Soil bacterial diversity varies across biomes with potential impacts on soil ecological functioning. Here, we incorporate key factors that affect soil bacterial abundance and diversity across spatial scales into a mechanistic modeling framework considering soil type, carbon inputs and climate towards predicting soil bacterial diversity. The soil aqueous-phase content and connectivity exert strong influence on bacterial diversity for each soil type and rainfall pattern. Biome-specific carbon inputs deduced from net primary productivity provide constraints on soil bacterial abundance independent from diversity. The proposed heuristic model captures observed global trends of bacterial diversity in good agreement with predictions by an individual-based mechanistic model. Bacterial diversity is highest at intermediate water contents where the aqueous phase forms numerous disconnected habitats and soil carrying capacity determines level of occupancy. The framework delineates global soil bacterial diversity hotspots; located mainly in climatic transition zones that are sensitive to potential climate and land use changes.
Soil hosts unparalleled bacterial diversity, ranking highest among all other compartments of the biosphere\textsuperscript{2–4}. The number of bacterial phylotypes ranges between 10\textsuperscript{2} and 10\textsuperscript{6} per gram of soil\textsuperscript{2,4,4}, with high values similar to the diversity in all of earths environments\textsuperscript{5}. This immense richness is often attributed to soil’s intrinsically heterogeneous physical and chemical micro-environments\textsuperscript{5–9}. The complex structure of soil pores offers numerous refugia for hosting diverse bacterial species\textsuperscript{9}. This wide range of microhabitats is particularly important for maintaining the rare components of the soil microbiome. Low abundance bacterial species play important roles in key biogeochemical processes\textsuperscript{10,11} and serve as a “seed bank” for species richness\textsuperscript{12}. Microbial diversity is manifested both at the scale of soil grains\textsuperscript{13} and at very large scales across climatic regions and terrestrial biomes\textsuperscript{13,14}. These observations often include variations in microbial biomass that responds to resource availability and affects bacterial diversity at all scales\textsuperscript{15–17}. For example, well-established observations of microbial abundance variations with soil depth\textsuperscript{18} could confound inferences of bacterial richness by promoting the detection of low abundant species in resource-rich environments.

Quantifying the roles of soil factors, such as soil texture, porosity and hydration conditions in relation to climate and vegetation cover, is an important step towards disentangling bacterial diversity and abundance as suggested by recent empirical evidence\textsuperscript{19}. Soil chemical properties such as pH\textsuperscript{2,14,17,19} and organic carbon content\textsuperscript{15–17} together with climatic attributes, such as aridity index\textsuperscript{15}, precipitation\textsuperscript{2,17} and temperature\textsuperscript{13}, have been identified as important explanatory variables. Yet, the rapid expansion of soil bacterial diversity datasets has not been met with similar development of predictive models for interpretation of the observed spatial patterns\textsuperscript{20}. Improved predictability of soil bacterial diversity could be essential for understanding soil bacterial functioning; from contributions to soil respiration\textsuperscript{11,21} to the resistance of bacterial communities to invasion by pathogens\textsuperscript{22}.

Such endeavors invariably require development of mechanistic frameworks for systematic incorporation of the various factors that affect soil bacterial diversity. In this study, we capitalize on recent empirical\textsuperscript{2,13,15,17,23} and theoretical developments\textsuperscript{7,24,25} to generalize the role of soil aqueous microhabitat fragmentation and its nearly universal role in mediating bacterial diversity across soil types and climatic conditions. To characterize the average conditions in soils and facilitate long-term predictions, we define a soil climatic water content that combines rainfall patterns and volumetric soil water holding capacity into a well-defined attribute. This measure considers the average duration between soil wetting events important for diversity maintenance (see Methods). Under a wide range of climatic conditions, soils remain unsaturated with the bacterial aqueous habitats fragmented to varying degrees based on soil type and rainfall dynamics (amount and frequency). A critical hypothesis is that the microscale arrangement of water retained in soil pores defines the size distribution and connectedness of aqueous bacterial habitats that, in turn, affect diffusion rates of substrates, the rates and spatial extents of cell motility\textsuperscript{25,26} and opportunities for cell-to-cell interactions\textsuperscript{27}. The objective of this study was to formalize the influence of these abiotic factors in a heuristic framework that enables quantitative representation of soil bacterial abundance and diversity at scales ranging from grains to watersheds and beyond.

The core of the model is the quantification of numbers and sizes of aqueous bacterial habitats considering climatic water contents and soil types. We use concepts of percolation theory to describe the size distribution of aqueous patches\textsuperscript{24} that could support bacterial cells. Soil organic carbon input flux, derived from the net primary productivity (NPP), and mean annual temperature (MAT) are used to estimate a soil-carrying capacity that defines limits for the abundance of bacterial cells (Fig. 1). For simplicity, we first assume that each isolated aqueous patch is inhabited by a single bacterial phylotype (hereafter referred to as "species"). This heuristically enables estimation of bacterial diversity based on the species abundance distribution (SAD) deduced from the size and number distribution of microscale aqueous habitats. The framework expresses soil bacterial diversity at two interlinked spatial scales: at the single aqueous habitat scale and at the soil sample scale that can contain many isolated aqueous habitats.

