AGRIVITA, Journal of Agricultural Science, 40, 328–337, 2018.

Yoon, S. W.: A measure of soil structure derived from water retention properties: A kullback-Leibler distance approach, Ph.D. thesis, Rutgers University-Graduate School-New Brunswick, 2009.

Youngs, E.: Infiltration measurements—a review, Hydrological processes, 5, 309–319, 1991.

Zakaria, S.: Water management in deep peat soils in Malaysia, Ph.D. thesis, Cranfield University, 1992.

Zhang, S., Xiahou, Y., Tang, H., Huang, L., Liu, X., and Wu, Q.: Study on the spatially variable saturated hydraulic conductivity and deformation behavior of accumulation reservoir landslide Based on surface nuclear magnetic resonance survey, Advances in civil engineering, 2018.

Zhao, H., Zeng, Y., Lv, S., and Su, Z.: Analysis of soil hydraulic and thermal properties for land surface modeling over the Tibetan Plateau, Earth system science data, 10, 1031, 2018.