Deeper topsoils enhance ecosystem productivity and climate resilience in arid regions, but not in humid regions

Yakun Zhang[1,*], Ankur R. Desai[2], Jingfeng Xiao[3], Alfred E. Hartemink[1]

[1] University of Wisconsin–Madison, Department of Soil Science, FD Hole Soils Lab, 1525 Observatory Drive, Madison, WI 53706, USA

[2] University of Wisconsin–Madison, Department of Atmospheric and Oceanic Sciences, 1225 W Dayton St, Madison, WI 53706, USA

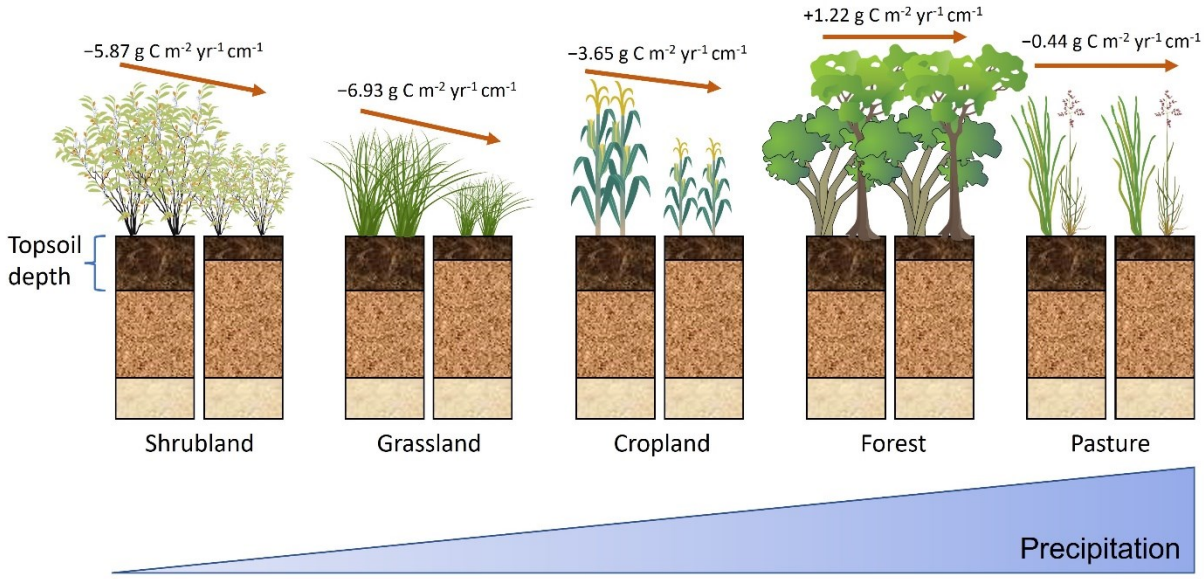
[3] Earth Systems Research Center, Institute for the Study of Earth, Oceans, and Space, University of New Hampshire, Durham, NH 03824, USA

* Corresponding author. Email: zhang878@wisc.edu

Highlight/Significance:

- Gross primary productivity (GPP) increased with deeper topsoils under cropland, grassland, and shrubland.
- Soils under forest have thin topsoils but the highest GPP.
- Topsoil depth controls GPP through water availability, which is vital for arid regions.
- Increases in GPP resilience to climatic extremes are associated to deeper topsoils in arid regions.

Graphical abstract



21

22

23 **Abstract**

24 Understanding the controlling mechanisms of soil properties on ecosystem productivity is
25 essential for sustaining productivity and increasing resilience under a changing climate. Here we
26 investigate the control of topsoil depth (e.g., A horizons) on long-term ecosystem productivity.
27 We used nationwide observations (n=2,401) of topsoil depth and multiple scaled datasets of
28 gross primary productivity (GPP) for five ecosystems (cropland, forest, grassland, pasture,
29 shrubland) over 36 years (1986–2021) across the conterminous USA. The relationship between
30 topsoil depth and GPP is primarily associated with water availability, which is particularly
31 significant in arid regions under grassland, shrubland, and cropland ($r=0.37$, 0.32, 0.15,
32 respectively, $p<0.0001$). For every 10 cm increase in topsoil depth, the GPP increased by 114 to
33 128 g C m^{-2} yr^{-1} in arid regions ($r=0.33$ and 0.45, $p<0.0001$). Paired comparison of relatively
34 shallow and deep topsoils while holding other variables (climate, vegetation, parent material, soil

type) constant showed that the positive control of topsoil depth on GPP occurred primarily in cropland (0.73, confidence interval of 0.57–0.84) and shrubland (0.75, confidence interval of 0.40–0.94). The GPP difference between deep and shallow topsoils was small and not statistically significant. Despite the positive control of topsoil depth on productivity in arid regions, its contribution (coefficients: 0.09–0.33) was similar to that of heat (coefficients: 0.06–0.39) but less than that of water (coefficients: 0.07–0.87). The resilience of ecosystem productivity to climate extremes varied in different ecosystems and climatic regions. Deeper topsoils increased stability and decreased the variability of GPP under climate extremes in most ecosystems, especially in shrubland and grassland. The conservation of topsoil in arid regions and improvements of soil depth representation and moisture-retention mechanisms are critical for carbon-sequestration ecosystem services under a changing climate. These findings and relationships should also be included in Earth system models.

Keywords: gross primary production, soil properties, structural equation modeling, climate change, climate extremes

1. Introduction

Terrestrial ecosystem productivity is essential for global food security and promoting carbon sequestration (Lorenz and Lal, 2009; Schmidhuber and Tubiello, 2007), but productivity is under pressure from climate change along with the increased frequency of fire, drought, floods, frost, and decreased biodiversity (Bellard et al., 2012; Grimm et al., 2013; Isbell et al., 2015; Wu et al., 2021; Xiao et al., 2016). Understanding different factors and mechanisms that control the variability of ecosystem primary productivity provides a scientific basis to sustain its productivity and increase its resilience. Soil is the main terrestrial reservoir of carbon, nutrients, and biota and serves as the core habitat for plant growth and the thriving of global ecosystems (Blum, 2005). Productive soils not only lead to higher crop yield (Bhardwaj et al., 2011) but also to enhanced resilience to climate change (Qiao et al., 2022). Most studies focused on biological aspects or were conducted in cropland. However, soil properties and mechanisms are important for all terrestrial ecosystems.

Among these properties, soil depth is important for regulating biogeochemical and hydrological cycles and ecosystem productivity (Shangguan et al., 2017). Soil formation from bedrock weathering is often accompanied by deepening soil regolith and the buildup of organic matter in the topsoil (Phillips, 2008), with formation rates controlled by climate, topography, and organisms (Jenny, 1941). However, soils are being lost since the past century from intensified land use change, agricultural practices, and climate extremes (Brown, 1984; Montgomery, 2007), which poses threats to ecosystem productivity (Berhe et al., 2018; Larson et al., 1983; Quinton et al., 2010). Deep topsoils have been related to high soil fertility and productivity. Studies have shown that for every 10 cm of soil loss, crop productivity dropped by 4.3% due to the loss of

nutrients and organic matter (Bakker et al., 2004). For every 2.5 cm of topsoil loss in the US, crop yield dropped up to 10% for wheat and corn (Lyles, 1975), and liming, fertilization, manure application, and other best management practices cannot compensate for the yield reduction from topsoil loss (Mielke and Schepers, 1986). Most of the studies on soil erosion have focused on cropland, but the capacity of soil depth and in particular the thickness of the A horizon to support productivity at other ecosystems and its effectiveness in different climate zones has not been extensively studied.

Topsoil depth, defined here as the thickness of the A horizon, is spatially variable (Francés and Lubczynski, 2011), due to soil forming factors (e.g., climate, vegetation, topography) and human-induced erosional and management effects (Zhang et al., 2023). The vulnerability of soil to erosion and the relationship between soil erosion and production differs by soil type (Larson et al., 1983). Plant productivity is also determined by many factors such as climate and varies for different vegetation types (O'Sullivan et al., 2020) but there are limited studies on the interactive effect of climate and soil on ecosystem productivity.

Here, we investigate the control of topsoil depth on ecosystem productivity using nationwide observations of topsoil depth and datasets of gross primary productivity (GPP). We aim to address the following questions: 1) Is the relationship of topsoil depth to GPP consistent for different ecosystems (e.g., natural vs. managed ecosystems)? 2) Is this relationship affected by different climatic conditions? 3) How strong is this relationship compared to other environmental controlling factors on GPP? 4) To what extent is the sensitivity of ecosystem productivity to climate change affected by topsoil depth? We hypothesize that the control of topsoil depth on ecosystem productivity is more significant in natural than in managed

ecosystems and in dry and cold environments. We hypothesize that deeper topsoils can increase the resilience of ecosystem productivity under climate change and climate extremes in arid regions. The findings of this study will advance our understanding of the role of soil in ecosystem productivity.