Modeled trends of soil bacterial carrying capacity and diversity are compared to empirical observations\textsuperscript{1,4,18} across terrestrial biomes and suggest a peak in bacterial diversity at intermediate climatic water contents. To evaluate predictions by this aqueous-phase fragmentation-based heuristic model (HM), we employ a detailed, spatially explicit individual-based model (SIM) that mechanistically simulates bacterial communities growing on hydrated soil surfaces\textsuperscript{22,23}. The SIM enables systematic variations of hydration conditions and tracks the growth and life history of multiple species interacting on soil grain surfaces (see Methods).
The simple HM does not differentiate between the roles of legacy and environmental conditions in shaping soil bacterial diversity. As evidenced from the choice of climatic averaging and the implicit representation of species with no taxonomic attribution, the focus lies on the role of aqueous habitats and their average connectivity. Other factors at play such as soil chemistry and the presence of larger organisms are not modeled. We refer to “microbes” for aspects that apply to all microbial life in soil (bacteria, fungi, protists and viruses), and specifically to bacteria for modeling and quantification of diversity and abundance. Summarizing, we propose a hydration-centered modeling framework that considers the interplay of climatic water content; carbon input flux and temperature in shaping soil microhabitats and thus bacterial diversity.

Results

Estimation of soil bacterial carrying capacity. We evaluated theoretical estimates of soil bacterial carrying capacity using previously published measurements of soil microbial carbon. The HM assumes that a certain proportion of the annual NPP-derived organic carbon input is allocated to bacteria (24% of NPP for bacterial respiration). We found that varying the range of expected values (14–30% of NPP) had little impact on estimates of carrying capacity. A constant value of this respiratory fraction was therefore considered based on mechanistic model simulations. We employ a basic estimate of bacterial cell maintenance rate of 1.5 gC gCcell⁻¹ y⁻¹ (=10⁻⁴ gC gCcell⁻¹ h⁻¹) and adjust it according to the local mean annual temperature (MAT) to account for different climatic regions. Combining local annual NPP and adjusted cell maintenance rate, we derive estimates of soil bacterial carrying capacity as upper bounds for soil bacterial cell density (Fig. 2a). Despite the many simplifying assumptions, we obtain reasonable estimates of potential soil bacterial carrying capacity that are comparable with observations of realized bacterial cell density across a range of environmental conditions. Model estimates of soil-carrying capacity for three values of MAT are depicted in Fig. 2a (representing the median of three groups: ≤0 °C, 0 °C–15 °C and >15 °C with –2, 9 and 19 °C, respectively). Observed cell densities tend to be higher for colder regions as considered by the HM. We note that soil bacterial cell density is expected to vary with soil depth due to the distribution of organic carbon flux from the soil surface and distribution by plant roots. Soil bacterial carrying capacity decreases steeply with depth and was represented parametrically by a lognormal distribution (μ = 0.18, σ = 1.00) (Fig. 2b). The lognormal distribution provided a better global representation of the average topsoil carrying capacity (upper 10 cm, Supplementary Fig. 1) over the previously reported exponential model. It is important to keep in mind that the estimated soil-carrying capacity was defined independently from bacterial diversity and values were calculated globally based on NPP, MAT and soil depth.

Modeling bacterial diversity considering climate and soil. The simple HM was developed in two conceptual steps. We first assumed only a single species per aqueous habitat. This approach, although useful as a heuristic, exhibited some limitations for large aqueous habitats under wet conditions (see comparison of species abundance distributions below). We thus adapted the model to allow multiple species in large habitats by assigning the number of species \( N_{sp} \) proportional to the length scale of a habitat of size \( s \) \( (N_{sp} \sim s^{1/d}, d = 2 \) or \( 3 \text{ dimensionality}) \). Hence, the HM links species richness to the soil aqueous-phase fragmentation via percolation theory and accommodates the possibility of multiple species per habitat. For most unsaturated conditions the refined formulation does not alter the prediction since small habitats are likely to host only a single species. In the following we refer to the multispecies HM if not stated otherwise. We have used median values of global soil-carrying capacity to describe trends in soil bacterial diversity across soil types and climatic regions. Comparisons of model estimates with empirical observations of bacterial richness obtained from the studies of Thompson et al. (EMP) and Delgado-Baquerizo et al. (DEL) are depicted in Fig. 3 along with the mechanistic predictions by the SIM. We have expressed mean soil hydration status via the climatic water content that is a proxy for average soil wetness and habitat connectivity. Soil and climatic variables were compiled from different sources (Supplementary Table 1) with matched geographical coordinates and soil depths for the samples. We present soil bacterial richness (total number of types) and note that taxonomic assignment was absent for the phylotypes detected in EMP. Bacterial richness was binned by water contents because some hydration conditions were over-represented (bin width: 0.05). Since richness in the EMP data was measured at different soil depths, they were also grouped to top and subsoil (≤25 cm and ≥25 cm). Exact number of samples per group are reported in Supplementary Table 2. The EMP data

![Fig. 2 Soil bacterial abundance varies in relation to carbon input, temperature and soil depth.](https://example.com/fig2.png)

