# Global patterns and drivers of soil total phosphorus concentration

Xianjin He[1]; Laurent Augusto[2]; Daniel S. Goll[3]; Bruno Ringeval[2]; Yingping Wang[4,5]; Julian Helfenstein[6]; Yuanyuan Huang[3,5]; Kailiang Yu[7]; Zhiqiang Wang[8]; Yongchuan Yang[1]; Enqing Hou[4]

1 Key Laboratory of the Three Gorges Reservoir Region's Eco-Environment, Ministry of Education, Chongqing University, Chongqing, 400045, China

2 INRAE, Bordeaux Sciences Agro, UMR 1391 ISPA, 33140 Villenave d'Ornon, France

3 Université Paris Saclay, CEA-CNRS-UVSQ, LSCE/IPSL, Gif sur Yvette, France

4 Key Laboratory of Vegetation Restoration and Management of Degraded Ecosystems, South China Botanical Garden, Chinese Academy of Sciences, Guangzhou, China

5 CSIRO Oceans and Atmosphere, Aspendale, Vic., Australia

6 Agroscope, 8046 Zürich, Switzerland

7 High Meadows Environmental Institute, Princeton University, Princeton, NJ 08544, USA

8 Institute for Advanced Study, Chengdu University, Chengdu, Sichuan, China

*Corresponding to*: Yongchuan Yang (ycyang@cqu.edu.cn); Enqing Hou (houeq@scbg.ac.cn).

**Abstract.** Soils represent the largest phosphorus (P) reserves on land and determining the amount is a critical first step for identifying sites where ecosystem functioning is potentially limited by soil P availability. However, global patterns and predictors of soil total P concentration remain poorly understood. To address this knowledge gap, we constructed a database of total P concentration of 5,275 globally distributed (semi-)natural soils from 761 published studies. We quantified the relative importance of 13 soil-forming variables in predicting soil total P concentration and then made further predictions at the global scale using a random forest approach. Soil total P concentration varied significantly among parent material types, soil orders, biomes, and continents, and ranged widely from 1.4 to 9,630.0 (median 430.0 and mean 570.0) mg kg$^{-1}$ across the globe. About two-thirds (65%) of the global variation was accounted for by the 13 variables that we selected, among which soil organic carbon concentration, parent material, mean annual temperature, and soil sand content were the most important. While global predictions of soil total P concentration increased significantly with latitude, they varied largely among regions with similar latitudes due to regional differences in parent material, topography, and/or climate conditions. Global soil P stocks (excluding Antarctica) were estimated to be 26.8 ± 3.1 (mean ± standard deviation) Pg and 62.2 ± 8.9 Pg (1 Pg=1×10$^{15}$ g) in the topsoil (0-30 cm) and subsoil (30-100 cm), respectively. Our global map of soil total P concentration as well as the underlying drivers of soil total P concentration can be used to constraint Earth system models that represent the P cycle and to inform quantification of global soil P availability. Raw datasets and global maps generated in this study are available at https://doi.org/10.6084/m9.figshare.14583375 (He et al., 2021).

## 1 Introduction

Soils represent the largest P reserves on land (Zhang et al., 2021). The amount and form of P determine the supply of soil

P to plants, which further regulate the structure and function of global terrestrial ecosystems (Vitousek et al., 2010; Hou et al., 2020; Elser et al., 2007; Hou et al., 2021). Moreover, soil P form depends on the amount or total concentration of P in soils

35    (Lang et al., 2016; Hou et al., 2018a; Turner and Engelbrecht, 2011). Therefore, it is important to determine the total concentration of P in soils, which varies up to three orders of magnitude across the globe (Yanai, 1998; Augusto et al., 2010; Zhang et al., 2021). Despite the large variation in soil total P concentration, its global patterns and drivers remain poorly resolved and improving this knowledge gap is needed to better represent the P cycle in Earth system models (Fleischer et al., 2019; Goll et al., 2017; Reed et al., 2015; Wang et al., 2015; Wieder et al., 2015; Zhang et al., 2011; Achat et al., 2016).

