

## CLIMATE CHANGE

## Climate legacies drive global soil carbon stocks in terrestrial ecosystems

Manuel Delgado-Baquerizo,<sup>1,2\*</sup> David J. Eldridge,<sup>3</sup> Fernando T. Maestre,<sup>4</sup> Senani B. Karunaratne,<sup>1</sup> Pankaj Trivedi,<sup>1,5</sup> Peter B. Reich,<sup>1,6</sup> Brajesh K. Singh<sup>1,7</sup>

2017 © The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. Distributed under a Creative Commons Attribution License 4.0 (CC BY).

Climatic conditions shift gradually over millennia, altering the rates at which carbon (C) is fixed from the atmosphere and stored in the soil. However, legacy impacts of past climates on current soil C stocks are poorly understood. We used data from more than 5000 terrestrial sites from three global and regional data sets to identify the relative importance of current and past (Last Glacial Maximum and mid-Holocene) climatic conditions in regulating soil C stocks in natural and agricultural areas. Paleoclimate always explained a greater amount of the variance in soil C stocks than current climate at regional and global scales. Our results indicate that climatic legacies help determine global soil C stocks in terrestrial ecosystems where agriculture is highly dependent on current climatic conditions. Our findings emphasize the importance of considering how climate legacies influence soil C content, allowing us to improve quantitative predictions of global C stocks under different climatic scenarios.

## INTRODUCTION

Soils store three times more carbon (C) than either the atmosphere or terrestrial vegetation (1). Moreover, soil C storage is one of the most important ecosystem processes for humans because it plays critical roles in supporting key ecosystem services such as climate regulation, soil fertility, and fiber and food production (2, 3). Short-term climate legacies (that is, days to decades) have been reported to strongly influence key ecosystem processes such as plant productivity and litter decomposition in terrestrial systems (4–6). However, much less is known about the extent to which long-term climate legacies (that is, centuries to millennia) control such processes. Projections into the future are conditional upon the past, but key ecosystem and biogeochemical variables, such as soil C, strongly reflect site history across centennial to millennial time scales (1). Thus, quantifying the influence of long-term climate legacies on soil C stocks can be of paramount importance to better understand soil C cycle–climate change feedbacks and to improve ecosystem and earth system simulation models, which are primary tools for predicting climate change impacts on soil C at the global scale (7). Although a theoretical framework based on ecological principles is emerging to explain long- versus short-term legacy impacts on ecosystem functions, such as plant productivity (5), we still lack empirical evidence to support this framework.

One of the most significant developments in the field of soil sciences in recent years involves the increased recognition that most soils are polygenetic, that is, archival products of pedogenic processes that vary widely over time (8). For example, we know from chronosequences that C accumulates in the soil over thousands of years during pedogenesis (at least as long as 10,000 years) (8–12), wherein the persistence of soil C is

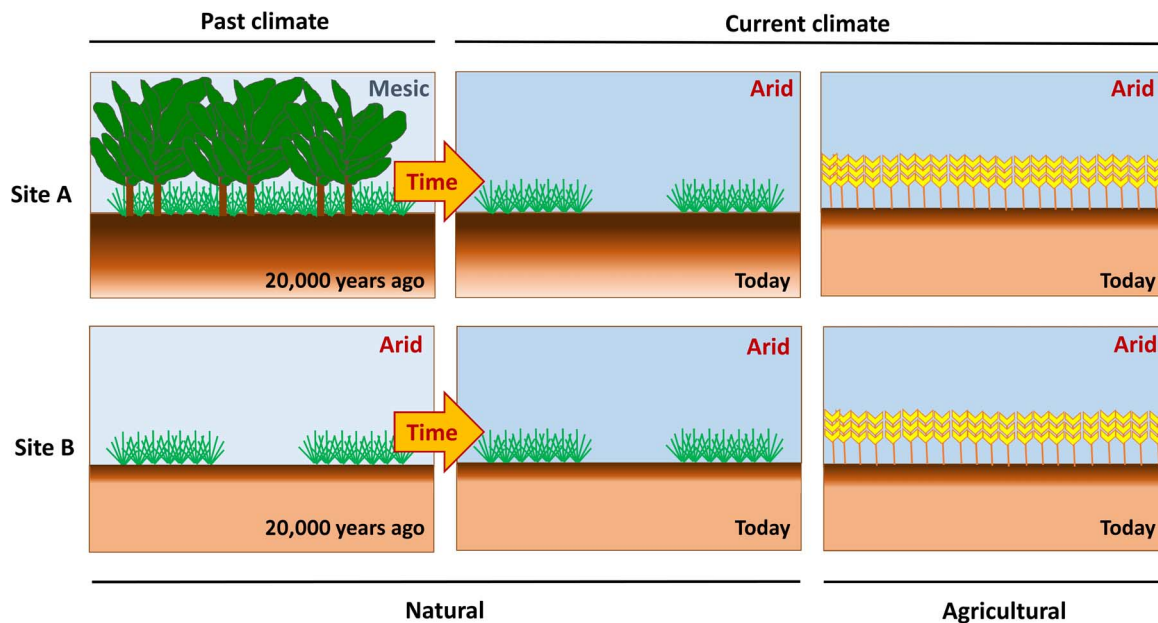
driven largely by abiotic and biotic factors (9, 12). However, the climate of a particular area changes over decades and centuries, resulting in large-scale biome migrations (5). Variations in vegetation type and plant productivity, both of which are linked to shifting climate and biome migrations, regulate the rates at which C is fixed from the atmosphere via photosynthesis and stored in the soil (13, 14). The long-term climate history of a region, rather than the current climatic conditions, could, in theory, be a better proxy of the amount of C that has been stored in a particular terrestrial ecosystem over many centuries. For example, a grassland under a current dry climate, which was previously a forest ecosystem under a wetter paleoclimate, may now have a greater amount of soil C than expected based on its current climate (Fig. 1) (15, 16). Despite the importance of soil C for human well-being, we lack empirical evidence on the relative importance of past climates in relation to current climatic conditions as drivers of global soil C stocks in terrestrial ecosystems.

Shifts in land use from natural systems to agriculture have been shown to markedly reduce the amount of soil C as a result of rapid C degradation and soil erosion linked to land clearing and cultivation (Fig. 1) (12, 17, 18). Thus, human-induced disturbances will have a greater impact on soil C stocks than those imposed only through climate legacies in natural systems, enhancing the importance of current climate in driving soil C stocks (Fig. 1). The global pressure on soils is expected to increase exponentially during this century due to the agricultural intensification needed to meet the increasing food demand to sustain a growing global population (19). Improving our understanding of the role of human disturbance in shifting the relative contribution of paleo- versus current climatic conditions in regulating soil C storage is critical to improve our ability to accurately predict soil C storage in terrestrial ecosystems.

