Global prediction of soil saturated hydraulic conductivity using random forest in a Covariate-based Geo Transfer Functions (CoGTF) framework

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

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

Key Points:

- Climate, vegetation and terrain affect spatial patterns of saturated hydraulic conductivity (Ksat)
- The effect of these covariates on Ksat is quantified using remote sensing data and machine learning
- We introduce geotransfer functions to improve Ksat predictions based on pedotransfer functions

---

Corresponding author: Surya Gupta, `surya.gupta@usys.ethz.ch`

Abstract

The saturated hydraulic conductivity (Ksat) is a key soil hydraulic parameter for representing infiltration and drainage in Earth system and land surface models. For large scale applications, Ksat is often estimated from pedotransfer functions (PTFs) based on easy-to-measure soil properties like soil texture and bulk density. The reliance of PTFs on data from uniform arable lands and omission of soil structure limits the applicability of texture-based predictions of Ksat in vegetated lands. A method to harness technological advances in machine learning and availability of remotely sensed surrogate information to derive a new global Ksat map at 1–km resolution using terrain, climate, vegetation, and soil covariates is proposed. For model training and testing, global compilation of 6,814 georeferenced Ksat measurements from the literature across the globe were used. The accuracy assessment results based on model cross-validations with re-fitting show a concordance correlation coefficient of 0.79 and root mean square error of 0.72 (in $\log_{10}$Ksat given in cm/day). The generated maps of Ksat represent spatial patterns of the vegetation-induced soil structure formation and clay mineralogy, more distinctly than previous global maps of Ksat such as computed with Rosetta 3 pedotransfer function. The validation of the model indicates that Ksat could be more accurately modeled using covariate-based Geo Transfer Functions (CoGTFs) that harness spatially distributed surface and climate attributes, compared to pedotransfer functions that rely only on soil information.

Plain Language Summary

The soil saturated hydraulic conductivity Ksat defines how fast water can infiltrate and percolate through the soil. To model water flow at large scale, accurate maps of Ksat are needed. Usually, Ksat is not measured directly but deduced from well known basic soil properties (soil texture, packing density). But these estimates neglect the influence of vegetation and climate on formation of soil structures that control Ksat. To improve predictions of Ksat, we use a new spatially referenced Ksat data collection and apply Machine Learning to find correlations between Ksat and other properties (soil information, terrain, climate and vegetation). These correlations are then implemented at global scale using maps of all relevant properties (so called *'covariates'* that were measured by remote sensing). We called this new approach to predictive soil mapping the *"Covariate-based Geotransfer functions"* (CoGTF) to highlight the difference to other maps that neglect spatial correlation with soil formatting properties and are based only on soil information (so called *"pedotransfer functions"* or PTFs). We show that the new maps based on CoGTF perform better than approaches based on PTFs.

1 Introduction

The description of water, energy, and carbon fluxes between the land surface and the atmosphere relies heavily on the availability of soil hydraulic data (Gutmann & Small, 2007; Fashi et al., 2016; Montzka et al., 2017). A prominent soil hydraulic property is the soil saturated hydraulic conductivity (Ksat) that affects the partitioning of rainfall between runoff and infiltration (Zimmermann et al., 2013), and plays a critical role in a variety of hydrological and climatological applications (Gutmann & Small, 2007; Or, 2019; Fatichi et al., 2020). At global scale, maps of soil hydraulic properties at ever increasing resolution are required for building Land Surface Models (LSMs) (Montzka et al., 2017).

For large scale applications (regional and global), soil hydraulic parameters are often estimated from easy-to-measure soil properties (e.g., texture, organic content, bulk density) by means of pedotransfer functions (PTFs) (Bouma, 1989; Santra & Das, 2008). PTFs are usually developed for specific geographic regions thus only representing local conditions of soil forming processes (e.g. Tomasella & Hodnett, 1998; Wösten et al., 1999; Nemes et al., 2005; Saxton & Rawls, 2006; Jorda et al., 2015; Khlosi et al., 2016). This hinders their transferability across large geographical regions (Vereecken et al., 2016). In addition, PTFs generally ignore soil structure and pedogenic information and rely heavily on soil textural information (Fatichi et al., 2020), limiting their applicability in soils characterized by aggregation and formation of biopores. Moreover, PTFs are generally defined as a function of clay content, without consideration of the effect of different clay minerals on soil hydraulic properties (Hodnett & Tomasella, 2002). Dai et al. (2019) have recently produced 1–km resolution global maps of soil hydraulic properties (and thermal soil conductivity) using the median values of multiple PTFs to estimate Ksat. Likewise, Y. Zhang and Schaap (2017) have developed a global map of van Genuchten parameters and Ksat based on the Rosetta 3 PTF (an extension of Schaap et al., 2001), making use of three data sets from North America and Europe (i.e., Rawls et al., 1982; Ahuja et al., 1989) and UNSODA (Unsaturated Soil Hydraulic Database) as described in Leij et al. (1996) and Nemes et al. (2001) and employing Artificial Neural Network and bootstrap sampling.

Maps produced by Dai et al. (2019) and Y. Zhang and Schaap (2017) are limited by the small number and unevenly distributed Ksat measurements ($N = 1306$) used for model training and large spatial gaps i.e. missing training points in tropics. Moreover, the training points used to produce estimates of Ksat were usually dominated by particular land use and land cover, mainly collected in arable land. Furthermore, only a limited set of basic soil variables (i.e., bulk density and texture) was employed in the derivation of the Rosetta 3 map (Y. Zhang & Schaap, 2019), while several studies have shown that also other soil properties such as organic carbon, soil depth and pH may increase accuracy of PTFs (Wösten et al., 1999; Mayr & Jarvis, 1999; Tóth et al., 2015). The availability of highly resolved remote sensing (RS) and landscape covariates offer new opportunities for injecting new and local information into the modeling of Ksat. Examples of the potential usefulness of such covariates are reported by Obi et al. (2014) that developed a PTF using terrain attributes for many soil hydraulic properties; Sharma et al. (2006) combined PTFs with vegetation and topography indices; Jana and Mohanty (2011) showed that the introduction of topographic attributes (i.e., Digital Elevation Model, DEM) and information on vegetation (i.e., Leaf Area Index, LAI) along with *in situ* soil basic properties could improve predictions of soil hydraulic properties.

Many of the recent PTFs use Machine Learning (ML) algorithms to quantify the relations between hydraulic properties and various covariates (Schaap et al., 2001; Jana & Mohanty, 2011; Araya & Ghezzehei, 2019). In this paper, we hypothesize that Ksat predictions could be improved using a combination of soil variables and remote sensing covariate layers integrated by using machine learning (ML) framework. We profit from the advancement in remote sensing techniques (providing spatial information on different ecological parameters with unprecedented resolution) to improve the predictions for soil hydraulic parameters and bridge the gap between site-specific soil properties and landscape variability. We merge concepts of predictive soil mapping with a large data set of Ksat measurements and local information (soil, vegetation, climate) into covariate-based *"Geo Transfer Functions"* (CoGTFs) to generate global estimates of Ksat values (to highlight the impact of Geo-referencing soil properties and RS-covariates we use the term GTF and not PTF). We compare mapping accuracy using global and local/regional assessment including visual interpretation of produced spatial predictions. We show how this method (providing novel covariate-based maps of Ksat) could be used to overcome some of the limitations of traditional PTFs.

Our specific objectives are:

1. to improve accuracy and spatial detail of global Ksat maps by harnessing the state-of-the-art global remote sensing data products at 1 km spatial resolution,
2. to generate global maps of Ksat at different soil depths (0, 30, 60 and 100 cm),
3. to identify the key environmental variables explaining the spatial distribution of Ksat.

We first describe the model training for Ksat mapping using a random forest ML algorithm, and then compare the results against maps generated with Rosetta 3 and the map shown in Dai et al. (2019). Note that for a detailed comparison of global maps, we focus on Rosetta 3 because the map in Dai et al. (2019) is heavily influenced by the application of a different soil textural map (see Supplementary Information file). Then, we validated the CoGTF map, Rosetta 3 map and the map of Dai et al. (2019) with independent dataset. We finally show the importance of using RS covariates to capture spatial patterns and improve the accuracy of soil hydraulic properties.