40    Soil total P concentration is the outcome of climatic, biotic, and landscape processes interacting over time on soil parent material (Dokuchaev, 1883; Jenny, 1941; Buendía et al., 2010). Each of these factors may be characterized by a few variables; for example, climate may be characterized by mean annual temperature (MAT) and precipitation (MAP). Relationships between soil total P concentration and variables such as parent material type and P concentration, MAT, MAP, site slope, and soil organic carbon (SOC) have been reported in previous studies, but mostly at local to regional scales (Brédoire et al., 2016;

45    Cheng et al., 2018; Li et al., 2019; Porder and Chadwick, 2009; Wang et al., 2009). Few studies have quantified the relative importance of these variables for predicting soil total P concentration at a global scale (Delgado-Baquerizo et al., 2020; Augusto et al., 2017; Yang et al., 2013). Such an understanding can guide the management of the soil P supply in agroecosystems of different regions (Ringeval et al., 2017) and is crucial for both mapping soil total P concentration in natural terrestrial ecosystems (Reed et al., 2015) and simulating ecosystem functioning (Achat et al., 2016).

50    While each soil-forming factor can determine soil total P concentration, the roles of some factors (e.g., climate and vegetation) are less understood than other factors (e.g., parent material and soil age). Since P in soil is derived mainly from parent materials, the control of parent material on soil total P concentration has been well recognized (Augusto et al., 2017; Porder and Ramachandran, 2013). Soil chronosequences provide a unique opportunity to isolate the effect of soil age from other soil-forming factors on soil P dynamics, and have shown that soil age negatively impacts soil total P concentration

55    (Wardle et al., 2004; Delgado-Baquerizo et al., 2020; Vitousek et al., 2010; Walker and Syers, 1976). Due to climate change, there is an increasing interest in how climate impacts soil total P concentration (Augusto et al., 2017; Vitousek and Chadwick, 2013; Hou et al., 2018a). Yet the effects of climate, vegetation, and topography on soil total P concentration remain largely unknown. Recently, Delgado-Baquerizo et al. (2020) surveyed 32 ecosystem properties, including soil total P concentration, in 16 soil chronosequences globally. They found that climate, vegetation, topography, and soil age together explained only

60    about 60% of the variation in soil total P concentration, despite examining 30 predictors and considering all possible interactions among predictors. This finding reflects our incomplete understanding of the controls of soil total P concentration.

Several pressing global issues such as mitigating climate change, increasing food security, and reducing nutrient run-off to bodies of water, rely on accurate soil P maps (Carpenter and Bennett, 2011; Steffen et al., 2015). While several maps of soil

total P concentration have been produced (Viscarra Rossel and Bui, 2016; Ballabio et al., 2019; Hengl et al., 2017; Delmas et

65     al., 2015), to our knowledge, there are only two published maps of soil total P concentration in natural terrestrial ecosystems

(Shangguan et al., 2014; Yang et al., 2013). These two maps have been used to explore global patterns of soil P supply (Yang

et al., 2013), estimate P limitation on future terrestrial C sequestration (Sun et al., 2017), and used as baseline information to

quantify P supply in agricultural ecosystems (Ringeval et al., 2017). They are also used frequently in land surface models to

benchmark soil P modules (Yang et al., 2014; Goll et al., 2012). However, the two maps may suffer from large uncertainties

70     due to limited numbers of predictors used and/or low spatial coverage of global soils. First, for example, Yang et al. (2013)

mapped soil total P concentration based only on parent material and soil chronosequence measurements. The map by

Shangguang et al. (2014) was based on a database that had poor coverage of many parts of the world (e.g. high latitude, Africa,

South America). Second, both maps only focus on the surface layers of soils, though subsoils are known to contribute to the P

nutrition of plants and P leaching to groundwater (Rodionov et al., 2020; Andersson et al., 2013). Third, both maps have a low

75     spatial resolution of 0.5° (e.g. about 50 km near the equator). Given the huge spatial heterogeneity of soil total P concentration

that is evident on a much finer scale (He et al., 2016; Kaňa and Kopáček, 2006; Garrett, 2009), developing a map with a finer

spatial resolution of total P distribution is important for addressing global issues related to ecosystem functioning.

