Water Resources Research

RESEARCH ARTICLE
10.1029/2023WR035510

Determinants of Soil Field-Saturated Hydraulic Conductivity Across Sub-Saharan Africa: Texture and Beyond

Aida Bargués-Tobella[1,2], Leigh Ann Winowiecki[2], Douglas Sheil[3,4], and Tor-Gunnar Vågen[2]

[1]Department of Forest Ecology and Management, Swedish University of Agricultural Sciences (SLU), Umeå, Sweden, [2]World Agroforestry (ICRAF), Nairobi, Kenya, [3]Forest Ecology and Forest Management Group, Wageningen University & Research, Wageningen, The Netherlands, [4]Center for International Forestry Research (CIFOR), Kota Bogor, Indonesia

Key Points:

- We present field infiltration measurements and accompanying indicators of soil and land health from 3,573 plots across sub-Saharan Africa
- Field-saturated hydraulic conductivity (K_{fs}) is associated with soil texture and factors related to soil structure
- Our results suggest that soil hydrological functioning can be enhanced through management practices that target soil structure

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

A. Bargués-Tobella,
aida.bargues.tobella@slu.se

Citation:

Bargués-Tobella, A., Winowiecki, L. A., Sheil, D., & Vågen, T.-G. (2024). Determinants of soil field-saturated hydraulic conductivity across sub-Saharan Africa: Texture and beyond. *Water Resources Research*, *60*, e2023WR035510. https://doi.org/10.1029/2023WR035510

Received 5 JUN 2023
Accepted 19 DEC 2023

Author Contributions:

Conceptualization: Aida Bargués-Tobella, Leigh Ann Winowiecki, Tor-Gunnar Vågen
Data curation: Aida Bargués-Tobella, Leigh Ann Winowiecki, Tor-Gunnar Vågen
Formal analysis: Aida Bargués-Tobella, Leigh Ann Winowiecki, Tor-Gunnar Vågen
Funding acquisition: Aida Bargués-Tobella, Leigh Ann Winowiecki, Tor-Gunnar Vågen
Methodology: Aida Bargués-Tobella, Leigh Ann Winowiecki, Tor-Gunnar Vågen

Abstract Soil infiltration is critical for water security and related ecosystem services. This infiltration, the ability of soils to absorb water at their surface, is controlled by the soil hydraulic conductivity. Despite recent efforts in assembling measurements of soil hydraulic conductivity, global databases and derived pedotransfer functions lack coverage in the tropics. Here, we present soil infiltration measurements and other indicators of soil and land health collected systematically in 3,573 plots from 83 100 km^2 sites across 19 countries in sub-Saharan Africa. We use these data to (a) determine field-saturated hydraulic conductivity (K_{fs}) and (b) explore which variables best predict variation in K_{fs}. Our results show that sand content, soil organic carbon (SOC), and woody cover had a positive relationship with K_{fs}, whereas grazing intensity and soil pH had a negative relationship. Our findings highlight that, despite soil texture being important, structure also plays a critical role. These results indicate considerable potential to improve soil hydrological functioning through management and restoration practices that target soil structure. Enhancing SOC content, limiting animal stocking, promoting trees, shrubs, and other vegetation cover, and preventing soil erosion can increase K_{fs} and improve water security. This data set can contribute to improving Earth system and land surface models for applications in Africa.

1. Introduction

Water is a universal concern. Two-thirds of the world's population experience seasonal water scarcity, with over half a billion enduring it year-round (Kummu et al., 2016; Mekonnen & Hoekstra, 2016). There has long been concern over declining water availability in Africa (Falkenmark, 1989). In sub-Saharan nations, an estimated 85%–90% of agriculture, 70% of employment, and 40% of its exports depend on reliable rain (Onwujekwe & Ezemba, 2021). Inadequate access to water impacts food security, nutrition, and development (Heino et al., 2018; Iizumi et al., 2014). By 2017, at least 310 million people in sub-Saharan Africa lacked access to safe and affordable water for domestic use, with various estimates suggesting that as many as 700 million people may be displaced by 2030 as water stress escalates (Leal Filho et al., 2022). The consequences include hardship (Owuor et al., 2016), conflicts, and migration (Kattel, 2019). Reliable access to sufficient freshwater is key to the UN Sustainable Development Goals, with Goal 6 - *Ensure availability and sustainable management of water and sanitation for all* - being a prerequisite for achieving the rest. The challenges are considerable. Global per-capita freshwater reserves halved between 1960 and 2016, while consumption is estimated to double approximately every two decades (Ripple et al., 2017), and groundwater reserves are being depleted at an accelerating rate (Wada et al., 2010). In much of sub-Saharan Africa, the replenishment of soil moisture and groundwater stores and water availability largely depend on whether and how water at the soil surface is captured—a process that is largely determined by infiltration.

Infiltration is a critical hydrological process controlling the partitioning of water at the soil surface into surface runoff and subsurface water recharge (Ferré & Warrick, 2004; Hillel, 1998; Lal & Shukla, 2004). When the rate of water supply at the soil surface exceeds the soil's ability to absorb it, infiltration proceeds at a maximal rate known as the soil's infiltration capacity (Horton, 1941), or infiltrability (Hillel, 1971) for the particular case in which water at the soil surface is at atmospheric pressure. This capacity influences the quantity, quality, location, and timing of freshwater supplies, as it controls the pathways and the time needed for water to reach the stream channel. It also governs the replenishment of soil moisture required to sustain primary production and rainfed agricultural systems, and is thus critical to support food security. Soil infiltration capacity thus governs the provision of several ecosystem services, including those related to primary production and water cycling (Brauman

Project Administration: Aida Bargués-Tobella, Leigh Ann Winowiecki, Tor-Gunnar Vågen
Supervision: Leigh Ann Winowiecki, Tor-Gunnar Vågen
Visualization: Aida Bargués-Tobella
Writing – original draft: Aida Bargués-Tobella
Writing – review & editing: Leigh Ann Winowiecki, Douglas Sheil, Tor-Gunnar Vågen

et al., 2007; MEA, 2005; Ponette-González et al., 2015; Sun et al., 2017), in particular the mitigation of floods and droughts, water purification and regulation, and erosion control. When soil infiltration capacity is high, more water can enter the soil surface and become available both for soil- and groundwater recharge. Enhanced infiltration capacity can eventually lead to an increase in groundwater recharge and dry season flow (Ilstedt et al., 2016; Krishnaswamy et al., 2013), which in turn can improve water security. In contrast, low infiltration capacity results in more surface runoff, thereby increasing the risk of flooding, soil erosion, and siltation of water bodies. Hence, understanding soil infiltration capacity is critical for soil and water management. Such understanding is particularly urgent under climate change, as it will likely increase the frequency and intensity of heavy precipitation and droughts (Loukas et al., 2008; Seneviratne et al., 2021).

Soil infiltrability into an initially unsaturated soil is typically high during the early stages of the infiltration process and tends to decrease asymptotically over time as infiltration proceeds (Hillel, 1998). As the soil saturates, soil infiltrability approaches a steady, gravity-driven rate known as the soil steady-state infiltrability or final infiltration capacity. Steady-state infiltrability is thus independent of initial soil water content and approximates saturated hydraulic conductivity (Hillel, 1998). Saturated hydraulic conductivity depends on soil texture and structure and is higher in soils with large conducting pores. Soil pores can be textural pores (i.e., the pores between the primary mineral particles) or structural pores such as interaggregate cracks, decayed root channels, or wormholes (Dexter, 2004). Coarse-textured soils with relatively large pores typically have a higher saturated hydraulic conductivity than clay soils with smaller pores (Ferré & Warrik, 2004; Saxton et al., 1986). Pedotransfer functions developed to predict soil hydraulic conductivity from soil properties are primarily based on texture or other easy-to-measure attributes such as bulk density or organic carbon content—for example, Rosetta (Schaap et al., 2001; Zhang & Schaap, 2017) or hydraulic properties of European soils (Wösten et al., 1999). These pedotransfer functions may be insufficient to adequately assess field-saturated hydraulic conductivity due to the importance of structural porosity, especially under real-world conditions in the field (Bonetti et al., 2021; Fatichi et al., 2020; Gupta, Lehmann, et al., 2021; Jarvis et al., 2013; Pachepsky & Rawls, 2004; Rahmati et al., 2018; Vereecken et al., 2010; Wösten et al., 2001; Zhang et al., 2020).

Several processes and practices can influence soil structural porosity and, thereby, soil hydrological functioning. Changes in land cover and land use can alter the infiltration capacity of soils through concurrent changes in the quality and quantity of soil organic matter inputs, the structure and abundance of soil biota communities, the characteristics of root systems, and the overall structure of soils. Soil organic matter promotes biological activity and aggregation in soils, which can improve infiltration capacity (Franzluebbers, 2002; Lado, Paz, & Ben-Hur, 2004), although in some cases increases in organic matter content can have the opposite effect (Araya & Ghezzehei, 2019; Jarvis et al., 2013; Larsbo et al., 2016; Nemes et al., 2005; Wang et al., 2009). Roots and soil biota, in particular macroinvertebrates, can further enhance infiltration through the creation of macropores and persistent soil aggregates (Beven & Germann, 2013; Lavelle et al., 2006; Weiler & Naef, 2003). Conversion of forests, woodlands, and natural grasslands to agriculture typically results in reduced infiltration capacity (Bormann et al., 2005; Lal, 1996; Lulandala et al., 2022; Nyberg et al., 2012; Zimmermann & Elsenbeer, 2008). This, in turn, can adversely affect various aspects of hydrological functioning, including reduced dry-season flows, increased surface runoff, and enhanced soil erosion (Bruijnzeel, 2004). Land management can potentially reverse such degradation and contribute to maintaining or improving hydrological functioning and derived ecosystem services. Adopting practices such as agroforestry, controlled grazing, and conservation tillage can improve soil infiltration capacity (Abdelkadir & Yimer, 2011; Bargués Tobella et al., 2014; Benegas et al., 2014; Ilstedt et al., 2007; Lulandala et al., 2022; Moreno et al., 1997; Mwendera & Saleem, 1997; Zhang et al., 2007). Similarly, tree planting or natural regeneration following the abandonment of croplands, pastures, or degraded grasslands can progressively restore soil infiltration capacity (Bonell et al., 2010; Colloff et al., 2010; Ilstedt et al., 2007; Lal, 1996; Zhao et al., 2013; Zimmermann et al., 2010). However, evidence is often restricted to observations and experiments conducted at the local level.

Understanding soil infiltration and surface water redistribution processes at scales relevant for management requires an assessment of saturated hydraulic conductivity over broad areas. However, soil saturated hydraulic conductivity is extremely variable in space, and individual measurements represent only a limited area (Nielsen et al., 1973; Usowicz & Lipiec, 2021; Warrick, 1998). Obtaining sufficient field observations to characterize saturated hydraulic conductivity at the landscape or regional level is difficult due to the cost, time, and effort associated with traditional field techniques. Using less complex methods such as single-ring infiltrometers (Bouwer, 1986; Reynolds & Elrick, 1990) allows for higher sampling intensity, which is usually preferred

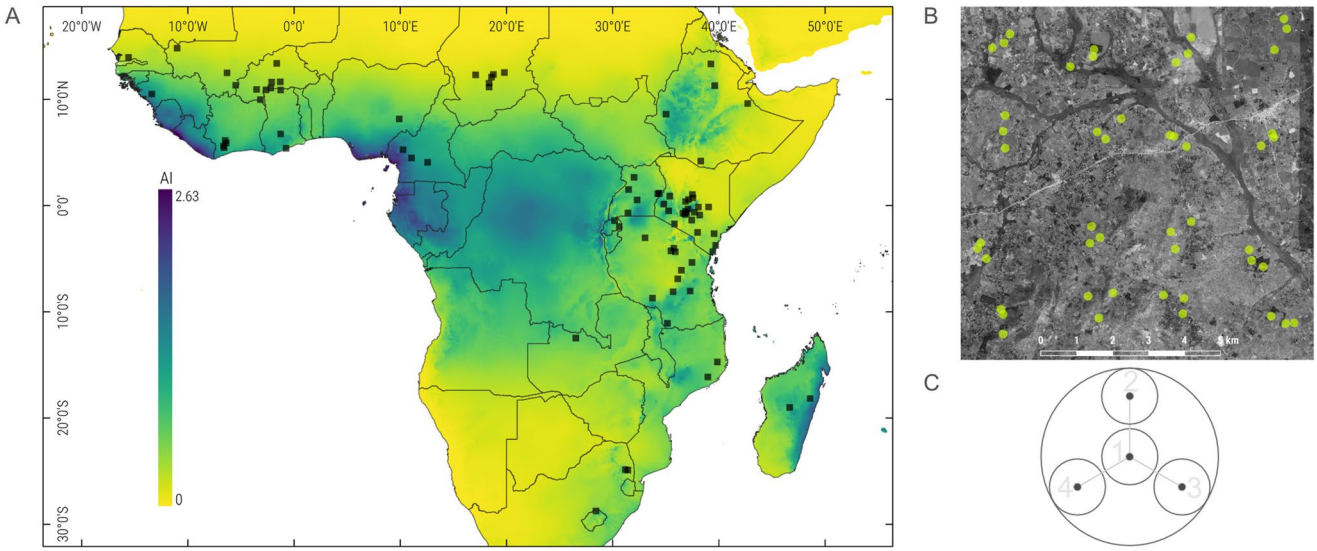

Figure 1. (a) Location of the 83 Land Degradation Surveillance Framework (LDSF) sites. The grid layer represents the annual average aridity index (AI = mean annual precipitation/mean annual potential evapotranspiration) over the period 1970–2000 (Trabucco & Zomer, 2019). (b) Example of an LDSF site (10 × 10 km) in Hoima, Uganda, showing the location of the sampling plots. (c) Plot outline, showing the four sub-plots.

even if at the expense of relatively higher uncertainty in the individual measurements (Bagarello et al., 2004; Nimmo et al., 2009). Pedotransfer functions for estimating soil hydraulic properties have also been developed to reduce the need for costly field measurements (Schaap et al., 2001; Wösten et al., 1999). However, these functions are often developed for specific regions, especially North America and Europe, which limits their broader geographical applicability, particularly in the tropics (Botula et al., 2012; Gupta, Hengl, et al., 2021; Hodnett & Tomasella, 2002; Tomasella et al., 2000; Young et al., 1999). Similarly, the soil data and resulting methods tend to neglect natural areas and focus on agricultural land. In an effort to address these geographical and land-use limitations, Rahmati et al. (2018) gathered published infiltration data to create the Soil Water Infiltration Global database (SWIG), which contains single infiltration curves and accompanying data, including data on soil properties and land management, from 215 sites across 53 countries. Efforts to compile data on soil hydraulic properties for specific regions in the tropics include the Hydrophysical Database for Brazilian Soils (HYBRAS) (Ottoni et al., 2018) and the soil infiltration and saturated hydraulic conductivity database of the soils of Rio de Janeiro (Martins et al., 2022). Africa lacks such databases, and the region remains sparsely represented in SWIG and other global databases.

Here, we present soil infiltration measurements and other soil and land health indicators collected systematically across sub-Saharan Africa. These detail 3,573 plots from 83 100 km^2 sites across 19 countries. We use these data to (a) determine field-saturated hydraulic conductivity (K_{fs}), and (b) explore which variables best explain the variation in K_{fs}. Variables include soil texture, pH, soil organic carbon (SOC), visible signs of erosion, vegetation structure, herbaceous cover, woody cover, and impact of grazing.

2. Materials and Methods

2.1. Study Sites and Field Sampling Design

The data presented and used in this study were collected following the Land Degradation Surveillance Framework (LDSF), an established sampling approach for landscape-level assessment of soil and land health (Vågen & Winowiecki, 2020). The data derive from 3,573 plots from 83 sites distributed across 19 countries in sub-Saharan Africa and covering diverse climates and contexts (Figure 1a; detailed site-specific information can be found in Table S1 in Supporting Information S1). These LDSF sites were selected at random across a region of interest or placed in a specific location representing an area of particular interest or planned activities (interventions).