2 MATERIALS AND METHODS

2.1 Covariate-based Geo Transfer Functions (CoGTF) framework

We propose here an integrated Predictive Soil Modeling (PSM) framework where soil variables are combined with RS-based covariates using random forest method (Figure 1). We refer to this approach as the *"Covariate-based Geo Transfer Functions"* (CoGTF) framework and envisage it as a combination of traditional PTF approach and purely data science approach where RS-based covariates are used to map patterns in soil properties. The CoGTF framework follows six principal steps:

1. Prepare georeferenced dataset of response variable (Ksat),
2. Overlay training points and covariates (including predictions of basic soil properties), and produce a regression matrix,
3. Optimize the hyper-parameters in the random forest approach (`mtry`),
4. Fit the random forest model,
5. Evaluate the performance of the Ksat model,
6. Produce spatial predictions of Ksat.

A central hypothesis in this study is that spatial and climatic covariates could be harnessed to improve the global mapping of Ksat (Jana & Mohanty, 2011). The basis for such hypothesis is the dominant role of climate, topography, and vegetation in soil formation and thus in shaping local hydraulic transport properties. For each location with Ksat measurement, the values of the remote sensing covariates were extracted together with modeled soil information from OpenLandMap.org. We implement the spatial predictions and the creation of Ksat maps in the R environment (R Core Team, 2013) for statistical computing and provide code examples via the `https://github.com/ETHZ-repositories/Ksat_mapping_2020/`.

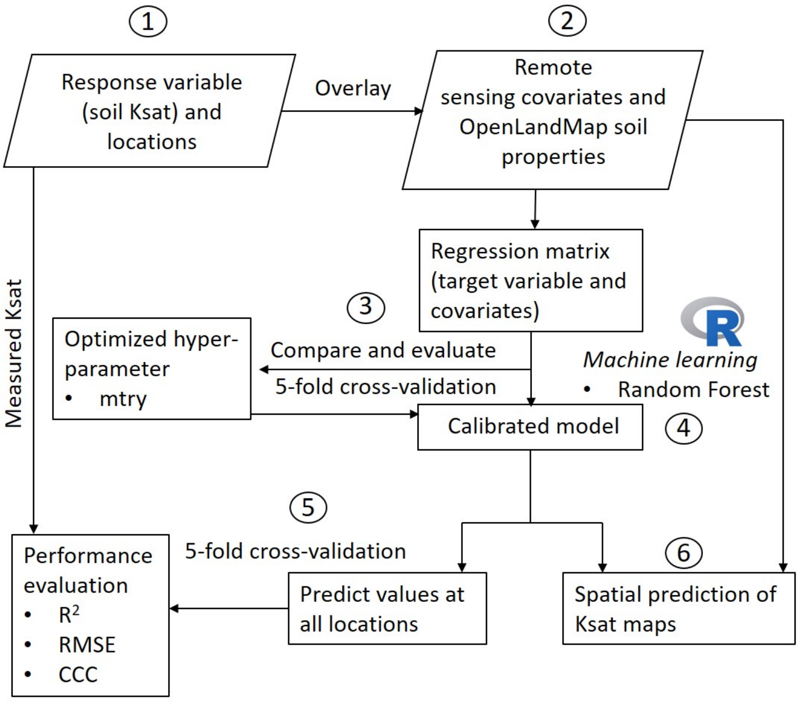

①
②
Response variable (soil Ksat) and locations
Overlay
Remote sensing covariates and OpenLandMap soil properties
Regression matrix (target variable and covariates)
③
Optimized hyper-parameter
• mtry
Compare and evaluate
5-fold cross-validation
Machine learning
• Random Forest
Calibrated model
④
Measured Ksat
⑤
5-fold cross-validation
⑥
Performance evaluation
• R2
• RMSE
• CCC
Predict values at all locations
Spatial prediction of Ksat maps

Figure 1. Computational workflow used to generate the soil Ksat map. See text for more details about the specific steps.

After extracting all covariates, a regression matrix was formed, and the best hyperparameter (`mtry`) was computed by five-fold cross-validation, using the R packages *'caret'* version 6.0-85 (Kuhn, 2012) and *'ranger'* version 0.12.1 (Wright & Ziegler, 2015). Then, log-transformed (log_{10}) Ksat was modeled as a function of depth using random forest (RF) algorithm.

2.2 Training point data

Our first task was to enlarge the Ksat measurement database beyond the ≈1,300 values used to train Rosetta 3 by compiling available and georeferenced Ksat values from the literature. The Ksat values were log-transformed (log_{10}Ksat) and cm/day was selected as a standardized unit. A detailed description of the data collection and processing is provided in Gupta et al. (2020). We managed to compile a total of 13,267 samples coming from 1,910 sites across the globe. Most training data are from the USA, followed by Europe, Asia, South America, Africa, and Australia as shown in Figure 2. The collected Ksat database (SoilKsatDB) includes both field ($N = 4,460$) and lab ($N = 8,807$) measurements.

To limit the over-representation of Florida (mainly arable land not representative of soils with natural vegetation), we randomly selected approximately only 1% of the 6,532 Florida samples, so that a total of 6,814 Ksat values were finally used for Ksat mapping. This resulted in geographical balance between other national data sets (the effect of this selection of Florida data is discussed in Supplementary information file).

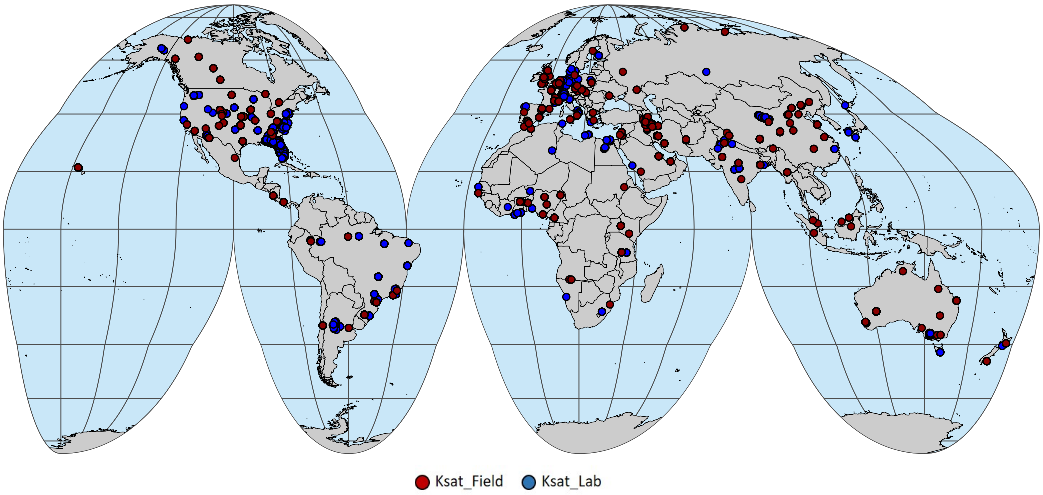


Figure 2. Spatial distribution of measured Ksat values (6,814 samples in total) used to produce the global Ksat map. Colors refer to laboratory (red) and field (blue) measurements. The map is presented in the Goode equal-area homolosine projection. For more details and access to the Ksat data see Gupta et al. (2020).

2.3 Soil and environmental covariates

As environmental and soil covariate layers for Ksat modeling at global scale, we used global maps of soil properties (sand, clay, and bulk density) and other 24 RS-based covariates available from `https://openlandmap.org/`. These were selected to represent ecological conditions essential in soil-forming processes according to Jenny (1994). The covariates can be divided into five groups:

1. *Climate-based covariates*, including mean annual precipitation, temperature, temperature seasonality, maximum temperature of warmest month, minimum temperature of coldest month, precipitation of wettest month, precipitation of driest month (Chelsa products, Karger et al., 2017), cloud fraction (Wilson & Jetz, 2016), diffuse irradiation, direct irradiation, annual land surface temperature, monthly precipitation and its standard deviation (Brocca et al., 2019).
2. *Digital terrain model (DTM)-based covariates* (Yamazaki et al., 2017), including landscape metrics (such as slope, aspect, topographic wetness index) derived from SAGA GIS (Conrad et al., 2015) and landform classification and lithological maps.
3. *Surface reflectance-based covariates*, including surface reflectance from Landsat and MODIS dataset for different wavelength bands (Hansen et al., 2013), snow probability (Buchhorn et al., 2017) and regularly flooded wetlands (Tootchi et al., 2019).