We tested the following hypotheses: (i) long-term climate legacies influence contemporary global soil C stocks in terrestrial ecosystems, and (ii) human disturbance alters the relative contribution of long-term (that is, paleoclimate) compared with current climate as drivers of these stocks. To do this, we gathered data from three independent large-scale surveys: the International Soil Reference and Information Centre's (ISRIC's) World Soil Information Service ("Global-WoSIS" hereafter; Materials and Methods) (20) including 4381 sites, data from a global field survey including 224 dryland sites from all continents except Antarctica ("Global-Drylands" hereafter) (21), and data from a regional survey including 450 sites scattered across a 400-km<sup>2</sup> area of eastern Australia ("Australia" hereafter; see Materials and Methods) (fig. S1).

<sup>1</sup>Hawkesbury Institute for the Environment, University of Western Sydney, Building L9, Locked Bag 1797, Penrith South, New South Wales 2751, Australia. <sup>2</sup>Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, CO 80309, USA. <sup>3</sup>Centre for Ecosystem Science, School of Biological, Earth and Environmental Sciences, University of New South Wales, Sydney, New South Wales 2052, Australia. <sup>4</sup>Departamento de Biología y Geología, Física y Química Inorgánica, Escuela Superior de Ciencias Experimentales y Tecnología, Universidad Rey Juan Carlos, Calle Tulipán Sin Número, Móstoles 28933, Spain. <sup>5</sup>Department of Bioagricultural Sciences and Pest Management, Colorado State University, Fort Collins, CO 80523, USA. <sup>6</sup>Department of Forest Resources, University of Minnesota, St. Paul, MN 55108, USA. <sup>7</sup>Global Centre for Land Based Innovation, University of Western Sydney, Penrith South, New South Wales 2751, Australia.

\*Corresponding author. Email: m.delgadobaquerizo@uws.edu.au



**Fig. 1. A theoretical framework explaining the effects of climatic legacies on soil C stocks in natural and agricultural areas.** Higher color intensity in soil represents more soil carbon. In the example, a grassland under a current dry climate, which was previously a forest ecosystem (site A) and developed under a wetter paleoclimate now, has a greater amount of soil C than expected based on its current climate or compared to a contemporary arid grassland subjected to arid paleoclimate (site B). Shifts in land use from natural systems to agriculture have been shown to markedly reduce the amount of soil C as a result of rapid C degradation and soil erosion linked to land clearing and cultivation.

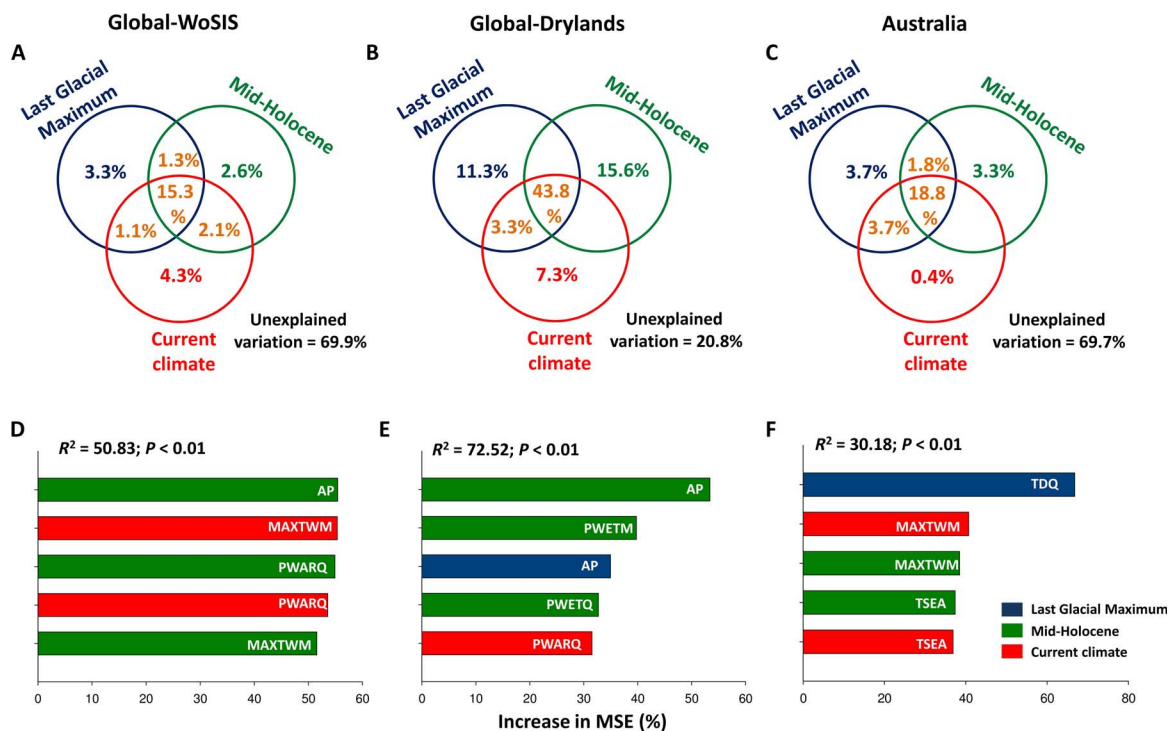
These data sets included information on soil carbon stocks ( $\text{kg C m}^{-2}$ ; Global-WoSIS and Australia) or concentrations (%) Global-Drylands) from the uppermost soil layer (top ~10 cm). These data sets include global and regional scale data comprising a wide range of paleo- and current climatic conditions. This characteristic allows sufficient climatic variation to test the relative importance of paleo- versus current climate in predicting soil C stocks. We obtained information on 19 bioclimatic variables (Table 1) from current and paleoclimates [climates experienced in mid-Holocene (about 6000 years ago) and Last Glacial Maximum (about 22,000 years ago)] using global climate models (see Materials and Methods) (22, 23). The reason for including more than one paleoclimate in our analyses is that changes in soil carbon stocks occur gradually with time as a consequence of changing conditions (8) rather than because of the climate of a certain time period.

## RESULTS AND DISCUSSION

We first used variation partitioning modeling (24) to quantify the relative contribution of current versus mid-Holocene and Last Glacial Maximum climates in controlling contemporary soil C stocks (top ~10 cm; Materials and Methods). Variation partitioning is a method specifically recommended to deal with multicollinearity because it partitions the variance in a given response variable that is attributed to a particular group of predictors from that variance shared among all predictors (24). Our models explained 30.1, 79.2, and 30.3% of the variance in soil C for the Global-WoSIS, Global-Drylands, and Australia data sets, respectively (Fig. 2). They indicate that paleoclimate (mid-Holocene and Last Glacial Maximum combined) explained a greater unique amount of the variance in soil C stocks than current climate in terrestrial ecosystems at regional and global scales (5.9 versus 4.3%, 26.9 versus 7.3%, and 7.0 versus 0.5% of the variance of soil carbon in the Global-WoSIS, Global-Drylands, and Australia data sets, respectively; Fig. 2). Bioclimatic variables from the different climatic periods evaluated were