2. Materials and methods

2.1. Topsoil dataset

The soil dataset was obtained from the National Cooperative Soil Survey (NCSS) Soil Characterization Database which contains profile descriptions and analytical data for over 30,000 pedons collected across the USA (National Cooperative Soil Survey). The sampling locations were selected to represent each mapping unit of the SSURGO map (Soil Survey Staff), and the final dataset covered half of all the mapping units; some mapping units contained more than one pedon. The samples of each pedon were collected by horizon and analyzed using standard analytical methods (Schoeneberger et al., 2012). We selected pedons that: 1) have records of longitude and latitude coordinates and are within the conterminous US (CONUS); 2) have explicit sampling year and were sampled after 1986 at which the GPP dataset became available; 3) have data from horizons collected from the ground surface down to at least a B horizon or a C horizon if a B horizon does not exist, and do not have reporting, buried, D (the previous letter of indicating rock), vesicular, or limnic horizons, or bi-sequences or formed in human-transported materials; 4) have measurements of soil organic carbon (SOC) content and texture for A horizons; 5) have consistent land use type since 1938 belonging to cropland, forest, grassland, pasture, or

115 shrubland (see below for details); 6) have continuous measurements of GPP from 1986 to 2021
116 (see below for details). In total 2,401 pedons (about 8% of the total) fitted these criteria.
117 We calculated the topsoil depth in each pedon by summing up the thickness of all the A
118 horizons. Transitional horizons (e.g., AB, AE) were not included. Topsoil SOC and texture were
119 weight-averaged for A horizons using horizon thickness as weights. Soil texture was determined
120 by pipette method for particles smaller than 2 mm and textural classes were based on the USDA
121 classification (Soil Survey Staff, 2014). The total carbon was determined by dry combustion
122 method and inorganic carbon was determined using gas chromatography with HCl addition. The
123 SOC content was estimated by subtracting the inorganic carbon from the total carbon (Soil Survey
124 Staff, 2014). We used a generalized additive model to explore the spatial-temporal variation of
125 topsoil depth and found that the temporal term (year) was non-significant ($p>0.05$, data not
126 shown here). This indicated that the topsoil depth was stable for the study period (1986–2021),
127 and we did not consider a temporal change of topsoil depth in this study.
128
129 *2.2. Productivity data*
130 Ecosystem productivity was characterized by GPP from three different datasets. Although eddy
131 covariance flux towers provide *in-situ* GPP data, these data are often location-specific and cannot
132 represent the productivity of our sample locations. Instead, we used raster GPP products which
133 were calibrated from and correlated with the flux GPP. The first GPP dataset was derived from
134 the Landsat satellite data with 30-m spatial resolution and 16-day temporal variation since 1986
135 (Robinson et al., 2018). GPP was calculated using Landsat Surface Reflectance and MOD17
136 algorithm (Running and Zhao, 2015). The principle of the MOD17 algorithm calculated GPP from

a biome-specific Light Use Efficiency (LUE) and Absorbed Photosynthetically Active Radiation (APAR), with the former optimized using eddy covariance flux tower measurements and the latter estimated from the gridded shortwave radiation (Abatzoglou, 2013) and Landsat's Normalized Difference Vegetation Index – NDVI (Robinson et al., 2018). The dynamic biome information was provided by the National Land Cover Database (NLCD) from 1992, 2001, 2006, and 2011 (Yang et al., 2018). Smoothing and gap-filling of the Landsat's NDVI data were conducted using a climatology driven approach (Robinson et al., 2017). For details on this GPP product, readers can refer to Robinson et al. (2018).

The GPP dataset from 1986 to 2021 was extracted to pedon locations using the Nearest Neighbor method with the Google Earth Engine. As the soil profile was selected to best represent the mapping unit, we assumed that the soil thickness of the single profile is representative within the 30-m pixel of Landsat data. This means that the short-scale variation of soil is not considered in this study. The annual accumulated GPP was calculated for each location and then averaged for 35 years to obtain the temporal mean GPP. The pedons that have missing GPP data were removed, and descriptive statistics of GPP for different ecosystems were calculated.

Given the uncertainty of the MOD17 algorithm and the Landsat data, a multi-model comparison was conducted to minimize the effect of the uncertain GPP estimates on the subsequent analysis. Here, we included two other GPP datasets that were generated from different principles for comparison.

Solar-induced chlorophyll fluorescence (SIF) has been recently proposed as a better proxy for GPP (Li et al., 2018). It measures the sunlight-induced photon emission from plant chlorophyll in the range from 600 to 800 nm (Baker, 2008). SIF can be retrieved from satellite observations

159 at certain wavelengths between 600 and 800 nm. When heat dissipation occurs at high light
160 levels, SIF is strongly correlated with photosynthesis (Baker, 2008) and has a better performance
161 than the traditionally used vegetation index (e.g., NDVI). The GOSIF GPP dataset used here (Li
162 and Xiao, 2019b) was derived from a global, OCO-2 based SIF product (GOSIF) (Li and Xiao, 2019a).
163 GOSIF consists of 0.05° and 8-day SIF estimates globally, and was based on discrete SIF
164 observations from the Orbiting Carbon Observatory-2 (OCO-2), meteorological reanalysis data
165 (PAR, air temperature, and vapor pressure deficit), and MODIS enhanced vegetation index using
166 a machine learning method (Li and Xiao, 2019a). GOSIF GPP is based on GOSIF and GPP-SIF
167 relationships. To estimate GPP from SIF, GPP-SIF relationships established using GPP data and
168 OCO-2 SIF at a number of eddy covariance flux sites were used (Li and Xiao, 2019b). GOSIF GPP
169 consists of global GPP maps with a 0.05° spatial resolution and an 8-day time step from 2000 to
170 2021.

171 FluxCom initiative provides another way to estimate GPP globally at fine spatial and
172 temporal resolutions. It uses an ensemble of machine learning algorithms to build relationships
173 between eddy covariance flux tower measured GPP and remote sensing satellite data (e.g.,
174 MODIS) with and without ancillary meteorological forcings (Jung et al., 2020). In this study, we
175 used the annual GPP from the remote sensing and meteorological data-based (RS+METEO)
176 FluxCom product with 0.5° spatial resolution. The GPP data was presented as daily GPP data (g C
177 m^{-2} d^{-1}) for a specific year from 1980 to 2013 and summed for annual values.

178 The GOSIF GPP maps (2001–2021) and FluxCom GPP maps (1986–2013) were
179 downloaded from https://globalecology.unh.edu/data/GOSIF.html and
180 https://www.fluxcom.org/, respectively, and extracted to the sampling locations using *extract*

function of raster package (Hijmans et al., 2013) in R version 4.1.0 (R Core Team, 2021). The soil profiles collected before 2001 were removed from the GOSIF GPP dataset and the soil profiles collected after 2013 were removed from the FluxCom GPP dataset, which resulted in a total of 1,657 samples and 2,208 samples, respectively. The temporal mean GOSIF GPP and temporal mean FluxCom GPP were calculated by averaging annual GOSIF GPP and annual FluxCom GPP for 21 years (2001–2021) and 28 years (1986–2013), respectively.

2.3. Land use data

We used two land cover databases to ascertain that the land use of selected sample locations was consistent over a long-term period. The USGS projected land use/land cover mosaics covered 1938 to 2021 with a 250-m spatial resolution (Sohl et al., 2016; Sohl et al., 2014). The USGS NLCD was available for eight years: 2001, 2004, 2006, 2008, 2011, 2013, 2016, and 2019 with a 30-m spatial resolution (Yang et al., 2018). Land use types of the sample locations were extracted for every year and the sample locations that had experienced land use change from 1938 to 2021 or had different land use types based on the two databases were removed from the dataset. The final dataset included 2,401 pedons belonging to five land uses: cropland (n=699), forest (n=802), grassland (n=324), pasture (n=273), and shrubland (n=303).

2.4. Environmental data

The environmental data used in this study include climate variables from TerraClimate, Köppen-Geiger climate classification, topographic variables, soil orders, soil parent materials, Watershed Boundary Dataset, and irrigation types.

TerraClimate provides monthly climate data since 1958 with a 4.6-km spatial resolution (Abatzoglou et al., 2018). We chose TerraClimate data instead of PRISM and Daymet datasets because the Terraclimate data provides both monthly climate data (precipitation, minimum and maximum temperature, solar radiation) and climatic water balance data (actual and potential evapotranspiration). Additionally, the annual (from 1986 to 2021) and long-term averaged climate data (precipitation, minimum and maximum temperature) were highly correlated (Pearson correlation > 0.95) among these three climate datasets. Monthly precipitation (pr), minimum temperature (tmmn), maximum temperature (tmmx), actual evapotranspiration (aet), potential evapotranspiration (pet), and downward surface shortwave radiation (srad) were downloaded for sample locations for every month from 1958 to 2021 using Google Earth Engine. The annual sum of pr, aet, and pet, and annual mean of tmmn, tmmx, and srad were calculated for every year and then averaged for 64 years. The Aridity Index (AI) was calculated (Equation 1), and it represents arid and humid conditions for AI < 1 and AI > 1, respectively (Seager et al., 2018). In our dataset, there were 1,461 locations in arid regions and 940 locations in humid regions.

$$AI = \frac{pr}{pet} \qquad (Equation\ 1)$$

The global map of Köppen-Geiger climate classification was downloaded from http://www.gloh2o.org/koppen/ for present day (1980–2016) at a 0.0083° resolution (Beck et al., 2018). It was derived using the method described in Peel et al. (2007) with three air temperature datasets (WorldClim V1 and V2, CHELSA V1.2) and four precipitation datasets (WorldClim V1 and V2, CHELSA V1.2, and CHPclim V1) with a 0.0083° resolution, in which the CHPclim V1.2 was downscaled from 0.05° to 0.0083° resolution using bilinear interpolation (Beck et al., 2018). The

Köppen-Geiger classification has a hierarchy structure with five classes at the highest level: tropical (A), arid (B), temperate (C), cold (D), and polar (E). The class B precedes other classes, and A, C, D, E classes are mutually exclusive but not with B (Beck et al., 2018). The B was identified by mean annual precipitation < 10 × $P_{threshold}$, in which $P_{threshold}$ was determined by mean annual temperature and annual precipitation pattern. The threshold to identify temperate (C) and cold (D) was 0 °C for the coldest month according to Russell (1931) and 10 °C for the warmest month. The Köppen-Geiger classification was extracted to the sample locations using *extract* function of raster package (Hijmans et al., 2013) in R. In our dataset, samples were classified as arid (n=564), temperate (n=748), cold (n=1,086), and polar (n=3).