4. *Vegetation-based covariate*, represented by the annual fraction of absorbed photosynthetically active radiation (FAPAR), averaged over the 2014-2019 period.
5. *Basic soil properties*, comprising sand, clay content and bulk density for different soil depths (matching the sampling depth of Ksat), which were obtained from OpenLandMap (Hengl et al., 2017). Soil depth is used as a covariate to model the change of Ksat with depth (the methodology to use depth as a covariate is described in Hengl & MacMillan, 2019).

A detailed list and description of all the covariates is provided in Table S1 in the Supplementary Information (SI). All covariate maps were resampled to the standard grid at a spatial resolution of 1 km covering latitudes between -62.0 and 87.37. We did not map Antarctica as this continent is dominantly covered with permanent ice and lacks training points.

2.4 Evaluating the performance of Ksat predictive models

The model-fitting results were evaluated using out of bag (OOB) error reported by the ranger package by default. A bootstrap sampling is used to construct each tree in the random forest and different bootstrap samples are used for each tree containing approximately 2/3 of the total observations. The samples not used in the bootstrap samples are called out-of-bag (OOB) samples (sub-dataset) (Peters et al., 2007; Rad et al., 2014). The relative importance of the covariates was assessed by the increase in node purity. It is calculated using gini criterion from all the splits (in our case 200 splits) in the forest based on a particular variable (Breiman, 2001; Rodrigues & de la Riva, 2014).

The performance of the Ksat model was evaluated using 5-fold cross-validation. This means that models were refitted 5 times using 80% of the data and the predictions for remaining 20% estimated using these models were compared with observations. The process was repeated three times to produce stable results. The final results are shown using hexbin plot with the LOWESS (Locally Weighted Scatterplot Smoothing) line to present the conditional bias of the Ksat values. The accuracy of the cross-validation predictions was evaluated using bias (mean error), root mean square error (RMSE), coefficient of determination (R^2) and concordance correlation coefficient (CCC).

Bias and RMSE are defined by:

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

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

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

R^2 is defined as:

$$R^2 = \left[1 - \frac{SSE}{SST}\right] \% \tag{3}$$

where SSE is the sum of squared errors between the cross-validation predictions $\widehat{y}$ and the measurements y, and SST is the total sum of squares (proportional to variance of measurements). A coefficient of determination equal to 1 indicates that variance of the prediction errors is equal to zero but the bias may differ from zero.

In addition, Concordance Correlation Coefficient (CCC) (as measure of the agreement between observed and predicted Ksat values) of cross validation (CV) (Lawrence & Lin, 1989) is given by:

$$CCC = \frac{2 \cdot \rho \cdot \sigma_{\hat{y}} \cdot \sigma_y}{\sigma_{\hat{y}}^2 + \sigma_y^2 + (\mu_{\hat{y}} - \mu_y)^2} \tag{4}$$

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

2.5 Comparision of accuracy of Ksat maps: CoGTF, Rosetta 3 and the map of Dai et al. (2019)

The accuracy of the predictions of Ksat by the three approaches was evaluated with a subset of the Ksat database that was selected in the following way: First, the surface of the Earth was divided into blocks of 5 degrees as shown in Figure S1 in the SI. For fair comparison, Ksat measurements in blocks in North America or Europe were dropped because Rosetta 3 was mostly calibrated with data from these regions (2525 Ksat values were outside of these regions). Then we randomly selected blocks until about 20% of the remaining Ksat measurements had been chosen. These 508 Ksat measurements formed the test set for which predictions were extracted from the Rosetta 3 and the Dai et al. (2019) maps. CoGTF predictions of Ksat were computed for these 508 test observations. The accuracy of the predictions by the three approaches was then evaluated with the same criteria as used for cross-validation.

3 Results

3.1 Model fitting and accuracy of modeled Ksat values

The CoGTF model fitted the logarithms of the Ksat measurements reasonably well (out-of-bag RMSE = 0.73 (log_{10}Ksat in cm/day) and R^2 = 0.66). Figure 3 shows the list of most important covariates for Ksat modelling. The x-axis displays the average increase in node purity. The higher the value, the more important is a covariate. Figure 3 shows that sand content was found the most important covariate followed by elevation (important for soil formation and water flow), clay content, and bulk density. Climate covariates are dominating after the fifth covariate.

The results of the 5-fold cross-validation are presented in Figure 4a using hexbin density plots. For predictions of Ksat greater than equal to 10 cm/day the line of LOWESS falls onto the 1:1-line, hence the predictions were conditionally unbiased here. A slight positive conditional bias is visible for predictions less than 10 cm/day where the LOWESS line is below the 1:1 line. CoGTF tended to overestimate small Ksat values, but this bias remains small. Hence, RF predictions were both marginally and conditionally approximately unbiased. Cross-validation re-

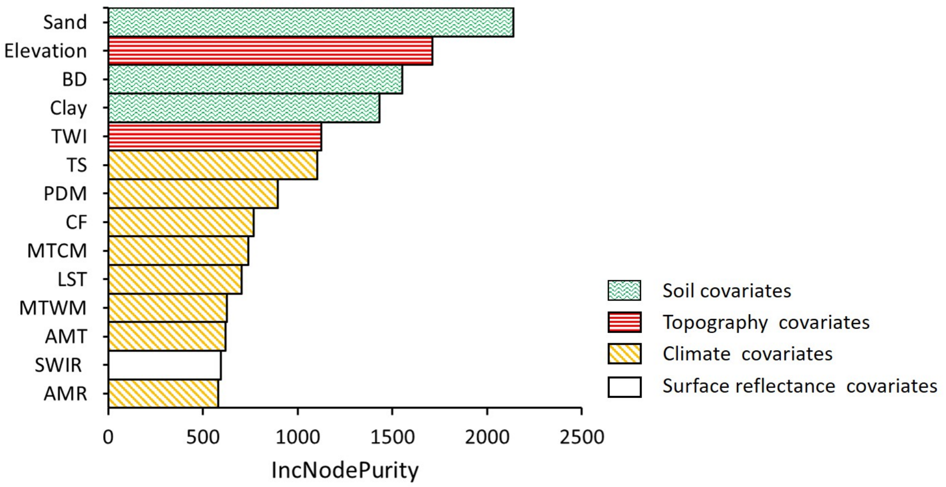


Figure 3. Importance of the covariates for modeling Ksat by a random forest model. The x-axis displays the average increase in node purity (the larger the value, the more important is a covariate). The 14 most important covariates are shown here: sand content, bulk density (BD), and clay content belong to soil covariates. Elevation and topographic wetness index (TWI) are topography covariates. Temperature seasonality (TS), precipitation of driest month (PDM), cloud fraction (CF), minimum temperature of coldest month (MTCM), annual average land surface temperature (LST), maximum temperature of warmest month (MTWM), mean annual temperature (AMT), and mean annual rainfall (AMR) belong to climate category. Shortwave infrared (SWIR) Landsat-7 band is from the surface reflectance group.

sults show a reasonable overall model accuracy, with R^2, CCC, and RMSE and bias equal to 0.66, 0.79, 0.72, and 0.0039 (log_{10} of Ksat in cm/day for RMSE and bias), respectively. The observations were also correlated with Rosetta 3 Ksat map (for this comparison, a total of 5,255 samples from shallow soil depth were selected out of 6,814 to compare with Rosetta 3 map for top 15 cm) as shown in Figure 4b. RMSE and CCC was observed 1.23 and 0.12 (log_{10} of Ksat in cm/day for RMSE), respectively.