**Table 1. Bioclimatic variables included in this study.**

Bioclimatic variable	Acronym
Annual mean temperature	AMT
Mean diurnal range	MDR
Isothermality	ISO
Temperature seasonality	TSEA
Maximum temperature of warmest month	MAXTWM
Minimum temperature of coldest month	MINTCM
Temperature annual range	TRANGE
Mean temperature of wettest quarter	TWETQ
Mean temperature of driest quarter	TDQ
Mean temperature of warmest quarter	TWARQ
Mean temperature of coldest quarter	TCQ
Annual precipitation	AP
Precipitation of wettest month	PWETM
Precipitation of driest month	PDM
Precipitation seasonality	PSEA
Precipitation of wettest quarter	PWETQ
Precipitation of driest quarter	PDQ
Precipitation of warmest quarter	PWARQ
Precipitation of coldest quarter	PCQ



**Fig. 2. Relative contribution of paleo- (mid-Holocene and Last Glacial Maximum) and current climate as drivers of soil carbon stocks.** Results from variation partitioning modeling aiming to identify the percentage of variance of soil carbon explained by past and current climate variables for the Global-WoSIS (A), Global-Drylands (B), and Australia (C) data sets are shown. Shared effects of these variable groups are indicated by the overlap of circles. (D to F) Results from random forest analyses aiming to identify the top five significant ( $P < 0.05$ ) bioclimatic variables regulating soil carbon for the three data sets used. Increase in the percentage of MSE is equal to the increase in the mean square error. Acronyms are available in Table 1.

highly correlated (table S1), which explains the high percentage of shared variance in predicting soil C stocks among climate periods observed (Fig. 2, A to C). Shared variation explaining C stock among different predictors (for example, current climate and mid-Holocene) cannot be attributed particularly to any of those climatic periods. Because of this, we only compared the unique portion of the variation in soil C stocks explained in a singular manner by either paleoclimate or current climate. Similar results were found after controlling for spatial influence in all cases (fig. S2) and also across bioclimatic regions (that is, tropical, temperate, continental, and arid) within the Global-WoSIS data set (fig. S3). Most importantly, these results were also maintained after explicitly accounting for other important predictors of soil C stocks such as space (latitude, longitude, and altitude), soil properties (soil pH, electrical conductivity, and sand content), and biotic factors (species richness and plant cover) simultaneously in the Global-Drylands data set, where these data were available (fig. S4). Only the Global-WoSIS data set contained soil C stock information for soil below 10-cm depth (www.isric.org/data/wosis) (25). Using this data set, we found a highly significant positive correlation between soil C stocks from the top 10 cm and those from all other soil depths (10 to 20, 20 to 50, 50 to 100, 0 to 20, 0 to 50, and 0 to 100 cm) (table S2). Our variation partitioning modeling provided evidence that paleoclimate still predicted a unique portion of the variance that cannot be predicted by current climate at 10 to 20, 20 to 50, 50 to 100, 0 to 20, 0 to 50, and 0 to 100 cm (figs. S5 and S6). Our findings suggest that paleoclimatic information could also be used to improve predictions of soil C stocks in the deep layers of soil (from 50 to 100 cm), where current model performance to predict soil organic C content is still largely limited (26).

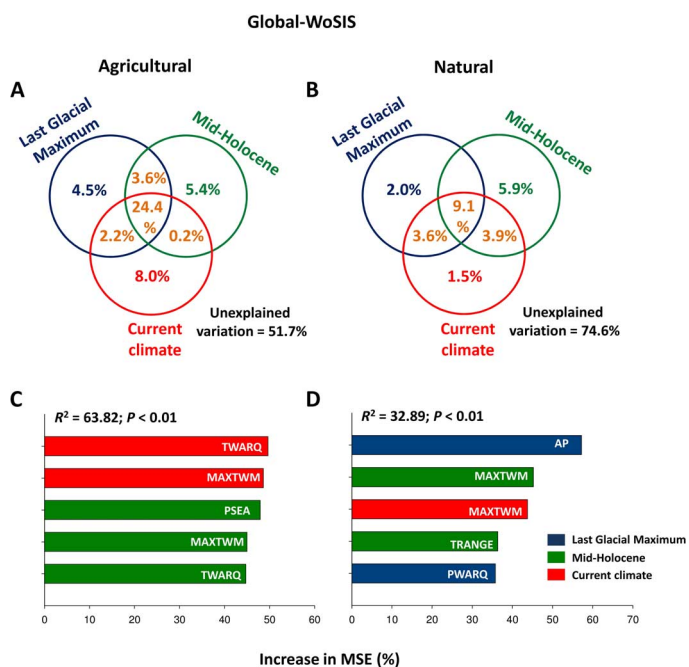
Within the Global-WoSIS data set, the footprint of paleoclimate legacies on soil C stocks was particularly noticeable in the midlatitudes compared with the tropics (fig. S7). Here, the interaction between paleo- and current climate was a stronger driver of soil C stocks than the climate of a particular period per se (fig. S7). These differences in the role of paleoclimate in driving soil C stocks observed between midlatitude and tropical areas may be related to the extent of the last glaciation in these regions. Middle latitudes were highly affected by the last glaciations, and their climatic conditions have substantially changed over the last 20,000 years (12). These climatic changes likely influenced the rate of soil C fixation in these regions, and our results suggest that they have left an imprint in soil C stocks that can still be identified nowadays. Conversely, climatic conditions have been more stable in the tropics (versus middle latitudes) during the last millennia due to the lack of effects from last glaciations (12). Therefore, the paleoclimatic footprint is difficult to observe in these areas. Together, our findings accord with recent studies highlighting the importance of past climate as a driver of ecosystem attributes, such as litter decomposition and biodiversity in terrestrial ecosystems (6, 27, 28), and provide, to the best of our knowledge, the first empirical demonstration that climatic legacies continue to drive the magnitude of contemporary soil C stocks at regional and global scales in terrestrial ecosystems.

We used random forest modeling (29) to identify the main past and/or current climatic drivers of soil C stocks and structural equation modeling (SEM) (30) to clarify their relative influence. These techniques are specially recommended to identify the main predictors of environmental response variables (see Materials and Methods). Our random forest models explained 50.8, 72.5, and 30.2% of the variance in soil carbon