To address these issues, we constructed a global database of total P concentration of 5,275 (semi-)natural soils from 761

published studies. We then used random forest algorithms to quantify the relative importance of soil-forming variables for

80     predicting soil total P concentration and further predicted it at the global scale with a 0.05° resolution (e.g., about 5 km near

the equator). With our enlarged dataset and our map of global soil P distribution, we addressed the following research questions:

(1) Which factors are the most important for predicting the spatial variation of soil total P concentration in the top 1 m of soil?

(2) How does soil total P concentration differ among regions and soil layers? and (3) How large is the global total P stock in

the top 1 m of soil?

85     **2 Material and Methods**

**2.1 Data source and processing**

Given massive measurements of soil total P concentration in literature, it is practically infeasible to collect all the

measurements in literature. Therefore, we collected soil total P concentration measurements in (semi-)natural terrestrial

ecosystems mainly from existing global or regional databases, and additionally from literature with focus on the

90     underrepresented regions identified in global databases, to ensure a good coverage of global terrestrial ecosystems. We defined

(semi-)natural ecosystems as ecosystems without any documented anthropogenic activities such as tillage, fertilization, and

heavily grazing. Forests with a stand age greater than 10 years were considered as (semi-)natural ecosystems. Despite our great

efforts to exclude soils significantly affected by anthropogenic activities, some soils in our database may be influenced by

undocumented anthropogenic activities (e.g., P fertilization in reforested lands), particularly in Western Europe and Eastern

95    USA (e.g., De Schrijver et al., 2012). We compiled the database in four steps, which are described as follows.

First, we searched existing global or regional databases that may include soil total P concentration measurements in (semi-)natural ecosystems in the Web of Science using key words "global OR terrestrial OR meta-analysis" AND "soil phosphorus" NOT "crop OR agriculture" in topic. This search returned 714 papers by 15th September, 2020. We screened each of these papers by looking at the title and abstract and picked out 163 potentially useful papers. We then checked the main

100    text, and the supplementary files, if available, of the 163 papers to identify databases with soil total P concentration measurements. Seven databases with soil total P measures from seven studies were selected. As observations in two databases (i.e., Li et al., 2014; Xu et al., 2013) were included in another database (i.e., Wang et al., 2021), we finally used five databases (i.e., Wang et al., 2021; Hou et al., 2020; Hou et al., 2018; Deng et al., 2017; Augusto et al., 2017) and found 2591 observations in this step, as described in detail in Table S1.

105    Second, we used "soil phosphorus" as keywords to search global or regional databases stored in public data repositories on 10th October, 2020, including Figshare (https://figshare.com/categories/Earth_and_Environmental_Sciences/33), Earthdata (https://earthdata.nasa.gov/), PANGAEA (https://www.pangaea.de/), Data.world (https://data.world/), Dryad (https://datadryad.org/stash/), and Zenodo (https://zenodo.org/). We firstly screened the databases by titles, and then picked out 80 potentially useful databases which were checked further by looking into the databases. There were nine databases with

110    soil total P concentration in (semi-)natural terrestrial ecosystems. Among the nine databases, five (Ji et al., 2018; Tipping et al., 2016; McGroddy, 2012a; Baribault et al., 2012; Cross, 1989) were excluded, due to a lack of specific site coordinates (i.e., longitude and latitude), which are needed to fill missing values of predictors from their global maps. In this step, 210 observations from four databases (i.e., Adams et al., 2020; Deiss et al., 2018; Yan et al., 2018; Gama-Rodrigues et al., 2014) were collected.

115    Third, we included 1693 measurements of soil total P concentration in a global database of soil extractable P concentration (Hou et al., unpublished), and 262 measurements of soil total P concentration in a global database of soil P fractions (He et al., unpublished). Original data sources of the two databases are given in Supplementary Text 1. After step 3, we combined measurements collected in steps 1-3 and deleted 22 duplicated ones (i.e., measurements with the same site coordinates and soil total P concentration), resulting in a total of 4734 site-level measurements of soil total P concentration from 11 databases listed