3.2 Global map of Ksat

Global Ksat maps were produced for four soil depths (0, 30, 60, and 100 cm). Figure 5a shows the CoGTF map of Ksat at 0 cm soil depth, while results for other soil depths are provided in Figure S2 (SI). Ksat values in the top layer (0 cm depth) vary between 0.05 to 31,600 cm/day. High Ksat values were predicted for the equatorial belt and for parts of Russia and Canada, while low Ksat values were produced in East America, Europe and parts of Asia (mainly India and North-East part of China). In general, Ksat value decreased with depth, with the most significant reduction observed in North America, South America, China, India, and Russia (see Figures S2-S3 in the SI). Figure 6 compares the probability distribution of the global Ksat map values with the distribution of measured and fitted Ksat values for the 6,814 Ksat samples. Results show a more peaked distribution of global Ksat map compared to the measured and fitted Ksat at the sampling locations. Both measured (red) and fitted log_{10} Ksat showed the same mean values of 1.64 with standard deviations 1.25 and 1.01, respectively, whereas the mean and standard deviation of global map were observed 1.99 and 0.30 respectively.

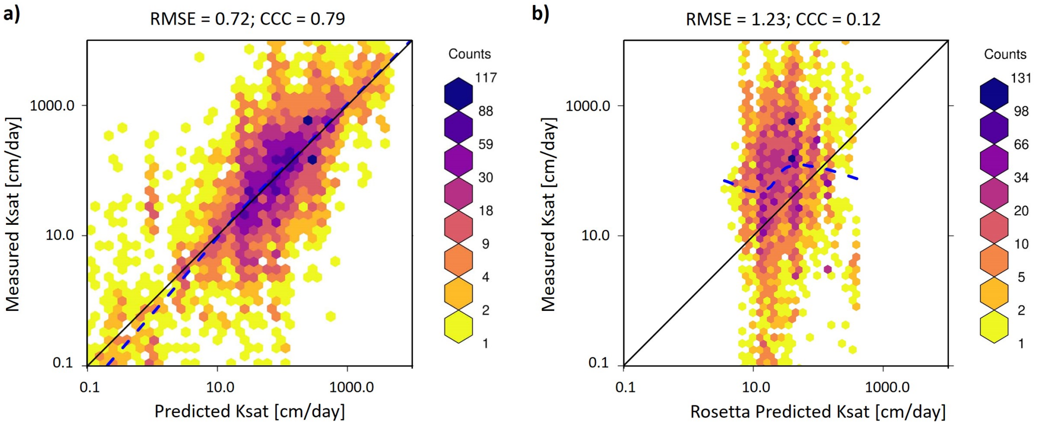


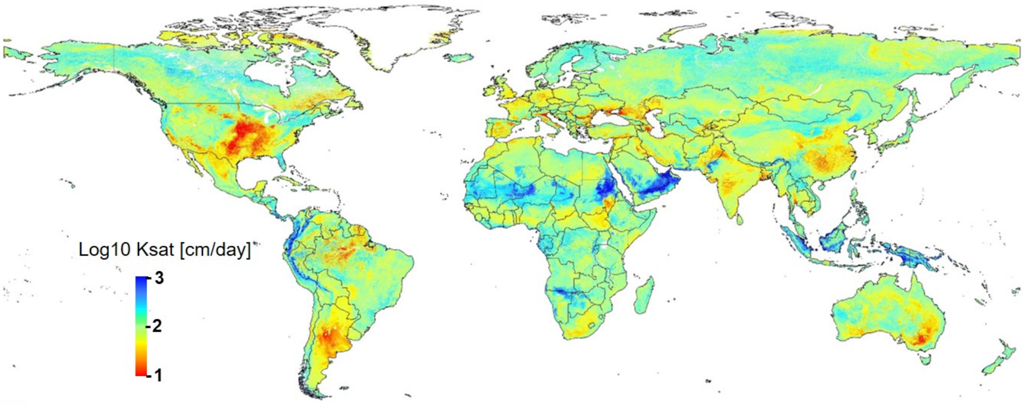
Figure 4. Accuracy plots based on cross-validation: (a) correlation between observations and cross-validation predictions of log_{10} Ksat based on CoGTF model, (b) correlation between observations (0-30 cm soil depth) and Rosetta 3 predicted values from 0-15 cm map. The color codes the number of observations in each hexagonal pixel. The solid black line is 1:1 line and the blue dashed line is LOWESS curve (locally weighted scatterplot smoothing). The model accuracy of CoGTF was assessed using CCC (0.79) and RMSE (0.72). The RMSE and CCC between observations and Rosetta 3 predicted Ksat values were observed 1.23 and 0.12, respectively. The unit of RMSE is log_{10} of Ksat in cm/day.

3.3 Comparison with Rosetta 3 global Ksat map

The CoGTF Ksat map is compared with the Rosetta 3 map (Y. Zhang & Schaap, 2019) in Figure 5. Note that there are different models of Rosetta 3 according to the soil information used to build the neural network: H1w (information on soil textural class), H2w (sand, silt, and clay percentage), H3w (sand, silt, and clay percentage plus bulk density), H4w (same information as H3w plus water content at 330 cm suction), and H5w (same as H3w plus water content at 330 cm and at 15,000 cm) (X. Zhang et al., 2019). As standard model H3w is often chosen (see map in Y. Zhang & Schaap, 2019) because information on water content at 330 cm and 15,000 cm is sparse at global scale compared to bulk density and soil texture information. For comparison with CoGTF, we chose H3w model as well.

The main differences between the CoGTF map and Rosetta 3 are the low Ksat values predicted by Rosetta 3 for tropical regions and the abrupt change in Rosetta 3 predictions in high latitude regions of Canada and Russia as a consequence of the strong sensitivity of Rosetta 3 predictions on bulk density. In general, lower Ksat values were observed in the Rosetta 3 map compared to the CoGTF map for most regions worldwide except the northern regions (Canada and Russia), while regions with coarser soils such as Sahara and middle East showed higher Ksat values in Rosetta 3. The lower values of Ksat in Rosetta 3 than the in CoGTF map is evident in Figure 6a. Medians of the common logarithm of Ksat (unit cm/day) were equal to 1.62 and 2.00, respectively (Figure 6b).

a) *CoGTF map (0 cm)*



b) *Rosetta 3 map (0 cm)*

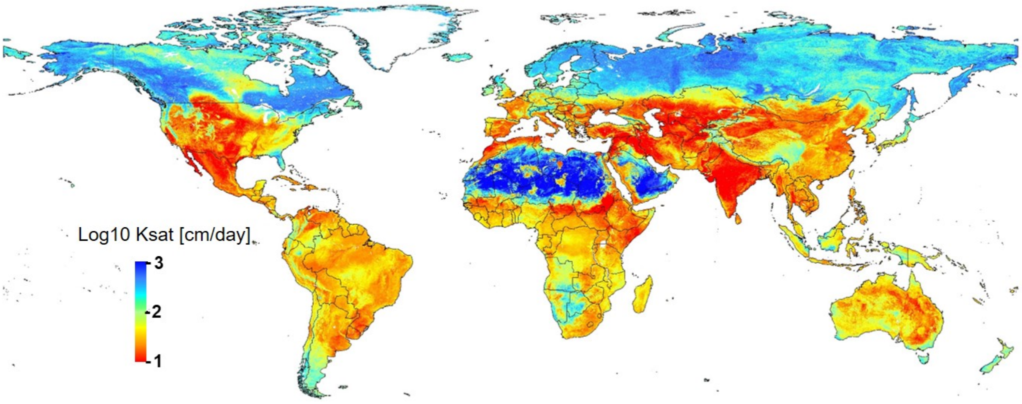


Figure 5. Visual comparison between (a) CoGTF Ksat map, and (b) map based on Rosetta 3 PTF. Ksat values predicted by Rosetta 3 were higher for sandy soils (Sahara) and in northern regions with smaller bulk density. The scale of the maps was truncated at minimum and maximum values of 10 and 1000 cm/day to show the significant variations in the maps

3.4 Validation of global Ksat maps

Table 1 shows the results of the comparison of the accuracy of Ksat predictions for the CoGTF, Rosetta 3 and Dai et al. (2019) maps (see Figures S7 and S8 for the map of Dai et al., 2019, with CoGTF map). A total of 372 Ksat samples out of the validation dataset with 508 samples (we selected samples with soil depth 0-30 cm) were compared with measured Ksat values and RMSE values of 1.02, 1.29, and 1.15 were computed (log_{10} of Ksat in cm/day) for the CoGTF map, Rosetta 3, and Dai et al. (2019) map, respectively. The RMSE illustrates that CoGTF map showed better performance than the other maps. However, RMSE of 1 also shows that the precision is limited for CoGTF as well. The better performance of CoGTF is manifested in the much lower bias compared to the two other models.