The LDSF employs a spatially stratified hierarchical (nested) sampling scheme with four levels corresponding to different spatial scales (Figure 1): sites (100 km^2), clusters (1 km^2) nested within sites, plots (1,000 m^2) nested

within clusters, and sub-plots (100 m^2) nested within plots. Each site is divided into 16 tiles, and cluster centroid locations are randomized within each tile. Within each cluster, 10 sampling plots are placed at random. Each plot contains four subplots. We only used data from the plots where soil infiltration was measured, which was typically in three plots per cluster (i.e., 48 plots per site. Figure 1b), although the number of plots with complete infiltration and soil data varied slightly among sites (Table S1 in Supporting Information S1).

2.2. Field Data Collection

Field data were collected both at the plot and sub-plot levels following the LDSF protocol (Vågen & Winowiecki, 2020). Plot coordinates were recorded at the center of subplot one with an accuracy better than 5 m. Soil infiltration measurements were conducted at the center of each plot using a single-ring infiltrometer (Bouwer, 1986) consisting of a metal cylinder with an inner radius of 78 mm and a height of 20–30 cm. The infiltrometer was inserted at least 20 mm into the soil, taking care to minimize soil surface disturbance. Before starting the infiltration measurements, the soil within the infiltration ring was pre-wetted to increase the chances of reaching steady-state conditions. Water was gently poured to fill the infiltration ring to the top, covering the soil surface to minimize surface disturbance, and ponding was maintained for 15 min. After pre-wetting, the infiltration measurements started and were carried out for 2.5 hr. A ruler was fixed on the inner side of the infiltration ring to measure the water level. The ring was then refilled, and the initial ponding depth was recorded. After 5 min, the final ponding depth was recorded, and the ring was immediately refilled to the start level. This procedure was repeated at 5-min intervals during the first 30 min of the experiment, 10-min intervals during the following 60 min, and 20-min intervals during the remaining 60 min. Intervals were adjusted as necessary when infiltration was too fast or too slow to obtain accurate readings.

In each plot, the dominant vegetation structure was classified into forest, woodland, bushland, shrubland, wooded grassland, grassland, or annual cropland following FAO's Land Cover Classification System (Di Gregorio & Jansen, 1998). Woody and herbaceous vegetation cover were visually assessed within each sub-plot and scored following a modified Braun-Blanquet cover-abundance rating scale (Braun-Blanquet, 1932) into either 0 (absent), 1 (<4%), 2 (4%–15%), 3 (15%–40%), 4 (40%–65%) and 5 (>65% vegetation cover). The four sub-plot ratings were averaged to obtain plot-level estimates of woody and herbaceous vegetation cover. Visible signs of soil erosion (rill, gully, or sheet) were recorded within sub-plot 1, which is where infiltration measurements were carried out. In addition, the impact of grazing was assessed at the plot level using a scale from 0 (no impact) to 3 (severe impact) based on visible signs of grazing and browsing.

Topsoil samples (0–20 cm) were collected from the center of each sub-plot using a soil auger and combined into a composite plot-level sample. These samples were air-dried before shipping them to the soil laboratory for analysis.

2.3. Determination of Topsoil Field-Saturated Hydraulic Conductivity (K_{fs})

We applied Nimmo et al. (2009)'s method to analyze ponded infiltration from within a single-ring infiltrometer under falling-head conditions and determine the field-saturated hydraulic conductivity (K_{fs}). The formula for K_{fs} (Equation 1) compensates for non-constant falling head and subsurface radial spreading typical of small-diameter infiltration rings like the ones used in this study:

$$K_{fs} = \frac{L_G}{t - t_0} \ln\left[\frac{L_G + \lambda + D_0}{L_G + \lambda + D}\right] \quad (1)$$

where t_0 and t are the initial and final time, λ is the soil macroscopic capillary length, D and D_0 are the initial and final ponding depths within the infiltration ring, and L_G is the ring-installation scaling length (Equation 2):

$$L_G = C_1 d + C_2 b \quad (2)$$

where b is the inner radius of the infiltration ring (78 mm), d is the ring insertion depth (20–40 mm), and C_1 and C_2 are empirically determined constants with recommended values of 0.316 π and 0.184 π, respectively (Reynolds et al., 2002).

The soil macroscopic capillary length parameter (λ), which here is the reciprocal of the soil macroscopic capillary length as defined by Reynolds and Elrick (1990) and Reynolds et al. (2002) (α^*), represents the relative importance

Table 1
Performance Metrics for the Prediction of Soil Properties Included in the Study Using Mid-Infrared Spectroscopy

Soil property	R^2	RMSEP*
Soil organic carbon (SOC; g kg^{-1})	0.98	2.03
Soil pH	0.98	0.14
Sand (%)	0.97	4.30
Clay (%)	0.97	3.91

Note. Performance was assessed by comparing predicted values to reference analysis methods, as outlined above.
*RMSEP = Root Mean Squared Error of Prediction.

of gravity and capillary forces during ponded infiltration (Reynolds et al., 2002). We used a value of α^* of 0.012 mm^{-1} (i.e., λ = 83.33 mm), as this value has previously been shown to be adequate for most soils with structural development (see Elrick et al. (1989) and Reynolds et al. (2002)).

We applied Equation 1 to all infiltration measurements, regardless of whether they were at (quasi-)steady state or not. We then fitted an asymptotic function to each individual curve of corrected ponded infiltration rates over time and extracted the asymptote value, which represents K_{fs} for that specific plot. Asymptote curves were fitted using the R package *nls.multstart*, which uses non-linear least square regression with the Levenberg-Marquardt algorithm (Padfield & Matheson, 2020).

2.4. Soil Laboratory Analyses

Soil samples were analyzed using mid-infrared (MIR) spectroscopy at the ICRAF Soil-Plant Spectral Diagnostics Laboratory in Nairobi, Kenya. Soil MIR spectroscopy is a well-established, non-destructive, rapid, and cost-effective method for predicting key soil properties such as SOC, pH and texture, enabling landscape-level assessments of soil health (Terhoeven-Urselmans et al., 2010; Vågen et al., 2016; Winowiecki et al., 2016). Prior to analysis, air-dried soil samples were ground and sieved through a 2 mm mesh. A subsample was further ground to attain a particle size between 20 and 53 µm. These subsamples were then scanned and analyzed in triplicate for MIR absorbance using a Tensor 27 HTS-XT MIR spectrometer (Bruker Optics, Karlsruhe, Germany), and the spectra were subsequently processed following Terhoeven-Urselmans et al. (2010).

Soil texture, pH, and SOC were simultaneously predicted from MIR spectra using Random Forest regression models (Breiman, 2001) built with data from ICRAF's global soil MIR spectral library following the methods developed by Vågen et al. (2016). Today, the library consists of spectra from about 162,000 samples and data on soil properties obtained using traditional analytical methods for 10% of these samples (including those used in this study). For this reference set of soil samples, soil pH was measured using an Eutech Cyberscan 1,100 pH meter on a 1:2 soil/water solution prepared by shaking 20 g of soil with 40 ml of deionized water for 30 min. SOC content was analyzed by dry combustion using an Elemental Analyzer Isotope Ratio Mass Spectrometer from Europa Scientific after removing inorganic C with 0.1 N HCl, at the IsoAnalytical Laboratory (United Kingdom). Finally, soil textural analyses were conducted using a LA-950 Laser Diffraction Particle Size Distribution Analyzer (HORIBA Scientific) after shaking each soil sample for 4 min in a dispersing solution of 1% sodium hexametaphosphate (calgon). The U.S. Department of Agriculture (USDA) soil particle-size classification system was used, which defines the size limit between sand and silt at 50 µm (USDA, 2017). No adjustment was applied to the soil particle-size distribution data obtained with the laser diffractometer to bring the values closer to those measured by classical sieve-sedimentation methods. Random Forest regression models were trained on the reference set of soil samples with both MIR spectra and associated data obtained using traditional analytical methods as outlined in Vågen et al. (2016). These models achieve a good prediction accuracy ($R^2 > 0.95$) for the soil properties that concern our investigation (Table 1).

2.5. Statistical Analysis—Modeling

2.5.1. Mixed-Effects Models

Given the nested structure of the LDSF data (Site:Cluster:Plot), we used hierarchical models (Zuur et al., 2013) with Site and Cluster as nested random effects. In other words, we applied a 2-way nested random structure with sites, and clusters within sites. Furthermore, since K_{fs} is positive, continuous, and skewed, and to take care of variance heterogeneity, we used a gamma generalized linear mixed-effects model (Zuur et al., 2013) with a log-link function and random intercept:

$$K_{\mathrm{fs}_{ijk}} \sim \text{gamma}\,(\mu_{ijk}, r)$$

$$E\left(K_{\mathrm{fs}_{ijk}}\right) = \mu_{ijk}$$

$$\mathrm{var}\left(K_{\mathrm{fs}_{ijk}}\right) = {\mu_{ijk}}^2/r$$

$$\log\left(K_{\mathrm{fs}_{ijk}}\right) = \text{fixed component} + \text{site}_i + \text{cluster}_{i,j}$$

where:

i is the Site. $i = 1, \ldots, 83$
j is the jth cluster within site i. $j = 1, \ldots, 16$
k is the kth observation (plot) within cluster j. $k = 1, 2, 3$
r is a shape parameter (gamma distribution)
site_i is the site-level random intercept
$\text{cluster}_{i,j}$ is the cluster-level random intercept

Mixed-effects models have become an increasingly popular tool for analyzing complex ecological data (Harrison et al., 2018; Zuur et al., 2009). Linear mixed-effects models and generalized linear mixed-effects models (GLMMs) are particularly suited to the analysis of highly structured data sets like ours containing discrete groupings of non-independent observational units that are hierarchical in nature. Mixed-effects models have advantages over regular linear models as they can account for non-independence and heterogeneity. Nevertheless, underlying assumptions and model selection and interpretation require care (Harrison et al., 2018). Here, we follow the best practices and guidelines for such analyses provided by Zuur et al. (2013, 2009) and Harrison et al. (2018).

Gamma GLMMs were fitted using the function *glmmTMB* within the *glmmTMB* package in R (Brooks et al., 2017). We fit four initial candidate models (Table 2). All four models included interactions between soil texture (either expressed as sand content or texture class) and the remaining covariates, as we expected it to influence the effect of the other covariates. We defined three soil texture classes, sandy (>50% sand content), clay (>50% clay content), and clay loam (<50% sand content and <50% clay content). Ordinal categorical variables (i.e., woody and herbaceous vegetation cover ratings and grazing impact rating) were treated as continuous covariates to simplify model structure and preserve ordering. Continuous covariates were standardized ($x - \mathrm{mean}(x)/\mathrm{sd}(x)$) to fit the models in order to improve model performance and interpretability (Schielzeth, 2010).

Model selection involved two steps. First, we applied model selection on the covariates, dropping non-significant interaction terms from the four initial candidate models using likelihood-ratio tests via the *drop1* function in the *stats* package in R (Chambers & Hastie, 1992). Second, the best model was selected from the four final candidate models (Table 2). The model with the lowest Akaike's Information Criterion (AIC) is considered to be the best in that it optimizes the trade-off between fit and complexity (Burnham & Anderson, 2004).

In the model validation, we followed Zuur et al. (2013) to check that model assumptions were not violated. Plots of Pearson residuals (residuals vs. fitted values, residuals vs. model covariates, and residuals vs. spatial coordinates) were used to visually assess the lack of heteroscedasticity, non-linear patterns, and spatial correlation. To aid with non-linear pattern detection, we fitted a smoother to the residuals versus model covariates plots via the *geom_smooth* function in the *ggplot2* package in R. We also used variograms of the residuals to confirm the lack of spatial correlation. Histograms of the random intercepts were used to check that the assumption of normality of random effects was fulfilled. Plots of random effects versus residuals and variograms of the random effects were also used to check the lack of spatial correlation. Models were checked for multicollinearity using the function *check_collinearity* in the *performance* package in R (Lüdecke et al., 2021), which uses the variance inflation factor. Model assumptions were not violated.

Model performance indices, including marginal and conditional R^2 (Nakagawa et al., 2017), Interclass Correlation Coefficient (ICC, Hox et al. (2017)) and AIC, were computed using the function *model_performance* within the *performance* package in R (Lüdecke et al., 2021). Marginal R^2 comprises variance explained by only the fixed effects, while conditional R^2 comprises the variance explained by both the fixed and random effects (Nakagawa et al., 2017).

2.5.2. Random Forest Regression Models

We used Random Forest regression models (Breiman, 2001) to gain further insight into the relative importance of the different covariates in predicting K_{fs}. Random Forests are a nonparametric method of statistical learning

Water Resources Research

10.1029/2023WR035510

Table 2
Candidate Fitted Gamma Generalized Linear Mixed-Effects Models

Model	Initial model formulation	Final model formulation	AIC	$R^2_{conditional}$	$R^2_{marginal}$	ICC_{site}	$ICC_{cluster}$
1	log (K_{fs}) = Sand * (pH + SOC + Erosion + VegStructure + ImpactGrazing + HerbCoverRate) + site_i + cluster_ij	log (K_{fs}) = Sand * (SOC + Erosion + VegStructure + ImpactGrazing) + pH + HerbCoverRate + site_i + cluster_ij	34,862	0.33	0.15	0.185	0.020
2	log (K_{fs}) = Sand * (pH + SOC + Erosion + WdCoverRate + ImpactGrazing + HerbCoverRate) + site_i + cluster_ij	log (K_{fs}) = Sand * (SOC + Erosion + ImpactGrazing) + pH + WdCoverRate + HerbCoverRate + site_i + cluster_ij	34,898	0.33	0.14	0.191	0.029
3	log (K_{fs}) = Texture * (pH + SOC + Erosion + VegStructure + ImpactGrazing + HerbCoverRate) + site_i + cluster_ij	log (K_{fs}) = Texture * (SOC + Erosion) + pH + VegStructure + ImpactGrazing + HerbCoverRate + site_i + cluster_ij	34,901	0.33	0.14	0.201	0.020
4	log (K_{fs}) = Texture * (pH + SOC + Erosion + WdCoverRate + ImpactGrazing + HerbCoverRate) + site_i + cluster_ij	log (K_{fs}) = Texture * (SOC + Erosion + ImpactGrazing) + pH + WdCoverRate + HerbCoverRate + site_i + cluster_ij	34,929	0.33	0.13	0.201	0.027

Note. Both the initial and final model formulations (before and after dropping non-significant interaction terms, respectively) are shown. For each final model, Akaike's AIC, conditional and marginal R^2, and Interclass Correlation Coefficients (ICC) are indicated. The best model, based on the AIC, is highlighted in gray. K_{fs}: Field-saturated hydraulic conductivity (mm h^{-1}), Sand: Sand content (%), SOC: Soil organic carbon content (g C kg^{-1}), pH: Soil pH, Erosion: Presence of erosion within subplot 1 (binary), VegStructure: Vegetation structure class, ImpactGrazing: Grazing impact rating, HerbCoverRate: Herbaceous cover rating, WdCoverRate: Woody cover rating, Texture: Soil texture class. The fixed component includes main effects of covariates and the interaction between sand content or texture class and the other covariates.

widely used across a broad range of disciplines, including ecology and soil science (Biau & Scornet, 2016). The popularity of Random Forest models is mainly due to their predictive performance and versatility, with few restrictions on the nature of the data. Their main drawback is that they can be complex to interpret. In addition to being a powerful prediction tool, a key application is the quantification of the relative importance of the candidate explanatory variables (Grömping, 2015; Molnar, 2022).

Model covariates included sand and clay content, SOC content, soil pH, erosion presence, vegetation structure class, grazing impact rating, and herbaceous and woody cover rating. Random Forest regression models were fitted using the *tidymodels* framework in R (Kuhn & Wickham, 2020), with the *ranger* package (Wright & Ziegler, 2017) as the underlying engine. First, the data set was split into training (80%) and test (20%) sets using simple random sampling. The training set was then used to fit the model and tune its parameters. We performed parameter tuning for *mtry*, that is, the number of variables randomly selected at each node, and *trees*, that is, the number of trees in the forest/ensemble. First, we conducted a grid search in which different combinations of the two parameters were explored (*mtry* = 1–10 and *trees* = 100, 250, and 500). Multiple models were trained on resampled data sets (k-fold cross-validation, $k = 10$) using the different combinations of *mtry and trees*. We then evaluated how well these models performed based on the cross-validated estimates of two standard performance metrics, Root Mean Squared Error (RMSE) and Mean Absolute Error. Based on this, we selected the best combination of *mtry* and *trees* (*mtry* = 2, *trees* = 500). The final model was then built with the selected values of the parameters. Overall model performance was evaluated on the test set using R^2 and RMSE as the performance metrics.