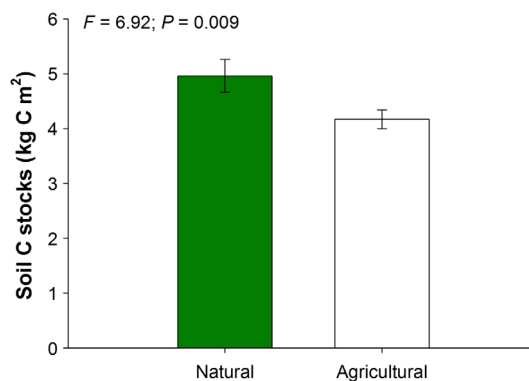
for the Global-WoSIS, Global-Drylands, and Australia data sets, respectively (Fig. 2, D to F). Precipitation from paleoclimates and maximum temperature in the warmest month from current and mid-Holocene climates were the main drivers of soil C stocks at the global scale (Fig. 2, D and E; see table S3 for a complete list of the soil C stock drivers evaluated). Structural equation models suggested that, in general, paleoclimate is more important than current climate in driving soil C stocks at the global scale (fig. S8, A and B). Precipitation variables were positively related to soil C content in the two global data sets studied (table S4). Precipitation is a key driver of plant growth and litter decomposition in terrestrial ecosystems and ultimately regulates the rates at which C is fixed from the atmosphere and stored in the soil (31, 32). As C accumulates in the soil over millennia in natural systems (8–10), the current soil C content reflects the climatic conditions experienced over this period. Conversely, temperature in the warmest month was negatively related to soil C content in both global data sets (table S4). Rising temperatures under climate change are predicted to promote soil C losses via soil respiration (12). Our study provides evidence that millennial precipitation and temperature legacies still influence soil C stocks today. Temperature in the driest quarter from the Last Glacial Maximum (positively related to soil C stocks) was the main driver of variations in soil C stocks at the regional scale in Australia (tables S4 and S5, Fig. 2F, and fig. S8C). Our random forest and SEM results suggest that paleoclimatic legacies linked to temperature regulated soil C stocks observed in eastern Australia, even within the limited precipitation range (280 to 648 mm) found in this data set

(Fig. 2F and fig. S8C). Together, our findings highlight the importance of considering climate legacies on soil C stocks to improve ecosystem and earth system and global change simulation models. Implications of our results can be considered to be minimal for global drylands because of their relatively low capacity to store C (33). However, drylands occupy ~45% of Earth’s land mass (34) and regulate the global C budget. Global drylands comprise 27% (431 GT of organic C) of the global soil organic C reserves (35). Thus, our findings further emphasize the need to consider climatic legacies when estimating soil C stocks in global drylands. This is critical for achieving global sustainability because drylands support 38% of global human populations that rely heavily on ecosystem services linked to soil C stocks, such as forage and food production (36).

To test our second hypothesis (that is, soil disturbance from agriculture alters the relative contribution of paleo- compared with current climate in controlling soil C stocks), we analyzed a subset of data from the Global-WoSIS where we were able to partition sites between natural and agricultural systems (see Materials and Methods). We restricted our analyses to the Global-WoSIS data set because all samples from the Global-Drylands and Australia data sets belong to natural or semi-natural ecosystems. Current climate showed a larger relative contribution to the prediction of soil C stocks for agricultural (8.0%) than natural (1.5%) systems (Fig. 3). Strikingly, in both cases, paleoclimate (mid-Holocene and Last Glacial Maximum) was still the best predictor of soil C stocks in agricultural (9.9%) and natural (7.9%) areas. Similar results were found after accounting for spatial influence in our models (fig. S9). However, paleoclimate explained five times more variance in soil C than current climate in natural compared to agricultural systems (Fig. 3 and figs. S5 and S6). These results support the hypothesis that the predictive power of current climate on soil C stocks increases with disturbances associated with agricultural practices. Reductions in soil C linked to agricultural intensification reported here (Fig. 4) and elsewhere (12) may promote the dependency of C stocks on current climatic conditions from a particular area after human disturbances. Although past precipitation and temperature regimes continue to be the main drivers of soil C at a global scale for natural systems (fig. S8E), current temperature (fig. S8D) and past precipitation regimes are major drivers of soil C stocks in agricultural systems (fig. S8D). These results further suggest that the greater contribution of current climate in agricultural than in natural systems may be driven more by temperature than precipitation. This is not fully unexpected because irrigation has removed the dependency of many agricultural regions on natural precipitation (37).



**Fig. 3. Relative contribution of paleo- (mid-Holocene and Last Glacial Maximum) and current climate as drivers of soil carbon in agricultural ( $n = 1167$ ) and natural ( $n = 814$ ) systems from the Global-WoSIS data set. (A and B) Variation partitioning modeling aiming to identify the percentage of variance of soil carbon explained by past and current climate variables for the identified agricultural and natural systems from the Global-WoSIS. Shared effects of these variable groups are indicated by the overlap of circles. (C and D) Results from the random forest analyses aiming to identify the top five bioclimatic variables regulating soil carbon for the three data sets used. Increase in the percentage of MSE is equal to the increase in the mean square error. Acronyms are available in Table 1.**



**Fig. 4. Soil carbon stocks for agricultural ( $n = 814$ ) and natural ( $n = 1167$ ) ecosystems from the Global-WoSIS data set. Analyses of variance (ANOVAs) were used to test for differences between natural and agricultural systems.**

## CONCLUSIONS

Together, our findings suggest that paleoclimatic legacies should be considered in predictions of global soil C stocks in agricultural and natural terrestrial systems. Including paleoclimatic conditions in contemporary climate models will help us improve our capacity to predict soil C stocks under changing environments over the next century. This knowledge will protect us against unreasonable expectations about the likely magnitudes of increase in C stocks with short-term management practices such as minimum tillage or low-intensity grazing. Ultimately, we will be in a better position to predict how past climates are shaping our current lives and how they affect the ability of terrestrial ecosystems to provide a range of essential services, including nutrient cycling, soil carbon storage, and fiber and food production. Our findings further emphasize the importance of considering climate legacies when assessing soil C stocks to improve ecosystem and earth system simulation models, which are primary tools for predicting climate change impacts on the global C cycle.

## MATERIALS AND METHODS

### Study sites and data collection

#### Global-WoSIS

We used available online data from the ISRIC's WoSIS ([www.isric.org/data/wosis](http://www.isric.org/data/wosis)) (20). Here, we used 4381 sites from the Global-WoSIS, including original information on soil carbon content and bulk density across different soil horizons, which allowed us to calculate soil C stocks. We extracted 0 to 10 cm of soil C stock data ( $\text{kg C m}^{-2}$ ) after fitting an equal area quadratic spline (38). We adopted a  $\lambda$  value of 0.01 in fitting equal area quadratic splines for each considered soil pedon, and splines were fitted using version 2 of "Spline Tool for estimating soil attributes at standard depths" developed by CSIRO (Australian Soil Resource Information System 2011, Canberra, Australia) ([www.asris.csiro.au/methods.html](http://www.asris.csiro.au/methods.html)). Similarly, we obtained information on soil C stocks at 10 to 20, 20 to 50, and 50 to 100 cm for this data set. Data on 0 to 20, 0 to 50, and 0 to 100 cm were then calculated as the sum of C stocks for the previous estimations.