The elevation was extracted to sample locations from the USGS 3D Elevation Program 10-meter resolution Digital Elevation Model (DEM) dataset, from which the slope was calculated using Google Earth Engine (United States Geological Survey (USGS)). Soil order for each pedon was determined by NRCS soil scientists at the time of sampling and if missing, the gSSURGO (30-m resolution, https://www.nrcs.usda.gov/resources/data-and-reports/description-of-gridded-soil-survey-geographic-gssurgo-database) and STATSGO (1:250,000, https://catalog.data.gov/dataset/u-s-general-soil-map-statsgo2) maps were used to determine soil orders. Soil temperature (1:7,500,000) and moisture (1:9,000,000) regimes maps were obtained from USDA-NRCS. The soil parent material was obtained from the Conservation Science Partners Ecologically Relevant Geomorphology map (90-m resolution) using Google Earth Engine (Soller et al., 2009; Theobald et al., 2015). The Watershed Boundary Dataset (WBD, 1:24,000-scale) provided hydrologic unit (HU) data with a scale of 1:24,000 (WBD). The watershed level (HU10) was downloaded for each sample location from the Google Earth Engine. The irrigation

247 types (irrigated or rainfed) of the cropland samples were extracted from the 2017 MODIS
248 irrigated agricultural dataset (250-m resolution) (Brown et al., 2019).
249
250 *2.5. Statistical analysis*
251 Four types of analysis were used to answer our research questions and test our hypotheses: 1)
252 the linear regressions and Pearson correlations were calculated to explore the topsoil depth-GPP
253 relationship in five ecosystems and five climatic regions; 2) paired comparison of relatively
254 shallow and deep soils was conducted in five ecosystems and five climatic regions by controlling
255 other environmental factors constant; 3) structural equation modeling (SEM) was used to
256 evaluate the effect of topsoil depth and other essential factors (light, heat, water, fertility) and
257 their relative contribution to GPP; 4) the effect of topsoil depth on the resilience of GPP was
258 investigated under four climatic extremes (dry, wet, hot, and cold). We used long-term averaged
259 GPP and climate variables for the first three analyses and annual GPP and climate variables for
260 the last analysis. Below are the detailed explanations.
261 Pearson correlation coefficients (r) were calculated between topsoil depth and temporal
262 mean GPP across the full dataset, for each ecosystem (cropland, forest, grassland, pasture, and
263 shrubland), and in each climatic region (arid and humid regions classified by AI and arid,
264 temperate, and cold regions classified by Köppen-Geiger classification). It was also used to
265 explore the relationships between temporal mean GPP and other climatic and topographic
266 variables. The relationships between temporal mean GPP and topsoil depth were fitted using
267 simple linear regression (Equation 2) for five ecosystems and five climatic regions using *lm*
268 function in R. The a represents the intercept which is the GPP when topsoil depth is zero. The b

represents the slope which is the increase of GPP for every cm increase in topsoil depth. We further evaluated the topsoil depth-GPP relationship for samples with depth <75 cm (n=2,384) in five ecosystems and five climatic regions using Pearson correlation and simple linear regression, as this covered over 99% of total samples.

$$y = a + bx \qquad (Equation\ 2)$$

As the topsoil depth-GPP relationship may be confounded by many other environmental factors (e.g., climate, topography) at a national scale, to evaluate such a relationship at a local scale with other soil, topographic, and climatic factors remaining similar, we selected one pair of relatively deep and shallow topsoils for each watershed to compare their GPP values. We used watershed as the smallest spatial unit rather than other broader classifications (e.g., ecoregions), as a watershed represents the spatial movement of water (rainfall and snowmelt) across the landscape and is directly related to soil erosion and deposition. As such, soils within the same watershed often have similar hydrological patterns (e.g., hydroclimatology) and the differences in plant productivity can be easily explained by the differences in soil properties when holding other environmental factors constant (see below). In each watershed, the deep and shallow topsoils were not determined by absolute depth but relative depth difference between them. The depth difference of the pair should be greater than 3 cm if they are shallow (≤15 cm) or greater than 5 cm if they are deep (>15 cm). The paired topsoils have 1) the same climate conditions (long-term averaged precipitation and temperature), land use, soil order, soil temperature and moisture regimes, parent material, and textural class; 2) similar SOC content, elevation, and slope; 3) different GPP values; and 4) are within 4-km distance. If multiple pairs

fulfilled the criteria for a watershed, the pair that has the most similar soil properties and the largest depth difference was selected. A total of 103 pairs were selected.

For the final selected pairs from different watersheds, two cases may exist: 1) the deeper topsoil aligns with greater mean GPP (our hypothesis), and 2) the shallower topsoil has greater mean GPP (alternative hypothesis). Summary statistics (e.g., the proportion of pair-comparisons where the site with deeper topsoil had greater mean GPP) were calculated for these two cases in each ecosystem and each climatic region. The Wald test was used to calculate the 95% confidence interval of the proportion of pair-comparisons where the site with deeper topsoil had greater mean GPP (our hypothesis). If the confidence interval of the proportion covers 0.5, it indicates that the proportion of our hypothesis is not statistically significant from that of the alternative hypothesis. The Wald test was conducted using *BinomCI* function in *DescTools* package (Signorell et al., 2023) in R. We also conducted a paired *t*-test to compare the mean values of GPP between deep and shallow topsoils in a specific ecosystem using *t.test* function in R. The normality of GPP difference was tested using Shapiro-Wilk test (Shapiro and Wilk, 1965) with *shapiro.test* function in R.

Since soil and environmental factors were greatly variable across watersheds, which may affect the GPP-topsoil depth relationship of the paired samples, we further developed linear mixed-effects models and multiple linear regressions using selected paired samples to account for other soil and environmental factors. In the linear mixed-effects models, we used SOC, clay content, precipitation, minimum temperature, and an interaction term of topsoil depth (both binary data – deep and shallow and numeric values were evaluated respectively) and ecosystem or climatic regions as the fixed effects to predict GPP (Equation 3). We also added the watershed

as the random effect to acknowledge the variation among watersheds. The linear mixed-effects models were developed using *lmer* function of the *lme4* package (Bates et al., 2009) in R. The significance level of the model coefficients was calculated using *lmerTest* (Kuznetsova et al., 2017) and *afex* (Singmann et al., 2015) packages in R.

$$GPP \sim Depth_{binary}(or\ Depth_{numeric}) * Ecosystem\ (or\ Climatic\ region) + SOC + Clay + Precipitation + Minimum\ temperature + (1|Watershed) \quad (Equation\ 3)$$

To further test whether the increase in GPP is proportional to the increase in topsoil depth of the paired samples. We developed multiple linear regressions between the change of GPP (ΔGPP) and the change of topsoil depth ($\Delta Depth$) (Equation 4). The absolute change (Equation 5) and relative change (Equation 6) were evaluated respectively. We also added mean depth of the paired samples, SOC, clay content, precipitation, and minimum temperature as we expected these variables to have an effect. The multiple linear regressions were developed for each of the five ecosystems and five climatic regions using *lm* function in R.

$$\Delta GPP\ (absolute\ or\ relative) \sim \Delta Depth\ (absolute\ or\ relative) + Depth_{mean} + SOC + Clay + Precipitation + Minimum\ temperature \quad (Equation\ 4)$$

$$\Delta GPP_{absolute} = GPP_{deep} - GPP_{shallow};\ \Delta Depth_{absolute} = Depth_{deep} - Depth_{shallow} \quad (Equation\ 5)$$

$$\Delta GPP_{relative} = \frac{GPP_{deep} - GPP_{shallow}}{GPP_{shallow}};\ \Delta Depth_{relative} = \frac{Depth_{deep} - Depth_{shallow}}{Depth_{shallow}} \quad (Equation\ 6)$$

Structural equation modeling (SEM) is a multivariate regression method to examine the causal relationships between multiple variables and uses graphics to represent the complex structure (Grace, 2006). It separates the direct and indirect causes, represents partial contributions, models the latent variables and model structure (Grace, 2006), and has been used in soil ecology to solve complex casual relationships (Eisenhauer et al., 2015). We used SEMs to understand the contributions of soil and environmental variables to ecosystem productivity. In statistics, a latent variable is the one that can only be inferred from other observed variables using a mathematical model (Dodge et al., 2003). In social or natural sciences, a latent variable can be used to represent a conceptual abstract (e.g., attitude, ability) or characterize a group, and similar to observed variables, it can be used as an independent or dependent variable in models (Bollen and Hoyle, 2012).

Here, we define productivity, light, heat, water, fertility, and topsoil as six latent variables, which were inferred from other observed or measured variables, including GPP, srad, tmmn and tmmx, pr and aet, SOC and clay content, and topsoil depth, respectively. In SEM, we define productivity as a function of five latent variables (i.e., light, heat, water, fertility, and topsoil), and the interactions between the five latent variables were not investigated here (Equation 7). The SEMs were fit for five ecosystems and five climatic regions using the *sem* function in *lavaan* package (Rosseel, 2012) in R. If the full model did not converge, some observed variables were dropped until the model converged.

$$Productivity \sim Light + Heat + Water + Fertility + Topsoil \qquad (Equation\ 7)$$

The model performance of SEMs was evaluated using Confirmatory Factor Index (CFI), Tucker Lewis Index (TLI), and Root Mean Square Error of Approximation (RMSEA). The CFI measures the percent decrease in model chi-square (corrected by degrees of freedom, $\delta = \chi^2 - df$) of the User Model compared to the Baseline Model and ranges from 0 to 1 (best fit) (Equation 8). The TLI measures the percent decrease in relative chi-square ($\frac{\chi^2}{df}$) of the User Model compared to the Baseline Model and when it is closer to 1 (rounded to 1 if it is greater than 1), the model is better (Equation 9). The RMSEA measures the absolute model fitting performance and when it is smaller, the model is better (Equation 10).

$$CFI = \frac{\delta(Baseline) - \delta(User)}{\delta(Baseline)} \qquad (Equation\ 8)$$

$$TLI = \frac{\frac{\chi^2(Baseline)}{df(Baseline)} - \frac{\chi^2(User)}{df(User)}}{\frac{\chi^2(Baseline)}{df(Baseline)} - 1} \qquad (Equation\ 9)$$

$$RMSEA = \sqrt{\frac{\delta}{df(N-1)}} \qquad (Equation\ 10)$$

To evaluate the contribution of topsoil depth to climate resilience of productivity, the annual accumulative GPP, accumulative precipitation, mean minimum temperature from 1986 to 2021 (36 years) were used. Four climatic extremes from 1986 to 2021 were considered: dry, wet, hot, and cold. To identify the dry extreme for each location, the year which received the lowest precipitation was first identified from the 36 years and the GPP of this year was obtained. Then the percent changes in GPP and precipitation to the 36-year averaged GPP and precipitation were calculated. Similarly, the wettest, the hottest, and the coldest years were identified for each

location based on the highest precipitation, the highest minimum temperature, and the lowest minimum temperature, respectively, and the percent changes of GPP, precipitation, and minimum temperature were calculated. Potential lag effects from climate were not considered in our study. The relationships between percent changes of GPP at four types of climate extremes and topsoil depth were evaluated for the five climatic regions and five ecosystems. We hypothesize that if the topsoil is deeper, ecosystem productivity is more stable and hence the percent change of GPP is smaller (closer to zero).