**Fig. 2** Soil bacterial abundance varies in relation to carbon input, temperature and soil depth. **a** Bacterial cell density at soil-carrying capacity as a function of net primary productivity (NPP) with model estimates sensitive to mean annual temperature (MAT) (solid lines). Estimates are compared with measured data of microbial biomass. **b** Variations of bacterial cell density with soil depth. The lognormal fit provides bounds on cell density (carrying capacity) for intermediate MAT (solid line) and for the central 95% of NPP (shaded area). Observed estimates of cell density are reported for their average sampling depth. Most samples were taken above 10 cm as shown in the boxplot. Source data are provided as a Source Data file.
model occupancy of potential habitats ($N_{\text{cell}} > 4000$). Lastly, we compared the aqueous-phase fragmentation-based HM to numerical simulations of the SIM. We simulated the spatially explicit growth and movement of individual cells in a diverse bacterial community on heterogeneous soil pore surfaces. Qualitatively, both HM and SIM predict similar trends of variations in bacterial richness with soil hydration conditions as estimated from the EMP and DEL datasets (Fig. 3a, b). In addition to removing single cells (singletons) from the simulated communities, the modeled species counts were rarefied to 5000 and 1000 for comparison with EMP and DEL, respectively. To compare with the DEL dataset, simulated bacterial richness is reported only for the 512 most abundant species and describes the observed invariance of richness towards low climatic water contents (Fig. 3b). The discrepancy in water contents where richness peaks (between HM and SIM) is attributed to the dimensionality of the models (three for HM, two for SIM) and is well captured by the percolation-based HM in two dimensions (Supplementary Fig. 2).

### Species abundance distribution varies with hydration status
We quantified variations in bacterial species abundance distribution (SAD) with soil attributes and climatic water contents in comparison with empirical estimates from the EMP and DEL datasets (Supplementary Fig. 3). Here we used soil properties and carrying capacity specific for each geographical location and soil depth. The results show good alignment of the single-species model predictions with observed relative SADs and resulted in Pearson correlation values of $0.84$ ($n = 230$) and $0.76$ ($n = 218$) for the EMP and DEL datasets, respectively (Supplementary Fig. 3a, b). Nevertheless, the single-species HM erroneously predicts a higher proportion of the most abundant species than observed. We attribute this systematic overestimation to the stringent assumption of one single species per aqueous (micro-) habitat. This discrepancy suggests that the single species per aqueous habitat assumption may not hold for very large aqueous habitats in wet soil that could host multiple species. To rectify this limitation, we considered a scenario where the number of species $N_{\text{sp}}$ is assumed proportional to the size $s$ of an aqueous habitat ($N_{\text{sp}} \sim s^{1.2}$). This relaxed occupancy assumption improved Pearson correlations to values of $0.88$ ($n = 230$) and $0.84$ ($n = 218$) for the EMP and DEL datasets, respectively (Supplementary Fig. 3c, d). Predictions by the HM for ranked SADs compare qualitatively with observations that were grouped by average hydration conditions (Supplementary Fig. 4). An increase in dominance of the most abundant bacterial species is visible in the ranked SADs of both datasets under sufficiently wet conditions (Supplementary Fig. 4b, c).

### Global patterns of soil bacterial habitat diversity
Motivated by the general agreement with observations of bacterial richness and the SADs produced by the HM, we used highly resolved global datasets for soil properties, NPP and precipitation as inputs to estimate global patterns of soil bacterial habitat richness (Fig. 4a). Recall that a central element of the model is the link between the number of distinct aqueous habitats per soil volume and soil bacterial richness. Additionally, we considered the sizes of aqueous habitats to yield spatially resolved distributions of the Shannon index of bacterial diversity patterns (Fig. 4b). We note that the modeled soil bacterial diversity follows constraints imposed by local soil-carrying capacity where high bacterial cell numbers are associated with locally high NPP and low cell maintenance requirements. Both diversity indices exhibit spatial patterns with distinct regions of increased diversity associated with climatic transition zones (e.g., the Sahel). This pattern is more pronounced when considering the Shannon index and suggests that soil bacterial community evenness, indicative of how equally habitats are

---

**Fig. 3** Observed and predicted variations in soil bacterial diversity with climatic water content. a, b Estimates of bacterial richness from two studies are binned by climatic water contents (bin width: 0.05) and the median and interquartile range are reported (circles and bars, respectively). The exact number of samples per group is listed in Supplementary Table 2. Individual data points are shown for bins containing less than ten samples (small circles). The solid black lines correspond to predictions by the fragmentation-based heuristic model (HM) for median carrying capacity specific to each dataset. The square symbols, thin solid line and shading (mean, rolling mean ± SD, $n = 12$) depict simulated bacterial richness using the spatially explicit individual-based model (SIM) for different water contents. a Bacterial richness from the Earth Microbiome Project (Thompson et al.—EMP) was reported for different soil depths and thus grouped accordingly ($<25$ and $\geq 25$ cm, top- and subsoil, respectively). The dashed line represents a model scenario with reduced carrying capacity by considering only the subsoil. b Soil bacterial richness from a recent study (Delgado-Baquerizo et al.—DEL). Colors indicate reported soil pH, which has been shown to be affected by climate$^{36}$. For comparison with the DEL dataset, only the top 512 species were considered in the SIM. Source data are provided as a Source Data file.