We also conducted analyses by splitting our data set into agricultural and natural ecosystems. To do so, we used four independent global land cover maps to identify the most likely land uses in each of the sites from the Global-WoSIS data set: global land cover map (GLC-2000) (39), GLC-SHARE (Global Land Cover Network; [www.glcn.org/databases/lc\\_glcshare\\_en.jsp](http://www.glcn.org/databases/lc_glcshare_en.jsp)), Globcover2009 from the European Space Agency ([http://due.esrin.esa.int/page\\_globcover.php](http://due.esrin.esa.int/page_globcover.php)), and global land cover maps obtained from the moderate resolution imaging spectroradiometer aboard NASA's Terra satellites ([http://neo.sci.gsfc.nasa.gov/view.php?datasetId=MCD12C1\\_T1](http://neo.sci.gsfc.nasa.gov/view.php?datasetId=MCD12C1_T1)). Only sites that were identified as agricultural or natural systems by all these global maps simultaneously (45.21% of the sites) were used for analyses, comparing agricultural to natural ecosystems. Soil C stocks were log-transformed to improve normality.

#### Global-Drylands

We used available online data from the EPES-BIOCOM global dryland survey (21) ([www.nature.com/nature/journal/v502/n7473/source\\_data/nature12670-f2.xls](http://www.nature.com/nature/journal/v502/n7473/source_data/nature12670-f2.xls)), which focuses on dryland ecosystems, defined as regions with an aridity index (precipitation/potential evapotranspiration) between 0.05 and 0.65 (40). This data set includes a wide variety of habitat types (grasslands, shrublands, and open woodlands) and environmental conditions (21). Field data were collected between 2006 and 2010 from 224 sites located in 16 countries from all continents except Antarctica (fig. S1B) according to a standardized sampling

protocol (21). At each site, a 30-m  $\times$  30-m plot was established under the most representative vegetation. A composite sample (that is, from five soil samples; 0 to 7.5 cm in depth) was randomly taken under the canopy of the dominant perennial plant species and in open areas devoid of perennial vegetation. After field collection, soil samples were taken to the laboratory, sieved ( $<2$  mm), and air-dried for 1 month. Soil C (percentage) was determined using the Walkley-Black chromic acid wet oxidation method (41). Bulk density information was not available for the Global-Drylands data set. However, in a subset of our data set where bulk density was available, soil C stocks ( $\text{kg C m}^{-2}$ ) were highly correlated with soil C concentration (percentage;  $\rho = 0.930$ ,  $P < 0.001$ ,  $n = 42$ ), providing further robustness to our approach. In the Global-Drylands data set, samples were collected in open areas and under the dominant vegetation patches; thus, all the soil variables in this data set were averaged to obtain site-level estimates by using the mean values observed in bare ground and vegetated areas, weighted by their respective cover at each site (21). Soil C was  $\log_{10}$ -transformed to improve normality.

#### Australia (regional scale)

This data set includes 450 sites scattered across a 400- $\text{km}^2$  area of eastern Australia (fig. S1C). Vegetation along this transect is dominated by dense woodlands to forests of blackbox (*Eucalyptus largiflorens*), white cypress pine (*Callitris glaucophylla*), and river red gum (*Eucalyptus camaldulensis*). This data set includes sites extensively used for livestock grazing, large areas dedicated for conservation (national parks and nature reserves), and smaller areas devoted to native forestry. Field data were collected in 2014. Each site comprised a 200-m-long transect running perpendicular to the nearest livestock watering point. Along this transect, we positioned five 0.25- $\text{m}^2$  (0.5 m  $\times$  0.5 m) plots every 50 m. Soils (0 to 5 cm in depth) were collected from the center of each quadrat, air-dried, ground, and passed through a 2-mm sieve to remove any roots or organic debris. Total soil C was determined using high-intensity combustion (LECO CNS-2000, LECO Corporation). Bulk density was determined with a 7-cm-diameter core to 5-cm depth, and the mass was determined after drying for 24 hours at 104°C. Soil C stocks ( $\text{kg C m}^{-2}$ ) (0 to 5 cm) were calculated from original total soil C and bulk density data. Soil C stocks were  $\log_{10}$ -transformed to improve normality.

#### Climate data

A total of 19 standardized climatic variables (Table 1) were obtained from all the sites surveyed from WorldClim ([www.worldclim.org](http://www.worldclim.org)) (22, 23). Average changes in mean annual temperature and precipitation from Last Glacial Maximum to current climate ranged from 955 to 1033 mm at 11.5° to 17.7°C, 426 to 421 mm at 11.0° to 15.5°C, and 407 to 428 mm at 12.9° to 16.6°C for the Global-WoSIS, Global-Drylands, and Australia data sets, respectively. In the case of mid-Holocene and Last Glacial Maximum climates, we used estimates provided by the Community Climate System Model (CCSM4; [www.cesm.ucar.edu/models/ccsm4.0/](http://www.cesm.ucar.edu/models/ccsm4.0/)) (42–44). We used data at a 2.5-min resolution (~4.5 km at equator) because this is the highest resolution available for the Last Glacial Maximum period. Bioclimatic data are also available for this resolution for current and mid-Holocene climates, allowing for the direct comparison among bioclimatic data at different periods. Climatic data are also available at the 30-s resolution for the current and mid-Holocene climates, which allowed us to compare the 2.5-min and 30-s resolution data for these two periods. Values calculated using a resolution of 2.5 min were identical to those calculated using a resolution of 30 s in all cases (Pearson's  $r > 0.99$ ;  $P < 0.001$ ).

## Statistical modeling

### Variation partitioning modeling

The main goal of this analysis was to quantify the relative importance of bioclimatic variables at different periods (Table 1) as predictors of soil C stocks. In particular, this analysis provides insights into whether climatic variables from current, mid-Holocene, and Last Glacial Maximum periods can explain a unique portion of the variance that is not explained by climate in other periods (24, 45). Analyses were done using C stocks ( $\text{kg C m}^{-2}$ ) for the Global-WoSIS and Australia data sets and C concentration (percentage) for the Global-Drylands data set; bulk density information was not available. Analyses in this section were repeated using the residuals from a multilinear regression between spatial variables (latitude and longitude) and soil C stocks as a response variable (that is, residuals of soil C stocks). The main goal of these analyses was to reduce the noise derived from spatial variables on soil C stocks because the residuals from these multilinear regressions are not influenced by either latitude or longitude. In all cases, variation partitioning analyses were conducted with the R package “vegan” (46). Because similar results were found across different soil depths for the Global-WoSIS data set (where this information was available) and for consistency with the other data sets included in this study, we conducted the rest of the statistical analyses using the top ~10-cm information.

### Random forest modeling

We conducted a classification random forest analysis (29) to identify the main bioclimatic predictors of soil C stocks. Contrary to the variation partitioning model described above, random forest analysis allowed us to identify the most important drivers of soil C among 19 bioclimatic variables from the different climatic periods studied (Table 1). This technique is a novel machine-learning algorithm that extends standard classification and regression tree (CART) methods by creating a collection of classification trees with binary divisions. Unlike traditional CART analyses, the fit of each tree is assessed using randomly selected cases (one of three of the data), which are withheld during its construction [out-of-bag (OOB) cases]. The importance of each predictor variable is determined by evaluating the decrease in prediction accuracy (that is, increase in the mean square error between observations and OOB predictions) when the data for that predictor is randomly permuted. This decrease is averaged over all trees to produce the final measure of importance. This accuracy importance measure was computed for each tree and averaged over the forest (9999 trees). Unlike multimodel inference using linear regressions or regression tree analyses, random forest analysis alleviates multicollinearity problems in multivariate analyses by building bagged tree ensembles and including a random subset of features for each tree (9999 trees). These analyses were conducted using the randomForest package (47) of the R statistical software, version 3.0.2 (<http://cran.r-project.org/>).