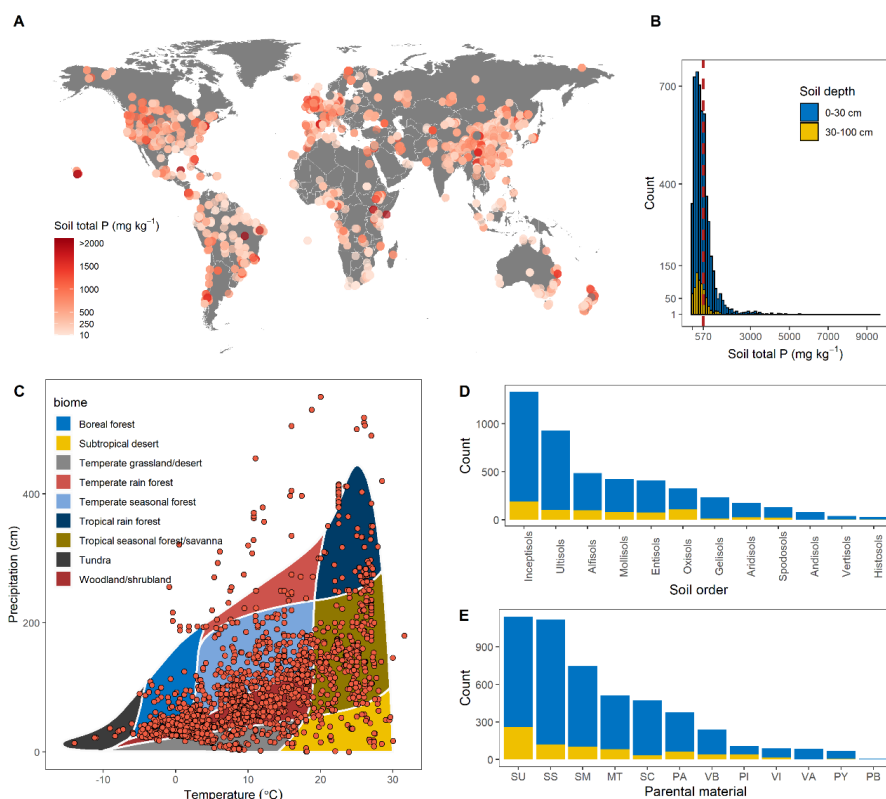
120    in Table S1.

Fourth, we searched additional soil total P concentration measurements from underrepresented regions identified in steps 1-3, from Web of Science using keywords of "soil phosphorus" along with the keywords of the underrepresented regions (listed in detail in Table S2). In this step, we collected 541 additional site-level measurements of soil total P concentrations from 60 additional papers (Table S2; Supplementary Text 1).

125     Following these steps, our database included 5,275 measurements of soil total P concentration at 1,894 sites from 761

studies (Supplementary Text 1 and Fig. S1), with 4,536 measurements in top 30 cm and 739 measurements in deeper soil

(depth > 30 cm). Besides soil total P concentration and site coordinates, we also included climate variables (i.e., MAT and

MAP), vegetation type, soil physiochemical properties (e.g., SOC, soil clay and sand contents, soil pH) in our database.

130     **Fig. 1 The distribution of our site-level training data.** The database contains 5,275 observations (A & B) covering all major terrestrial

biomes (C), 12 soil orders (D), and 12 parent materials (E). Red dashed line in figure (B) indicates the arithmetic mean of the soil total P

concentration (570 mg kg$^{-1}$). The abbreviations in figure (E) represent the following: SS: Siliciclastic sedimentary; SU: Unconsolidated

sediments; SM: Mixed sedimentary; MT: Metamorphics; SC: Carbonate sedimentary; PA: Acid plutonic; VB: Basic volcanic; VI:

Intermediate volcanic; PI: Intermediate plutonic; VA: Acid volcanic; PY: Pyroclastics; PB: Basic plutonic.

135     Soil total P concentration is thought to be influenced by five soil-forming factors, which are parent material, climate,

vegetation productivity, topography, and soil age (Delgado-Baquerizo et al., 2020; Jenny, 1941; Dokuchaev, 1883). Four of

the five factors were directly considered here (Table 1): parent material, climate (i.e. mean annual temperature (MAT), mean

annual precipitation (MAP), and biome), vegetation (i.e. net primary production (NPP)), and topography (e.g. slope and

elevation). As soil age was rarely reported, we used USDA soil orders as a proxy for age with 3 classes: slightly, intermediately,

140     and strongly weathered (Yang et al., 2013; Smeck, 1985). Among the 12 USDA soil orders, Entisols, Inceptisols, Histosols,

Andisols, and Geilsols are classified as slightly weathered soils. Alfisols, Mollisols, Aridisols, and Vertisols are classified as

intermediately weathered soils. Oxisols, Ultisols, and Spodosols are classified as strongly weathered soils (Yang et al., 2013; Smeck, 1985).