To compare the percent change in GPP of shallow and deep topsoils, we calculated the mean topsoil depth in each climatic and ecosystem region as a threshold value and then split the data into shallow topsoils and deep topsoils. We compared the mean and variance of the percent change in GPP of shallow and deep topsoils. The mean value represents an absolute comparison of the percent GPP change, and we hypothesize that the mean percent change in GPP was closer to zero (no matter positive or negative) in deeper topsoils. The variance value represents the spread of the percent GPP change, and we hypothesize that deeper topsoils had more stable ecosystem productivity with climate change and thus the variance is smaller than that in shallower topsoils. The *t*-test and Levene's test were used to assess the homogeneity of mean and variance respectively in these two groups (shallow and deep topsoils) using *t.test* function and *leveneTest* function in *car* package (Fox et al., 2012) in R. If the homogeneity was not rejected in these tests, it indicates that the topsoil depth had non-significant effect on the percent change in GPP under climate extremes.

Given the uncertainty and noises of the Landsat GPP data, we used another two datasets (GOSIF GPP from 2001 to 2021 and FluxCom GPP from 1986 to 2013) to repeat two analyses: 1)

the simple linear regression between temporal mean GPP and topsoil depth in five different ecosystems and climatic regions was calculated; 2) the percent changes of GPP in four climatic extremes (dry, wet, hot, and cold) were calculated and their relationships with topsoil depth were evaluated for five climatic regions and five ecosystems. The *t*-test and Levene's test were also used to assess the homogeneity of mean and variance of percent GPP change for two soil depth groups (shallow and deep) in five climatic regions and five ecosystems.

3. Results

3.1. Topsoil depth and GPP in different ecosystems and climate zones

On the national scale, areas with deeper topsoil were not identical to the areas with higher GPP (Fig. 1a and 1b). Deeper topsoils primarily occurred in the Midwest (Fig. 1a) under cropland (mean depth=27 cm, mean GPP=1,249 g C m^{-2} yr^{-1}, Fig. 1c and 1d, Supplementary Table S1), while the highest GPP occurred in the east and along the West Coast (Fig. 1b) under forest (mean GPP=1,466 g C m^{-2} yr^{-1}, Fig. 1c and 1e) with topsoils of about 15 cm (Fig. 1a and 1d). The western CONUS was dominated by shrubland and grassland (Fig. 1c) which had shallow topsoils (mean depth=14 and 18 cm, Fig. 1a and 1d) and lower GPP (mean GPP=243 and 563 g C m^{-2} yr^{-1}, Fig. 1b and 1e). A weak positive correlation was observed between the GPP and topsoil depth for all data, but it varied for different ecosystems (Fig. 1f). A stronger positive correlation existed for grassland ($r=0.37$, $p=7e{-}12$) and shrubland ($r=0.32$, $p=2e{-}8$) with a weak correlation in cropland ($r=0.15$, $p=4e{-}5$), and no clear relationships for forest and pasture ($r=-0.04$ and 0.02 respectively, $p>0.05$).

As GPP varied significantly with climate and topography (Supplementary Figs. S1, S2, S3, and Tables S2 and S3), it confounds the effect of topsoil depth on GPP. The relationship between GPP and topsoil depth was examined for different climatic regions (Fig. 2). For the same topsoil depth, humid regions had a higher GPP (mean=1,493 g C m^{-2} yr^{-1}) than arid regions (mean=827 g C m^{-2} yr^{-1}, Fig. 2c, Supplementary Table S4). The correlation of topsoil depth and GPP was more evident in arid regions (AI<1, p<2e−16, r=0.33), while in humid regions (AI>1), the topsoil depth was slightly negatively correlated with the GPP (p=0.005, r=0.09, Fig. 2c). For every 10 cm increase in topsoil thickness, GPP increased by 114 g C m^{-2} yr^{-1} in arid regions. As for Köppen-Geiger classification, the correlation of topsoil depth and GPP was not statistically significant in the temperate and cold regions (Fig. 2d), but it was statistically significant in arid regions (p<2e−16, r=0.45). For every 10 cm increase in topsoil thickness, GPP increased by 128 g C m^{-2} yr^{-1} in arid regions of the Köppen-Geiger classification. This may indicate that the correlation of topsoil depth and GPP was primarily associated with water availability instead of temperature, and it was stronger in dry regions.

The samples with depth <75 cm showed similar topsoil depth-GPP relationship to that of the whole dataset (Supplementary Figs. S4 and S5). The correlations in shrubland (r=0.35 vs. 0.32) and dry regions (r=0.35 vs. 0.33 and r=0.49 vs. 0.45) were slightly stronger than that of the whole dataset. This is likely because in shallower topsoils, the changes of topsoil depth tend to be associated more with the GPP due to lack of available water and nutrients, while in deeper topsoils, such association is low.

438 *3.2. Association of topsoil depth with GPP at a local scale*

439 The association of topsoil depth and GPP at a national scale is impacted by mixing effects of
440 climate, vegetation type, topography, and soil type. To uncover the direct association, we
441 conducted a paired comparison of GPP of relatively deep and shallow topsoils in different
442 watersheds by holding other variables (climate, vegetation, parent material, soil type) constant
443 (Fig. 3a, Supplementary Fig. S6). The paired comparison contradicted our hypothesis that deeper
444 topsoils tend to have greater GPP (Fig. 3, Supplementary Fig. S7). The positive association of
445 topsoil depth with GPP occurred primarily in cropland (0.73, 95% confidence interval of 0.57–
446 0.84, 29 watersheds) and shrubland (0.75, 95% confidence interval of 0.40–0.94, 6 watersheds),
447 while in forest, grassland, and pasture, over half of the shallower topsoils had higher GPP than
448 paired deeper topsoils (Table 1). The percent increase of GPP in deeper topsoil over paired
449 shallower topsoil was greater in shrubland (mean=8%) followed by pasture (mean=5%, Fig. 3,
450 Table 1). However, the paired *t*-test showed that the difference of GPP in deep and shallow
451 topsoils was marginal and non-significant (Table 1). In the forest, the deeper topsoil was
452 associated with a decreased average GPP change (−1.3%). At different climatic regions, the
453 positive association of topsoil depth with GPP was marginal and not statistically significant, but
454 it was slightly higher in arid regions (0.61, 95% confidence interval of 0.48–0.72 and 0.41–0.78)
455 (Supplementary Table S5).

456 The linear mixed-effects models showed that the shallow topsoil was negatively
457 associated with GPP (coefficient=−0.01), and topsoil depth as a numeric variable was positively
458 associated with GPP (coefficient=0.0018) after accounting for the other environmental factors,
459 but such associations were not statistically significant (Table 2). In addition, the interaction of

460 depth and forest was statistically negatively associated with GPP (coefficient=−0.01) (Table 2),
461 which indicates that deeper topsoil was associated with smaller GPP in forest. The association of
462 GPP with ecosystem types and climatic variables (precipitation and temperature) was stronger
463 than that with soil characteristics (Table 2). When interacted with climatic regions, the shallow
464 topsoil was positively associated with GPP, while depth as a numeric variable was positively
465 associated with GPP (Supplementary Table S6), but such associations were not statistically
466 significant. The precipitation was solely associated with GPP in AI classification, while both
467 precipitation and temperature were associated with GPP in the Köppen-Geiger classification.
468 The absolute and relative changes of GPP did not show clear pattern with changes of
469 topsoil depth in five ecosystems and climatic regions (Supplementary Fig. S8). This was also
470 shown in the multiple linear regressions with non-statistically significant coefficients of $\Delta Depth$
471 (Supplementary Fig. S9). But such coefficients were slightly positive in most ecosystems and
472 climatic regions except for grassland. We also observed that the absolute changes of GPP and the
473 coefficient of $\Delta Depth_{absolute}$ were small in shrubland, but their relative changes and the
474 coefficient of $\Delta Depth_{relative}$ were markedly higher (Fig. 3, Supplementary Fig. S9). This may
475 indicate that relative changes of topsoil depth and GPP on small absolute numbers can be more
476 significant than they are in shrubland. The mean depth also showed statistically significant
477 negative association with GPP change in forest.
478
479 *3.3. Association of topsoil depth and other environmental factors with GPP*
480 To comprehensively examine the contributions of topsoil depth (topsoil) and other soil and
481 environmental factors (light, heat, water, fertility) to ecosystem productivity, SEMs

(Supplementary Fig. S10) were developed for each ecosystem (Fig. 4) and climatic region (Supplementary Fig. S11). The statistically significant variables and their coefficients are shown in the SEM results (Fig. 4, Supplementary Fig. S11). Consistent with our linear regressions, topsoil depth was positively associated with the productivity in cropland, grassland, and shrubland (dry regions, coefficients: 0.09 to 0.13), but the association was not statistically significant in soils under forest and pasture (humid regions) (Fig. 4). Likewise, a positive association of topsoil depth with productivity was observed in arid regions (coefficients: 0.17 and 0.33), but it was negative in humid regions (coefficient=−0.26) (Supplementary Fig. S11). Soil fertility indicated by SOC and clay content was positively associated with the productivity in cropland and shrubland (coefficients=0.06), but negatively associated with the pasture and humid regions (coefficients: −0.23 and −0.06). In pasture, the soils with higher clay content can be more easily compacted from field traffic and form restrictive layers for root development, which may lead to lower productivity. In humid regions, the SOC and clay content was higher towards the north (Supplementary Fig. S2), which was slightly opposite to the increasing GPP towards the southeast. This may indicate that other factors (e.g., climate) may play a more important role in affecting GPP than soil fertility at this scale.