### Structural equation modeling

Using random forest analysis, we identified the top five predictors of soil C for each of the data sets and periods studied. We used SEM to further clarify, using an independent analysis, the relative importance of these soil C predictors. A useful characteristic of SEM for our purposes lies on its utility for partitioning the effects that a variable may have on another and for estimating the strengths of these multiple effects. Unlike regression or ANOVA, SEM offers the ability to separate multiple pathways of influence and view them as parts of a system and thus is useful for investigating the complex relationships among predictors commonly found in natural ecosystems (30). Thus, we included in our SEMs the best climatic predictor, selected by random forest analyses, at the three

studied periods and soil C as our response variable. We allowed climatic predictors to covariate in these models. All the SEM analyses were conducted using AMOS 20.0.

## SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at <http://advances.sciencemag.org/cgi/content/full/3/4/e1602008/DC1>

- table S1. Correlation (Pearson's  $r$ ) among bioclimatic variables across different time periods.  
 table S2. Correlations (Spearman  $\rho$ ) among soil C stocks estimated at 0 to 10, 10 to 20, 20 to 50, 50 to 100, 0 to 20, 0 to 50, and 0 to 100 cm of soil depth for the Global-WoSIS data set.  
 table S3. Results from random forest analyses aiming to identify the most important bioclimatic variables regulating soil C stocks for the three data sets used.  
 table S4. Correlations (Spearman  $\rho$ ) among bioclimatic variables across different time periods (current climate, mid-Holocene, and Last Glacial Maximum) and soil C contents for the Global-WoSIS, Global-Drylands, and Australia data sets.  
 table S5. Direct effects of current and past climate on soil C stocks and correlations among exogenous variables (that is, climate from different periods) extending results of the structural equation models shown in fig. S6.  
 fig. S1. Location of the sites included in the Global-WoSIS ( $n = 4381$ ), Global-Drylands ( $n = 224$ ), and Australia ( $n = 450$ ) data sets.  
 fig. S2. Relative contribution of paleo- (mid-Holocene and Last Glacial Maximum) and current climate of the residuals of soil C stocks (from a multilinear regression with latitude and longitude as predictors of soil C stocks).  
 fig. S3. Relative contribution of paleo- versus current climate in driving soil C in tropical ( $n = 1354$ ), temperate ( $n = 1566$ ), continental ( $n = 655$ ), and arid ( $n = 775$ ) ecosystems from the Global-WoSIS data set.  
 fig. S4. Relative contribution of paleoclimate, current climate, and other factors including space (latitude, longitude, and altitude), soil properties (soil pH, electrical conductivity, and sand content), and biotic features (total plant cover and species richness) in driving soil C stocks in the Global-Drylands data set.  
 fig. S5. Relative contribution of paleo- versus current climate in driving soil C across different soil depths: 10 to 20 cm (all sites,  $n = 4234$ ; agricultural sites,  $n = 1134$ ; and natural sites,  $n = 790$ ), 20 to 50 cm (all sites,  $n = 3797$ ; agricultural sites,  $n = 1046$ ; and natural sites,  $n = 670$ ), and 50 to 100 cm (all sites,  $n = 2400$ ; agricultural sites,  $n = 610$ ; and natural sites,  $n = 448$ ) for all sites available and also for the identified agricultural and natural systems from the Global-WoSIS.  
 fig. S6. Relative contribution of paleo- versus current climate in driving soil C across different soil depths: 0 to 20 cm (all sites,  $n = 4234$ ; agricultural sites,  $n = 1134$ ; and natural sites,  $n = 790$ ), 0 to 50 cm (all sites,  $n = 3786$ ; agricultural sites,  $n = 1046$ ; and natural sites,  $n = 674$ ), and 0 to 100 cm (all sites,  $n = 2349$ ; agricultural sites,  $n = 604$ ; and natural sites,  $n = 435$ ) for all sites available and also for the identified agricultural and natural systems from the Global-WoSIS.  
 fig. S7. Relative contribution of paleo- versus current climate in driving soil C stocks in middle latitudes ( $n = 2080$ ) and tropics ( $n = 2301$ ) for the Global-WoSIS data set.  
 fig. S8. Structural equation modeling aiming to identify the relative influence of the main bioclimatic variables from current, mid-Holocene, and land maximum climate (as identified by random forest analyses) on soil C stocks.  
 fig. S9. Relative contribution of paleo (mid-Holocene and Last Glacial Maximum) and current climate as drivers of the residuals of soil C stocks (from a multilinear regression with latitude and longitude as predictors of soil C stocks) in agricultural ( $n = 1167$ ) and natural ( $n = 814$ ) systems from the Global-WoSIS data set.

## REFERENCES AND NOTES

- M. W. I. Schmidt, M. S. Torn, S. Abiven, T. Dittmar, G. Guggenberger, I. A. Janssens, M. Kleber, I. Kögel-Knabner, J. Lehmann, D. A. C. Manning, P. Nannipieri, D. P. Rasse, S. Weiner, S. E. Trumbore, Persistence of soil organic matter as an ecosystem property. *Nature* **5**, 49–56 (2011).
- H. Tiessen, E. Cuevas, P. Chacon, The role of soil organic matter in sustaining soil fertility. *Nature* **371**, 783–785 (1994).
- United Nations Environment Programme (UNEP), *Year Book: Emerging Issues in our Global Environment* (United Nations Environment Programme, 2012).
- K. Ogle, J. J. Barber, G. A. Barron-Gafford, L. P. Bentley, J. M. Young, T. E. Huxman, M. E. Loik, D. T. Tissue, Quantifying ecological memory in plant and ecosystem processes. *Ecol. Lett.* **18**, 221–235 (2015).
- C. Monger, O. E. Sala, M. C. Duniway, H. Goldfuss, I. A. Meir, R. M. Poch, H. L. Throop, E. R. Vivoni, Legacy effects in linked ecological-soil-geomorphic systems of drylands. *Front. Ecol. Environ.* **13**, 13–19 (2015).
- M. S. Strickland, A. D. Keiser, M. A. Bradford, Climate history shapes contemporary leaf litter decomposition. *Biogeochemistry* **122**, 165–174 (2015).