**Table 1 Summary of training data used to predict soil total P concentration.** P10 and P 90 indicate the percentile rank of 10% and 90%.

| Group | Variables | Unit | Min | P10 | Mean | P90 | Max | PFL[*] | PFGM[#] |
|---|---|---|---|---|---|---|---|---|---|
| Climate | MAT | °C | -14 | 0.7 | 11.9 | 25.1 | 31.6 | 91% | 9% |
| | MAP | mm | 10 | 356 | 1146 | 2337 | 6576 | 91% | 9% |
| Soil property | SOC | g kg$^{-1}$ | 0.1 | 2.6 | 41.3 | 92.8 | 545 | 81% | 19% |
| | Soil pH | | 2.5 | 4.2 | 5.9 | 8.1 | 10.5 | 77% | 23% |
| | Soil clay | g kg$^{-1}$ | 0.3 | 50 | 222 | 435 | 954 | 48% | 52% |
| | Soil sand | g kg$^{-1}$ | 10 | 135 | 497 | 862 | 997 | 29% | 71% |
| | Depth | cm | 0.5 | 5 | 19.4 | 50 | 100 | 100% | 0% |
| | Soil order | | 12 USDA soil orders | | | | | 64% | 36% |
| Parent material | | | 13 parent materials | | | | | 0% | 100% |
| Vegetation | Biomes | | 6 major biomes | | | | | 91% | 9% |
| | NPP | kg C m$^{-2}$ | 13 | 1517 | 5678 | 10117 | 21681 | 0% | 100% |
| Topography | Slope | ° | 0 | 0 | 8.28 | 22 | 72 | 0% | 100% |
| | Elevation | m | -41 | 34 | 861 | 2141 | 5175 | 67% | 33% |

145    MAT: Mean annual temperature; MAP: Mean annual precipitation; SOC: Soil organic carbon; NPP: Net primary production. * PFL: Proportion from literature; # PFGM: Proportion from gridded map. PFL and PFGM indicate proportions of measurements from literature and extracted from global gridded maps, respectively.

In addition to predictors of soil total P concentration related to soil-forming factors, we collected information about the properties of the soils (e.g. soil organic carbon (SOC), soil pH, soil clay content (Clay) and soil sand content (Sand), and soil

150    depth (Depth); Table 1). These soil properties were used as additional predictors. We extracted predictors from each original publication when available. In cases where information on predictors were not reported, we extracted the missing data from gridded datasets (Table S3) based on the geographical coordinates of the measurement sites.

In the random forest model, correlated predictors can be substituted for each other, so that the importance of correlated predictors will be shared, making the estimated importance smaller than the true value (Strobl et al., 2008). Thus, we did not

155    include soil total nitrogen content as it is correlated with SOC ($r = +0.84$), nor did we include aridity index as it is strongly correlated with MAP ($r = +0.82$). We also did not include variables that were rarely reported in the referenced studies (e.g. soil extractable aluminum and iron concentrations).

## 2.2 Statistical modelling

Among the 5,275 soil total P measurements, there were 15 extremely high values (> 4000 mg kg$^{-1}$) (Fig. 1B). These high

160    values were likely derived in exceptional geological contexts (Porder and Ramachandran, 2013), or special soils (e.g. very young volcanic soils). We reported these extremely high values while summarizing the database, for example, in Table 2 and

Table 3. However, we excluded these 15 measurements from model training and correlation analyses to avoid their possible large influences on the overall relationships between soil total P concentration and other variables.