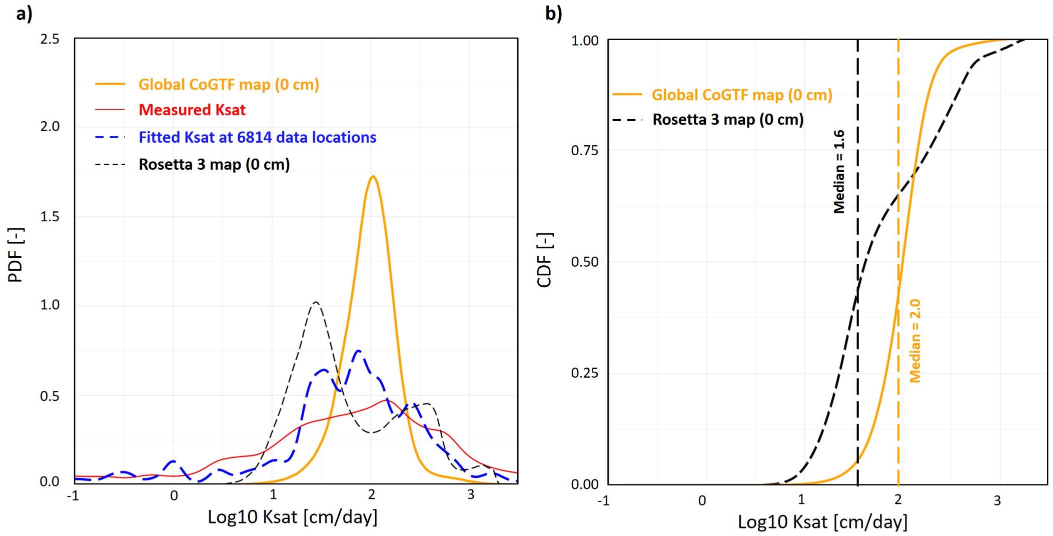


Figure 6. Difference in probability density functions (PDF): (a) between global CoGTF map (yellow) and Rosetta 3 (black) Ksat values at 0 cm depth, measured (red) and fitted (blue) Ksat values at the sampling sites, (b) cumulative distribution functions for Rosetta 3 map (black) and CoGTF map (yellow) for soil depth 0 cm.

Table 1. Root mean square error (RMSE) and bias of predictions of log_{10}(Ksat) (units cm/day) for test data. A total of 372 Ksat sample points were selected to investigate the accuracy of Ksat predictions (0-30 cm soil depth were used). The negative signs in bias demonstrate that all three models underestimated Ksat values. The range shows the minimum and maximum values of 372 samples.

Models	Samples used	RMSE	bias	Extracted points range
CoGTF (0 -15 cm)	372	1.02	-0.19	0.85-2.60
Rosetta 3 (0 -15 cm)	372	1.29	-0.75	0.83-2.64
Dai et al. (2019) (0 -15 cm)	372	1.15	-0.51	0.68-2.30

4 Discussion

4.1 Characteristics of the CoGTF global Ksat maps

In this paper we have produced global estimates of Ksat by linking terrain, climate, vegetation and soil spatial covariates to measured Ksat values, thus injecting local information usually ignored by traditional PTFs. We refer to this approach as the Covariate-based Geo Transfer Functions (CoGTF) framework. The newly developed global CoGTF map of Ksat (Figure 5) shows high values in the Northern part of South America, the central part of Africa and Southeast Asia (mainly Indonesia, Malaysia, Myanmar (Burma), Philippines, Singapore, and Thailand), most likely due to high rainfall, temperature, and vegetation. Our results shows (Figure 3) that rainfall, temperature and their variation are the most important climate covariates for the Ksat mapping (Shoji et al., 2006). This indicates that these climatic factors not only act as catalyst in soil chemical reactions but also determine the type and biomass of vegetation that is important for soil structure formation. This impact of vegetation on soil Ksat is in line with the research by Niemeyer et al. (2014) who compared the leaf area index with Ksat and observed that high leaf area index increases the Ksat (with R-square = 0.33).

The central part of India, eastern part of Australia, and parts of China showed low Ksat values due to the presence of high clay content that reduces the soil permeability (see as well discussion on role of clay mineral type in section 4.2). The west part of North America, middle east countries (Tibet, Iran, Turkey), and northern parts of Algeria have low Ksat values that may be related to high elevation, low rainfall, less vegetation and thus less structure formation processes. Many studies have recognized the indirect influence of elevation on soil proprieties (Leij et al., 2004; Carter & Ciolkosz, 1991). Similarly, different land-use (forest or pasture) directly impact Ksat. Chandler et al. (2018) showed that forests had larger soil hydraulic conductivity than pastures.

Likewise, high values of Ksat up to around 100 to 300 cm/day are observed in desert regions such as Thar desert in India, northern and southern Africa, and central Australia, where dominating fractions of sand cause high water permeability. Similarly, Colombia and Peru showed high Ksat values due to high organic carbon content (Allison, 1973). Furthermore, high Ksat values were observed in parts of Brazil that strongly decreased with depth. Similar results were reported by Belk et al. (2007). They conducted a study in the tropical forest of Brazil and measured the Ksat at various depths for different sites. The authors found that Ksat values at surface were mainly between 100 to 1000 cm/day and decreased with depth.

4.2 Effect of clay type — active and inactive clay minerals

Pedotransfer functions like Rosetta 3 and the ensemble of PTFs used in (Dai et al., 2019) to estimate soil hydraulic properties based on clay fraction and do not take into account the large differences in microstructure and hydration of different clay minerals. The remarkable spatial segregation in climatic regions of different clay minerals (see Ito & Wagai, 2017) and the different hydraulic properties of the clay minerals indicate that PTFs built for temperate regions with swelling clays cannot be applied for tropical regions with non-swelling clays (see Ottoni et al., 2018). In tropical soils, dense vegetation, and non-swelling (*'inactive'*) kaolinite clay minerals result in higher conductivities (Hodnett & Tomasella, 2002) in contrast to PTFs that are trained with data from temperate soils with swelling (more 'active') clays. This is further discussed for estimates relevant to Brazil shown in Figure 7.

In Figure 7, the CoGTF and Rosetta 3 Ksat maps are shown together with six covariates and clay mineral map. The Ksat values predicted with CoGTF are one order of magnitude higher than based on Rosetta 3. The difference stems from the dominant role of soil texture for Rosetta 3 as illustrated with a black polygon in Figure 7: the polygon marks a region of high sand content and low clay content that is manifested in relatively high values of Ksat for Rosetta 3, with values typical for temperate regions. For CoGTF, the conductivity in this *'sand band'* is relatively low because other covariates and processes are more important. These lower values coincidence with low elevation. The important role of elevation in CoGTF is also manifested in the high Ksat values in the mountainous region in the south and the low Ksat values in the Amazon region. Another reason for the lack of correlation between Ksat and texture for CoGTF in Brazil is the inactive clay mineral type (kaolinite) that does not limit Ksat the same way as in case of more ac-