Water (pr and aet) played the most important role in enhancing productivity in all the ecosystems with coefficients ranging from 0.07 to 0.87, except for temperate regions where water was abundant due to high precipitation and was distributed uniformly in this region (Supplementary Fig. S2). Heat (tmmn and tmmx) was positively associated with the productivity in forest, shrubland, and cold regions (coefficients: 0.06 to 0.36). This is probably because the main limiting factor for forests (located mostly in mountainous and humid regions) is

504 temperature. The shrubland occurred in cold and high-elevation regions (Supplementary Fig. S3),
505 so increasing temperature promoted plant productivity. Light (srad) was negatively associated
506 with the productivity in the forest, shrubland, arid, temperate, and cold regions (coefficients:
507 −0.32 to −0.15), which may be associated with drought and heat stress in arid and semi-arid
508 regions. But light was positively associated with it in humid regions (coefficient=0.11), because
509 humid regions are often radiation-limited, and an increase in light will increase annual GPP.
510
511 *3.4. Effect of topsoil depth on climate resilience of ecosystem productivity*
512 The control of topsoil depth on the resilience of GPP was investigated under four climatic
513 extremes in five climatic regions and five ecosystems. The distribution of ecosystems was strongly
514 dependent on the climatic regions, in which grassland and shrubland dominated the arid regions,
515 while forest and pasture mainly occurred in humid regions (Fig. 5). In dry years, the precipitation
516 was 20–80% lower than the average, and it was more severe in arid regions (Supplementary Fig.
517 S12). Accordingly, the GPP dropped in most locations, especially in the grassland and shrubland
518 (Fig. 5). In wet years, precipitation was 20–100% higher (Supplementary Fig. S12), and GPP
519 increased (Fig. 5). The changes of GPP due to climatic extremes were relatively small in forest
520 and pasture (relative changes within −50% to 50%) (Fig. 5). As shown in grassland and shrubland
521 in arid regions, the percent change of GPP was closer to zero and the productivity was more
522 stable when the topsoil was deeper. On the contrary, when the topsoil was shallower, the GPP
523 change was substantial. (Fig. 5). In some cases (e.g., pasture in humid regions), the percent
524 change of GPP was evenly distributed with topsoil depth, which may indicate that topsoil did not
525 affect the GPP change in climatic extremes in these regions.

Levene's test for deep and shallow topsoils was non-significant in most cases and thus did not strongly support the hypothesis that deeper topsoils had significantly smaller changes in GPP (Supplementary Figs. S13). It was likely due to that threshold values were selected from mean topsoil depth (ranging from 13 to 37 cm) and may not be able to reflect the deep topsoil cases (>40 cm). The deeper topsoils had a significantly smaller variance of the percent GPP change than shallower topsoils in cropland in arid years (Supplementary Fig. S13) and contributed to more stable productivity in dry extremes. However, the variance was significantly larger in deeper topsoils in forest in wet years and pasture in humid regions in wet years (Supplementary Fig. S13). This may suggest that forest with deeper topsoils promoted hydraulic redistribution within the deep vadose zone and fractured rocks leading to increased variation of GPP and its behavior under climate extremes (Montaldo and Oren, 2022). In some cases, the variance difference between shallow and deep topsoils was substantial (e.g., soils of pasture in arid regions, Supplementary Fig. S13), but Levene's test was non-significant, which was likely affected by a small sample size. Additionally, although the percent change decreased with topsoil depth in shrubland in arid regions (Fig. 5), the deeper topsoils had a larger variance than shallower topsoils (Supplementary Fig. S13), which was likely affected by a few extreme samples at 30-cm topsoil depths (Fig. 5).

When using Köppen-Geiger classification, a similar pattern was observed in dry and wet extremes (Supplementary Figs. S14, S15, S16). The percent changes of GPP decreased with deeper topsoils in cropland in cold regions, forest in temperate regions, and shrubland in arid regions, while such a pattern was less evident in grassland and pasture, and cropland in arid and temperate regions (Supplementary Fig. S15). As to the hot and cold extremes, the temperature

548 increased in hot years and decreased in cold years, which was more severe in arid and cold
549 regions (Supplementary Fig. S17). Similar to that in moisture extremes, the percent changes of
550 GPP decreased with deeper topsoils in cropland in cold regions, forest in temperate and cold
551 regions, grassland and shrubland in arid regions, but it was less evident in pasture
552 (Supplementary Fig. S18). Similarly, Levene's test was non-significant in most cases
553 (Supplementary Fig. S19). In cropland in arid regions, the coefficients were evenly distributed
554 with topsoil depth (Supplementary Fig. S15 and S18), which was likely due to extensive irrigation,
555 so that cropland water stress may be alleviated and showed no difference from rainfed fields in
556 terms of their relationship with topsoil depth (Supplementary Fig. S20).
557
558 *3.5. Assessment using GOSIF and FluxCom GPP*
559 The same analysis was conducted using the GOSIF and FluxCom GPP (Supplementary Figs. S21–
560 S38). The GOSIF and FluxCom GPP showed similar spatial distribution to the Landsat GPP
561 (Supplementary Figs. S21 and S30), but the GPP of pasture was significantly higher in these two
562 datasets than that of Landsat GPP. The relationships between GPP and topsoil depth across
563 climatic regions remained largely the same in GOSIF and FluxCom GPP datasets, except that the
564 GOSIF and FluxCom GPP decreased significantly with topsoil depth in cold regions
565 (Supplementary Figs. S22 and S31). The percent change of GPP under climatic extremes was also
566 evaluated using GOSIF and FluxCom GPP. For GOSIF GPP, similarly, in dry years, it decreased,
567 while in wet years, it increased; such a change was stronger in arid regions than in humid regions
568 (Supplementary Figs. S24). The percent change of GPP decreased with topsoil depth in many
569 ecosystems, especially grassland and shrubland (Supplementary Figs. S24 and S27). However, for

570 FluxCom GPP, it decreased in dry years and increased in wet years in nearly all the locations. It
571 was right-triangled with topsoil depth and the percent change of GPP was closer to zero with
572 increasing topsoil depth, particularly when it was deeper than 40 cm (Supplementary Figs. S33
573 and S35). Under temperature extremes, the percent changes of GPP can be both positive and
574 negative and decrease with topsoil depth in grassland and shrubland (Supplementary Fig. S37).
575
576 **4. Discussion**
577 *4.1. Environmental controls on topsoil depth–productivity relationship*
578 While earlier work showed positive control of topsoil depth on crop productivity, we found that
579 these relationships varied in different ecosystems and climatic regions. By analyzing the
580 nationwide dataset, we found that the control of topsoil depth on plant productivity was only
581 statistically significant in drier areas, and it was not statistically significant in wetter areas. In drier
582 regions, GPP decreased with a decreasing soil water content, but it did not change in wetter areas
583 (Fu et al., 2022). Therefore, when plants are under water stress, deeper topsoils aligned with
584 higher water storage can positively contribute to plant productivity. In our dataset, humid regions
585 were distributed in the eastern CONUS and West Coast, where deeper topsoils occurred under
586 cropland and pasture and had lower GPP than shallower soils under forest (Fig. 1), and therefore
587 a negative association of topsoil depth with productivity was observed in humid regions (Fig. 2).
588 The control of topsoil depth on productivity was not statistically significant in temperate
589 regions, but it was negatively associated with the GPP in cold regions when using GOSIF and
590 FluxCom GPP (Supplementary Fig. S22 and S31), which contradicts our hypothesis. It was
591 assumed that deeper topsoils with generally more SOC have a higher thermal buffering capacity

(Werner et al., 2020), which might be vital in cold regions to sustain plant productivity. However, in our study, cold regions were distributed mainly in the northeastern US, Midwest, and the Rocky Mountains with predominantly forest and cropland (Fig. 1), where forest had a higher GPP but shallower topsoils than cropland. Therefore, a negative relationship of topsoil depth and productivity was observed in cold regions. In temperate regions without cold stress, the relationship between topsoil depth and ecosystem productivity was not statistically significant.

The topsoil depth-GPP relationship differed for natural (e.g., forest, grassland, shrubland) and managed (e.g., cropland, pasture) ecosystems. Previous studies and topsoil removal experiments have focused mostly on cropland and demonstrated the negative effects of topsoil reduction on crop production (Zhang et al., 2021). Our results showed a clear positive association of topsoil depth with GPP in cropland, but such a relationship was not as strong as in grassland or shrubland. Grassland and shrubland root systems are often within 2-m depth compared to trees and thus more affected by topsoil properties. Moreover, grassland and shrubland were distributed mainly in drier regions (Supplementary Fig. S3), where topsoil depth was more strongly associated with ecosystem productivity. Under cropland and pasture, fertilization and irrigation may mask the effect of soil nutrients and water limitation due to deeper topsoils.

It is noteworthy that soils under forest generally had shallow topsoils (mean thickness=15 cm) but the highest GPP. There was no relation between topsoil depth and forest productivity (Fig. 1). This is likely due to sufficient water and nutrient supply in the forest ecosystem and its unique root architecture and functionality. Precipitation and topsoil SOC were high under the forest (Supplementary Fig. S3), and hence productivity was not restricted by water or nutrient supply. Moreover, woody forest roots include primary roots going deep into the soil and fine

roots spreading laterally (Danjon et al., 2013). Trees can use bedrock water for transpiration apart from soil water, which mainly occurred in western and southern US (e.g., California and Texas, dry regions) and the Appalachian Mountains (high elevation and slope, shallow soils) (McCormick et al., 2021). In California, over 50% of aboveground biomass production in woody forest can be attributed to bedrock water (McCormick et al., 2021). Therefore, in these cases, topsoil depth would have little influence on forest GPP.