The importance of each explanatory variable was assessed using the Mean Decrease in Accuracy (MDA) or Permutation Importance measure (Breiman, 2001). The MDA is a measure of variable importance, and it is constructed from the Random Forest model prediction accuracy (the Mean Squared Error, MSE, in the case of regression models), which is assessed using the out-of-bag (OOB) observations. The MDA for a particular variable represents the normalized average (among trees in a forest) increase in the MSE made by a tree when randomly permuting the observations for that variable in the OOB samples. The more important a variable is as a predictor, the higher this increase will be.

To clarify how predicted K_{fs} varied across the observed ranges of the model covariates, we constructed partial dependence plots (Friedman, 2001) using the *pdp* package in R (Greenwell, 2017). Partial dependence plots illustrate the marginal effect a specific feature of interest has on the predicted outcome of a machine learning model (Friedman, 2001; Molnar, 2022). These plots visualize the direction, strength, and form of relationships, although they might not capture the full complexity of the model (Molnar, 2022).

All data analyses were conducted using R version 4.0.0 (R Core Team, 2015).

3. Results

3.1. Overview

The sample set consists of data from 3,573 plots across 83 sites. Of these plots, 45% were under cropland, whereas the vegetation structure in the remaining plots was classified as either shrubland (17%), grassland (10%), bushland (9%), wooded grassland (7%), woodland (7%), or forest (5%) (Table S2 in Supporting Information S1). The distribution of vegetation structure classes varied among sites, but cropland was the predominant class (≥50% of the plots) in 36 out of 83 sites (Figure S1 in Supporting Information S1). Plot-average woody and herbaceous vegetation cover ratings (0–5) varied within and among sites, with an overall mean value of 1.7 and 2.8, respectively (Figure S2 in Supporting Information S1). On average, half of the plots per site had visible signs of erosion, but erosion prevalence varied significantly among sites (Figure S3 in Supporting Information S1). The overall mean impact of grazing rating (0–3) was 1.1, but it varied within and among sites (Figure S4 in Supporting Information S1). Overall, 57% of the plots presented visible signs of grazing and browsing (impact of grazing rating ≥1). This proportion was highest for plots classified as bushland (84%), followed by wooded grassland (75%), grassland (75%), shrubland (69%), cropland (46%), and forest (26%) (Figure S5 in Supporting Information S1).

Topsoil field-saturated hydraulic conductivity (K_{fs}) for the 3,573 measures followed a right-skewed distribution with a median value of 41.2 mm hr^{-1} and an interquartile range of 57.7 mm hr^{-1} (Figure 2, Table S3 in Supporting Information S1). K_{fs} varied within and among sites; Median K_{fs} for the 83 sites ranged from 6.8 (at Temki, Chad) to 160.9 mm hr^{-1} (Mbinga, Tanzania), and, in general, sites with higher median K_{fs} also exhibited higher variability (Figure 2).

10.1029/2023WR035510

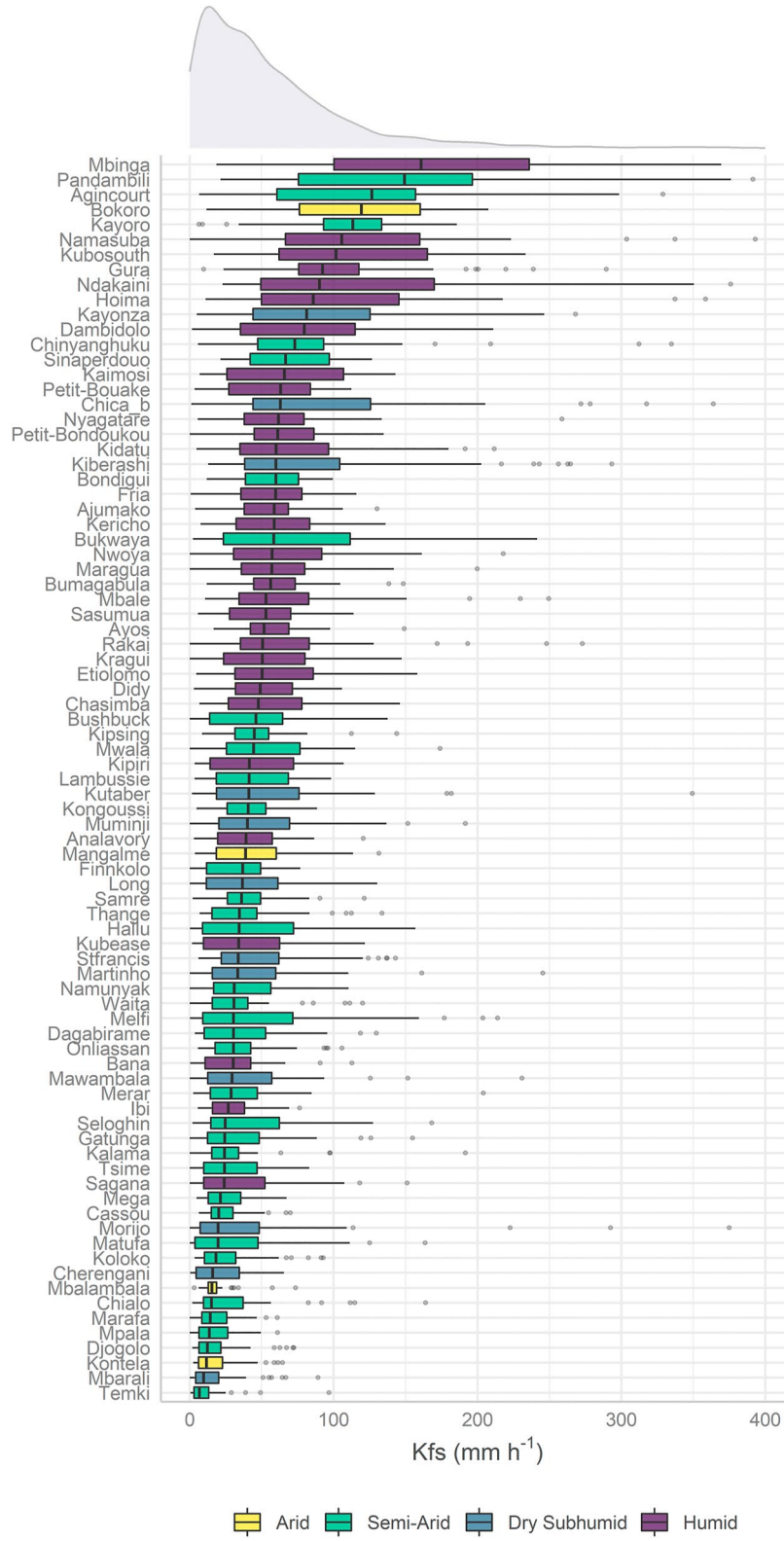

Figure 2. Boxplots (Q1, median, Q3) of topsoil field-saturated hydraulic conductivity (K_{fs}) for the 83 sites included in the study. The color indicates the average site aridity class (UNEP, 1992) for the period 1970–2000, based on Trabucco and Zomer (2019). The overall distribution of K_{fs} is also shown with a marginal density plot. The *x*-axis is limited to 400 mm hr^{-1} to improve data visualization (only seven observations were above this threshold).

10.1029/2023WR035510

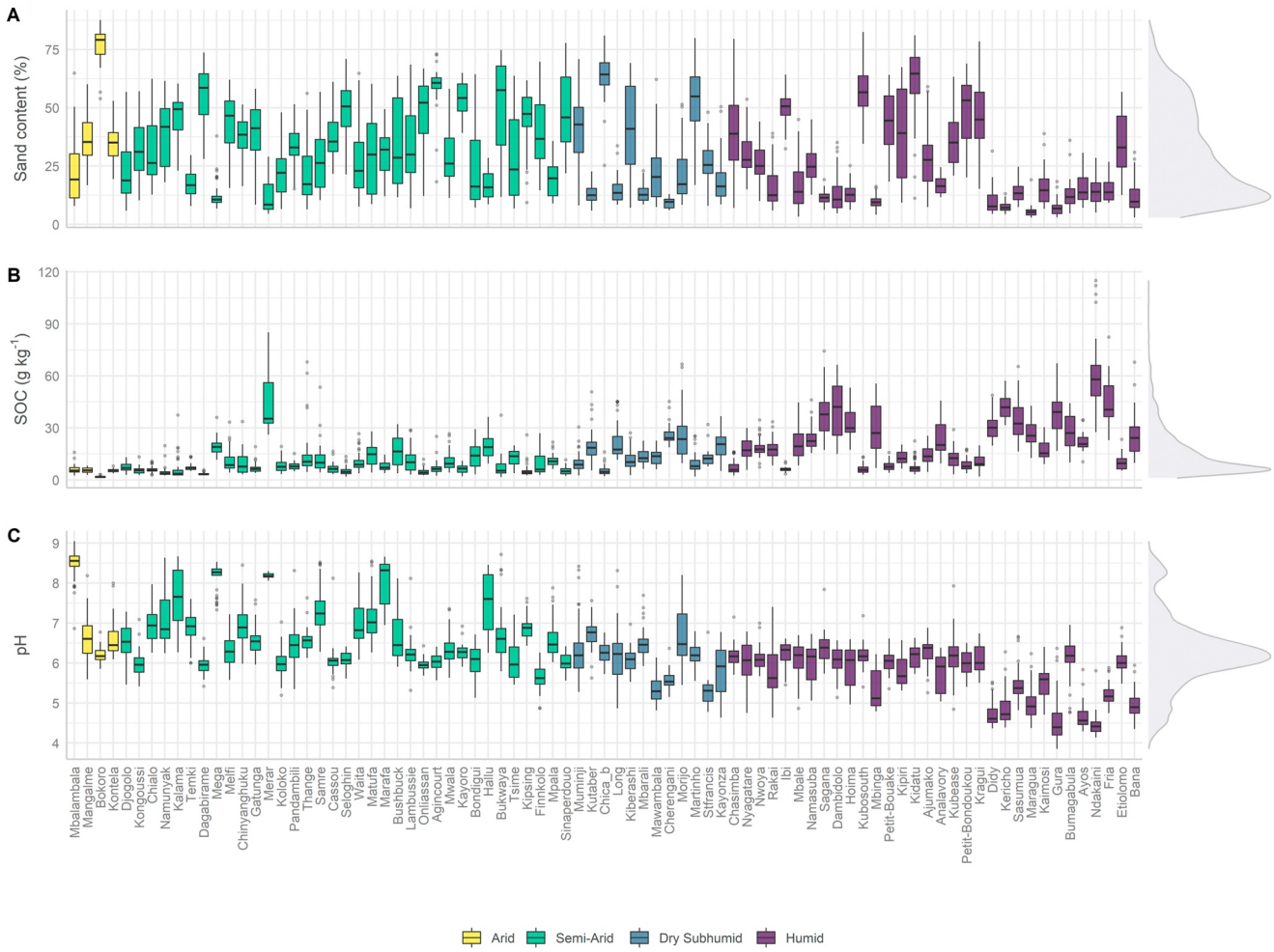

Figure 3. Boxplots (Q1, median, Q3) and marginal density plots of sand content (a), soil organic carbon (b), and pH (c) for the 83 study sites included in the study. The color indicates the average site aridity class (UNEP, 1992) for the period 1970–2000, based on Trabucco and Zomer (2019). The sites are ordered by their aridity index, from more arid to more humid sites.

Soil sand content, SOC content, and pH also varied within and among sites (Figure 3). In general, SOC content increased, and soil pH decreased with humidity. Sites with lower median SOC also had lower variability in SOC. The sampled soils had relatively low silt content (below 50%); 53% of them had more than 50% clay content, while only 19% had more than 50% sand content (Table S4 in Supporting Information S1 and Figure 4b).

Overall, there was a moderate and positive correlation of K_{fs} with SOC content ($r = 0.18$), woody cover rating ($r = 0.13$), soil sand content ($r = 0.1$), and herbaceous cover rating ($r = 0.08$), and a negative correlation of K_{fs} with soil pH ($r = -0.22$), the impact of grazing rating ($r = -0.15$), and soil clay content ($r = -0.09$) (Figure S6 in Supporting Information S1). Sites with a more humid climate tended to have higher K_{fs} than drier sites (Figure 2), as indicated by the positive correlation between site mean aridity index (AI = mean annual precipitation/mean annual potential evapotranspiration) and median K_{fs} ($r = 0.3$). Median K_{fs} per aridity class across all plots also decreased with increasing aridity (Figure 4a). Overall, coarser soils had higher K_{fs} than more fine-textured ones (Figure 4). However, as humidity increased, the effect of soil texture on K_{fs} was less pronounced (Figure 4a). Interestingly, more extreme K_{fs} values (above the overall 95th percentile) occurred not only in sandy soils but also in soils with more than 50% clay content (Figure 4b).

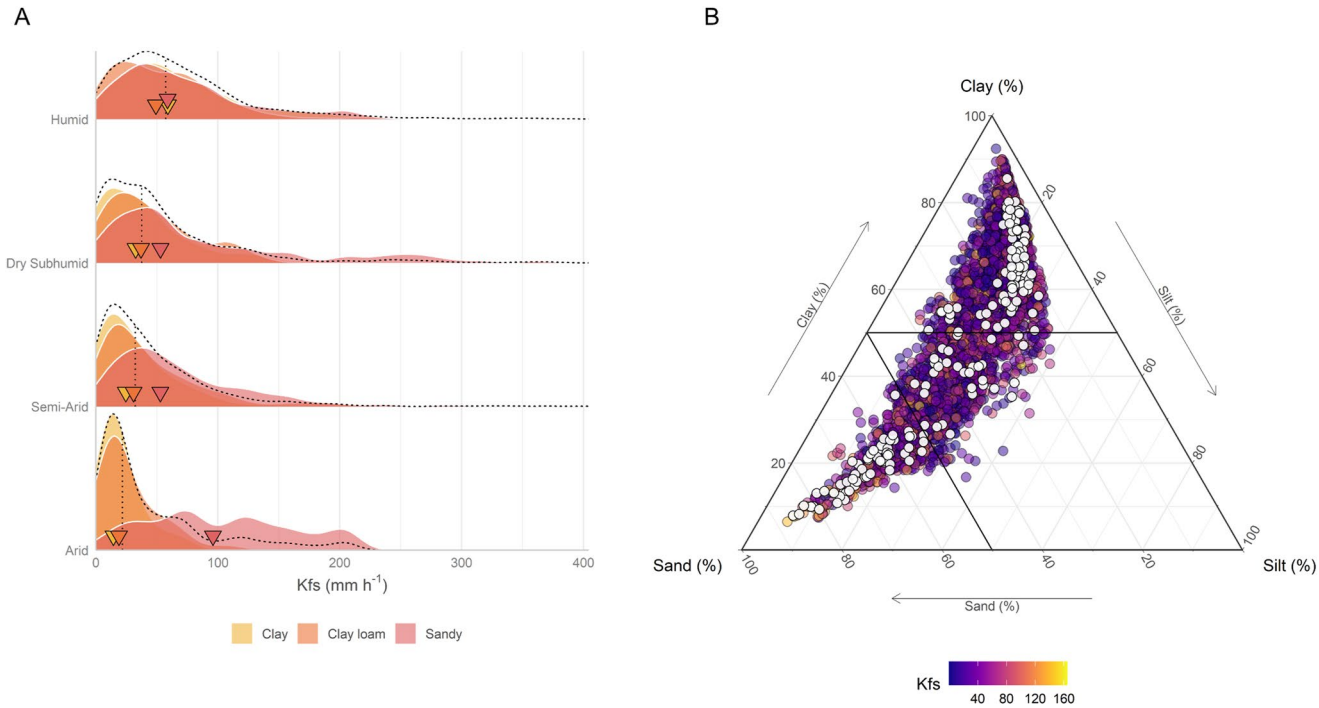

Figure 4. (a) Density plots showing the distribution of field-saturated hydraulic conductivity (K_{fs}) values per aridity class (UNEP, 1992) and soil texture class (sandy: sand content >50%; clay: clay content >50%; clay loam: sand content <50% and clay content <50%). The dotted lines show the overall distribution of K_{fs} per aridity class, and the dotted vertical lines correspond to median values. The triangles along the *x* axes indicate median values per textural class (b) Soil texture ternary diagram of the samples. The color scale is related to the magnitude of K_{fs}. Extreme K_{fs} values are highlighted in white (i.e., greater than the 95% percentile of 165 mm hr^{-1}).