display a tendency of lower values of richness in the subsoil (Fig. 3a). In the DEL dataset, measurements were taken at the same soil depth, and soil pH is reported instead (Fig. 3b). We observe a strong tendency of lower soil pH in climatically wetter soils. The results depict an average decrease in bacterial richness where the soil becomes saturated as also predicted by the HM for median soil-carrying capacity (Fig. 3a, b). The modeled sensitivity to soil-carrying capacity is shown for a scenario of reduced cell densities (e.g. less carbon input to deeper soil layers; Fig. 3a dashed line). We emphasize that parameters were not fitted to observed diversity data, but rather are based on mean values for soil properties (porosity $\theta_s = 0.49$ and 0.47; sample length $L = 5$ and 6 mm; textural length $\delta = 0.07$ and 0.1 mm; for EMP and DEL, respectively). Additionally, we used a fixed value for the critical water content ($\theta_c \approx 0.15$) and a threshold for the number of cells $N_{\text{cell}}$ needed to
strain the HM for evenness prediction (Fig. 5). Considering the values for carrying capacity and climatic water contents to climatic water contents. This could result in unrealistic values based HM, predicted bacterial cell densities are independent of (Supplementary Fig. 6). In the aqueous-phase fragmentation-modeling (SIM) for small spatial and short temporal scales being associated with an increase in cell densities and was confirmed (with no prior assumptions) using detailed mechanistic modeling (SIM) for small spatial and short temporal scales (Supplementary Fig. 6). In the aqueous-phase fragmentation-based HM, predicted bacterial cell densities are independent of climatic water contents. This could result in unrealistic values relative to empirical observations. We therefore used pairs of values for carrying capacity and climatic water contents to constrain the HM for evenness prediction (Fig. 5). Considering the relation between climatic water content and soil-carrying capacity highlights the sensitivity of HM predictions to bacterial cell density as also observed in the mechanistic simulation results of the SIM. The dependency of cell density on climatic water content in the SIM results in a persistent decrease of evenness with increasing water content (Supplementary Fig. 7). When considering paired values of water content and cell densities obtained partitioned, is sensitive to soil wetness. Such an association is also observed empirically where evenness decreases with increasing climatic water contents (Pearson $r = -0.17$ and $-0.43$ for EMP and DEL, respectively; Supplementary Fig. 5a).

Disentangling soil bacterial abundance and diversity. To address the challenge of disentangling bacterial abundance and diversity, we compared bacterial community evenness with climatic water content and carrying capacity (Fig. 5). Evenness decreases gradually with climatic water content and with increasing soil-carrying capacity (Fig. 5, Supplementary Fig. 5b). The results are consistent with the tendency of wetter conditions being associated with an increase in cell densities and was confirmed (with no prior assumptions) using detailed mechanistic modeling (SIM) for small spatial and short temporal scales (Supplementary Fig. 6). In the aqueous-phase fragmentation-based HM, predicted bacterial cell densities are independent of climatic water contents. This could result in unrealistic values relative to empirical observations. We therefore used pairs of values for carrying capacity and climatic water contents to constrain the HM for evenness prediction (Fig. 5). Considering the relation between climatic water content and soil-carrying capacity highlights the sensitivity of HM predictions to bacterial cell density as also observed in the mechanistic simulation results of the SIM. The dependency of cell density on climatic water content in the SIM results in a persistent decrease of evenness with increasing water content (Supplementary Fig. 7). When considering paired values of water content and cell densities obtained

from the SIM, the simpler HM captures the simulated trends reasonably well (Supplementary Fig. 7). Although beyond the scope of this study, we observed that pre-processing measurements of relative species abundance may affect diversity metrics such as richness and evenness, which alters the apparent tendencies (Supplementary Fig. 8).

Discussion

The heuristic nature of the aqueous-phase fragmentation-based model (HM) precludes comparison of bacterial richness and abundance on a per sample basis, as climatic assumptions and associated large-scale variables are not likely to apply at a particular sampling location and time. Nonetheless, the proposed HM captures the salient features of global trends in bacterial richness related to climate, biome and soil type. Our estimate of soil bacterial cell density represents an upper bound on soil bacterial abundance (carrying capacity) and shows general agreement with measurements of soil bacterial biomass carbon\(^{18}\). It tracks the temperature dependency of reaction rates\(^{30}\) and provides an independent measure of maximal cell density that is sensitive to climate and organic carbon input by vegetation. Bacterial diversity increases towards lower values of climatic water contents (i.e., with increased aridity\(^{15}\)), as long as soil bacterial life is not limited by low organic carbon input. Assuming a constant soil bacterial carrying capacity, we can attribute much of the variations in bacterial richness to the microscale behavior of soil hydration conditions (Fig. 3). Surprisingly, the trends of bacterial richness for both surveys EMP\(^{3}\) and DEL\(^4\) were very similar despite their different objectives and processing protocols of the genetic information, namely the use of amplicon sequence variants in EMP and operational taxonomic units in DEL (Fig. 3a, b). We note that the values of bacterial richness in the DEL dataset saturate towards lower values of climatic soil hydration (Fig. 3b). This is likely due to the truncation of species richness used in that study which focuses on the most abundant soil bacteria\(^4\). These, highly abundant species, might be
the last to disappear under reduced carrying capacity and therefore do not show a decline towards dry conditions. The data available at low climatic water contents are sparse and do not provide support for the predicted steep decline of bacterial diversity as soil becomes dry that was previously reported with increased aridity at large scales23. However, a significant decrease in bacterial richness was also observed in a recent statistical meta-analysis for climatic scales31 and could be confirmed using the SIM (Fig. 3b). Additionally, it has been reported that bacterial diversity declines sharply with moisture in dry soils of Antarctica23 and decreases with soil relative humidity along transects of the Atacama desert32. Microcosm experiments revealed an increase in richness with moisture that peaks at intermediate water contents that promote rare bacterial species23. Similarly, bacterial richness was highest at intermediate climatic water contents where isolated aqueous habitats are numerous and sufficiently well supplied by diffusion to realize the soil-carrying capacity (Fig. 3). This observation is supported by the mechanistic simulation results of the SIM, which explicitly considers the dynamics and spatial structure of the bacterial community (Fig. 3). The generality of the aqueous-phase fragmentation-based approach permits comparison of systems with different dimensionality and can account for the shift of maximal richness towards higher water contents when comparing the HM with the two-dimensional simulation of bacterial life on hydrated surfaces by the SIM (Supplementary Fig. 2).