We compared a suite of algorithms against the afore mentioned 13 predictors which included three generalized linear models, Cubist model, Boosted tree model, and random forest model (Table S4). Model performance was compared in terms of $R^2$ and Root Mean Square Error (RMSE) (Minasny et al., 2017). A five-fold cross-validation method was used to evaluate the performance of the models. In this method, the whole dataset was randomly split into five folds, each of which contained 20% of the data. One fold of data was used as test data, while the other four folds were used as training data. Then another fold of data was used as test data, and the remaining ones as training data, and so on and so forth for a total of five times. Averages of five sets of $R^2$ and RMSE were used as the model $R^2$ and RMSE, respectively. Based on the five-fold cross-validation method, the random forest algorithm performed the best ($R^2 = 0.65$) among all five algorithms (Table S4) and was therefore selected for follow-up analyses. Five-fold cross-validation was performed using the R package *caret* (v. 6.0-86) (Kuhn, 2020). Random forest analysis was performed with the R package *caret* by applying the embedded R package *randomForest* version 3.1 (Liaw and Wiener, 2002) with an automated *mtry* parameter. The mean decrease in accuracy (%IncMSE) was used to indicate the relative importance of each variable for predicting soil total P concentration. Partial dependence plots showed the marginal effect of each continuous predictor on soil total P concentration.

Finally, we applied the above trained model to global databases of the 13 predictors to generate a global map of soil total P concentration. The gridded driver variables used for the global prediction were all re-gridded to a spatial resolution of 0.05° × 0.05° (the original resolution can be found in the Table S3). We did not mask out cropland or other heavily influenced areas (e.g. cities, roads, etc.), so the predicted map can be used to represent an initial state before anthropogenic activities.

Soil depth was used as a covariate, so that the models could predict soil total P concentration for any given depth (Hengl et al., 2017). The partial dependence plot indicated that soil total P concentration approximately linearly decreased with soil depth in the top 30 cm and there was no apparent trend with depth in the subsoil (~30-100 cm). Given this, we predicted global soil total P concentration at 5 cm, 15 cm, 25 cm, and 65 cm to represent the soil total P concentration in the 0-10 cm, 10-20 cm, 20-30 cm, and 30-100 cm layers, respectively. Averages in other depth intervals (e.g. 0-30 cm or 0-100 cm), can be derived by taking a weighted average of the predictions within the depth interval (Hengl et al., 2017). We used global gridded soil depth data (Shangguan et al., 2017) to correct the soil depth when it was less than 100 cm in any cell. The global soil P stock maps for 0-10 cm, 10-20 cm, 20-30 cm, and 30-100 cm layers were calculated from the soil total P concentrations predicted here and the soil bulk density in corresponding layers predicted by Hengl et al. (2017).

Prediction uncertainty of each cell in global gridded map was assessed using bootstrap samples with the quantile regression forest technique (Meinshausen, 2006). Standard deviation was calculated to represent the uncertainty using *quantregForest* function in the *quantregForest* R package (Meinshausen, 2017). Individual predictions of each tree in the random forest model

(n=500) were returned to assess the variation of predicted global mean soil total P concentration and these results were used

to assess the standard deviation of the estimated global soil P stock.

195     All statistical analyses and plotting were performed in the R environment (v. 4.0.2) (R Core Team, 2018).

## 3 Results

### 3.1 Characteristics of soil total P concentration across the world

Our soil total P concentration database included 5,275 measurements from 1,894 geographically distinct sites and covered

6 continents, all major biomes, and all 12 USDA soil orders in terrestrial ecosystems (Fig. 1A-D & Table S5). The database

200     was highly right-skewed (Fig.1B) and revealed that the soil total P concentration in natural soils of terrestrial ecosystems varied

from 1.4 to 9,636.0 mg kg$^{-1}$, with a mean, median and standard deviation of 570.0, 430.0, and 646.5 mg kg$^{-1}$, respectively

(Table 2). The database included soil total P concentration measurements from topsoil to 100 cm depth, with 84.4 % of the

measurements from the topsoil (e.g. 0-30 cm). The average soil total P concentration in our database was 583.7 and 495.2 mg

kg$^{-1}$ in the topsoil (0-30cm) and subsoil (30-100 cm), respectively.