3.2. Modeling Results

3.2.1. General Linear Mixed-Effects Model

The best model explaining K_{fs} (see Table 2 for the final model formulation) had a marginal and conditional R^2 of 0.15 and 0.33, respectively (Table 3). Sand and SOC content had a significant positive effect on K_{fs} (Figures 5 and 6). In contrast, the impact of grazing and soil pH had a significant negative effect. The effect of vegetation structure was also significant, with shrublands, wooded grasslands, croplands, woodlands, bushlands, and forests having an increasingly positive effect on K_{fs} compared to grasslands. On the other hand, herbaceous cover and the presence of erosion had a positive and negative effect, respectively, on K_{fs}, although not statistically significant.

The relationship between field-saturated hydraulic conductivity and SOC content, erosion, vegetation structure, and grazing impact varied with soil texture. The positive effect of SOC on K_{fs} (Figure 6b) was stronger in fine-textured soils as opposed to soils with higher sand content (Figure 6c). According to the 95% confidence intervals from the selected gamma GLMM model, there were no significant differences in K_{fs} between plots with or without visible signs of erosion for the average sand content; nonetheless, the presence of erosion in coarser soils had a negative relationship with K_{fs} (Figure 6d). The relationship with vegetation structure also varied depending on soil texture. Overall, K_{fs} increased with increasing tree cover, from grasslands to forests (Figure 6f). Higher sand content was associated with higher K_{fs} across all vegetation structure classes, particularly in the case of woodlands, bushlands, and forests (Figure 6g). Finally, grazing was associated with lower values of K_{fs} (Figure 6h). This negative pattern was more pronounced for finer soils with little sand than for soils with higher sand content, although sand content increased overall K_{fs}.

Table 3
Summary of the Optimal Gamma Generalized Linear Mixed-Effects Model Explaining Variation in Field-Saturated Hydraulic Conductivity (K_{fs})

Predictors	Estimates	CI	P
Intercept	3.56	3.42–3.71	<0.001
Sand	0.36	0.24–0.47	<0.001
SOC	0.17	0.10–0.24	<0.001
Erosion (presence)	−0.02	−0.09–0.05	0.613
VegStructure (cropland)	0.26	0.15–0.38	<0.001
VegStructure (shrubland)	0.08	−0.04–0.21	0.197
VegStructure (wooded grassland)	0.22	0.07–0.38	0.004
VegStructure (woodland)	0.30	0.15–0.46	<0.001
VegStructure (bushland)	0.33	0.18–0.48	<0.001
VegStructure (forest)	0.49	0.29–0.68	<0.001
ImpactGrazing	−0.07	−0.11 to −0.02	0.003
pH	−0.14	−0.19 to −0.08	<0.001
HerbCovRate	0.03	−0.01–0.07	0.129
Sand * SOC	−0.15	−0.22 to −0.08	<0.001
Sand * Erosion	−0.08	−0.15 to −0.01	0.029
Sand * VegStructure (cropland)	−0.18	−0.28 to −0.07	0.001
Sand * VegStructure (shrubland)	−0.19	−0.32 to −0.06	0.004
Sand * VegStructure (wooded grassland)	−0.20	−0.35 to −0.05	0.009
Sand * VegStructure (woodland)	−0.13	−0.28–0.01	0.072
Sand * VegStructure (bushland)	−0.04	−0.18–0.10	0.579
Sand * VegStructure (forest)	−0.08	−0.28–0.12	0.457
Sand * ImpactGrazing	0.05	0.01–0.09	0.014
Random effects			
ICC	0.21		
$N_{Cluster:Site}$	16		
N_{Site}	83		
Observations	3,573		
Marginal R^2/Conditional R^2	0.15/0.33		

Note. The standardized estimates of the fixed effects (beta coefficients), their associated 95% confidence intervals (CI) and corresponding p-values are shown (both based on Wald Z statistics). The random effects, number of observations, sites and clusters within sites, the marginal and conditional R^2, and the Interclass Correlation Coefficient (ICC) are included. The estimates of the fixed effects are based on the standardized values of continuous covariates. Sand: Sand content (%), SOC: Soil organic carbon content (g C kg^{-1}), Erosion: Presence of erosion (binary), VegStructure: Vegetation structure class, ImpactGrazing: Grazing impact rating, pH: Soil pH, HerbCoverRate: Herbaceous cover rating. The effects and corresponding p-values for the factor variables Erosion and VegStructure are in relation to their baseline levels (absence of erosion and grassland, respectively).

3.2.2. Random Forest Regression Model

Sand and clay content were the most important covariates for predicting K_{fs}, followed by SOC content, whereas the presence of erosion was the least important (Figure 7).

Partial dependence plots showed a positive relationship between K_{fs} and sand content, SOC, herbaceous and woody cover ratings, and a negative relationship with clay content, pH and grazing (Figure 8). The relationship between sand content and predicted K_{fs} was most pronounced between sand content values of 50% and 75% (Figure 8a). On the other hand, the relationship with clay content was strongest for values below 25%, while increases in clay content above this threshold did not have a substantial effect on predicted K_{fs} (Figure 8b). There was a positive relationship between SOC and predicted K_{fs} up to around 60 g kg^{-1}, and thereafter predicted K_{fs} leveled off and appeared to be unaffected by SOC (Figure 8c). However, SOC content above 60 g kg^{-1} was rare (only 1.5%

Figure 5. Fixed effects standardized estimates of the optimal gamma generalized linear mixed-effects model explaining field-saturated hydraulic conductivity (K_{fs}) and associated 95% confidence intervals. Asterisks indicate the significance level of the corresponding p-values (***$p < 0.001$, **$p < 0.01$, *$p < 0.05$). Sand: Sand content (%), SOC: Soil organic carbon content (g C kg^{-1}), Erosion: Presence of erosion (binary), VegStructure: Vegetation structure class, ImpactGrazing: Grazing impact rating, HerbCoverRate: Herbaceous cover rating. The effects for Erosion and VegStructure are in relation to their baseline levels (absence of erosion and grassland, respectively).

of samples were above this threshold); hence, this relationship should be interpreted cautiously. Predicted K_{fs} declined with increasing soil pH up to values around 7, while it remained low and relatively unchanged in the case of alkaline soils (Figure 8d). In the case of herbaceous cover, predicted K_{fs} varied little over the ratings 0–4, but increased between 4 and 5 (i.e., between 40%–65% and >65% herbaceous cover) (Figure 8i). The effect of grazing on predicted K_{fs} was strongest between ratings of 0 and 2 (Figure 8e). Predicted K_{fs} varied with vegetation structure class, and it was higher in classes with more trees (such as forest or bushland) compared to grasslands, shrublands or croplands (Figure 8g). Similarly, woody cover had a positive relationship with predicted K_{fs} (Figure 8h). Finally, predicted K_{fs} was slightly lower (57 vs. 59 mm hr^{-1}) in plots with erosion than in plots without erosion (Figure 8f).

10.1029/2023WR035510

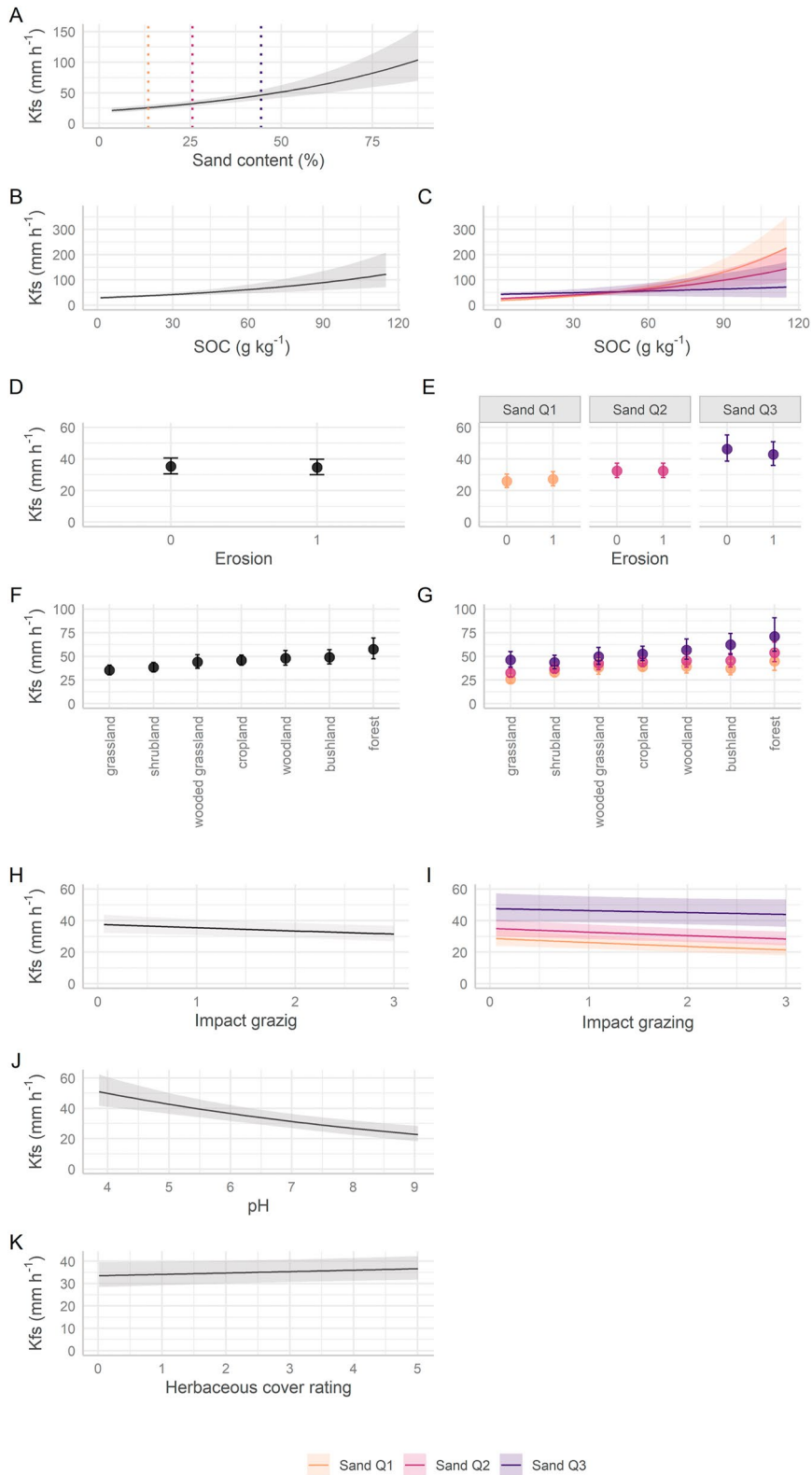

Figure 6.

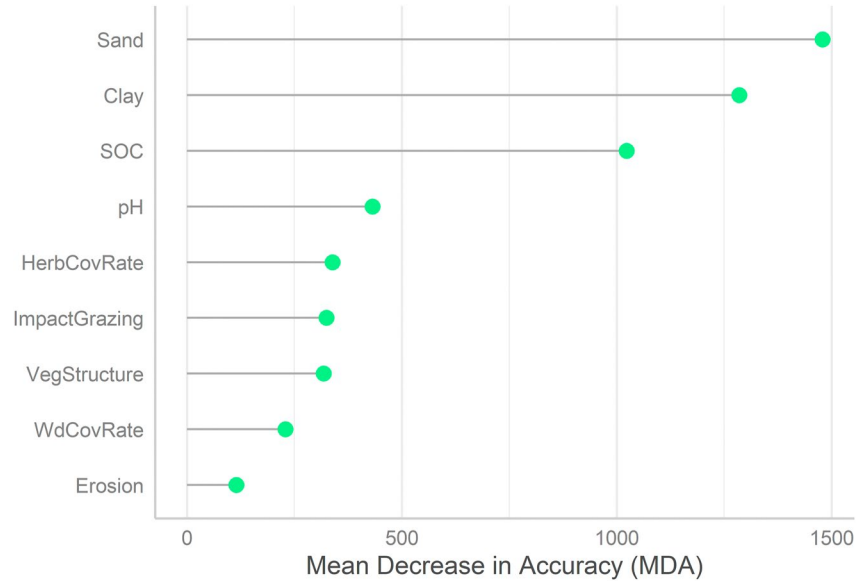

Figure 7. Random Forest variable importance plot showing the relative importance (Mean Decrease in Accuracy or Permutation Importance) of each covariate in predicting field-saturated hydraulic conductivity (K_{fs}) using a Random Forest regression model ($R^2 = 0.21$). Sand: Sand content (%), Clay: Clay content (%), SOC: Soil organic carbon (g C kg^{-1}), ImpactGrazing: Grazing impact rating, HerbCoverRate: Herbaceous cover rating, VegStructure: Vegetation structure class, WdCovRate: Woody cover rating, Erosion: Presence of erosion (binary).

4. Discussion

Previously, representative observations of soil hydraulic properties in Africa were few and of mixed form and quality (Gupta, Hengl, et al., 2021; Rahmati et al., 2018; Schaap et al., 2001; Wösten et al., 1999; Zhang & Schaap, 2017). Our data set expands existing regional and global databases of soil hydraulic properties, improving coverage for Africa and providing field data for underrepresented land uses and soils. As such, we envision that our data set can contribute to improved understanding and prediction of soil hydraulic properties and to improved Earth system and land surface models. In particular, we see great potential to validate and improve the performance of regional and global predictions of K_s. In our analyses of these unique data, we used established methods (Nimmo et al., 2009) to estimate topsoil K_{fs} from field measurements of soil infiltration in 3,573 plots across 83 sites in sub-Saharan Africa. Median K_{fs} across these sites spanned a similar range to median K_{fs} values for different soil texture classes from the SWIG database (6.8–160.9 vs. 2.8–150 mm hr^{-1}) (Rahmati et al., 2018). The positive correlation between within-site variability in K_{fs} and median K_{fs} is also consistent with previous findings (Nyberg et al., 2012; Takoutsing et al., 2022). In general, K_{fs} tended to be higher in more humid versus drier sites. This likely results from differences in inherent soil properties on the one hand and differences in net primary production and, subsequently, in SOC and soil structure on the other; these factors are discussed in detail below.

Results from the GLMM and Random Forest models show that soil sand content, SOC content, and woody cover had a positive relationship with K_{fs}, whereas grazing impact and soil pH had a negative relationship. Our findings also indicate that soil texture (sand and clay content) and SOC were the most important covariates for predicting K_{fs}. As expected, K_{fs} increased with sand content, with some variation depending on climate. The positive effect of the coarse-textured fractions on K_{fs} agrees with previous findings (Lehmann et al., 2021; Saxton et al., 1986; Tomasella et al., 2003). This relationship was less marked in more humid sites, likely due to potential differences in clay type and mineralogy between drier and more humid sites. While low-activity (non-swelling) clay

Figure 6. Marginal effects of the optimal gamma generalized linear mixed-effects model model explaining field-saturated hydraulic conductivity (K_{fs}) and associated 95% confidence intervals. Predicted values of K_{fs} are shown as a function of the main terms (left column) and the interaction terms (right column). Observe that the *y*-axes have different scales. The plots show the predicted K_{fs} as a function of the predictor variable(s) on the *x*-axis while considering mean values of all other continuous predictor variables and baseline levels of factor variables Erosion and Vegetation structure (absence of erosion and grassland, respectively). Interactions between sand content and soil organic carbon (c), erosion (e), vegetation structure class (g), and Impact Grazing (i) are illustrated based on sand content first (Q1), second (Q2) and third quartile (Q3) values, which are shown as vertical lines in (a).