Increasing the organic carbon input and thus soil bacterial abundance seems to support higher diversity of soil microorganisms15. This is in line with the observation of decreasing bacterial richness with soil depth (Fig. 3a) that is often attributed to diminishing carbon inputs with depth (Fig. 2b). However, considering the various interacting factors at play, the general picture might be more complicated. An increase in soil-carrying capacity may not necessarily translate to increased bacterial diversity as evidenced by declining community evenness (Fig. 5, Supplementary Fig. 5). This could be due to dominance of a few species that may cluster near nutrient hotspots34, or loss of oligotrophic species that would be outcompeted in well-connected and dense communities9. We observe sensitivity of bacterial evenness to climatic water contents (Fig. 5), also in relation to the soil-carrying capacity (Supplementary Fig. 5). However, care should be taken regarding the interpretation of bacterial richness and evenness, since biases introduced by data processing and sampling could depend on the shape of the underlying SAD (Supplementary Fig. 8). Mechanistic models, such as HM and SIM, are valuable tools to investigate such dependencies as illustrated by considering only the most abundant species (Fig. 3b) or increasing sampling effort and removing species present at low abundance (Supplementary Fig. 8a, b, respectively). Nonetheless, an inherent tradeoff between availability of nutrients and protection by spatial isolation appears to play an important role in the establishment and maintenance of high soil bacterial diversity17,31,34. In other words, the relation between bacterial abundance and diversity is only positive when the aqueous phase is fragmented and spatial isolation suppresses the dominance of few species. As aqueous microhabitats become connected following soil rewetting by rainfall or irrigation, competition and other trophic interactions between bacterial cells are likely to reduce soil bacterial diversity (Fig. 3a, b) by reducing the communities evenness (Fig. 5). Many other factors such as pH14,19, nutrient composition19, carbon sources distribution6,34, stoichiometric constraints14,18 and metabolic dependencies35 shape soil bacterial abundance and diversity and could contribute to the discrepancy between our HM and empirical observations. Our study suggests that some of those factors might be associated with climatic hydration conditions. Interestingly, we find that soil samples exhibiting high bacterial diversity at intermediate climatic water contents coincide with near neutral pH values. In contrast, samples at low and high climatic water contents show high (basic) and low (acidic) pH tendencies, respectively (Fig. 3b).

This is supported by studies that relate soil pH with differences in soil water balance at climatological timescales20. We consequently expect soil pH to result from differences between precipitation and evapotranspiration as described by climatic water contents (Supplementary Fig. 9). Teasing apart such confounding associations requires detailed statistical analysis and experimental validation, which are best conducted in dedicated studies.

Using a single parameter set, largely based on standard percolation theory combined with data on soil properties, our HM predicts SADs that closely resemble empirical observations (Supplementary Figs. 3, 4). Nevertheless, the increased aqueous-phase connectedness in climatically wet soils may also promote interactions that are suppressed under spatial isolation of dry conditions23. Processes that support bacterial species coexistence across small distances are not captured by the present model and would result in persistent underestimation of bacterial diversity (unless provisions are introduced for very large aqueous habitats—see Supplementary Figs. 2, 3). Another inherent limitation of the analyses presented here is the focus on soil bacteria ignoring the interplay with other soil microorganisms that comprise Earth’s microbiome20. For example, fungi could play an important role in modifying soil bacterial habitats8 and are currently only considered in the partitioning of microbial carbon.

The framework presented in this study captures the salient spatial trends in soil bacterial diversity at climatic timescales and provides insights into effects of habitat fragmentation on the prevalence of bacterial interactions in natural soil. This is particularly important for the interpretation of species co-occurrence and interspecific interactions35. Such interactions between different species become possible only for conditions supported by the soil aqueous-phase connectedness23. This promotes diversity by enabling macroscopic coexistence5,7,24 in soil bacterial communities competing for space and a common resource.

A unique aspect of the HM is the ability to bridge scales from soil pores to biomes where information at both scales is preserved. Further investigations are required to test some of the model implications at different scales. For example, elucidating the dependency of cell microscale distribution on soil type and hydration conditions could provide insights into the processes shaping bacterial interactions in soil. Additionally taking into account factors affecting the partitioning of carbon at the ecosystem scale could refine model estimates of bacterial abundance beyond potential carrying capacity. Nonetheless, modeling climate and soil-specific bacterial diversity offers a useful reference for comparing effects of climatic shifts (e.g. in temperature, precipitation) or land use change (e.g. in intensity of agricultural management or restoration to natural ecosystems) on soil bacterial communities that could guide future exploration of the soil bacterial micro- and macro geography.