The paired comparison showed that the topsoil depth-GPP relationship was stronger in shrubland, while in cropland, most deeper topsoils had greater GPP, although the GPP increase was small. The GPP difference between deep and shallow topsoils and the GPP-topsoil depth relationship of the paired samples were not statistically significant across ecosystems and climatic regions. The change of GPP was generally positively related to change of topsoil depth, but such a relationship was not statistically significant. The ecosystem type and climatic regions were more related to GPP of the paired samples. One drawback of the paired comparison was that the topsoil depth and GPP data had different sample support (i.e., the length, area, or volume associated with a measurement (Goovaerts, 2014)). The topsoil depth was measured on point-based sampling locations, while the GPP was obtained from raster images with 30-m spatial resolution. Soil thickness has a large local-scale variation, and therefore the point-based measurements may not be able to represent its averaged depth in the 30-m raster pixel and may increase the randomness of the comparison at local scales, which can partially explain the negative relationship of GPP and topsoil depth at some watersheds. Additionally, other factors (e.g., ecosystem types and climatic factors) were more related with GPP, which may affect the GPP-topsoil depth relationship of the paired samples in different watersheds.

Although topsoil depth was positively associated with the GPP in cropland, grassland, shrubland, and arid regions (Fig. 4, Supplementary Fig. S11), its association (coefficients: 0.09–0.33) was smaller than that of water (coefficients: 0.07–0.87) and similar to that of heat (coefficients: 0.06–0.39). The association of topsoil depth with plant productivity was higher in arid regions. For example, in arid regions of the Köppen-Geiger classification, the coefficient of topsoil depth to productivity (0.33) reached 60% of the coefficient of water (0.55). Our study also found a stronger correlation between GPP and precipitation (r=0.8) than that between GPP and topsoil depth (r=−0.04–0.37 for five ecosystems). Despite the weaker relationship, the topsoil depth played an important role in storing water and maintaining plant productivity in arid regions.

4.2. Topsoil depth and climate resilience

Our results showed that topsoil depth not only was associated with increased long-term averaged GPP (i.e., ecosystem productivity) in some ecosystems (cropland, grassland, shrubland) and in dry regions, but it was also associated with increased resilience of ecosystem productivity to climatic variation and extremes. In arid regions, the changes in GPP under climatic extremes were more severe, while in humid regions, the GPP was more stable with smaller changes under climatic extremes. The changes in GPP were smaller in forest and pasture (wetter regions) than that in cropland, grassland, and shrubland (drier regions). These indicate that the GPP was less stable and more easily affected by climatic extremes in drier regions. However, in arid regions (especially shrubland and grassland), as topsoil was deeper, the percent change of GPP was closer to zero and the productivity was more stable. This indicates that topsoil was associated with increased climate resilience of plant productivity especially in arid regions.

Deeper topsoils aligned with higher SOC content and water storage capacity tend to increase buffering capacity to climate change and extremes and maintain ecosystem productivity. Similarly, a recent study showed that high-quality soils reduced crop yield variability and its sensitivity to climate change by over 15% (Qiao et al., 2022). Other studies have shown that biodiversity also enhanced the resilience of ecosystem productivity to climate extremes (Isbell et al., 2015) and defined it as insurance effects which included both a buffering effect and a performance-enhancing effect (Yachi and Loreau, 1999). Accordingly, deeper topsoils tend to increase both productivity and their buffering capacity to climate change.

4.3. Implications, limitations, and prospects

Understanding the control of topsoil depth on ecosystem productivity is important for crop production, erosion control, wasteland reclamation, ecosystem restoration, and maintaining ecosystem resilience. In this study, a comprehensive analysis was conducted to study the interrelationships between topsoil depth and ecosystem productivity in various ecosystems and different climatic conditions and results showed that such relationship was stronger in drier regions and grassland and shrubland. Deeper topsoil was also associated with improved resilience of ecosystem productivity to climatic extremes in these regions. The results can improve our understanding of the control of topsoil depth to ecosystem functioning and lead to a better representation of the role of soil in earth system modeling and climate modeling. It also provides evidence for natural resources management under climate change.

However, limitations exist in the analysis and results interpretation. First, our analysis was based on existing observational data from long-term soil surveys, and we did not conduct any

controlled experiments or explicitly select sampling locations based on edaphic, ecological, and climatic factors. The results and interpretation cannot completely eliminate multiple effects from other soil and environmental factors and solely investigate the effects of topsoil depth. Second, the large short-scale variation of topsoil depth and mismatch of the dimensions between soil data (point-based measurements) and GPP data (30-m resolution) may increase the randomness of the results at local scales. Third, this study investigated only the one-way effects of topsoil depth on ecosystem productivity and its resilience to climate extremes. Higher ecosystem productivity and low disturbance rates over the long run also contributed to greater SOC accumulation in forests and deeper topsoils in grassland. But this is a relatively slow process, especially in terms of the annual variation of GPP and its interaction with annual climate.

Soil variables are increasingly used in earth system modeling with the increasing availability, accessibility, and accuracy of national and global soil maps (Chaney et al., 2019; Poggio et al., 2021). A constant value has been used to represent soil depth in many cases in the past, but recently more global products have been available for soil depth (Pelletier et al., 2016; Shangguan et al., 2017). Topsoil depth is an important variable, as it reflects the carbon-rich and most microbially active layer of the soil and directly affects nutrient availability and ecosystem productivity. However, there is currently no available large-scale map of topsoil depth, which would be important for understanding biogeochemical cycling and earth system modeling. Future work is needed for creating national and global maps of topsoil depth. In addition, soil carbon stabilization and enhancement are considered as an important nature-based climate solution. Other soil factors (e.g., soil depth, bulk density) which are essential parameters to calculate SOC stock and directly related to soil erosion and compaction have been under-studied.

Similarly, the role of other soil physical, chemical, and biological properties on ecosystem functioning under climate change should be further investigated.

5. Conclusions

The relationship between topsoil depth and plant productivity (GPP) was investigated for different ecosystems and climatic regions using a nationwide dataset across the CONUS. A weak positive correlation was observed between the GPP and topsoil depth in soils under grassland and shrubland. The control of topsoil depth on GPP was primarily associated with water availability, which was more significant in arid regions. Forest productivity was less associated with topsoil depth due to its higher SOC content, high precipitation, and deeper root architecture and functionality. However, the pairing of deep and shallow topsoils showed a small but non-statistically significant relationship between GPP and topsoil depth. The lack of a significant relationship may be due to different sample support of soil and GPP data (i.e., point-based and raster data), or the effects of other soil and environmental factors across watersheds. Moreover, the association of GPP with topsoil depth was smaller than that with water and similar to that with heat. The topsoil depth was also related to increased stability of ecosystem productivity to climate change in arid regions and shrubland and grassland. We conclude that topsoil depth affects ecosystem productivity and its stability and resilience to climate extremes in dry regions. Such relationship does not exist in humid regions.

Acknowledgements

We are grateful for Marvin Beatty for creating a professorship and providing us funding to do this research. J.X. was supported by the National Science Foundation (NSF) (Macrosystem Biology & NEON-Enabled Science program: DEB-2017870). We are grateful to USDA-NRCS for collecting and providing the open-access soil data, and Rich Ferguson, Steve Monteith, Scarlett Murphy for advice about the dataset. We thank Steven Moen and Caleb Stevens at the University of Wisconsin-Madison for statistical advice.

Author Contributions

Y.Z. and A.E.H. conceived the work. Y.Z. performed the data analysis, interpreted the results, and wrote the draft. A.R.D., J.X., and A.E.H. contributed to the interpretation of the results and writing.

Competing Interest Statement

The authors declare no competing interests.

Data accessibility

All datasets supporting the findings of this study can be found in the supplementary file.

References

Abatzoglou, J.T., 2013. Development of gridded surface meteorological data for ecological applications and modelling. International Journal of Climatology 33(1), 121-131.

743 Abatzoglou, J.T., Dobrowski, S.Z., Parks, S.A., Hegewisch, K.C., 2018. TerraClimate, a high-
744 resolution global dataset of monthly climate and climatic water balance from 1958–2015.
745 Scientific data 5(1), 1-12.
746 Baker, N.R., 2008. Chlorophyll fluorescence: a probe of photosynthesis in vivo. Annual review of
747 plant biology 59, 89.
748 Bakker, M.M., Govers, G., Rounsevell, M.D.A., 2004. The crop productivity–erosion relationship:
749 an analysis based on experimental work. CATENA 57(1), 55-76.
750 Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R.H.B., Singmann, H., Dai, B., Scheipl,
751 F., Grothendieck, G., Green, P., 2009. Package 'lme4'. https://github.com/lme4/lme4/.
752 Beck, H.E., Zimmermann, N.E., McVicar, T.R., Vergopolan, N., Berg, A., Wood, E.F., 2018. Present
753 and future Köppen-Geiger climate classification maps at 1-km resolution. Scientific Data
754 5(1), 180214.
755 Bellard, C., Bertelsmeier, C., Leadley, P., Thuiller, W., Courchamp, F., 2012. Impacts of climate
756 change on the future of biodiversity. Ecology Letters 15(4), 365-377.
757 Berhe, A.A., Barnes, R.T., Six, J., Marin-Spiotta, E., 2018. Role of soil erosion in biogeochemical
758 cycling of essential elements: carbon, nitrogen, and phosphorus. Annual Review of Earth
759 and Planetary Sciences 46, 521-548.
760 Bhardwaj, A.K., Jasrotia, P., Hamilton, S.K., Robertson, G.P., 2011. Ecological management of
761 intensively cropped agro-ecosystems improves soil quality with sustained productivity.
762 Agriculture, Ecosystems & Environment 140(3), 419-429.
763 Blum, W.E., 2005. Functions of soil for society and the environment. Reviews in Environmental
764 Science and Bio/Technology 4(3), 75-79.