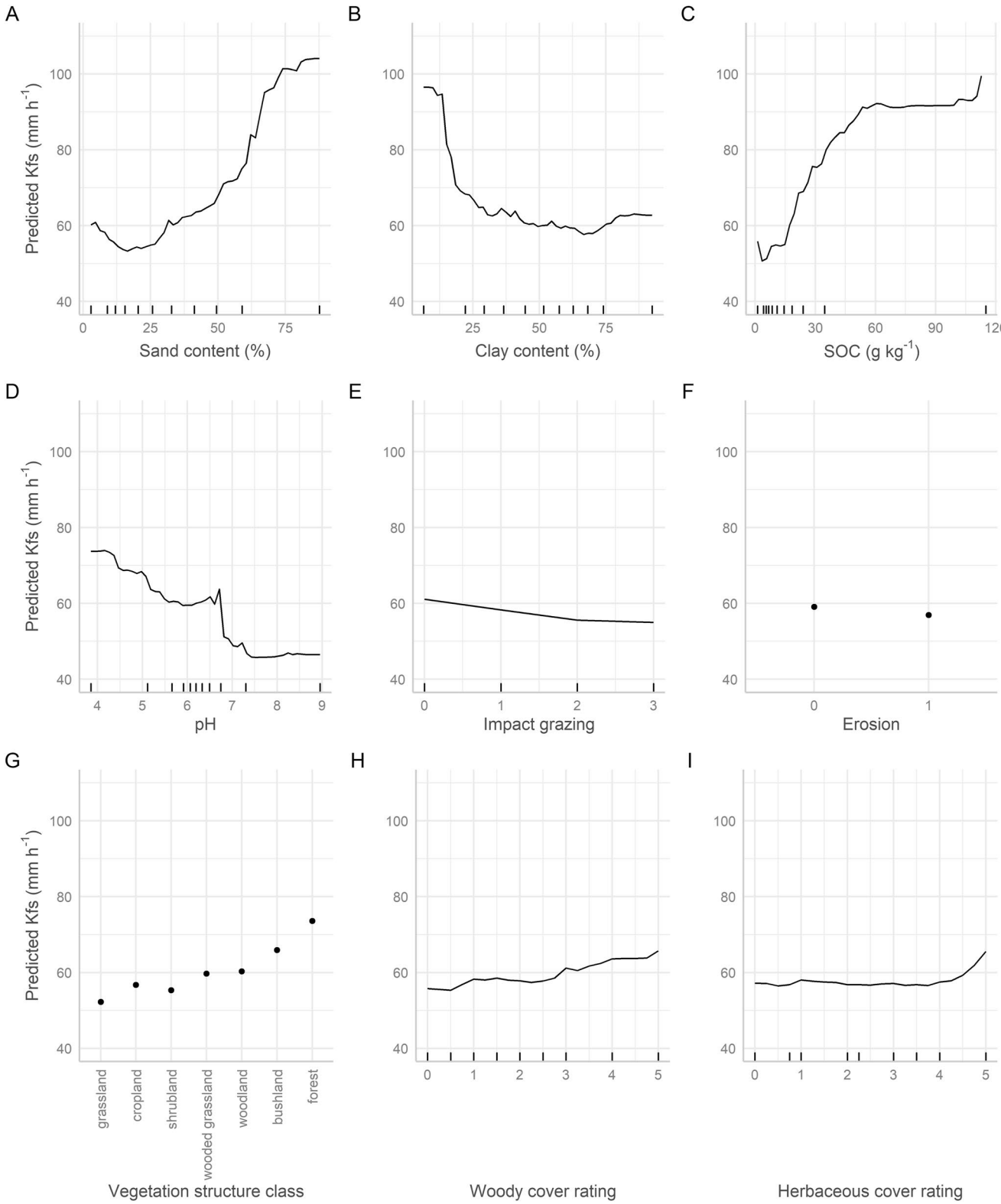

Figure 8. Partial dependence plots showing the marginal effect of the model covariates (A–J) on the predicted field-saturated hydraulic conductivity (K_{fs}) using a Random Forest regression model. Marks on the *x*-axis indicate the data distribution (deciles, minimum and maximum values).

minerals predominate in the humid tropics, more active (swelling) clay minerals are more common in drylands (Lehmann et al., 2021). Soils dominated by swelling clay minerals (e.g., smectite) are dispersive and, when saturated, have very low hydraulic conductivity. In contrast, weathered kaolinite-dominated soils such as Ferralsols (Oxisols) are highly permeable due to their stable micro-aggregated structure and low dispersibility (Hodnett & Tomasella, 2002; Lado & Ben-Hur, 2004; Lehmann et al., 2021; van den Berg et al., 1997). Our observations of extreme (>95th percentile) K_{fs} values occurring in soils with more than 50% clay content were surprising but may reflect such micro-aggregated structure or the presence of macropores such as those created by roots and soil fauna and cracks in some vertic soils. Our results also show a negative relationship between soil pH and $K_{fs,}$ with a substantial drop at around 7. Soil pH can indicate the degree of weathering and can thus also be related to clay mineralogy and exchangeable base cation concentration (Hodnett & Tomasella, 2002). Soils from drier sites tended to have higher pH values, which may reflect the dominance of more active and dispersive clays as well as a higher concentration of monovalent base cations that contribute to the dispersion of clay aggregates and soil structure degradation (Rengasamy et al., 2016) and, therefore, lower K_{fs}. Mills et al. (2006) also found a similar negative trend between pH and infiltrability for soils in Namibia and South Africa, which they ascribed to an increase in clay dispersibility with increasing pH. Future work needs to investigate further the effect of clay mineralogy in K_{fs}. Current soil databases used to train pedotransfer functions of soil hydraulic properties are strongly biased toward temperate regions with higher silt content and different clay minerals than tropical soils, which highlights the importance of increasing the amount of data from tropical soils in these databases when used for global predictions (Botula et al., 2012; Gupta, Hengl, et al., 2021; Tomasella et al., 2000). Our data set helps fill this gap. Our results also support the importance of considering dominant clay mineral types, not just clay fraction, when estimating soil hydraulic properties in the tropics (Lehmann et al., 2021).

Topsoil organic carbon content had a positive relationship with K_{fs}. SOC is crucial in supporting soil's physical quality (Dexter, 2004; Rawls et al., 2004) and is a major indicator of overall soil health (Bagnall et al., 2023; Cardoso et al., 2013; Lal, 2016). In general, SOC is considered to improve soil aggregation (Beare et al., 1994; Castro Filho et al., 2002) and porosity (Adams, 1973) and, as a result, saturated hydraulic conductivity (Lado, Paz, & Ben-Hur, 2004; Mbagwu & Auerswald, 1999). Our findings support this understanding and provide a broader evidence base. We also found that the implied effect of SOC on K_{fs} was texture-dependent, with a more substantial positive relationship for fine-textured soils than for coarser ones. This is consistent with the results of Araya and Ghezzehei (2019) for U.S. soils in the USKSAT database (Pachepsky & Park, 2015), who found that K_s increased with SOC across all textural classes except for sand and loamy sand. These results might be explained by the formation of organic matter and clay complexes that enhance soil porosity and reduce clay dispersibility in water (Dexter et al., 2008). Our results also indicate that, while at low to moderate SOC contents coarser soils exhibit higher K_{fs} than more fine-textured soils, this trend is reversed for high SOC contents. Our finding that K_{fs} is positively associated with SOC contrasts with other studies that found a negative relationship (Jarvis et al., 2013; Larsbo et al., 2016; Nemes et al., 2005; Wang et al., 2009). This negative relationship has been linked to greater soil water repellency and soil tortuosity as soil organic matter content increases. While our results indicate that the positive relationship between K_{fs} and SOC is less strong for coarser compared to more fine-textured soils, they do not support any negative K_s-SOC relationship. This suggests that, overall, in our data set, the positive effect of SOC on soil aggregation outweighs any negative effects of SOC on K_{fs}. Nevertheless, given that other studies found evidence of a negative K_{s-}SOC relationship, we recognize that this may still occur under specific conditions. Future studies should further explore these relationships and identify the conditions under which the relationship might be negative. Taken together, our results and interpretations suggest considerable potential to enhance soil hydrological functioning through the adoption of land use and management practices that increase SOC, particularly in the case of fine-textured soils. Such practices include agroforestry, cover crops, crop residue retention, reduced tillage, and optimal livestock stocking rates (Minasny et al., 2017; Paustian et al., 2016; Vågen et al., 2005). Increasing SOC can improve soil hydrological functioning while yielding many other benefits, from increased food and nutrition security to climate change mitigation (Johnston et al., 2009; Lal, 2006, 2010; Vågen & Winowiecki, 2013; Vågen et al., 2005), and should hence be key in land restoration efforts.

Vegetation structure was strongly associated with K_{fs}, with grasslands and forests exhibiting the lowest and highest K_{fs} values, respectively. These results and those from the Random Forest model suggest that woody cover, especially tree cover, has a positive impact on K_{fs}. These findings are consistent with a meta-analysis conducted by Ilstedt et al. (2007) showing that tree planting in the tropics, both in afforestation and agroforestry, improves soil infiltration capacity across a wide range of conditions and humidity levels. More recent studies

in the tropics further support these findings (Abaker et al., 2018; Benegas et al., 2014; Bonnesoeur et al., 2019; Filoso et al., 2017; Leite et al., 2018; Lozano-Baez et al., 2019, 2021; Mens et al., 2023; Niemeyer et al., 2014; Nyamadzawo et al., 2007; Zwartendijk et al., 2017). The positive impacts appear due to increased litter inputs, improved soil biological activity, and enhanced aggregation and macropore formation (e.g., root and faunal channels) (Bargués Tobella et al., 2014; Belsky et al., 1989, 1993; Niemeyer et al., 2014; Zwartendijk et al., 2017). Our results suggest that the effect of vegetation structure on K_{fs} is texture-dependent, with a stronger effect of tree cover on K_{fs} for more coarse-textured soils. This finding somehow disagrees with previous observations showing that the recovery of infiltration capacity following tree planting in the tropics is slower in the case of sandy soils (Lozano-Baez et al., 2019), although our data set does not allow testing this. Along the same lines, Niemeyer et al. (2014) found that the relationship between woody cover and K_{fs} along a vegetation gradient in dry tropical Nicaragua was greater for fine-textured soils than for coarser soils, which they ascribed to the impact of root-macropore formation on K_{fs} in fine-textured soils. Our results, which differ from these observations, may partly be explained by the so-called "inverse texture effect" observed in dryland ecosystems and by which coarser soils appear to support taller and denser perennial vegetation than finer soils (Gupta et al., 2022; Noy-Meir, 1973). Such apparently contradictory observations indicate that the combined relationship of woody vegetation and soil texture on K_{fs} is complex and requires further investigation. Overall, our findings support the importance of incorporating some measure of woody vegetation cover into models to predict K_{fs} (Bonetti et al., 2021; Niemeyer et al., 2014). Our results also show a smaller, but still positive, effect of herbaceous cover on K_{fs}. Results from the Random Forest model suggest that increases in herbaceous cover between 0% and 40% (ratings 0–4) have little relationship with K_{fs}, whereas increases above 40% have a larger positive effect, underscoring the value of maintaining a high herbaceous cover. Grasslands exhibited the lowest K_{fs} values of all vegetation structure classes, even compared to croplands. These results contrast with findings from a global meta-analysis on the impact of land use on soil hydraulic properties showing that grasslands had higher K values than comparable soil under cultivation (Robinson et al., 2022). Our vegetation structure classes do not fully reflect the land use intensity nor the land use history; many plots classified as grasslands were degraded, and some corresponded to abandoned croplands or land under fallow, which may help explain our results. Future studies using this data set might benefit from a more detailed analysis of the impacts of land use and land use change on K_{fs}. Different plant species will likely have different influences on infiltration via litter quality and quantity, root traits, microclimate, and associated soil fauna (Mens et al., 2023). This also raises important research questions for future studies.

Increasing grazing impact had a negative effect on K_{fs}, especially in fine-textured soils. This is in line with previous local studies showing that soil infiltration capacity decreases with grazing intensity across different land cover classes (Lulandala et al., 2022; Mwendera & Saleem, 1997; Savadogo et al., 2007; Talore et al., 2016; Vandandorj et al., 2017). The negative impact of grazing has been attributed primarily to soil compaction through animal trampling and reduced structural pore space, as indicated by higher soil bulk density values in more heavily grazed areas (Dudley et al., 2002; Lai & Kumar, 2020; Lulandala et al., 2022). In addition, heavy grazing can lead to reductions in plant cover and biomass, and SOC (Dlamini et al., 2016; Eldridge et al., 2016; Kikoti et al., 2015; Lai & Kumar, 2020; Lulandala et al., 2022; Savadogo et al., 2007), which can also contribute to decreasing infiltration capacity. Heavy livestock grazing and other anthropogenic activities can dilute or override the positive impacts of woody vegetation on K_{fs} (Ghimire et al., 2013, 2014; Lulandala et al., 2022). In this first study of this data set, we did not consider the interaction between grazing intensity and vegetation structure class due to the limited number of observations, which did not allow us to fit more complex models. This and other potentially relevant interactions should be investigated in future studies. Livestock grazing is a widespread land use in Sub-Saharan Africa, where millions rely on livestock for their livelihood (Robinson et al., 2011; Thornton, 2002). Our results highlight the importance of actively managing livestock stocking rates to prevent and reverse the degradation of soil hydrological functioning.

The presence of visible signs of erosion had a weak to no relationship with $K_{fs,}$ though for coarse-textured soils, K_{fs} tended to be lower in such plots. Erosion can lead to the formation of seals at the soil surface that reduce infiltration capacity, thus increasing overland flow and soil erosion (Singer & Le Bissonnais, 1998; Singer & Shainberg, 2004). Previous studies show that soils with intermediate clay content (*ca.* 10%–40%) have less stable aggregates and are more susceptible to seal formation and pore-clogging compared to soils with higher clay content (>40%), which may contribute to our results (Ben-Hur et al., 1985; Lado, Ben-Hur, & Shainberg, 2004). Similarly, Yair (1992) found that surface sealing in fine-textured soils was low due to the high stability and strong flocculation of clay-rich aggregates. Removal of SOC and reduced vegetation cover following erosion (Lal, 2005)

can also reduce infiltration capacity. In turn, decreased infiltration capacity can increase infiltration-excess overland flow, and, as a result, increase erosion further, especially under high rainfall intensities. This positive (self-reinforcing) feedback loop amplifies land degradation and can lead to a degraded state that is hard to reverse (D'Odorico et al., 2013; Scheffer & Carpenter, 2003). Erosion is one of the most critical and widespread land degradation processes in Sub-Saharan Africa and elsewhere (Montanarella et al., 2016; Vågen & Winowiecki, 2019). Improving soil infiltration capacity can help prevent and curb such processes.

Our findings support the view that soil texture is a major determinant of K_{fs}, while also underlining the additional role of soil structure and soil structure-altering processes. These additional factors are indicated by the positive relationships between K_{fs} and SOC and vegetation cover (in particular woody vegetation) and the negative relationship between K_{fs} and grazing and erosion. These relationships suggest significant opportunities to improve soil hydrological functioning by adopting land management practices that protect and enhance soil structure. Practices that reduce compaction and erosion, and/or enhance soil organic matter, vegetation cover, and soil microbial and faunal activity are likely to prove beneficial (Bronick & Lal, 2005; Meurer et al., 2020; Pagliai et al., 2004). The application of compost and manure, as well as mulching, alternative tillage systems, crop rotation, and improved grazing management, may all contribute (Bronick & Lal, 2005; Conant et al., 2017; Minasny et al., 2017; Pagliai et al., 2004). Protecting and promoting tree cover inside and outside forests, for instance, through agroforestry and maintaining trees on field boundaries, steeper slopes, and around drainage features, will likely be beneficial as well (Bargués Tobella et al., 2014; Ilstedt et al., 2016). Subsequent improvements in soil hydrological functioning can bolster the delivery of related ecosystem services — including water provisioning and purification, the regulation of hydrological flows, erosion control, biomass production, and biodiversity conservation (Bünemann et al., 2018; Keesstra et al., 2016). We thus recommend that land management practices that enhance soil structure and hydrological functioning be widely promoted across the region.

Our data and analyses represent a major improvement on previous studies but can be further improved. Our data set encompasses various climates and settings, although it may not fully represent sub-Saharan Africa. Notably, data remains limited in Southern and Central Africa, including the Congo Basin, and approximately 95% of our sites are not in arid climates. In terms of land use and cover, forests and woodlands remain underrepresented. In terms of measurements, we were constrained by the soil and land health indicators measured as part of the LDSF, which were selected in part for their cost-effectiveness, while other aspects that we have not considered could influence soil structure and K_{fs}. Variables that identify and capture variation in clay mineralogy, the soil biota, vegetation, and herbivory all warrant further attention. We also see a need to explore the role and form of SOC and its relationship to other soil and site properties in greater detail. In addition to increasing the data set and coverage, alternative statistical approaches, like Bayesian hierarchical modeling, may offer insights in future analyses. Our field surveys followed a systematic approach to minimize measurement errors and ensure consistency among the data collection teams. Remaining uncertainties relate to field measurements, particularly soil infiltration (Nimmo et al., 2009; Reynolds, 2008), and the estimation of K_{fs} (Bagarello et al., 2014; Nimmo et al., 2009; Reynolds, 2008). In our data set, we provide both the measured and estimated K_{fs} values. This permits others to use and assess different estimation methods beyond the scope of this study.