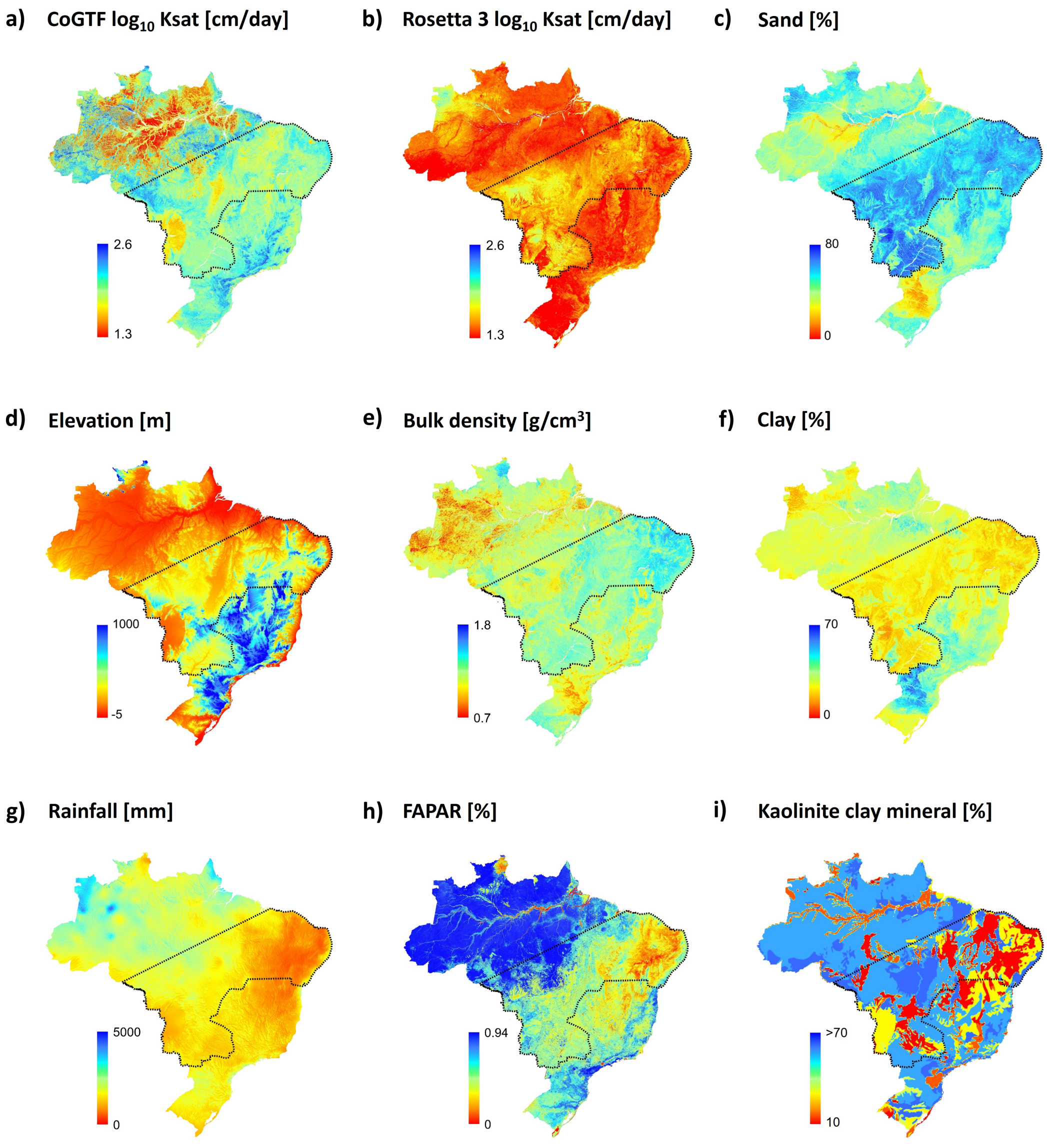


Figure 7. Predicted Ksat values for Brazil (a), spatial patterns of the Rosetta 3 Ksat map in $\log_{10}$ cm/day (b) and of the first four most important covariates (c-f, see Figure 3): sand fraction (%), elevation (meters above sea level), bulk density (g/cm^3) and clay fraction (%). Other covariates that are related to soil formation to link with Ksat are shown as well (g-i): mean annual rainfall (mm), Copernicus fraction of absorbed photosynthetically active radiation (FAPAR, values in %) and kaolinite (in %) clay mineral. The region with black polygon marks a region with high sand and low clay content that is expressed in Rosetta 3 as band of relatively high Ksat values. In contrast to Rosetta 3, CoGTF is not dominated by soil texture but takes into account covariates that are important for soil formation (here mainly the elevation).

tive clay in temperate regions. Precipitation and temperature are the main reasons for the strong chemical weathering of the rock and the formation of the non-swelling, kaolinite clay minerals (Montes et al., 2002). It is evident in Figure 7(g to i) that in the region with low rainfall and vegetation kaolinite percentage is lower than other regions with high rainfall and vegetation.

In contrast to Brazil, India is a region with active (swelling) clay minerals. In contrast to the inactive kaolinite in Brazil, for the active clays in India, low Ksat values can be expected. Figure S4a and S4i show the correlation between low Ksat values and high contents of smectite clay mineral. The low values of Ksat in central India directly relate to high clay content, low vegetation biomass and low mean annual rainfall (see Figure S4 for covariates in SI). Figure S4b illustrates the Ksat values from Rosetta 3 for India. The patterns of high active clay fraction in India is not captured by Rosetta 3 model. This might be the effect of considering only soil basic properties or using samples from only temperate region.

4.3 Effects of information clustering — The Florida database example

Out of 13,267 Ksat values, only 6,814 values were used for the Ksat mapping to avoid a distortion of the Ksat predictions by the many data from Florida. The dataset contained 6,532 Ksat values from Florida but we used only 1% of these points for mapping. Figure S5a and S5b compares the map computed with all 13,267 Ksat measurements with the map trained on 6,814 measurements. The difference between these maps (Figure S5c) showed a large impact on the sandy regions such as Sahara and center part of Africa and middle east with significantly higher Ksat values when all Florida points are included in the fitting. A similar effect was observed in parts of South America and Australia. On the other hand, the south of Africa and the higher northern latitudes showed higher Ksat values when only 1% of the Florida data was used.

4.4 Improved model performance using remote sensing covariates

As we described above, the RS (vegetation, topography, climatic) covariates could be used to harness the heterogeneity produced by these environmental variables as these factors shape clay activity and soil-forming processes that control saturated hydraulic conductivity (Ottoni et al., 2018; Hao et al., 2019). To investigate this effect of RS covariates in the predictions, we fitted the RF model only with soil properties or remote sensing covariates. The maps are shown in Figure S6a and S6b in the SI.

Table 2. Root mean square error (RMSE) and coefficient of determination (R^2) for different models.

Models	RMSE	R^2	Total covariates	Best mtry
CoGTF	0.72	0.66	28	6
Only soil covariates	0.75	0.63	4	2
Only RS covariates	0.73	0.65	24	16

Table 2 shows the RMSE and R^2 using different models where we used only soil covariates, only remote sensing covariates, and the CoGTF model. Remote sensing (RMSE = 0.73; R^2 = 0.65) predicted the Ksat better than only soil covariates (RMSE = 0.75; R^2 = 0.63). Similarly, the CoGTF model showed lower RMSE (0.72) and higher R^2 (0.66) than only RS covariate. Hence, consideration of RS covariates in predicting hydraulic properties could increase the accuracy of the predictions of soil hydraulic properties compared to a model that is based only on soil information.

4.5 Usage of the global CoGTF Ksat maps and future developments

We observed that RMSE in the model validation for the CoGTF map was better than the other maps. However, RMSE with 1 (log_{10} Ksat in units of cm/day) also shows that the precision of even the CoGTF map is not so good. On the other hand, the bias for CoGTF map was much better than for the other maps. Although, the predictions are not so accurate, it shows the one step ahead in terms of improvement in the predictions using distributed Ksat dataset and consideration of RS covariates.

The global CoGTF maps can be used to extract the information of Ksat at different depths for local, regional, and global scale studies. On the local scale, these maps can be helpful in agronomic processes such as soil interpretation, water-plant relationships, and assessing soil suitability for agriculture. For regional and global scale, the maps could provide unique values to each pixel in watershed scale and Earth surface models and would enhance the heterogeneity and accuracy in the area. The maps could also be useful for the soil water management policies as guideline to show where soil reclamation is required to reduce and enhance the hydraulic conductivity.