**Methods**

In the following, we provide a detailed overview of the methods used in the study and list key assumptions. Although the HM uses a yearly timescale for climatic averaging, the framework could be applied to finer and more resolved datasets. The global predictions of soil bacterial diversity were based on a 0.1°×0.1° grid to harmonize raster layers. For a description of data sources, see Supplementary Table 1. Variables added to the datasets of point measurements are taken at the native, highest spatial resolution of the respective property. Where necessary and not explicitly stated, missing values were imputed using the mean value of the corresponding variable.

**Soil bacterial carrying capacity derived from NPP**

The flux of carbon into the soil is taken from the MODIS NPP dataset27. We have used mean annual values (2000–2015). Missing values (e.g. desert) were imputed with values obtained from the Miami model38 using parameters fitted to the nonmissing values of MODIS...
NPP. Only an average fraction (ε = 0.24) of the total NPP entering the soil column is available for bacterial respiration18,22. The vertical distribution of microbial carbon in the soil column follows the distribution of plant roots18. This allowed us to impose the depth z at which most of the carbon is released by integrating over the sampled interval dz and calculating the fraction of NPP available for bacteria at a particular depth (NPPr,θ = ε(z)F,θ = ε(z)f(z)/dz). The factor F,θ denotes the fraction of carbon available at a particular depth and is described by f(z) for the entire depth of the soil profile considered (Ld = 1 m). Assuming no net growth of the bacterial community so that only energy requirements for maintenance metabolism are satisfied permits computation of maximal bacterial cell density ρc,max (m−3). This soil-carrying capacity supported by the input flux of carbon is calculated using Eq. (1).

\[
ρ_{c,\text{calc}}(L, T) = \frac{\text{NPP}_{\text{avail}}}{f(T)M_r}
\]  

Using a constant mass of carbon per cell Mr and by fitting maintenance rate m, we calculated the bacterial cell density ρc,θ. Temperature dependency was implemented as a factor fT based on the Schoolfield model30 using mean annual temperature (MAT) from the WorldClim dataset39.

Soil bacterial abundance dataset. Xu et al. (XU)38 compiled a dataset for the abundance of soil carbon associated with microbial biomass. This was used here as a reference for bacterial abundance for a range of geographical locations. We considered the relation between the soils carbon to nitrogen (C:N) ratio and the climatic water contents that allow the depth of the soil profile. We used the empirical relation between the soils carbon to nitrogen (C:N) ratio and the climatic water contents that allow the depth of the soil profile, as calculated using Eq. (1).

\[
θ = \frac{θ_1}{\epsilon(z)}e^{-\alpha \frac{S}{V_{water}}} + \frac{S}{\alpha} \frac{\text{PET}_{\text{drying}}}{\text{PET}}
\]  

The significance of the parameter is that it is the combination patterns, PET, and soil properties over climatic timescales and provides a measure of the average hydration conditions experienced by soil bacteria in a particular geographical location (Supplementary Fig. 9).

Estimation of aqueous habitat size distribution. We estimated the size distribution of distinct aqueous habitats based on soil properties and hydration conditions (e.g., climatic water content). Soil water content was treated as the aqueous-phase occupancy probability (the probability of finding a water filled pore or roughness element) that, in turn, enabled the application of standard percolation theory to the representation of the size distribution of bacterial habitats (sizes and number) and their properties (size and number). The soil-carrying capacity is thus defined by Eq. (1) and could be expressed as a critical saturation Sθ (4) to remove the dependency on θ.

\[
S_θ = \frac{θ}{θ_c(\tau_θ)} ≈ \frac{1}{2} - \frac{s}{\beta s}
\]  

The size distribution of aqueous patches n(s) was assumed to follow general probability of percolation theory (S−7/4)

\[
n(s) ∝ s^−ξ \tau_θ^s
\]  

With the patch size s (number of sites/pores) for s > 1, Fisher exponent τ = 2.18 (2D: τ = 1.87/91), cutoff exponent θ = 0.45 (2D: θ = 36/91) and cutoff size sθ = 46

\[
P(s) ∝ (s - p_c)^b
\]  

for the fraction of the domain occupied by a spanning (algorithmically infinite) patch with exponent β = 0.41 (2D: β = 5/36). The patch sizes follow a power law distribution at p = p_c. Away from this critical point when the cutoff size is smaller, patches shrink with decreasing water content (p < p_c) or merge and grow when approaching saturation (p > p_c) as patches of size s > sθ become exponentially scarcer. Although the prediction is strictly valid only for p close to p_c, we assume such relations to hold for the range of conditions considered. The occupancy probability p was thus substituted with climatic water content θ and p_c with a critical water content θ_c = 0.15 that correspond to a simple cubic lattice with porosity θ_c = 0.5 (triangular lattice in 2D: θ_c = 0.25).