5. Conclusions

Taken together, results from this study indicate that, across the 83 study sites in sub-Saharan Africa, K_{fs} is linked with soil texture and soil structure, which highlights the importance of considering both when estimating K_{fs}, as suggested by others (e.g., Fatichi et al., 2020 or Bonetti et al., 2021). Incorporating attributes that influence soil structure when estimating K_{fs} is especially urgent for applications in the tropics, where the impacts of soil structure on hydrologic response are more prominent due to the prevalence of high-intensity rainfall events and soils with high contents of clay (Bonell, 1993; Bonetti et al., 2021; Tomasella et al., 2000; Wohl et al., 2012). We provide evidence for the importance of structure-altering processes, including compaction, aggregation, and burrowing by roots and soil fauna, in determining K_{fs}, as indicated by the relationships between K_{fs} and SOC, grazing impact, vegetation cover, pH, and erosion. Our findings also suggest that soil texture regulates many of these relationships. Hence, while inherent soil properties influence K_{fs}, dynamic properties affected by land use and management are also crucial. This underscores the potential of management and restoration practices to maintain and promote soil hydrological functioning. Effort must be placed on promoting practices that improve soil

structure and avoiding those that worsen it. Enhancing SOC content, limiting livestock stocking rates, promoting vegetation cover, particularly woody vegetation, and preventing and halting soil erosion, can thus maintain and enhance K_{fs}. This evidence can guide sustainable land management practices and restoration interventions for improved soil health and water security.

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

Data Availability Statement

Field and lab data used in this study are freely available on the Figshare repository at https://doi.org/10.6084/m9.figshare.23284913. Associated plot coordinates can be made available upon request.

Acknowledgments
This study was conducted using data collected with the financial support from various donors, including: the European Union Directorate General for International Partnerships through the Regreening Africa Programme (Grant DCI-ENV/2017/387–627); The Bill and Melinda Gates Foundation through the Africa Soil Information Service (AfSIS) project, phase one (Grant 51353); The Kenya Cereal Enhancement Programme - Climate Resilient Agricultural Livelihoods (KCEP-CRAL) supported by the Government of Kenya and the International Fund for Agricultural Development (IFAD); the Adaptation for Smallholder Agriculture Program (Grant 2000000927) supported by the Government of the Republic of Chad; the International Fund for Agricultural Development (IFAD) (Grant 2000000925); the International Fund for Agricultural Development (IFAD) through the project Beyond the Static (Grant 2000001302); the International Fund for Agricultural Development (IFAD) and the European Commission through the project on Restoration of degraded lands for food security and poverty reduction in East Africa and the Sahel: taking successes to scale (Grants 2000000520 and 2000000976); The Nature Conservancy; the Northern Rangeland Trust; Mpala Research Center; and the CGIAR Research Programmes on Forests, Trees and Agroforestry and Water, Land and Ecosystems. This research was supported by a grant from the Swedish Research Council for Sustainable Development, Formas (Grant 2017-00430); We also acknowledge funding from the Swedish Research Council, VR (Grant 2017-05566). We are grateful for the contributions of the field teams that collected these data. We would like to thank the farmers, pastoralists, and local authorities for their collaboration. We also thank the staff of the CIFOR-ICRAF Soil and Land Health Laboratories for the pre-processing and analysis of soil samples. We are grateful to Muhamad Nabi Ahmed and Benard Onkware at CIFOR-ICRAF for setting up the electronic data entry forms.

References

Abaker, W. E., Berninger, F., & Starr, M. (2018). Changes in soil hydraulic properties, soil moisture and water balance in Acacia Senegal plantations of varying age in Sudan. *Journal of Arid Environments*, *150*, 42–53. https://doi.org/10.1016/j.jaridenv.2017.12.004

Abdelkadir, A., & Yimer, F. (2011). Soil water property variations in three adjacent land use types in the Rift Valley area of Ethiopia. *Journal of Arid Environments*, *75*(11), 1067–1071. https://doi.org/10.1016/j.jaridenv.2011.06.012

Adams, W. (1973). The effect of organic matter on the bulk and true densities of some uncultivated podzolic soils. *Journal of Soil Science*, *24*(1), 10–17. https://doi.org/10.1111/j.1365-2389.1973.tb00737.x

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. https://doi.org/10.1029/2018wr024357

Bagarello, V., Di Prima, S., Iovino, M., & Provenzano, G. (2014). Estimating field-saturated soil hydraulic conductivity by a simplified Beerkan infiltration experiment. *Hydrological Processes*, *28*(3), 1095–1103. https://doi.org/10.1002/hyp.9649

Bagarello, V., Iovino, M., & Elrick, D. (2004). A simplified falling-head technique for rapid determination of field-saturated hydraulic conductivity. *Soil Science Society of America Journal*, *68*(1), 66–73. https://doi.org/10.2136/sssaj2004.6600

Bagnall, D. K., Rieke, E. L., Morgan, C. L. S., Liptzin, D. L., Cappellazzi, S. B., & Honeycutt, C. W. (2023). A minimum suite of soil health indicators for North American agriculture. *Soil Security*, *10*, 100084. https://doi.org/10.1016/j.soisec.2023.100084

Bargués Tobella, A., Reese, H., Almaw, A., Bayala, J., Malmer, A., Laudon, H., & Ilstedt, U. (2014). The effect of trees on preferential flow and soil infiltrability in an agroforestry parkland in semiarid Burkina Faso. *Water Resources Research*, *50*(4), 3342–3354. https://doi.org/10.1002/2013WR015197

Beare, M. H., Hendrix, P. F., & Coleman, D. C. (1994). Water-stable aggregates and organic matter fractions in conventional- and no-tillage soils. *Soil Science Society of America Journal*, *58*(3), 777–786. https://doi.org/10.2136/sssaj1994.03615995005800030020x

Belsky, A. J., Amundson, R. G., Duxbury, J. M., Riha, S. J., Ali, A. R., & Mwonga, S. M. (1989). The effect of trees on their physical, chemical and biological environments in a semi-arid savanna in Kenya. *Journal of Applied Ecology*, *26*(3), 1005–1024. https://doi.org/10.2307/2403708

Belsky, A. J., Mwonga, S. M., Amundson, R. G., Duxbury, J. M., & Ali, A. R. (1993). Comparative effects of isolated trees on their undercanopy environments in high- and low-rainfall savannas. *Journal of Applied Ecology*, *30*(1), 143–155. https://doi.org/10.2307/2404278

Benegas, L., Ilstedt, U., Roupsard, O., Jones, J., & Malmer, A. (2014). Effects of trees on infiltrability and preferential flow in two contrasting agroecosystems in Central America. *Agriculture, Ecosystems & Environment*, *183*, 185–196. https://doi.org/10.1016/j.agee.2013.10.027

Ben-Hur, M., Shainberg, I., Bakker, D., & Keren, R. (1985). Effect of soil texture and CaCO3 content on water infiltration in crusted soil as related to water salinity. *Irrigation Science*, *6*(4), 281–294. https://doi.org/10.1007/BF00262473

Beven, K., & Germann, P. (2013). Macropores and water flow in soils revisited. *Water Resources Research*, *49*(6), 3071–3092. https://doi.org/10.1002/wrcr.20156

Biau, G., & Scornet, E. (2016). A random forest guided tour. *TEST*, *25*(2), 197–227. https://doi.org/10.1007/s11749-016-0481-7

Bonell, M. (1993). Progress in the understanding of runoff generation dynamics in forests. *Journal of Hydrology*, *150*(2), 217–275. https://doi.org/10.1016/0022-1694(93)90112-m

Bonell, M., Purandara, B. K., Venkatesh, B., Krishnaswamy, J., Acharya, H. A. K., Singh, U. V., et al. (2010). The impact of forest use and reforestation on soil hydraulic conductivity in the Western Ghats of India: Implications for surface and sub-surface hydrology. *Journal of Hydrology*, *391*(1–2), 47–62. https://doi.org/10.1016/j.jhydrol.2010.07.004

Bonetti, S., Wei, Z., & Or, D. (2021). A framework for quantifying hydrologic effects of soil structure across scales. *Communications Earth & Environment*, *2*(1), 107. https://doi.org/10.1038/s43247-021-00180-0

Bonnesoeur, V., Locatelli, B., Guariguata, M. R., Ochoa-Tocachi, B. F., Vanacker, V., Mao, Z., et al. (2019). Impacts of forests and forestation on hydrological services in the Andes: A systematic review. *Forest Ecology and Management*, *433*, 569–584. https://doi.org/10.1016/j.foreco.2018.11.033

Bormann, H., Fass, T., Giertz, S., Junge, B., Diekkruger, B., Reichert, B., & Skowronek, A. (2005). From local hydrological process analysis to regional hydrological model application in Benin: Concept, results and perspectives. *Physics and Chemistry of the Earth*, *30*(6–7), 347–356. Proceedings Paper. https://doi.org/10.1016/j.pce.2005.06.005

Botula, Y. D., Cornelis, W. M., Baert, G., & Van Ranst, E. (2012). Evaluation of pedotransfer functions for predicting water retention of soils in Lower Congo (D.R. Congo). *Agricultural Water Management*, *111*, 1–10. https://doi.org/10.1016/j.agwat.2012.04.006

Bouwer, H. (1986). Intake rate: Cylinder infiltrometer. In A. Klute (Ed.), *Methods of soil analysis. Part 1. Physical and mineralogical methods* (pp. 825–844). American Society of Agronomy, Inc.

Brauman, K. A., Daily, G. C., Duarte, T. K. e., & Mooney, H. A. (2007). The nature and value of ecosystem services: An overview highlighting hydrologic services. *Annual Review of Environment and Resources*, *32*(1), 67–98. https://doi.org/10.1146/annurev.energy.32.031306.102758

Braun-Blanquet, J. (1932). Plant sociology. The study of plant communities. In *Plant sociology. The study of plant communities* (1st ed.).

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/a:1010933404324

Bronick, C. J., & Lal, R. (2005). Soil structure and management: A review. *Geoderma*, *124*(1), 3–22. https://doi.org/10.1016/j.geoderma.2004.03.005
Brooks, M. E., Kristensen, K., Van Benthem, K. J., Magnusson, A., Berg, C. W., Nielsen, A., et al. (2017). glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *The R journal*, *9*(2), 378–400. https://doi.org/10.32614/rj-2017-066
Bruijnzeel, L. A. (2004). Hydrological functions of tropical forests: Not seeing the soil for the trees? *Agriculture, Ecosystems & Environment*, *104*(1), 185–228. Proceedings Paper. https://doi.org/10.1016/j.agee.2004.01.015
Bünemann, E. K., Bongiorno, G., Bai, Z., Creamer, R. E., De Deyn, G., de Goede, R., et al. (2018). Soil quality – A critical review. *Soil Biology and Biochemistry*, *120*, 105–125. https://doi.org/10.1016/j.soilbio.2018.01.030
Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference: Understanding AIC and BIC in model selection. *Sociological Methods & Research*, *33*(2), 261–304. https://doi.org/10.1177/0049124104268644
Cardoso, E., Vasconcellos, R. L. F., Bini, D., Miyauchi, M. Y. H., dos Santos, C. A., Alves, P. R. L., et al. (2013). Soil health: Looking for suitable indicators. What should be considered to assess the effects of use and management on soil health? *Scientia Agricola*, *70*(4), 274–289. https://doi.org/10.1590/s0103-90162013000400009
Castro Filho, C., Lourenço, A., de F Guimarães, M. M., & Fonseca, I. C. B. (2002). Aggregate stability under different soil management systems in a red latosol in the state of Parana, Brazil. *Soil and Tillage Research*, *65*(1), 45–51. https://doi.org/10.1016/s0167-1987(01)00275-6
Chambers, J., & Hastie, T. (1992). Linear models. Chapter 4. In *Statistical models in S*. Wadsworth & Brooks/Cole.
Colloff, M. J., Pullen, K. R., & Cunningham, S. A. (2010). Restoration of an ecosystem function to revegetation communities: The role of Invertebrate macropores in enhancing soil water infiltration. *Restoration Ecology*, *18*(s1), 65–72. https://doi.org/10.1111/j.1526-100x.2010.00667.x
Conant, R. T., Cerri, C. E. P., Osborne, B. B., & Paustian, K. (2017). Grassland management impacts on soil carbon stocks: A new synthesis. *Ecological Applications*, *27*(2), 662–668. https://doi.org/10.1002/eap.1473
Dexter, A. R. (2004). Soil physical quality: Part I. Theory, effects of soil texture, density, and organic matter, and effects on root growth. *Geoderma*, *120*(3), 201–214. https://doi.org/10.1016/j.geoderma.2003.09.004
Dexter, A. R., Richard, G., Arrouays, D., Czyż, E. A., Jolivet, C., & Duval, O. (2008). Complexed organic matter controls soil physical properties. *Geoderma*, *144*(3), 620–627. https://doi.org/10.1016/j.geoderma.2008.01.022
Di Gregorio, A., & Jansen, L. J. (1998). A new concept for a land cover classification system. *The Land*, *2*(1), 55–65.
Dlamini, P., Chivenge, P., & Chaplot, V. (2016). Overgrazing decreases soil organic carbon stocks the most under dry climates and low soil pH: A meta-analysis shows. *Agriculture, Ecosystems & Environment*, *221*, 258–269. https://doi.org/10.1016/j.agee.2016.01.026
D'Odorico, P., Bhattachan, A., Davis, K. F., Ravi, S., & Runyan, C. W. (2013). Global desertification: Drivers and feedbacks. *Advances in Water Resources*, *51*, 326–344. https://doi.org/10.1016/j.advwatres.2012.01.013
Dudley, D. M., Tate, K. W., McDougald, N. K., & George, M. R. (2002). Factors influencing soil-surface bulk density on oak savanna rangeland in the southern Sierra Nevada foothills. Paper presented at the Standiford, RB; McCreary, D.; Purcell, KL, technical coordinators. In *Proceedings, fifth symposium on oak woodlands: Oaks in California's changing landscape. Gen. Tech. Rep. PSW-GTR-184*. Pacific Southwest Research Station, Forest Service, US Department of Agriculture.
Eldridge, D. J., Poore, A. G. B., Ruiz-Colmenero, M., Letnic, M., & Soliveres, S. (2016). Ecosystem structure, function, and composition in rangelands are negatively affected by livestock grazing. *Ecological Applications*, *26*(4), 1273–1283. https://doi.org/10.1890/15-1234
Elrick, D., Reynolds, W., & Tan, K. (1989). Hydraulic conductivity measurements in the unsaturated zone using improved well analyses. *Groundwater Monitoring & Remediation*, *9*(3), 184–193. https://doi.org/10.1111/j.1745-6592.1989.tb01162.x
Falkenmark, M. (1989). The massive water scarcity now threatening Africa: Why isn't it being addressed? *Ambio*, 112–118. https://www.jstor.org/stable/4313541
Fatichi, S., Or, D., Walko, R., Vereecken, H., Young, M. H., Ghezzehei, T. A., et al. (2020). Soil structure is an important omission in Earth System Models. *Nature Communications*, *11*(1), 522. https://doi.org/10.1038/s41467-020-14411-z
Ferré, T. P. A., & Warrik, A. W. (2004). Infiltration. In D. Hillel (Ed.), *Encyclopedia of soils in the environment* (Vol. 4, pp. 254–260). Elsevier Ltd.
Filoso, S., Bezerra, M. O., Weiss, K. C. B., & Palmer, M. A. (2017). Impacts of forest restoration on water yield: A systematic review. *PLoS One*, *12*(8), e0183210. https://doi.org/10.1371/journal.pone.0183210
Franzluebbers, A. J. (2002). Water infiltration and soil structure related to organic matter and its stratification with depth. *Soil and Tillage Research*, *66*(2), 197–205. Proceedings Paper. https://doi.org/10.1016/s0167-1987(02)00027-2
Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, *29*(5), 1189–1232. https://doi.org/10.1214/aos/1013203451
Ghimire, C. P., Bonell, M., Bruijnzeel, L. A., Coles, N. A., & Lubczynski, M. W. (2013). Reforesting severely degraded grassland in the Lesser Himalaya of Nepal: Effects on soil hydraulic conductivity and overland flow production. *Journal of Geophysical Research: Earth Surface*, *118*(4), 2528–2545. https://doi.org/10.1002/2013jf002888
Ghimire, C. P., Bruijnzeel, L. A., Bonell, M., Coles, N., Lubczynski, M. W., & Gilmour, D. A. (2014). The effects of sustained forest use on hillslope soil hydraulic conductivity in the Middle Mountains of Central Nepal. *Ecohydrology*, *7*(2), 478–495. https://doi.org/10.1002/eco.1367
Greenwell, B. M. (2017). pdp: An R package for constructing partial dependence plots. *R J*, *9*(1), 421. https://doi.org/10.32614/rj-2017-016
Grömping, U. (2015). Variable importance in regression models. *WIREs Computational Statistics*, *7*(2), 137–152. https://doi.org/10.1002/wics.1346
Gupta, S., Bonetti, S., Lehmann, P., & Or, D. (2022). Limited role of soil texture in mediating natural vegetation response to rainfall anomalies. *Environmental Research Letters*, *17*(3), 034012. https://doi.org/10.1088/1748-9326/ac5206
Gupta, S., Hengl, T., Lehmann, P., Bonetti, S., & Or, D. (2021). SoilKsatDB: Global database of soil saturated hydraulic conductivity measurements for geoscience applications. *Earth System Science Data*, *13*(4), 1593–1612. https://doi.org/10.5194/essd-13-1593-2021
Gupta, S., Lehmann, P., Bonetti, S., Papritz, A., & Or, D. (2021). Global prediction of soil saturated hydraulic conductivity using random forest in a Covariate-Based GeoTransfer Function (CoGTF) framework. *Journal of Advances in Modeling Earth Systems*, *13*(4), e2020MS002242. https://doi.org/10.1029/2020ms002242
Harrison, X. A., Donaldson, L., Correa-Cano, M. E., Evans, J., Fisher, D. N., Goodwin, C. E., et al. (2018). A brief introduction to mixed effects modelling and multi-model inference in ecology. *PeerJ*, *6*, e4794. https://doi.org/10.7717/peerj.4794
Heino, M., Puma, M. J., Ward, P. J., Gerten, D., Heck, V., Siebert, S., & Kummu, M. (2018). Two-thirds of global cropland area impacted by climate oscillations. *Nature Communications*, *9*(1), 1257. https://doi.org/10.1038/s41467-017-02071-5
Hillel, D. (1971). Soil and water: Physical principles and processes. In *Physiological ecology: A series of monographs and treatises*. Academic Press.
Hillel, D. (1998). *Environmental soil physics: Fundamentals, applications, and environmental considerations*. Elsevier.
Hodnett, M. G., & 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), 155–180. https://doi.org/10.1016/s0016-7061(02)00105-2