The actual CoGTF map has a resolution of 1 km. This resolution may be improved in the near future considering various initiatives to estimate soil and RS information with higher resolution. But independent of improved resolutions, subgrid information on Ksat may be required for a catchment when specific information on soil texture or vegetation type is available. For such applications, we are actually developing a parametric model of CoGTF so that Ksat can be estimated as a linear combination of most important covariates.

5 Conclusions

Soil saturated hydraulic conductivity is an important soil property for the parameterization of Earth system and land surface models. The major limitations of currently available maps are that (1) they are developed using a limited number of Ksat measurements mainly from temperate regions, (2) they are derived only from basic soil properties thus ignoring the effect of biologically-induced soil structure as well as clay mineralogy, and (3) they are not benefiting from the wealth of local remote sensing (RS) covariates. Therefore, we proposed a new global map of Ksat obtained by linking the measured Ksat values (6,814 samples) with 24 remote sensing covariates and 3 soil properties (sand content, clay content and bulk density) to add local information on vegetation, climate, and topography. The new map combines georeferenced information of soil properties and remote sensing covariates and is called covariate-based Geo Transfer Functions (CoGTF) map. We used the random forest machine-learning algorithm to fit the Ksat models and the performance was assessed using CCC and RMSE which was computed using 5 fold cross-validation. The CCC and RMSE (in log_{10} Ksat given in cm/day) were observed 0.79 and 0.72, respectively. The CoGTF global Ksat map was compared with the map calculated with the well known Rosetta 3 PTF and major differences between the two maps were found. Firstly, Ksat values in Rosetta 3 were much lower for tropical regions compared to the CoGTF map. The tropical regions are expected to have rather high Ksat values due to intense soil formation processes and presence of more conductive clay minerals (kaolinite). The effects of active and inactive clay minerals on Ksat are captured in CoGTF map as formation of clay minerals are linked to precipitation, temperature and dense vegetation. Secondly, in CoGTF there is no abrupt change in Ksat as shown in Rosetta 3 map for the higher latitude regions such as Canada and Russia. This large contrast is related to a change in bulk density that is dominant in Rosetta 3. In CoGTF, RS covariates pattern cover this contrast. Furthermore, the CoGTF map, Rosetta 3 map, and the map of Dai et al. (2019) were validated using test data that were not used to calibrate the models, and the result showed that the CoGTF map performed better than the other models. Consequently, we propose to transition from PTFs based only on soil texture and bulk density to spatial-association of climate and vegetation covariates ("GTFs") to estimate Ksat. The study provides a blueprint for how georeferenced covariates could be used within the machine learning framework to improve Ksat predictive mapping. Moreover, the resulting CoGTF global maps are readily updatable as more information becomes available (covariates of measured Ksat).

Acknowledgments

The study was supported by ETH Zurich (Grant ETH-18 18-1). We would like to thank Zhongwang Wei, Samuel Bickel and Simone Fatichi (ETH Zurich) for insightful discussions.The data sets produced in this study are available at `https://doi.org/10.5281/zenodo.3934853`.

References

Ahuja, L., Cassel, D., Bruce, R., & Barnes, B. (1989). Evaluation of spatial distribution of hydraulic conductivity using effective porosity data. *Soil Science*, *148*(6), 404–411.

Allison, F. E. (1973). *Soil organic matter and its role in crop production*. Elsevier.

Araya, S. N., & Ghezzehei, T. A. (2019). Using machine learning for prediction of saturated hydraulic conductivity and its sensitivity to soil structural perturbations. *Water Resources Research*, *55*(7), 5715–5737.

Belk, E. L., Markewitz, D., Rasmussen, T. C., Carvalho, E. J. M., Nepstad, D. C., & Davidson, E. A. (2007). Modeling the effects of throughfall reduction on soil water content in a brazilian oxisol under a moist tropical forest. *Water Resources Research*, *43*(8).

Bouma, J. (1989). Using soil survey data for quantitative land evaluation. In *Advances in soil science* (pp. 177–213). Springer.

Breiman, L. (2001). Random forests. *Machine learning*, *45*(1), 5–32.

Brocca, L., Filippucci, P., Hahn, S., Ciabatta, L., Massari, C., Camici, S., … Wagner, W. (2019). SM2RAIN-ASCAT (2007–2018): Global daily satellite rainfall from ASCAT soil moisture. *Earth Syst. Sci. Data Discuss*, 1–31.

Buchhorn, M., Bertels, L., Smets, B., Lesiv, M., & Wur, N. (2017). Copernicus Global Land Operations “Vegetation and Energy”. *Copernicus Global Land Operations “Vegetation and Energy*.

Carter, B. J., & Ciolkosz, E. J. (1991). Slope gradient and aspect effects on soils developed from sandstone in pennsylvania. *Geoderma*, *49*(3-4), 199–213.

Chandler, K., Stevens, C., Binley, A., & Keith, A. (2018). Influence of tree species and forest land use on soil hydraulic conductivity and implications for surface runoff generation. *Geoderma*, *310*, 120–127.

Conrad, O., Bechtel, B., Bock, M., Dietrich, H., Fischer, E., Gerlitz, L., … Böhner, J. (2015). System for automated geoscientific analyses (saga) v. 2.1.4. *Geoscientific*

Model Development, *8*(7), 1991–2007. doi: 10.5194/gmd-8-1991-2015

Dai, Y., Xin, Q., Wei, N., Zhang, Y., Shangguan, W., Yuan, H., … Lu, X. (2019). A global high-resolution dataset of soil hydraulic and thermal properties for land surface modeling. *Journal of Advances in Modeling Earth Systems*.

Fashi, F. H., Gorji, M., & Shorafa, M. (2016). Estimation of soil hydraulic parameters for different land-uses. *Modeling Earth Systems and Environment*, *2*(4), 1–7.

Fatichi, S., Or, D., Walko, R., Vereecken, H., Young, M. H., Ghezzehei, T. A., … Avissar, R. (2020). Soil structure is an important omission in earth system models. *Nature Communications*, *11*.

Gupta, S., Hengl, T., Lehmann, P., Bonetti, S., & Or, D. (2020, April). SoilKsatDB: a global compilation of soil saturated hydraulic conductivity measurements. doi: 10.5281/zenodo.3752721

Gutmann, E. D., & Small, E. E. (2007). A comparison of land surface model soil hydraulic properties estimated by inverse modeling and pedotransfer functions. *Water Resources Research*, *43*(5).

Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., … Townshend, J. R. G. (2013). High-resolution global maps of 21st-century forest cover change. *Science*, *342*(6160), 850–853. doi: 10.1126/science.1244693

Hao, M., Zhang, J., Meng, M., Chen, H. Y., Guo, X., Liu, S., & Ye, L. (2019). Impacts of changes in vegetation on saturated hydraulic conductivity of soil in subtropical forests. *Scientific reports*, *9*(1), 1–9.

Hengl, T., de Jesus, J. M., Heuvelink, G. B., Gonzalez, M. R., Kilibarda, M., Blagotić, A., … others (2017). Soilgrids250m: Global gridded soil information based on machine learning. *PLoS one*, *12*(2), e0169748.

Hengl, T., & MacMillan, R. A. (2019). *Predictive Soil Mapping with R*. Lulu. com.

Hodnett, M., & Tomasella, J. (2002). 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*(3-4), 155–180.

Ito, A., & Wagai, R. (2017). Global distribution of clay-size minerals on land surface for biogeochemical and climatological studies. *Scientific data*, *4*, 170103.

Jana, R. B., & Mohanty, B. P. (2011). Enhancing ptfs with remotely sensed data for multiscale soil water retention estimation. *Journal of hydrology*, *399*(3-4), 201–211.

Jenny, H. (1994). *Factors of soil formation: a system of quantitative pedology*. Courier Corporation.

Jorda, H., Bechtold, M., Jarvis, N., & Koestel, J. (2015). Using boosted regression trees to explore key factors controlling saturated and near-saturated hydraulic conductivity. *European Journal of Soil Science*, *66*(4), 744–756.