To account for different soil types, a characteristic length scale θ is estimated based on the geometric mean diameter of soil particles46. This length scale is used for normalization of the aqueous patch size distribution in the range of water content and patches sizes relevant for bacterial life. The soil type length scale θ and the system size L were considered (soil domain or sample size); here we used the mass of soil sampled m_{soil} and bulk density ρ_{soil} specific to soil type (8). The total number of candidate sites N_{soil} in the sampled soil was then determined from simple geometry considering the dimensionality d = 2 or 3 (9).

\[
N_{soil} = \frac{L}{\delta d}
\]  

We approximated the behavior of the percolation transition using a bounded logistic curve that provides a smooth function P^θ

\[
P^θ = \frac{θ}{1 + e^{-β(P - θ)}}
\]
where $k$ describes the “sharpness” of the transition ($k = 16$ for all calculations). The total size of aqueous clusters or potential habitats $N_{s}$ was normalized as follows:

$$N_{s}(\theta, N_{s}) = \frac{\theta - \theta^{*}}{\sum_{j} s_{j}(\theta)}.$$  

(11)

$$N_{k}(\theta, s) = N_{s}(\theta, N_{s}) s_{k}(\theta).$$  

(12)

Thus requiring, by pre-factor $N_{k}$, that the total volume of aqueous patches conserves the volume of soil water at a given state of hydration. For practical reasons, subsequent calculations of aqueous patches proceed by removing the largest patch after normalization (this large patch biases the counting of habitats in a sample).

**Calculation of bacterial species diversity.** The distribution of aqueous patches derived from percolation theory and their properties defined the degree of spatial isolation and restricted the number of potential habitats. Both aspects were expected to alter the bacterial diversity patterns observed in natural soils. The isolation and restricted the number of potential habitats. Both aspects were derived from percolation theory and their properties de

**Calculation of bacterial species diversity.** The distribution of aqueous patches derived from percolation theory and their properties defined the degree of spatial isolation and restricted the number of potential habitats. Both aspects were expected to alter the bacterial diversity patterns observed in natural soils. The isolation and restricted the number of potential habitats. Both aspects were derived from percolation theory and their properties de

**Calculation of bacterial species diversity.** The distribution of aqueous patches derived from percolation theory and their properties defined the degree of spatial isolation and restricted the number of potential habitats. Both aspects were expected to alter the bacterial diversity patterns observed in natural soils. The isolation and restricted the number of potential habitats. Both aspects were derived from percolation theory and their properties de