765 Bollen, K.A., Hoyle, R.H., 2012. Latent variables in structural equation modeling. Handbook of
766 structural equation modeling, 56-67.
767 Brown, J.F., Howard, D.M., Shrestha, D., Benedict, T.D., 2019. Moderate Resolution Imaging
768 Spectroradiometer (MODIS) Irrigated Agriculture Datasets for the Conterminous United
769 States (MIrAD-US). US Geological Survey Data Release.
770 Brown, L.R., 1984. The global loss of topsoil. Journal of Soil and Water Conservation 39(3), 162-
771 165.
772 Chaney, N.W., Minasny, B., Herman, J.D., Nauman, T.W., Brungard, C.W., Morgan, C.L.,
773 McBratney, A.B., Wood, E.F., Yimam, Y., 2019. POLARIS soil properties: 30-m probabilistic
774 maps of soil properties over the contiguous United States. Water Resources Research
775 55(4), 2916-2938.
776 Danjon, F.F., Stokes, A., Bakker, M.R.M., 2013. Root systems of woody plants. CRC Press.
777 Dodge, Y., Cox, D., Commenges, D., 2003. The Oxford dictionary of statistical terms. Oxford
778 University Press on Demand.
779 Eisenhauer, N., Bowker, M.A., Grace, J.B., Powell, J.R., 2015. From patterns to causal
780 understanding: structural equation modeling (SEM) in soil ecology. Pedobiologia 58(2-3),
781 65-72.
782 Fox, J., Weisberg, S., Adler, D., Bates, D., Baud-Bovy, G., Ellison, S., Firth, D., Friendly, M., Gorjanc,
783 G., Graves, S., 2012. Package 'car'. Vienna: R Foundation for Statistical Computing 16.
784 Francés, A.P., Lubczynski, M.W., 2011. Topsoil thickness prediction at the catchment scale by
785 integration of invasive sampling, surface geophysics, remote sensing and statistical
786 modeling. Journal of Hydrology 405(1), 31-47.

787 Fu, Z., Ciais, P., Prentice, I.C., Gentine, P., Makowski, D., Bastos, A., Luo, X., Green, J.K., Stoy, P.C.,
788 Yang, H., Hajima, T., 2022. Atmospheric dryness reduces photosynthesis along a large
789 range of soil water deficits. Nature Communications 13(1), 989.
790 Goovaerts, P., 2014. Sample Support. In: N. Balakrishnan, T. Colton, B. Everitt, W. Piegorsch, F.
791 Ruggeri, J.L. Teugels (Eds.), Wiley StatsRef: Statistics Reference Online.
792 Grace, J.B., 2006. Structural equation modeling and natural systems. Cambridge University Press.
793 Grimm, N.B., Chapin III, F.S., Bierwagen, B., Gonzalez, P., Groffman, P.M., Luo, Y., Melton, F.,
794 Nadelhoffer, K., Pairis, A., Raymond, P.A., Schimel, J., Williamson, C.E., 2013. The impacts
795 of climate change on ecosystem structure and function. Frontiers in Ecology and the
796 Environment 11(9), 474-482.
797 Hijmans, R.J., Van Etten, J., Mattiuzzi, M., Sumner, M., Greenberg, J., Lamigueiro, O., Bevan, A.,
798 Racine, E., Shortridge, A., 2013. Raster package in R. Version 3.6-3.
799 https://rspatial.org/raster/raster/RasterPackage.pdf.
800 Isbell, F., Craven, D., Connolly, J., Loreau, M., Schmid, B., Beierkuhnlein, C., Bezemer, T.M., Bonin,
801 C., Bruelheide, H., de Luca, E., Ebeling, A., Griffin, J.N., Guo, Q., Hautier, Y., Hector, A.,
802 Jentsch, A., Kreyling, J., Lanta, V., Manning, P., Meyer, S.T., Mori, A.S., Naeem, S., Niklaus,
803 P.A., Polley, H.W., Reich, P.B., Roscher, C., Seabloom, E.W., Smith, M.D., Thakur, M.P.,
804 Tilman, D., Tracy, B.F., van der Putten, W.H., van Ruijven, J., Weigelt, A., Weisser, W.W.,
805 Wilsey, B., Eisenhauer, N., 2015. Biodiversity increases the resistance of ecosystem
806 productivity to climate extremes. Nature 526(7574), 574-577.
807 Jenny, H., 1941. Factors of soil formation: a system of quantitative pedology. McGraw-Hill.

Jung, M., Schwalm, C., Migliavacca, M., Walther, S., Camps-Valls, G., Koirala, S., Anthoni, P., Besnard, S., Bodesheim, P., Carvalhais, N., 2020. Scaling carbon fluxes from eddy covariance sites to globe: synthesis and evaluation of the FLUXCOM approach. Biogeosciences 17(5), 1343-1365.

Kuznetsova, A., Brockhoff, P.B., Christensen, R.H., 2017. lmerTest package: tests in linear mixed effects models. Journal of statistical software 82, 1-26.

Larson, W.E., Pierce, F.J., Dowdy, R.H., 1983. The threat of soil erosion to long-term crop production. Science 219(4584), 458-465.

Li, X., Xiao, J., 2019a. A global, 0.05-degree product of solar-induced chlorophyll fluorescence derived from OCO-2, MODIS, and reanalysis data. Remote Sensing 11(5), 517.

Li, X., Xiao, J., 2019b. Mapping photosynthesis solely from solar-induced chlorophyll fluorescence: A global, fine-resolution dataset of gross primary production derived from OCO-2. Remote Sensing 11(21), 2563.

Li, X., Xiao, J., He, B., Altaf Arain, M., Beringer, J., Desai, A.R., Emmel, C., Hollinger, D.Y., Krasnova, A., Mammarella, I., Noe, S.M., Ortiz, P.S., Rey-Sanchez, A.C., Rocha, A.V., Varlagin, A., 2018. Solar-induced chlorophyll fluorescence is strongly correlated with terrestrial photosynthesis for a wide variety of biomes: First global analysis based on OCO-2 and flux tower observations. Global Change Biology 24(9), 3990-4008.

Lorenz, K., Lal, R., 2009. Carbon sequestration in forest ecosystems. Springer Dordrecht.

Lyles, L., 1975. Possible effects of wind erosion on soil productivity. Journal of Soil and Water Conservation 30(6), 279-283.

McCormick, E.L., Dralle, D.N., Hahm, W.J., Tune, A.K., Schmidt, L.M., Chadwick, K.D., Rempe, D.M., 2021. Widespread woody plant use of water stored in bedrock. Nature 597(7875), 225-229.

Mielke, L.N., Schepers, J.S., 1986. Plant response to topsoil thickness on an eroded loess soil. Journal of Soil and Water Conservation 41(1), 59-63.

Montaldo, N., Oren, R., 2022. Rhizosphere water content drives hydraulic redistribution: Implications of pore-scale heterogeneity to modeling diurnal transpiration in water-limited ecosystems. Agricultural and Forest Meteorology 312, 108720.

Montgomery, D.R., 2007. Soil erosion and agricultural sustainability. Proceedings of the National Academy of Sciences 104(33), 13268-13272.

National Cooperative Soil Survey, National Cooperative Soil Survey Soil Characterization Database. http://ncsslabdatamart.sc.egov.usda.gov/. Accessed Tuesday, November 10, 2020.

O'Sullivan, M., Smith, W.K., Sitch, S., Friedlingstein, P., Arora, V.K., Haverd, V., Jain, A.K., Kato, E., Kautz, M., Lombardozzi, D., Nabel, J.E.M.S., Tian, H., Vuichard, N., Wiltshire, A., Zhu, D., Buermann, W., 2020. Climate - driven variability and trends in plant productivity over recent decades based on three global products. Global Biogeochemical Cycles 34(12), e2020GB006613.

Peel, M.C., Finlayson, B.L., McMahon, T.A., 2007. Updated world map of the Köppen-Geiger climate classification. Hydrology and earth system sciences 11(5), 1633-1644.

Pelletier, J.D., Broxton, P.D., Hazenberg, P., Zeng, X., Troch, P.A., Niu, G.Y., Williams, Z., Brunke, M.A., Gochis, D., 2016. A gridded global data set of soil, intact regolith, and sedimentary

deposit thicknesses for regional and global land surface modeling. Journal of Advances in Modeling Earth Systems 8(1), 41-65.

Phillips, J.D., 2008. Soil system modelling and generation of field hypotheses. Geoderma 145(3), 419-425.

Poggio, L., De Sousa, L.M., Batjes, N.H., Heuvelink, G., Kempen, B., Ribeiro, E., Rossiter, D., 2021. SoilGrids 2.0: producing soil information for the globe with quantified spatial uncertainty. Soil 7(1), 217-240.

Qiao, L., Wang, X., Smith, P., Fan, J., Lu, Y., Emmett, B., Li, R., Dorling, S., Chen, H., Liu, S., Benton, T.G., Wang, Y., Ma, Y., Jiang, R., Zhang, F., Piao, S., Müller, C., Yang, H., Hao, Y., Li, W., Fan, M., 2022. Soil quality both increases crop production and improves resilience to climate change. Nature Climate Change 12(6), 574-580.

Quinton, J.N., Govers, G., Van Oost, K., Bardgett, R.D., 2010. The impact of agricultural soil erosion on biogeochemical cycling. Nature Geoscience 3(5), 311-314.

R Core Team, 2021. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Robinson, N.P., Allred, B.W., Jones, M.O., Moreno, A., Kimball, J.S., Naugle, D.E., Erickson, T.A., Richardson, A.D., 2017. A dynamic Landsat derived normalized difference vegetation index (NDVI) product for the conterminous United States. Remote sensing 9(8), 863.

Robinson, N.P., Allred, B.W., Smith, W.K., Jones, M.O., Moreno, A., Erickson, T.A., Naugle, D.E., Running, S.W., 2018. Terrestrial primary production for the conterminous United States derived from Landsat 30 m and MODIS 250 m. Remote Sensing in Ecology and Conservation 4(3), 264-280.

Rosseel, Y., 2012. lavaan: An R package for structural equation modeling. Journal of Statistical Software 48(2), 1 - 36.