Karger, D. N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R. W., … Kessler, M. (2017). Climatologies at high resolution for the earth's land surface areas. *Scientific data*, *4*, 170122.

Khlosi, M., Alhamdoosh, M., Douaik, A., Gabriels, D., & Cornelis, W. (2016). Enhanced pedotransfer functions with support vector machines to predict water retention of calcareous soil. *European Journal of Soil Science*, *67*(3), 276–284.

Kuhn, M. (2012). The caret package. *R Foundation for Statistical Computing, Vienna, Austria. URL https://cran. r-project. org/package= caret*.

Lawrence, I., & Lin, K. (1989). A concordance correlation coefficient to evaluate reproducibility. *Biometrics*, 255–268.

Leij, F., Alves, W., Van Genuchten, M. T., & Williams, J. (1996). The unsoda unsaturated soil hydraulic database; user's manual, version 1.0. *Rep. EPA/600/R-96*, *95*, 103.

Leij, F., Romano, N., Palladino, M., Schaap, M. G., & Coppola, A. (2004). Topographical attributes to predict soil hydraulic properties along a hillslope transect. *Water Resources Research*, *40*(2).

Mayr, T., & Jarvis, N. (1999). Pedotransfer functions to estimate soil water retention parameters for a modified brooks–corey type model. *Geoderma*, *91*(1-2), 1–9.

Montes, C. R., Melfi, A. J., Carvalho, A., Vieira-Coelho, A. C., & Formoso, M. L. (2002). Genesis, mineralogy and geochemistry of kaolin deposits of the jari river, amapá state, brazil. *Clays and Clay Minerals*, *50*(4), 494–503.

Montzka, C., Herbst, M., Weihermüller, L., Verhoef, A., & Vereecken, H. (2017). A global data set of soil hydraulic properties and sub-grid variability of soil water retention and hydraulic conductivity curves. *Earth System Science Data*, *9*(2), 529–543.

Nemes, A., Rawls, W. J., & Pachepsky, Y. A. (2005). Influence of organic matter on the estimation of saturated hydraulic conductivity. *Soil Science Society of America Journal*, *69*(4), 1330–1337.

Nemes, A., Schaap, M., Leij, F., & Wösten, J. (2001). Description of the unsaturated soil hydraulic database unsoda version 2.0. *Journal of Hydrology*, *251*(3-4), 151–162.

Niemeyer, R., Fremier, A. K., Heinse, R., Chávez, W., & DeClerck, F. A. (2014). Woody vegetation increases saturated hydraulic conductivity in dry tropical Nicaragua. *Vadose Zone Journal*, *13*(1).

Obi, J., Ogban, P., Ituen, U., & Udoh, B. (2014). Development of pedotransfer functions for coastal plain soils using terrain attributes. *Catena*, *123*, 252–262.

Or, D. (2019). The tyranny of small scales–on representing soil processes in global land surface models. *Water Resources Research*.

Ottoni, M. V., Ottoni Filho, T. B., Schaap, M. G., Lopes-Assad, M. L. R., & Rotunno Filho, O. C. (2018). Hydrophysical database for Brazilian soils (HYBRAS) and pedotransfer functions for water retention. *Vadose Zone Journal*, *17*(1).

Peters, J., De Baets, B., Verhoest, N. E., Samson, R., Degroeve, S., De Becker, P., & Huybrechts, W. (2007). Random forests as a tool for ecohydrological distribution modelling. *ecological modelling*, *207*(2-4), 304–318.

R Core Team. (2013). R: A Language and Environment for Statistical Computing [Computer software manual]. Vienna, Austria. Retrieved from `http://www.R-project.org/`

Rad, M. R. P., Toomanian, N., Khormali, F., Brungard, C. W., Komaki, C. B., & Bogaert, P. (2014). Updating soil survey maps using random forest and conditioned latin hypercube sampling in the loess derived soils of northern iran. *Geoderma*, *232*, 97–106.

Rawls, W. J., Brakensiek, D. L., & Saxtonn, K. (1982). Estimation of soil water properties. *Transactions of the ASAE*, *25*(5), 1316–1320.

Rodrigues, M., & de la Riva, J. (2014). An insight into machine-learning algorithms to model human-caused wildfire occurrence. *Environmental Modelling & Software*, *57*, 192–201.

Santra, P., & Das, B. S. (2008). Pedotransfer functions for soil hydraulic properties developed from a hilly watershed of eastern india. *Geoderma*, *146*(3-4), 439–448.

Saxton, K. E., & Rawls, W. J. (2006). Soil water characteristic estimates by texture and organic matter for hydrologic solutions. *Soil science society of America Journal*, *70*(5), 1569–1578.

Schaap, M. G., Leij, F. J., & Van Genuchten, M. T. (2001). Rosetta: A computer program for estimating soil hydraulic parameters with hierarchical pedotransfer functions. *Journal of hydrology*, *251*(3-4), 163–176.

Sharma, S. K., Mohanty, B. P., & Zhu, J. (2006). Including topography and vegetation attributes for developing pedotransfer functions. *Soil Science Society of America Journal*, *70*(5), 1430–1440.

Shoji, S., Nanzyo, M., & Takahashi, T. (2006). Factors of soil formation: climate. as exemplified by volcanic ash soils. *Soils: Basic concepts and future challenges. Cambridge University Press, Cambridge, UK*, 131–149.

Tomasella, J., & Hodnett, M. G. (1998). Estimating soil water retention characteristics from limited data in brazilian amazonia. *Soil science*, *163*(3), 190–202.

Tootchi, A., Jost, A., & Ducharne, A. (2019). Multi-source global wetland maps combining surface water imagery and groundwater constraints. *Earth System Science Data*, *11*, 189–220.

Tóth, B., Weynants, M., Nemes, A., Makó, A., Bilas, G., & Tóth, G. (2015). New generation of hydraulic pedotransfer functions for europe. *European journal of soil science*, *66*(1), 226–238.

Vereecken, H., Schnepf, A., Hopmans, J. W., Javaux, M., Or, D., Roose, T., … others (2016). Modeling soil processes: Review, key challenges, and new perspectives. *Vadose zone journal*, *15*(5).

Wilson, A. M., & Jetz, W. (2016, 03). Remotely sensed high-resolution global cloud dynamics for predicting ecosystem and biodiversity distributions. *PLOS Biology*, *14*(3), 1-20. Retrieved from `http://dx.doi.org/10.1371%2Fjournal.pbio.1002415` doi: 10.1371/journal.pbio.1002415

Wösten, J., Lilly, A., Nemes, A., & Le Bas, C. (1999). Development and use of a database of hydraulic properties of european soils. *Geoderma*, *90*(3-4), 169–185.

Wright, M. N., & Ziegler, A. (2015). ranger: A fast implementation of random forests for high dimensional data in c++ and r. *arXiv preprint arXiv:1508.04409*.

Yamazaki, D., Ikeshima, D., Tawatari, R., Yamaguchi, T., O'Loughlin, F., Neal, J. C., … Bates, P. D. (2017). A high-accuracy map of global terrain elevations. *Geophysical Research Letters*, *44*(11), 5844–5853.

Zhang, X., Zhu, J., Wendroth, O., Matocha, C., & Edwards, D. (2019). Effect of macroporosity on pedotransfer function estimates at the field scale. *Vadose Zone Journal*, *18*(1).

Zhang, Y., & Schaap, M. G. (2017). Weighted recalibration of the Rosetta pedotransfer model with improved estimates of hydraulic parameter distributions and summary statistics (Rosetta3). *Journal of Hydrology*, *547*, 39–53.

Zhang, Y., & Schaap, M. G. (2019). Estimation of saturated hydraulic conductivity with pedotransfer functions: A review. *Journal of Hydrology*, *575*, 1011–1030.

Zimmermann, A., Schinn, D. S., Francke, T., Elsenbeer, H., & Zimmermann, B. (2013). Uncovering patterns of near-surface saturated hydraulic conductivity in an overland flow-controlled landscape. *Geoderma*, *195*, 1–11.