**Calculation of bacterial species diversity.** The distribution of aqueous patches derived from percolation theory and their properties defined the degree of spatial isolation and restricted the number of potential habitats. Both aspects were expected to alter the bacterial diversity patterns observed in natural soils. The isolation and restricted the number of potential habitats. Both aspects were derived from percolation theory and their properties de
7. Wang, G. & Or, D. Hydration dynamics promote bacterial coexistence on rough surfaces. ISME J. 7, 393–404 (2013).
8. Bai, E. M., Williams, D. T., Hargrove, W. L., Yang, F. & Hofmockel, K. S. Greatest soil microbial diversity found in micro-habitats. Soil Biol. Biochem. 118, 217–226 (2018).
9. Vos, M., Wolf, A. B., Jennings, S. J. & Kowalchuk, G. A. Micro-scale determinants of bacterial diversity in soil. FEMS Microbiol. Rev. 37, 936–954 (2013).
10. Jørgensen, A. et al. Where less may be more: how the rare biosphere pulls ecosystems strings. ISME J. 11, 853–862 (2017).
11. Bell, T., Newman, J. A., Silverman, B. W., Turner, S. L. & Lilley, A. K. The contribution of species richness and composition to bacterial services. Nature 436, 1157–1160 (2005).
12. Lennon, J. T. & Jones, S. E. Microbial seed banks: the ecological and evolutionary implications of dormancy. Nat. Rev. Microbiol. 9, 119–130 (2011).
13. Zhou, J. et al. Temperature mediates continental-scale diversity of microbes in forest soils. Nat. Commun. 7, 12083 (2016).
14. Kaiser, K. et al. Driving forces of soil bacterial community structure, diversity, and function in temperate grasslands and forests. Sci. Rep. 6, 33669 (2016).
15. Maestre, F. T. et al. Increasing aridity reduces soil microbial diversity and abundance in global drylands. Proc. Natl. Acad. Sci. USA 112, 15684–15689 (2015).
16. Siciliano, S. D. et al. Soil fertility is associated with fungal and bacterial richness, whereas pH is associated with community composition in polar soil microbial communities. Soil Biol. Biochem. 78, 10–20 (2014).
17. Groffman, P. B. & Liptzin, D. Divergent national-scale trends of microbial and animal biodiversity revealed across diverse temperate soil ecosystems. Nat. Commun. 10, 1107 (2019).
18. Xu, X., Thornton, P. E. & Post, W. M. A global analysis of soil microbial biomass carbon, nitrogen and phosphorus in terrestrial ecosystems. Glob. Ecol. Biogeogr. 22, 737–749 (2013).
19. Fierer, N. & Jackson, R. B. The diversity and biogeography of soil bacterial communities. Proc. Natl. Acad. Sci. USA 103, 626–631 (2006).
20. Blaser, M. J. et al. Toward a predictive understanding of Earth’s microbiomes to address 21st century challenges. mBio 7, e00174-16 (2016).
21. Yu, X., Polz, M. F. & Alm, E. J. Interactions in self-assembled microbial communities saturate with diversity. ISME J. https://doi.org/10.1038/ismej.2019.0356-5 (2019).
22. Elass, J. Dvan et al. Microbial diversity determines the invasion of soil by a bacterial pathogen. Proc. Natl. Acad. Sci. USA 109, 1159–1164 (2012).
23. Lee, K. C. et al. Stochastic and deterministic effects of a moisture gradient on soil microbial communities in the McMurdo dry valleys of Antarctica. Front. Microbiol. 9, 2619 (2018).
24. Wang, G. & Or, D. A hydration-based biophysical index for the onset of soil microbial coexistence. Sci. Rep. 2, 881 (2012).
25. Kim, M. & Or, D. Individual-based model of microbial life on hydrated rough soil surfaces. PLoS ONE 11, e0147394 (2016).
26. Wang, G. & Or, D. Aqueous films limit bacterial cell motility and colony expansion on partially saturated rough surfaces: aqueous films limit bacterial motion. Environ. Microbiol. 12, 1363–1373 (2010).
27. Tecon, R., Ebrahimi, A., Kleyer, H., Levi, S. E. & Or, D. Cell-to-cell bacterial interactions promoted by drier conditions on soil surfaces. Proc. Natl. Acad. Sci. USA 115, 9791–9796 (2018).
28. Fatchi, S., Mannoni, S., Or, D. & Paschalos, A. A mechanistic model of microbially mediated soil biogeochemical processes—a reality check. Glob. Biogeochem. Cycles. https://doi.org/10.1029/2018GB006077 (2019).
29. Fierer, N., Strickland, M. S., Liptzin, D., Bradford, M. A. & Cleveland, C. C. Global patterns in belowground communities. Ecol. Lett. 12, 1238–1249 (2009).
30. Schoolfield, R. M., Sharpe, P. J. H. & Magnuson, C. E. Non-linear regression of biogeochemical temperature-dependent rate models based on absolute reaction rate theory. J. Theor. Biol. 88, 719–731 (1981).
31. Bickel, S., Chen, X., Papritz, A. & Or, D. A hierarchy of environmental covariates control the global biogeochemistry of soil bacterial richness. Sci. Rep. 9, 1–10 (2019).
32. Neilson, J. W. et al. Significant impacts of increasing aridity on the arid soil microbiome. mSystems 2, e01915–16 (2017).
33. Banerjee, S. et al. Legacy effects of soil moisture on microbial community structure and N2O emissions. Soil Biol. Biochem. 95, 40–50 (2016).
34. Non, N., Leopoud, J., Ruempel, S., Pouteau, V. & Chen, C. Effects of habitat constraints on soil microbial community function. Sci. Rep. 7, 4280 (2017).
35. Zelezniak, A. et al. Metabolic dependencies drive species co-occurrence in diverse microbial communities. Proc. Natl. Acad. Sci. USA 112, 6449–6454 (2015).
36. Slasaere, E. W. et al. Water balance creates a threshold in soil pH at the global scale. Nature 540, 567–569 (2016).
37. Zhao, M., Heinsch, F. A., Nemani, R. R. & Running, S. W. Improvements of the MODIS terrestrial gross and net primary production global data set. Remote Sens. Environ. 95, 164–176 (2003).
38. Lieth, H. in Primary Productivity of the Biosphere (eds Lieth, H. & Whittaker, R. H.) 257–263 (Springer Berlin Heidelberg, 1975).
39. Tick, S. E. & Hijmans, R. J. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas: new climate surfaces for global land areas. Int. J. Climatol. 37, 4302–4315 (2017).
40. Waring, B. G., Averill, C. & Hawkes, C. V. Differences in fungal and bacterial physiology alter soil carbon and nitrogen cycling: insights from meta-analysis and theoretical models. Ecol. Lett. 16, 887–894 (2013).
41. Whitman, W. B., Coleman, D. C. & Wiebe, W. J. Prokaryotes: the unseen majority. Proc. Natl. Acad. Sci. USA 95, 6578–6583 (1998).
42. Töth, B. et al. New generation of hydraulic pedotransfer functions for Europe: New hydraulic pedotransfer functions for Europe. Eur. J. Soil Sci. 66, 228–238 (2015).
43. Hengl, T. et al. SoilGrids250m: Global gridded soil information based on machine learning. PLoS ONE 12, e0169748 (2017).
44. Beck, H. E. et al. MSWEP V2 global 3-hourly 0.1° precipitation: methodology and quantitative assessment. Bull. Amer. Meteor. Soc. 100, 473–500 (2019).
45. Jensen, M. E. & Haise, R. H. Estimating evapotranspiration from solar radiation. Proc. Am. Soc. Civil Eng. J. Irrig. Drain. Div. 89, 15–41 (1963).
46. Stauffer, D. Scaling theory of percolation clusters. Phys. Rep. 54, 1–74 (1979).
47. Scher, H. & Zallen, R. Critical density in percolation processes. J. Chem. Phys. 53, 3759–3761 (1970).
48. Shirazi, M. A. & Boersma, L. A unifying quantitative analysis of soil texture. Soil Sci. Soc. Am. J. 48, 142–147 (1984).
49. Hill, M. O. Diversity and evenness: a unifying notation and its consequences. Ecology 54, 427–432 (1973).
50. Keller, E. & Segel, L. Model for chemotaxis. J. Theor. Biol. 30, 225–234 (1971).