205     **Table 2 Soil total P concentration (mg kg$^{-1}$) in natural ecosystems for major biomes at the global scale.** Results based on our site-level
database in the top 1 m of soil. P10, P25, P75, and P 90 indicate the percentile rank of 10%, 25%, 75%, and 90%.

| Biome | Min | P10 | P25 | Median | Mean | P75 | P90 | Max |
|---|---|---|---|---|---|---|---|---|
| Tundra | 35.0 | 104.7 | 254.4 | 551.0 | 986.1 | 1000.0 | 2216.0 | 9630.0 |
| Boreal | 3.0 | 121.0 | 327.5 | 556.5 | 715.5 | 851.6 | 1384.6 | 5520.0 |
| Temperate | 3.0 | 125.6 | 246.3 | 458.0 | 537.2 | 678.2 | 1002.7 | 4086.6 |
| Mediterranean | 4.8 | 96.0 | 252.3 | 443.3 | 554.4 | 621.5 | 873.0 | 4433.0 |
| Desert | 5.0 | 33.7 | 63.0 | 337.5 | 381.4 | 566.9 | 717.0 | 4800.0 |
| Tropics | 1.4 | 63.4 | 137.8 | 283.1 | 416.8 | 526.3 | 919.3 | 3898.0 |
| Global | 1.4 | 90.0 | 212.8 | 430.0 | 570.0 | 685.4 | 1060.9 | 9630.0 |

**Table 3 Soil total P concentration (mg kg$^{-1}$) in 12 USDA soil orders and three weathering stages.** Results based on our database in the
top 1 m of soil. P10, P25, P75, and P 90 indicate the percentile rank of 10%, 25%, 75%, and 90%.

| | Min | P10 | P25 | Median | Mean | P75 | P90 | Max |
|---|---|---|---|---|---|---|---|---|
| *Slightly weathered* | 11.0 | 143.1 | 290.0 | 540.0 | 645.6 | 820.0 | 1234.9 | 9630.0 |
| Andisols | 175.0 | 321.9 | 534.4 | 883.2 | 1042.4 | 1459.8 | 2040.9 | 3548.0 |
| Gelisols | 35.0 | 175.1 | 321.8 | 584.1 | 1145.8 | 1047.7 | 3013.0 | 9630.0 |
| Entisols | 14.1 | 67.8 | 240.0 | 502.8 | 541.3 | 789.5 | 1020.0 | 2321.3 |
| Inceptisols | 11.0 | 150.8 | 293.4 | 530.0 | 665.0 | 802.7 | 1235.6 | 5520.0 |
| Histosols | 66.0 | 350.5 | 575.0 | 800.0 | 882.4 | 1069.2 | 1680.0 | 1996.0 |
| *Intermediately weathered* | 1.4 | 41.5 | 191.0 | 390.0 | 470.9 | 610.0 | 918.7 | 4800.0 |
| Aridisols | 5.0 | 36.7 | 152.1 | 392.4 | 445.6 | 591.5 | 865.5 | 4800.0 |
| Alfisols | 1.4 | 29.2 | 147.0 | 327.5 | 468.3 | 602.5 | 1057.0 | 4243.0 |
| Mollisols | 9.8 | 76.8 | 270.0 | 458.5 | 470.8 | 619.0 | 831.9 | 3199.0 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Vertisols | 112.7 | 199.6 | 242.0 | 421.8 | 871.5 | 1012.5 | 2825.0 | 3680.0 |
| *Strongly weathered* | 3.4 | 103.9 | 210.0 | 376.5 | 468.6 | 599.6 | 899.1 | 4086.6 |
| Oxisols | 5.1 | 94.0 | 141.0 | 358.5 | 433.2 | 625.0 | 938.0 | 2000.0 |
| Ultisols | 3.4 | 111.7 | 218.8 | 380.0 | 466.3 | 582.8 | 837.0 | 4086.6 |
| Spodosols | 14.5 | 146.0 | 252.5 | 401.8 | 584.8 | 722.9 | 1119.3 | 3444.2 |

210

The database revealed that soil total P concentration varied within and among biomes. The soil from tundra and boreal biomes had the highest soil total P concentrations. Mediterranean and temperate soils had intermediate soil total P concentrations. Soils in the desert and tropics had relatively lower soil total P concentrations (Table 2 & Fig. 2B). Soil total P concentration also varied with different soil orders (Table 3). Strongly and intermediately weathered soils (mean values of

215    468.6 mg kg$^{-1}$ and 470.9 mg kg$^{-1}$, respectively) had lower soil total P concentrations than slightly weathered soils (mean value 645.6 mg kg$^{-1}$) (Fig. 2C).