Horton, R. E. (1941). An approach toward a physical interpretation of infiltration-capacity 1. *Soil Science Society of America Journal*, *5*(C), 399–417. https://doi.org/10.2136/sssaj1941.036159950005000c0075x
Hox, J. J., Moerbeek, M., & Van de Schoot, R. (2017). *Multilevel analysis: Techniques and applications*. Routledge.
Iizumi, T., Luo, J.-J., Challinor, A. J., Sakurai, G., Yokozawa, M., Sakuma, H., et al. (2014). Impacts of El Niño Southern Oscillation on the global yields of major crops. *Nature Communications*, *5*(1), 3712. https://doi.org/10.1038/ncomms4712
Ilstedt, U., Bargués Tobella, A., Bazié, H. R., Bayala, J., Verbeeten, E., Nyberg, G., et al. (2016). Intermediate tree cover can maximize groundwater recharge in the seasonally dry tropics. *Scientific Reports*, *6*(1), 21930. https://doi.org/10.1038/srep21930
Ilstedt, U., Malmer, A., Verbeeten, E., & Murdiyarso, D. (2007). The effect of afforestation on water infiltration in the Tropics: A systematic review and meta-analysis. *Forest Ecology and Management*, *251*(1–2), 45–51. https://doi.org/10.1016/j.foreco.2007.06.014
Jarvis, N., Koestel, J., Messing, I., Moeys, J., & Lindahl, A. (2013). Influence of soil, land use and climatic factors on the hydraulic conductivity of soil. *Hydrology and Earth System Sciences*, *17*(12), 5185–5195. https://doi.org/10.5194/hess-17-5185-2013
Johnston, A. E., Poulton, P. R., & Coleman, K. (2009). Chapter 1. Soil organic matter: Its importance in sustainable agriculture and carbon dioxide fluxes. In D. L. Sparks (Ed.), *Advances in agronomy* (Vol. 101, pp. 1–57). Academic Press.
Kattel, G. R. (2019). State of future water regimes in the world's river basins: Balancing the water between society and nature. *Critical Reviews in Environmental Science and Technology*, *49*(12), 1107–1133. https://doi.org/10.1080/10643389.2019.1579621
Keesstra, S. D., Bouma, J., Wallinga, J., Tittonell, P., Smith, P., Cerdà, A., et al. (2016). The significance of soils and soil science towards realization of the United Nations Sustainable Development Goals. *Soil*, *2*(2), 111–128. https://doi.org/10.5194/soil-2-111-2016
Kikoti, I. A., Mligo, C., & Kilemo, D. B. (2015). The impact of grazing on plant natural regeneration in northern slopes of Mount Kilimanjaro, Tanzania. *Open Journal of Ecology*, *5*(06), 266–273. https://doi.org/10.4236/oje.2015.56021
Krishnaswamy, J., Bonell, M., Venkatesh, B., Purandara, B. K., Rakesh, K. N., Lele, S., et al. (2013). The groundwater recharge response and hydrologic services of tropical humid forest ecosystems to use and reforestation: Support for the "infiltration-evapotranspiration trade-off" hypothesis. *Journal of Hydrology*, *498*, 191–209. https://doi.org/10.1016/j.jhydrol.2013.06.034
Kuhn, M., & Wickham, H. (2020). Tidymodels: A collection of packages for modeling and machine learning using tidyverse principles.
Kummu, M., Guillaume, J., de Moel, H., Eisner, S., Flörke, M., Porkka, M., et al. (2016). The world's road to water scarcity: Shortage and stress in the 20th century and pathways towards sustainability. *Scientific Reports*, *6*(1), 38495. https://doi.org/10.1038/srep38495
Lado, M., & Ben-Hur, M. (2004). Soil mineralogy effects on seal formation, runoff and soil loss. *Applied Clay Science*, *24*(3), 209–224. https://doi.org/10.1016/j.clay.2003.03.002
Lado, M., Ben-Hur, M., & Shainberg, I. (2004). Soil wetting and texture effects on aggregate stability, seal formation, and erosion. *Soil Science Society of America Journal*, *68*(6), 1992–1999. https://doi.org/10.2136/sssaj2004.1992
Lado, M., Paz, A., & Ben-Hur, M. (2004). Organic matter and aggregate size interactions in infiltration, seal formation, and soil loss contribution from the Agricultural Research Organization, The Volcani Center, no. 623/02, 2002 series. *Soil Science Society of America Journal*, *68*(3), 935–942. https://doi.org/10.2136/sssaj2004.9350
Lai, L., & Kumar, S. (2020). A global meta-analysis of livestock grazing impacts on soil properties. *PLoS One*, *15*(8), e0236638. https://doi.org/10.1371/journal.pone.0236638
Lal, R. (1996). Deforestation and land-use effects on soil degradation and rehabilitation in western Nigeria. I. Soil physical and hydrological properties. *Land Degradation & Development*, *7*(1), 19–45. https://doi.org/10.1002/(sici)1099-145x(199603)7:1<19::aid-ldr212>3.3.co;2-d
Lal, R. (2005). Soil erosion and carbon dynamics. *Soil and Tillage Research*, *81*(2), 137–142. https://doi.org/10.1016/j.still.2004.09.002
Lal, R. (2006). Enhancing crop yields in the developing countries through restoration of the soil organic carbon pool in agricultural lands. *Land Degradation & Development*, *17*(2), 197–209. https://doi.org/10.1002/ldr.696
Lal, R. (2010). Enhancing eco-efficiency in agro-ecosystems through soil carbon sequestration. *Crop Science*, *50*(S1), S120–S131. https://doi.org/10.2135/cropsci2010.01.0012
Lal, R. (2016). Soil health and carbon management. *Food and Energy Security*, *5*(4), 212–222. https://doi.org/10.1002/fes3.96
Lal, R., & Shukla, M. K. (2004). *Principles of soil physics*. CRC Press.
Larsbo, M., Koestel, J., Kätterer, T., & Jarvis, N. (2016). Preferential transport in macropores is reduced by soil organic carbon. *Vadose Zone Journal*, *15*(9), 1–7. https://doi.org/10.2136/vzj2016.03.0021
Lavelle, P., Decaëns, T., Aubert, M., Barot, S., Blouin, M., Bureau, F., et al. (2006). Soil invertebrates and ecosystem services. *European Journal of Soil Biology*, *42*, S3–S15. https://doi.org/10.1016/j.ejsobi.2006.10.002
Leal Filho, W., Totin, E., Franke, J. A., Andrew, S. M., Abubakar, I. R., Azadi, H., et al. (2022). Understanding responses to climate-related water scarcity in Africa. *Science of the Total Environment*, *806*, 150420. https://doi.org/10.1016/j.scitotenv.2021.150420
Lehmann, P., Leshchinsky, B., Gupta, S., Mirus, B. B., Bickel, S., Lu, N., & Or, D. (2021). Clays are not created equal: How clay mineral type affects soil parameterization. *Geophysical Research Letters*, *48*(20), e2021GL095311. https://doi.org/10.1029/2021GL095311
Leite, P. A. M., de Souza, E. S., dos Santos, E. S., Gomes, R. J., Cantalice, J. R., & Wilcox, B. P. (2018). The influence of forest regrowth on soil hydraulic properties and erosion in a semiarid region of Brazil. *Ecohydrology*, *11*(3), e1910. https://doi.org/10.1002/eco.1910
Loukas, A., Vasiliades, L., & Tzabiras, J. (2008). Climate change effects on drought severity. *Advances in Geosciences*, *17*, 23–29. https://doi.org/10.5194/adgeo-17-23-2008
Lozano-Baez, S. E., Cooper, M., Meli, P., Ferraz, S. F. B., Rodrigues, R. R., & Sauer, T. J. (2019). Land restoration by tree planting in the tropics and subtropics improves soil infiltration, but some critical gaps still hinder conclusive results. *Forest Ecology and Management*, *444*, 89–95. https://doi.org/10.1016/j.foreco.2019.04.046
Lozano-Baez, S. E., Domínguez-Haydar, Y., Zwartendijk, B. W., Cooper, M., Tobón, C., & Di Prima, S. (2021). Contrasts in top soil infiltration processes for degraded vs. restored lands. A case study at the Perijá range in Colombia. *Forests*, *12*(12), 1716. https://doi.org/10.3390/f12121716
Lüdecke, D., Ben-Shachar, M. S., Patil, I., Waggoner, P., & Makowski, D. (2021). Performance: An R package for assessment, comparison and testing of statistical models. *Journal of Open Source Software*, *6*(60), 3139. https://doi.org/10.21105/joss.03139
Lulandala, L., Bargués-Tobella, A., Masao, C. A., Nyberg, G., & Ilstedt, U. (2022). Excessive livestock grazing overrides the positive effects of trees on infiltration capacity and modifies preferential flow in dry miombo woodlands. *Land Degradation & Development*, *33*(4), 581–595. https://doi.org/10.1002/ldr.4149
Martins, A. L. d. S., Teixeira, W. G., Ottoni, M. V., Lima, M. M. d. G. d., Pimentel, L. G., & Reis, A. M. H. d. (2022). Infiltração e condutividade hidráulica saturada nos solos do Estado do Rio de Janeiro [Survey data; Experimental data]. https://doi.org/10.48432/YBZTGJ
Mbagwu, J. S. C., & Auerswald, K. (1999). Relationship of percolation stability of soil aggregates to land use, selected properties, structural indices and simulated rainfall erosion. *Soil and Tillage Research*, *50*(3), 197–206. https://doi.org/10.1016/s0167-1987(99)00006-9
MEA. (2005). *Ecosystems and human well-being: Synthesis*. Island Press.