7. P. M. Cox, R. A. Betts, C. D. Jones, S. A. Spall, I. J. Totterdell, Acceleration of global warming due to carbon-cycle feedbacks in a coupled climate model. *Nature* **408**, 184–187 (2000).
8. D. B. Richter, D. H. Yaalon, “The Changing Model of Soil” revisited. *Soil Sci. Soc. Am. J.* **76**, 766–778 (2012).
9. W. H. Schlesinger, Evidence from chronosequence studies for a low carbon-storage potential of soils. *Nature* **348**, 232–234 (1990).
10. J. W. Harden, R. K. Mark, E. T. Sundquist, R. F. Stallard, Dynamics of soil carbon during deglaciation of the Laurentide ice sheet. *Science* **258**, 1921–1924 (1992).
11. Y. Wang, R. Amundson, S. Trumbore, Radiocarbon dating of soil organic matter. *Quaternary Res.* **45**, 282–288 (1996).
12. W. H. Schlesinger, *Biogeochemistry, An Analysis of Global Change* (Academic Press, 1996).
13. J. M. Melillo, T. V. Callaghan, F. I. Woodward, E. Salati, E. S. K. Sinha, Effects on ecosystems, in *Climate change: The IPCC scientific assessment*, J. T. Houghton, G. T. Jenkins, J. J. Ephraums, Eds. (Cambridge Univ. Press, 1990), pp. 283–310.
14. M. T. Sykes, I. Colin Prentice, B. Smith, W. Cramer, S. Venevsky, An introduction to the European terrestrial ecosystem modelling activity. *Global Ecol. Biogeogr.* **10**, 581–593 (2001).
15. X. Xu, P. E. Thornton, W. M. Post, A global analysis of soil microbial biomass carbon, nitrogen and phosphorus in terrestrial ecosystems. *Global Ecol. Biogeogr.* **22**, 737–749 (2013).
16. W. R. Wieder, G. B. Bonan, S. D. Allison, Global soil carbon predictions are improved by modelling microbial processes. *Nat. Clim. Chang.* **3**, 909–912 (2013).
17. P. Smith, J. I. House, M. Bustamante, J. Sobocká, R. Harper, G. Pan, P. C. West, J. M. Clark, T. Adhya, C. Rumpel, K. Paustian, P. Kuikman, M. F. Cotrufo, J. A. Elliott, R. McDowell, R. I. Griffiths, S. Asakawa, A. Bondeau, A. K. Jain, J. Meersmans, T. A. M. Pugh, Global change pressures on soils from land use and management. *Glob. Chang. Biol.* **22**, 1008–1028 (2016).
18. A. Don, J. Schumacher, A. Freibauer, Impact of tropical land-use change on soil organic carbon stocks—A meta-analysis. *Glob. Chang. Biol.* **17**, 1658–1670 (2011).
19. World Development Report, *Agriculture for Development* (World Bank, 2008).
20. E. Ribeiro, N. H. Batjes, J. G. B. Leenaars, A. J. M. Van Oostrum, J. Mendes de Jesus, *Towards the Standardization and Harmonization of World Soil Data: Procedures Manual ISRIC World Soil Information Service (WoSIS version 2.0)* (ISRIC Report 2015-03, 2015).
21. F. T. Maestre, J. L. Quero, N. J. Gotelli, A. Escudero, V. Ochoa, M. Delgado-Baquerizo, M. García-Gómez, M. A. Bowker, S. Soliveres, C. Escolar, P. García-Palacios, M. Berdugo, E. Valencia, G. Gozalo, A. Gallardo, L. Aguilera, T. Arredondo, J. Blones, B. Boeken, D. Bran, A. A. Conceição, O. Cabrera, M. Chaieb, M. Derak, D. J. Eldridge, C. I. Espinosa, A. Florentino, J. Gaitán, M. G. Gatica, W. Ghiloufi, S. Gómez-González, J. R. Gutiérrez, R. M. Hernández, X. Huang, E. Huber-Sannwald, M. Jankju, M. Miriti, J. Moneris, R. L. Mau, E. Morici, K. Naseri, A. Ospina, V. Polo, A. Prina, E. Pucheta, D. A. Ramírez-Collantes, R. Romão, M. Tighe, C. Torres-Díaz, J. Val, J. P. Veiga, D. Wang, E. Zaady, Plant species richness and ecosystem multifunctionality in global drylands. *Science* **335**, 214–218 (2012).
22. R. J. Hijmans, S. E. Cameron, J. L. Parra, P. G. Jones, A. Jarvis, Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatol.* **25**, 1965–1978 (2005).
23. T. W. Wu, W. P. Li, J. J. Ji, X. G. Xin, L. Li, Z. Z. Wang, Y. W. Zhang, J. L. Li, F. Zhang, M. Wei, X. L. Shi, F. H. Wu, L. Zhang, M. Chu, W. H. Jie, Y. M. Liu, F. Wang, X. W. Liu, Q. P. Li, M. Dong, X. Y. Liang, Y. Gao, J. Zhang, Global carbon budgets simulated by the Beijing Climate Center Climate System Model for the last century. *J. Geophys. Res. Atmos.* **118**, 4326–4347 (2013).
24. P. Legendre, D. Borcard, D. W. Roberts, Variation partitioning involving orthogonal spatial eigenfunction submodels. *Ecology* **93**, 1234–1240 (2012).
25. T. Hengl, J. Mendes de Jesus, R. A. MacMillan, N. H. Batjes, G. B. M. Heuvelink, E. Ribeiro, A. Samuel-Rosa, B. Kempen, J. G. B. Leenaars, M. G. Walsh, M. RUIPerez Gonzalez, SoilGrids1km — Global soil information based on automated mapping. *PLOS ONE* **9**, e105992 (2014).
26. K. Adhikari, A. E. Hartemink, B. Minasny, R. Bou Kheir, M. B. Greve, M. H. Greve, Digital mapping of soil organic carbon contents and stocks in Denmark. *PLOS ONE* **9**, e105519 (2014).
27. S. K. Lyons, K. L. Amatangelo, A. K. Behrensmeyer, A. Bercovici, J. L. Blois, M. Davis, W. A. DiMichele, A. Du, J. T. Eronen, J. Tyler Faith, G. R. Graves, N. Jud, C. Labandeira, C. V. Looy, B. McGill, J. H. Miller, D. Patterson, S. Pineda-Munoz, R. Potts, B. Riddle, R. Terry, A. Tóth, W. Ulrich, A. Villaseñor, S. Wing, H. Anderson, J. Anderson, D. Waller, N. J. Gotelli, Holocene shifts in the assembly of plant and animal communities implicate human impacts. *Nature* **529**, 80–83 (2016).
28. P. Weigelt, M. Jonas Steinbauer, J. Sarmento Cabral, H. Kreft, Late Quaternary climate change shapes island biodiversity. *Nature* **532**, 99–102 (2016).
29. L. Breiman, Random forests. *Mach. Learn.* **45**, 5–32 (2001).
30. J. B. Grace, *Structural Equation Modeling Natural Systems* (Cambridge Univ. Press, 2006).
31. A. T. Austin, L. Yahdjian, J. M. Stark, J. Belpas, A. Porporato, U. Norton, D. A. Ravetta, S. M. Schaeffer, Water pulses and biogeochemical cycles in arid and semiarid ecosystems. *Oecologia* **141**, 221–235 (2004).
32. J. M. Briggs, A. K. Knapp, Interannual variability in primary production in tallgrass prairie: Climate, soil moisture, topographic position, and fire as determinants of above ground biomass. *Am. J. Bot.* **82**, 1024–1030 (1995).
33. S. M. F. Rabbi, M. Tighe, M. Delgado-Baquerizo, A. Cowie, F. Robertson, R. Dalal, K. Page, D. Crawford, B. R. Wilson, G. Schwenke, M. Mcleod, W. Badger, Y. P. Dang, M. Bell, G. O’Leary, D. Li Liu, J. Baldock, Climate and soil properties limit the positive effects of land use reversion on carbon storage in Eastern Australia. *Sci. Rep.* **5**, 17866 (2015).
34. R. Právělie, Drylands extent and environmental issues. A global approach. *Earth-Sci. Rev.* **161**, 259–278 (2016).
35. Millennium Ecosystem Assessment, Dryland systems, in *Millennium Ecosystem Assessment, Ecosystems and human well-being: Current state and trends* (Island Press, 2005), vol. 1, pp. 623–662.
36. D. V. Powell, D. M. Agnew, Assessing agricultural literacy elements of project food land and people in K–5 using the food and fiber systems literacy standards. *J. Agric. Educ.* **52**, 155–170 (2011).
37. J. A. Entry, R. E. Sojka, G. E. Shewmaker, Management of irrigated agriculture to increase organic carbon storage in soils. *Soil Sci. Soc. Am. J.* **66**, 1957–1964 (2002).
38. T. F. A. Bishopa, A. B. McBratney, G. M. Laslett, Modelling soil attribute depth functions with equal-area quadratic smoothing splines. *Geoderma* **91**, 27–45 (1999).
39. S. Fritz, E. Bartholomé, A. Belward, *Harmonization, Mosaicking, and Production of the Global Land Cover 2000 Database* (Joint Research Center, 2003).
40. United Nations Environment Programme, *World Atlas of Desertification* (Edward Arnold, ed. 2, 2002).
41. J. M. Anderson, J. S. I. Ingrassia, *Tropical Soil Biology and Fertility: A Handbook of Methods* (CABI, ed. 2, 1993).
42. P. R. Gent, G. Danabasoglu, L. J. Donner, M. M. Holland, E. C. Hunke, S. R. Jayne, D. M. Lawrence, R. B. Neale, P. J. Rasch, M. Vertenstein, P. H. Worley, Z.-L. Yang, M. Zhang, The community climate system model version 4. *J. Climate* **24**, 4973–4991 (2011).
43. N. Bystráková, S. W. Ansell, S. J. Russell, M. Grundmann, J. C. Vogel, H. Schneider, Present, past and future of the European rock fern *Asplenium fontanum*: combining distribution modelling and population genetics to study the effect of climate change on geographic range and genetic diversity. *Ann. Bot.* **113**, 453–465 (2013).
44. M. Tallavaara, M. Luoto, N. Korhonen, H. Järvinen, H. Seppä, Human population dynamics in Europe over the Last Glacial Maximum. *Proc. Natl. Acad. Sci. USA* **112**, 8232–8237 (2015).
45. D. Borcard, P. Legendre, P. Drapeau, Partialling out the spatial component of ecological variation. *Ecology* **73**, 1045–1055 (1992).
46. J. Oksanen, F. G. Blanchet, M. Friendly, R. Kindt, P. Legendre, D. McGlenn, P. R. Minchin, R. B. O’Hara, G. L. Simpson, P. Solymos, M. H. H. Stevens, E. Szocs, H. Wagner, *vegan: Community ecology package, R package version 2.3-0* (2015); <https://cran.r-project.org/web/packages/vegan/>.
47. A. Liaw, M. Wiener, Classification and regression by randomForest. *R news*, **2**, 18–22 (2002).