Running, S.W., Zhao, M., 2015. Daily GPP and annual NPP (MOD17A2/A3) products NASA Earth Observing System MODIS land algorithm. MOD17 User's Guide 2015, 1-28.

Russell, R.J., 1931. Dry climates of the United States: I. Climatic map, 1. University of California Press.

Schmidhuber, J., Tubiello, F.N., 2007. Global food security under climate change. Proceedings of the National Academy of Sciences 104(50), 19703-19708.

Schoeneberger, P.J., Wysocki, D.A., Benham, E.C., 2012. Field book for describing and sampling soils, version 3.0. USDA Natural Resources Conservation Service, National Soil Survey Center, Lincoln, NE.

Seager, R., Lis, N., Feldman, J., Ting, M., Williams, A.P., Nakamura, J., Liu, H., Henderson, N., 2018. Whither the 100th Meridian? The Once and Future Physical and Human Geography of America's Arid–Humid Divide. Part I: The Story So Far. Earth Interactions 22(5), 1-22.

Shangguan, W., Hengl, T., Mendes de Jesus, J., Yuan, H., Dai, Y., 2017. Mapping the global depth to bedrock for land surface modeling. Journal of Advances in Modeling Earth Systems 9(1), 65-88.

Shapiro, S.S., Wilk, M.B., 1965. An analysis of variance test for normality (complete samples). Biometrika 52(3/4), 591-611.

Signorell, A., Aho, K., Alfons, A., Anderegg, N., Aragon, T., Arachchige, C., Arppe, A., Baddeley, A., Barton, K., Bolker, B., 2023. DescTools: tools for descriptive statistics. R package version 0.99. 49. R Foundation for Statistical Computing Vienna, Austria.

Singmann, H., Bolker, B., Westfall, J., Aust, F., Ben-Shachar, M., Højsgaard, S., 2015. Package 'afex'. http://www.psychologie.uni-freiburg.de/Members/singmann/R/afex.

Sohl, T., Reker, R., Bouchard, M., Sayler, K., Dornbierer, J., Wika, S., Quenzer, R., Friesz, A., 2016. Modeled historical land use and land cover for the conterminous United States. Journal of Land Use Science 11(4), 476-499.

Sohl, T.L., Sayler, K.L., Bouchard, M.A., Reker, R.R., Friesz, A.M., Bennett, S.L., Sleeter, B.M., Sleeter, R.R., Wilson, T., Soulard, C., 2014. Spatially explicit modeling of 1992–2100 land cover and forest stand age for the conterminous United States. Ecological Applications 24(5), 1015-1036.

Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture. Soil Survey Geographic (SSURGO) Database. Available online. Accessed [01/10/2023].

Soil Survey Staff, 2014. Kellogg Soil Survey Laboratory Methods Manual. Soil Survey Investigations Report No. 42, Version 5.0. R. Burt and Soil Survey Staff (ed.). U.S. Department of Agriculture, Natural Resources Conservation Service.

Soller, D.R., Reheis, M.C., Garrity, C.P., Van Sistine, D.R., 2009. Map database for surficial materials in the conterminous United States. US Geological Survey Data Series 425.

Theobald, D.M., Harrison-Atlas, D., Monahan, W.B., Albano, C.M., 2015. Ecologically-relevant maps of landforms and physiographic diversity for climate adaptation planning. PloS one 10(12), e0143619.

United States Geological Survey (USGS), 3D Elevation Program 10-Meter Resolution Digital Elevation Model.

WBD, Coordinated effort between the United States Department of Agriculture-Natural Resources Conservation Service (USDA-NRCS), the United States Geological Survey (USGS), and the Environmental Protection Agency (EPA). The Watershed Boundary Dataset (WBD) was created from a variety of sources from each state and aggregated into a standard national layer for use in strategic planning and accountability. Watershed Boundary Dataset for HUC10 [Online WWW]. Available URL: (https://datagateway.nrcs.usda.gov) [Accessed 14/09/2022].

Werner, W.J., Sanderman, J., Melillo, J.M., 2020. Decreased soil organic matter in a long-term soil warming experiment lowers soil water holding capacity and affects soil thermal and hydrological buffering. Journal of Geophysical Research: Biogeosciences 125(4), e2019JG005158.

Wu, G.-L., Cheng, Z., Alatalo, J.M., Zhao, J., Liu, Y., 2021. Climate warming consistently reduces grassland ecosystem productivity. Earth's Future 9(6), e2020EF001837.

Xiao, J., Liu, S., Stoy, P.C., 2016. Preface: Impacts of extreme climate events and disturbances on carbon dynamics. Biogeosciences 13(12), 3665-3675.

Yachi, S., Loreau, M., 1999. Biodiversity and ecosystem productivity in a fluctuating environment: the insurance hypothesis. Proceedings of the National Academy of Sciences 96(4), 1463-1468.

Yang, L., Jin, S., Danielson, P., Homer, C., Gass, L., Bender, S.M., Case, A., Costello, C., Dewitz, J., Fry, J., Funk, M., Granneman, B., Liknes, G.C., Rigge, M., Xian, G., 2018. A new generation of the United States National Land Cover Database: Requirements, research priorities,

design, and implementation strategies. ISPRS Journal of Photogrammetry and Remote Sensing 146, 108-123.

Zhang, L., Huang, Y., Rong, L., Duan, X., Zhang, R., Li, Y., Guan, J., 2021. Effect of soil erosion depth on crop yield based on topsoil removal method: a meta-analysis. Agronomy for Sustainable Development 41(5), 63.

Zhang, Y., Hartemink, A.E., Vanwalleghem, T., Bonfatti, B.R., Moen, S., 2023. Soil thickness change in the Anthropocene across the USA: natural and human causes. In review.

Fig. 1. Topsoil depth and gross primary productivity (GPP) across the conterminous US (CONUS) and in five ecosystems. a. Topsoil depth measurements across the CONUS. b. Landsat GPP dataset across the CONUS. c. The distribution of sampled locations in five ecosystems. d. The distribution of topsoil depth across five ecosystems. Numbers in parentheses indicate the sample size for each ecosystem. e. The distribution of Landsat GPP across five ecosystems. The black dots in d and e indicate the mean values in each ecosystem. The black lines in d and e indicate the mean values across all the ecosystems. f. The relationship between GPP and topsoil depth for the five ecosystems. The Pearson correlations between GPP and topsoil depth are 0.15, −0.04, 0.37, 0.02, 0.32 for cropland, forest, grassland, pasture, and shrubland, respectively. Shadows indicate the 95% confidence intervals.

Fig. 2. The relationships between gross primary productivity (GPP) and topsoil depth in different climatic regions. a. Distribution of sample locations in arid and humid regions which were

960 distinguished by Aridity Index (AI). b. Distribution of sample locations in three climatic regions
961 (arid, cold, temperate) based on the Köppen-Geiger climate classification. c, d. The linear
962 relationships between GPP and topsoil depth in different climatic regions. The Pearson
963 correlations between GPP and topsoil depth are 0.33 and −0.09 for arid and humid regions
964 classified by AI, and 0.45, −0.02, and −0.06 for arid, cold, and temperate regions of Köppen-Geiger
965 climate classification.
966
967 Fig. 3. Paired comparison of gross primary productivity (GPP) for 103 watersheds. a. In each
968 watershed, one pair of relatively deep (orange dots) and shallow (blue dots) topsoils were
969 selected to compare their GPP values. The deep and shallow topsoils were not determined by
970 absolute depth but relative depth difference between them. The depth difference of the pair
971 should be greater than 3 cm if they are shallower (<15 cm) or greater than 5 cm if they are deeper
972 (>15 cm). The paired topsoils have 1) the same climate conditions, ecosystem type, soil order,
973 soil temperature and moisture regimes, parent material, and soil textural class; 2) similar organic
974 carbon content, elevation, and slope; 3) different GPP values; and 4) are within 4-km distance.
975 Each pair is connected by a gray line, in which the line width indicates the depth difference of the
976 pair. If an orange dot is on top of a blue dot, it indicates the deeper topsoil aligns with greater
977 GPP; otherwise, a shallower topsoil aligns with greater GPP. The pairs from different watersheds
978 in each ecosystem were ranked from left to right in the x-axis by their mean annual precipitation.
979 b. The distribution of absolute change of GPP between deep and shallow topsoils in each
980 watershed. c. The distribution of relative change of GPP between deep and shallow topsoils in
981 each watershed. The colors of the dots in b and c indicate five ecosystems, while the sizes of the

dots indicate the GPP difference between deep and shallow topsoils. The two maps in b and c represent the two scenarios: 1) the deeper topsoil aligns with greater GPP (GPP differences are positive); 2) a shallower topsoil aligns with greater GPP (GPP differences are negative).

Fig. 4. Structural equation models (SEMs) for predicting productivity from several latent variables in different ecosystems. The blue solid lines indicate positive contributions from latent variables to productivity with coefficients provided next to the line. The red solid lines indicate negative contributions from latent variables to productivity with coefficients provided next to the line. Only the statistically significant coefficients are provided in the figure. The conceptual structure of the SEMs is shown in Supplementary Fig. S10. CFI: Comparative Fit Index; TLI: Tucker-Lewis Index; RMSEA: Root Mean Square Error of Approximation.

Fig. 5. The relationships between percent change in GPP of a dry year (minimum precipitation) and a wet year (maximum precipitation) from 1986 to 2021 and topsoil depth in two climatic regions (arid and humid) and five ecosystems. Dashed horizontal lines indicate no change of GPP in specific conditions (y=0). Dashed vertical lines indicate the mean topsoil depth in each region, which were used as threshold values to separate shallow and deep topsoils in Supplementary Fig. S13.

Table 1. Summary statistics of the paired comparison of GPP in different watersheds of five ecosystems in Fig. 3.

1004 Table 2. Unstandardized coefficients of fixed and random effect variables in the linear mixed-
1005 effect models with paired data. Topsoil depth was used as a binary variable (deep and shallow)
1006 and a numeric variable respectively in two models and interacted with ecosystem.
1007