Mekonnen, M. M., & Hoekstra, A. Y. (2016). Four billion people facing severe water scarcity. *Science Advances*, *2*(2), e1500323. https://doi.org/10.1126/sciadv.1500323
Mens, L. P., Bargués-Tobella, A., Sterck, F., Vågen, T.-G., Winowiecki, L. A., & Lohbeck, M. (2023). Towards effectively restoring agricultural landscapes in East African drylands: Linking plant functional traits with soil hydrology. *Journal of Applied Ecology*, *60*(1), 91–100. https://doi.org/10.1111/1365-2664.14311
Meurer, K., Barron, J., Chenu, C., Coucheney, E., Fielding, M., Hallett, P., et al. (2020). A framework for modelling soil structure dynamics induced by biological activity. *Global Change Biology*, *26*(10), 5382–5403. https://doi.org/10.1111/gcb.15289
Mills, A. J., Fey, M. V., Gröngröft, A., Petersen, A., & Medinski, T. V. (2006). Unravelling the effects of soil properties on water infiltration: Segmented quantile regression on a large data set from arid south-west Africa. *Soil Research*, *44*(8), 783–797. https://doi.org/10.1071/SR05180
Minasny, B., Malone, B. P., McBratney, A. B., Angers, D. A., Arrouays, D., Chambers, A., et al. (2017). Soil carbon 4 per mille. *Geoderma*, *292*, 59–86. https://doi.org/10.1016/j.geoderma.2017.01.002
Molnar, C. (2022). Interpretable machine learning: A guide for making black box models explainable (2nd ed.).
Montanarella, L., Pennock, D. J., McKenzie, N., Badraoui, M., Chude, V., Baptista, I., et al. (2016). World's soils are under threat. *Soil*, *2*(1), 79–82. https://doi.org/10.5194/soil-2-79-2016
Moreno, F., Pelegrín, F., Fernández, J. E., & Murillo, J. M. (1997). Soil physical properties, water depletion and crop development under traditional and conservation tillage in southern Spain. *Soil and Tillage Research*, *41*(1), 25–42. https://doi.org/10.1016/s0167-1987(96)01083-5
Mwendera, E. J., & Saleem, M. A. M. (1997). Infiltration rates, surface runoff, and soil loss as influenced by grazing pressure in the Ethiopian highlands. *Soil Use & Management*, *13*(1), 29–35. https://doi.org/10.1111/j.1475-2743.1997.tb00553.x
Nakagawa, S., Johnson, P. C., & Schielzeth, H. (2017). The coefficient of determination R 2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. *Journal of The Royal Society Interface*, *14*(134), 20170213. https://doi.org/10.1098/rsif.2017.0213
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. https://doi.org/10.2136/sssaj2004.0055
Nielsen, D. R., Biggar, J. W., & Erh, K. T. (1973). Spatial variability of field-measured soil-water properties. *Hilgardia*, *42*, 215–260.
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), 1–11. https://doi.org/10.2136/vzj2013.01.0025
Nimmo, J. R., Schmidt, K. M., Perkins, K. S., & Stock, J. D. (2009). Rapid measurement of field-saturated hydraulic conductivity for areal characterization. *Vadose Zone Journal*, *8*(1), 142–149. https://doi.org/10.2136/vzj2007.0159
Noy-Meir, I. (1973). Desert ecosystems: Environment and producers. *Annual Review of Ecology and Systematics*, *4*(1), 25–51. https://doi.org/10.1146/annurev.es.04.110173.000325
Nyamadzawo, G., Chikowo, R., Nyamugafata, P., & Giller, K. E. (2007). Improved legume tree fallows and tillage effects on structural stability and infiltration rates of a kaolinitic sandy soil from central Zimbabwe. *Soil and Tillage Research*, *96*(1–2), 182–194. https://doi.org/10.1016/j.still.2007.06.008
Nyberg, G., Bargués Tobella, A., Kinyangi, J., & Ilstedt, U. (2012). Soil property changes over a 120-yr chronosequence from forest to agriculture in western Kenya. *Hydrology and Earth System Sciences*, *16*(7), 2085–2094. https://doi.org/10.5194/hess-16-2085-2012
Onwujekwe, E. C., & Ezemba, C. C. (2021). Food security and safety: Africans perspectives A review. *Archives of Current Research International*, 14–20. https://doi.org/10.9734/acri/2021/v21i830260
Ottoni, M. V., Ottoni Filho, T. B., Schaap, M. G., Lopes-Assad, M. L. R. C., & Rotunno Filho, O. C. (2018). Hydrophysical database for Brazilian soils (HYBRAS) and pedotransfer functions for water retention. *Vadose Zone Journal*, *17*(1), 170095. https://doi.org/10.2136/vzj2017.05.0095
Owuor, S. O., Butterbach-Bahl, K., Guzha, A. C., Rufino, M. C., Pelster, D. E., Díaz-Pinés, E., & Breuer, L. (2016). Groundwater recharge rates and surface runoff response to land use and land cover changes in semi-arid environments. *Ecological Processes*, *5*(1), 16. https://doi.org/10.1186/s13717-016-0060-6
Pachepsky, Y., & Park, Y. (2015). Saturated hydraulic conductivity of US soils grouped according to textural class and bulk density. *Soil Science Society of America Journal*, *79*(4), 1094–1100. https://doi.org/10.2136/sssaj2015.02.0067
Pachepsky, Y. & Rawls, W. J. (Eds.). (2004). *Development of pedotransfer functions in soil hydrology* (Vol. 30). Elsevier.
Padfield, D., & Matheson, G. (2020). nls.multstart: Robust Non-Linear Regression using AIC scores. R package version 1.2.0. Retrieved from https://CRAN.R-project.org/package=nls.multstart
Pagliai, M., Vignozzi, N., & Pellegrini, S. (2004). Soil structure and the effect of management practices. *Soil and Tillage Research*, *79*(2), 131–143. https://doi.org/10.1016/j.still.2004.07.002
Paustian, K., Lehmann, J., Ogle, S., Reay, D., Robertson, G. P., & Smith, P. (2016). Climate-smart soils. *Nature*, *532*(7597), 49–57. https://doi.org/10.1038/nature17174
Ponette-González, A. G., Brauman, K. A., Marín-Spiotta, E., Farley, K. A., Weathers, K. C., Young, K. R., & Curran, L. M. (2015). Managing water services in tropical regions: From land cover proxies to hydrologic fluxes. *Ambio*, *44*(5), 367–375. https://doi.org/10.1007/s13280-014-0578-8
Rahmati, M., Weihermüller, L., Vanderborght, J., Pachepsky, Y. A., Mao, L., Sadeghi, S. H., et al. (2018). Development and analysis of the Soil Water Infiltration Global database. *Earth System Science Data*, *10*, 1237-1263. https://doi.org/10.5194/essd-10-1237-2018
Rawls, W. J., Nemes, A., & Pachepsky, Y. (2004). Effect of soil organic carbon on soil hydraulic properties. In *Developments in soil science* (Vol. 30, pp. 95–114). Elsevier.
R Core Team. (2015). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/
Rengasamy, P., Tavakkoli, E., & McDonald, G. K. (2016). Exchangeable cations and clay dispersion: Net dispersive charge, a new concept for dispersive soil. *European Journal of Soil Science*, *67*(5), 659–665. https://doi.org/10.1111/ejss.12369
Reynolds, W. (2008). Saturated hydraulic properties: Ring infiltrometer. *Soil Sampling and Methods of Analysis*, *58*.
Reynolds, W., Elrick, D., & Youngs, E. (2002). Ring or cylinder infiltrometers (vadose zone). In J. Dane & G. Topp (Eds.), *Methods of soil analysis: Part 4 physical methods* (pp. 818–826). Soil Science Society of America.
Reynolds, W. D., & Elrick, D. E. (1990). Ponded infiltration from a single ring: I. Analysis of steady flow. *Soil Science Society of America Journal*, *54*(5), 1233–1241. https://doi.org/10.2136/sssaj1990.03615995005400050006x
Ripple, W. J., Wolf, C., Newsome, T. M., Galetti, M., Alamgir, M., Crist, E., et al. (2017). World scientists' warning to humanity: A second notice. *BioScience*, *67*(12), 1026–1028. https://doi.org/10.1093/biosci/bix125
Robinson, D. A., Nemes, A., Reinsch, S., Radbourne, A., Bentley, L., & Keith, A. M. (2022). Global meta-analysis of soil hydraulic properties on the same soils with differing land use. *Science of the Total Environment*, *852*, 158506. https://doi.org/10.1016/j.scitotenv.2022.158506
Robinson, T. P., Thornton, P. K., Francesconi, G. N., Kruska, R., Chiozza, F., Notenbaert, A. M. O., et al. (2011). *Global livestock production systems*. FAO and ILRI.

Savadogo, P., Sawadogo, L., & Tiveau, D. (2007). Effects of grazing intensity and prescribed fire on soil physical and hydrological properties and pasture yield in the savanna woodlands of Burkina Faso. *Agriculture, Ecosystems & Environment*, *118*(1–4), 80–92. https://doi.org/10.1016/j.agee.2006.05.002

Saxton, K. E., Rawls, W. J., Romberger, J. S., & Papendick, R. I. (1986). Estimating generalized soil-water characteristics from texture. *Soil Science Society of America Journal*, *50*(4), 1031–1036. https://doi.org/10.2136/sssaj1986.03615995005000040054x

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. https://doi.org/10.1016/s0022-1694(01)00466-8

Scheffer, M., & Carpenter, S. R. (2003). Catastrophic regime shifts in ecosystems: Linking theory to observation. *Trends in Ecology & Evolution*, *18*(12), 648–656. https://doi.org/10.1016/j.tree.2003.09.002

Schielzeth, H. (2010). Simple means to improve the interpretability of regression coefficients. *Methods in Ecology and Evolution*, *1*(2), 103–113. https://doi.org/10.1111/j.2041-210x.2010.00012.x

Seneviratne, S. I., Zhang, X., Adnan, M., Badi, W., Dereczynski, C., Di Luca, A., et al. (2021). Weather and climate extreme events in a changing climate. In V. Masson-Delmotte, P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, et al. (Eds.), *Climate change 2021: The physical science basis. Contribution of working group I to the sixth assessment report of the intergovernmental panel on climate change* (pp. 1513–1766). Cambridge University Press.

Singer, M. J., & Le Bissonnais, Y. (1998). Importance of surface sealing in the erosion of some soils from a mediterranean climate. *Geomorphology*, *24*(1), 79–85. https://doi.org/10.1016/s0169-555x(97)00102-5

Singer, M. J., & Shainberg, I. (2004). Mineral soil surface crusts and wind and water erosion. *Earth Surface Processes and Landforms*, *29*(9), 1065–1075. https://doi.org/10.1002/esp.1102

Sun, G., Hallema, D., & Asbjornsen, H. (2017). Ecohydrological processes and ecosystem services in the Anthropocene: A review. *Ecological Processes*, *6*(1), 35. https://doi.org/10.1186/s13717-017-0104-6

Takoutsing, B., Winowiecki, L. A., Bargués-Tobella, A., & Vågen, T.-G. (2022). Determination of land restoration potentials in the semi-arid areas of Chad using systematic monitoring and mapping techniques. *Agroforestry Systems*, *97*(7), 1289–1305. https://doi.org/10.1007/s10457-021-00720-9

Talore, D. G., Tesfamariam, E. H., Hassen, A., Du Toit, J., Klampp, K., & Jean-Francois, S. (2016). Long-term impacts of grazing intensity on soil carbon sequestration and selected soil properties in the arid Eastern Cape, South Africa. *Journal of the Science of Food and Agriculture*, *96*(6), 1945–1952. https://doi.org/10.1002/jsfa.7302

Terhoeven-Urselmans, T., Vagen, T.-G., Spaargaren, O., & Shepherd, K. D. (2010). Prediction of soil fertility properties from a globally distributed soil mid-infrared spectral library. *Soil Science Society of America Journal*, *74*(5), 1792–1799. https://doi.org/10.2136/sssaj2009.0218

Thornton, P. (2002). *Mapping poverty and livestock in the developing world* (Vol. 1). ILRI (aka ILCA and ILRAD).

Tomasella, J., Hodnett, M. G., & Rossato, L. (2000). Pedotransfer functions for the estimation of soil water retention in Brazilian soils. *Soil Science Society of America Journal*, *64*(1), 327–338. https://doi.org/10.2136/sssaj2000.641327x

Tomasella, J., Pachepsky, Y., Crestana, S., & Rawls, W. J. (2003). Comparison of two techniques to develop pedotransfer functions for water retention. *Soil Science Society of America Journal*, *67*(4), 1085–1092. https://doi.org/10.2136/sssaj2003.1085

Trabucco, A., & Zomer, R. J. (2019). Global aridity index and potential evapotranspiration (ET0) climate database v2. *figshare*. Fileset.

UNEP. (1992). In N. Middleton & D. Thomas (Eds.), *World atlas of desertification*. Edward Arnold.

USDA. (2017). Soil survey manual, handbook No. 18.

Usowicz, B., & Lipiec, J. (2021). Spatial variability of saturated hydraulic conductivity and its links with other soil properties at the regional scale. *Scientific Reports*, *11*(1), 8293. https://doi.org/10.1038/s41598-021-86862-3

Vågen, T., & Winowiecki, L. (2020). The land degradation surveillance framework field guide. v2020. Retrieved from http://landscapeportal.org/documents/2477/download

Vågen, T.-G., Lal, R., & Singh, B. R. (2005). Soil carbon sequestration in sub-Saharan Africa: A review. *Land Degradation & Development*, *16*(1), 53–71. https://doi.org/10.1002/ldr.644

Vågen, T. G., & Winowiecki, L. A. (2013). Mapping of soil organic carbon stocks for spatially explicit assessments of climate change mitigation potential. *Environmental Research Letters*, *8*(1), 015011. https://doi.org/10.1088/1748-9326/8/1/015011

Vågen, T.-G., & Winowiecki, L. A. (2019). Predicting the spatial distribution and severity of soil erosion in the global tropics using satellite remote sensing. *Remote Sensing*, *11*(15), 1800. https://doi.org/10.3390/rs11151800

Vågen, T. G., Winowiecki, L. A., Tondoh, J. E., Desta, L. T., & Gumbricht, T. (2016). Mapping of soil properties and land degradation risk in Africa using MODIS reflectance. *Geoderma*, *263*, 216–225. https://doi.org/10.1016/j.geoderma.2015.06.023

van den Berg, M., Klamt, E., van Reeuwijk, L. P., & Sombroek, W. G. (1997). Pedotransfer functions for the estimation of moisture retention characteristics of Ferralsols and related soils. *Geoderma*, *78*(3), 161–180. https://doi.org/10.1016/s0016-7061(97)00045-1

Vandandorj, S., Eldridge, D. J., Travers, S. K., Val, J., & Oliver, I. (2017). Microsite and grazing intensity drive infiltration in a semiarid woodland. *Ecohydrology*, *10*(4), e1831. https://doi.org/10.1002/eco.1831

Vereecken, H., Weynants, M., Javaux, M., Pachepsky, Y., Schaap, M. G., & Genuchten, M. T. v. (2010). Using pedotransfer functions to estimate the van Genuchten–Mualem soil hydraulic properties: A review. *Vadose Zone Journal*, *9*(4), 795–820. https://doi.org/10.2136/vzj2010.0045

Wada, Y., van Beek, L. P. H., van Kempen, C. M., Reckman, J. W. T. M., Vasak, S., & Bierkens, M. F. P. (2010). Global depletion of groundwater resources. *Geophysical Research Letters*, *37*(20). https://doi.org/10.1029/2010GL044571

Wang, T., Wedin, D., & Zlotnik, V. A. (2009). Field evidence of a negative correlation between saturated hydraulic conductivity and soil carbon in a sandy soil. *Water Resources Research*, *45*(7). https://doi.org/10.1029/2008WR006865

Warrick, A. W. (1998). Spatial variability. In D. Hillel (Ed.), *Environmental soil physics* (pp. 655–675). Academic Press.

Weiler, M., & Naef, F. (2003). An experimental tracer study of the role of macropores in infiltration in grassland soils. *Hydrological Processes*, *17*(2), 477–493. Proceedings Paper. https://doi.org/10.1002/hyp.1136

Winowiecki, L., Vagen, T. G., Massawe, B., Jelinski, N. A., Lyamchai, C., Sayula, G., & Msoka, E. (2016). Landscape-scale variability of soil health indicators: Effects of cultivation on soil organic carbon in the Usambara Mountains of Tanzania. *Nutrient Cycling in Agroecosystems*, *105*(3), 263–274. https://doi.org/10.1007/s10705-015-9750-1

Wohl, E., Barros, A., Brunsell, N., Chappell, N. A., Coe, M., Giambelluca, T., et al. (2012). The hydrology of the humid tropics. *Nature Climate Change*, *2*(9), 655–662. https://doi.org/10.1038/nclimate1556

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. https://doi.org/10.1016/s0016-7061(98)00132-3

Wösten, J. H. M., Pachepsky, Y. A., & Rawls, W. J. (2001). Pedotransfer functions: Bridging the gap between available basic soil data and missing soil hydraulic characteristics. *Journal of Hydrology*, *251*(3), 123–150. https://doi.org/10.1016/s0022-1694(01)00464-4

Wright, M. N., & Ziegler, A. (2017). Ranger: A fast implementation of random forests for high dimensional data in C++ and R. *Journal of Statistical Software*, *77*(1), 1–17. https://doi.org/10.18637/jss.v077.i01

Yair, A. (1992). The control of headwater area on channel runoff in a small arid watershed. In A. Parsons & A. Abrahams (Eds.), *Overland flow: Hydraulics and erosion mechanics* (pp. 53–68). UCL Press.

Young, M. D. B., Gowing, J. W., Hatibu, N., Mahoo, H. M. F., & Payton, R. W. (1999). Assessment and development of pedotransfer functions for semi-arid sub-Saharan Africa. *Physics and Chemistry of the Earth - Part B: Hydrology, Oceans and Atmosphere*, *24*(7), 845–849. https://doi.org/10.1016/s1464-1909(99)00091-x

Zhang, G. S., Chan, K. Y., Oates, A., Heenan, D. P., & Huang, G. B. (2007). Relationship between soil structure and runoff/soil loss after 24 years of conservation tillage. *Soil and Tillage Research*, *92*(1), 122–128. https://doi.org/10.1016/j.still.2006.01.006

Zhang, X., Wendroth, O., Matocha, C., Zhu, J., & Reyes, J. (2020). Assessing field-scale variability of soil hydraulic conductivity at and near saturation. *Catena*, *187*, 104335. https://doi.org/10.1016/j.catena.2019.104335

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. https://doi.org/10.1016/j.jhydrol.2017.01.004

Zhao, Y., Wu, P., Zhao, S., & Feng, H. (2013). Variation of soil infiltrability across a 79-year chronosequence of naturally restored grassland on the Loess Plateau, China. *Journal of Hydrology*, *504*, 94–103. https://doi.org/10.1016/j.jhydrol.2013.09.039

Zimmermann, B., & Elsenbeer, H. (2008). Spatial and temporal variability of soil saturated hydraulic conductivity in gradients of disturbance. *Journal of Hydrology*, *361*(1), 78–95. https://doi.org/10.1016/j.jhydrol.2008.07.027

Zimmermann, B., Papritz, A., & Elsenbeer, H. (2010). Asymmetric response to disturbance and recovery: Changes of soil permeability under forest–pasture–forest transitions. *Geoderma*, *159*(1), 209–215. https://doi.org/10.1016/j.geoderma.2010.07.013

Zuur, A., Hilbe, J., & Leno, E. (2013). *A beginner's guide to GLM and GLMM with R: A frequentist and Bayesian perspective for ecologists* (p. 270). Highland Statistics Ltd.

Zuur, A. F., Ieno, E. N., Walker, N. J., Saveliev, A. A., & Smith, G. M. (2009). *Mixed effects models and extensions in ecology with R* (Vol. 574). Springer.

Zwartendijk, B. W., van Meerveld, H. J., Ghimire, C. P., Bruijnzeel, L. A., Ravelona, M., & Jones, J. P. G. (2017). Rebuilding soil hydrological functioning after swidden agriculture in eastern Madagascar. *Agriculture, Ecosystems & Environment*, *239*, 101–111. https://doi.org/10.1016/j.agee.2017.01.002