**Acknowledgments:** We thank T. Hengl (ISRIC’s WoSIS, Wageningen, Netherlands) for his helpful comments on a previous version of this manuscript. We acknowledge the use of data from ISRIC’s WoSIS, which are made available subject to the ISRIC data policy ([www.isric.org/data/wosis](http://www.isric.org/data/wosis)). We also thank all the members of the EPES-BIOCOM network for providing the Global-Drylands data. **Funding:** This study was supported by the European Research Council (ERC) under the European Community’s Seventh Framework Programme (FP7/2007-2013)/ERC Grant Agreement 242658 (BIOCOM), by the Australian Research Council (projects DP13010484, DP15010419, and DP170104634), and by the Grains Research and Development Corporation (UWS00008). F.T.M. acknowledges support from the ERC Grant Agreement 647038 (BIODESERT). D.J.E. was supported by a grant from the Hermon-Slade Foundation. M.D.-B. acknowledges support from the Marie Skłodowska-Curie Actions of the Horizon 2020 Framework Programme (H2020-MSCA-IF-2016) under a Research Executive Agency grant agreement (no. 702057). **Author contributions:** M.D.-B. designed this study and analyzed data from the GlobalWoSIS, Global-Drylands, and Australia data sets. The Global-Drylands and Australia data sets were originally designed by F.T.M. and D.J.E., respectively. S.B.K., B.K.S., and P.T. provided data on the GlobalWoSIS data set. The first draft of this paper was written by M.D.-B., edited by D.J.E., F.T.M., B.K.S., and P.B.R., and all co-authors significantly contributed to improve it. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Additional data related to this paper may be requested from the authors.

Submitted 24 August 2016

Accepted 14 February 2017

Published 12 April 2017

10.1126/sciadv.1602008

**Citation:** M. Delgado-Baquerizo, D. J. Eldridge, F. T. Maestre, S. B. Karunaratne, P. Trivedi, P. B. Reich, B. K. Singh, Climate legacies drive global soil carbon stocks in terrestrial ecosystems. *Sci. Adv.* **3**, e1602008 (2017).

## **Climate legacies drive global soil carbon stocks in terrestrial ecosystems**

Manuel Delgado-Baquerizo, David J. Eldridge, Fernando T. Maestre, Senani B. Karunaratne, Pankaj Trivedi, Peter B. Reich and Brajesh K. Singh

*Sci Adv* 3 (4), e1602008.  
DOI: 10.1126/sciadv.1602008

### ARTICLE TOOLS

<http://advances.sciencemag.org/content/3/4/e1602008>

### SUPPLEMENTARY MATERIALS

<http://advances.sciencemag.org/content/suppl/2017/04/10/3.4.e1602008.DC1>

### REFERENCES

This article cites 36 articles, 3 of which you can access for free  
<http://advances.sciencemag.org/content/3/4/e1602008#BIBL>

### PERMISSIONS

<http://www.sciencemag.org/help/reprints-and-permissions>

Use of this article is subject to the [Terms of Service](#)

---

*Science Advances* (ISSN 2375-2548) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. 2017 © The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. The title *Science Advances* is a registered trademark of AAAS.