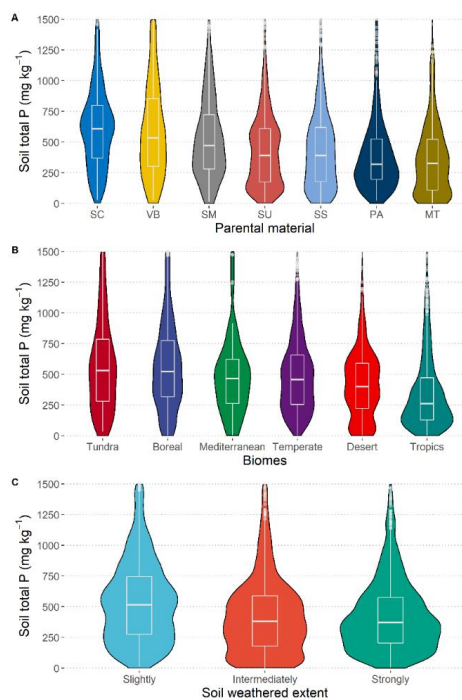



**Fig. 2 Soil total P concentration in relation to parent material, biome, and soil weathering extent.** For visualization, we chose to limit
220    the y-axis to 1500 mg kg$^{-1}$; and in panel A, only parent material types with more than 100 measurements in our database were shown; the abbreviations in figure (A) represent the following: SC: Carbonate sedimentary; VB: Basic volcanic; SM: Mixed sedimentary; SU: Unconsolidated sediments; SS: Siliciclastic sedimentary; PA: Acid plutonic; MT: Metamorphics.

**3.2 Model performance and drivers of soil total P concentration**

The random forest regression model explained 65% of soil total P concentration variability across all sites, with an RMSE
of 288.8 mg kg$^{-1}$ (Fig. 3B & Table S4). The random forest model revealed that the two most important predictors of soil total
P concentration were SOC content and parent material. The remaining predictors showed a lower, but non-negligible influence,
with MAT and soil sand content having the most noticeable influence (Fig. 3A). Although soil order, biome, elevation, slope,
depth, NPP, and pH showed significant influences on soil total P concentration (Fig. 2 and Fig. S3), their relative importance
was lower than the above four predictors. Partial dependent plots (Fig. 4) revealed similar results to Pearson correlation analysis
(Fig. S3). The partial dependent plots indicated a significant and positive relationship between soil total P concentration and
SOC at a global scale; soil total P concentration was significantly and negatively correlated with MAT and soil sand content
(Fig. S4). The Pearson correlation indicated the correlation coefficients between soil total P concentration and the top three
continuous predictors MAT, SOC, and soil sand content were -0.23, 0.19, and -0.18, respectively (Fig. S3).

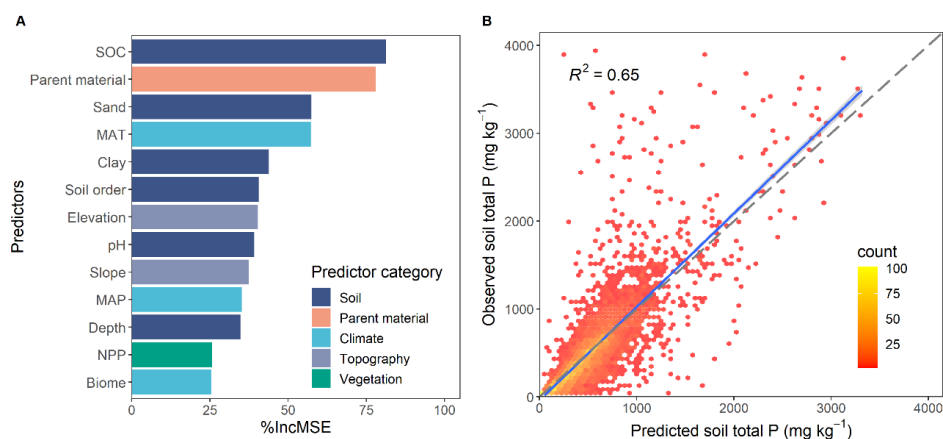



**Fig. 3 Results of the random forest model predicting soil total P concentration.** (A) The relative importance of predictors in the model.
(B) Predicted vs. observed soil total P concentration; the dashed line indicates the 1:1 line; the blue line indicates the regression line between
predicted and observed values.

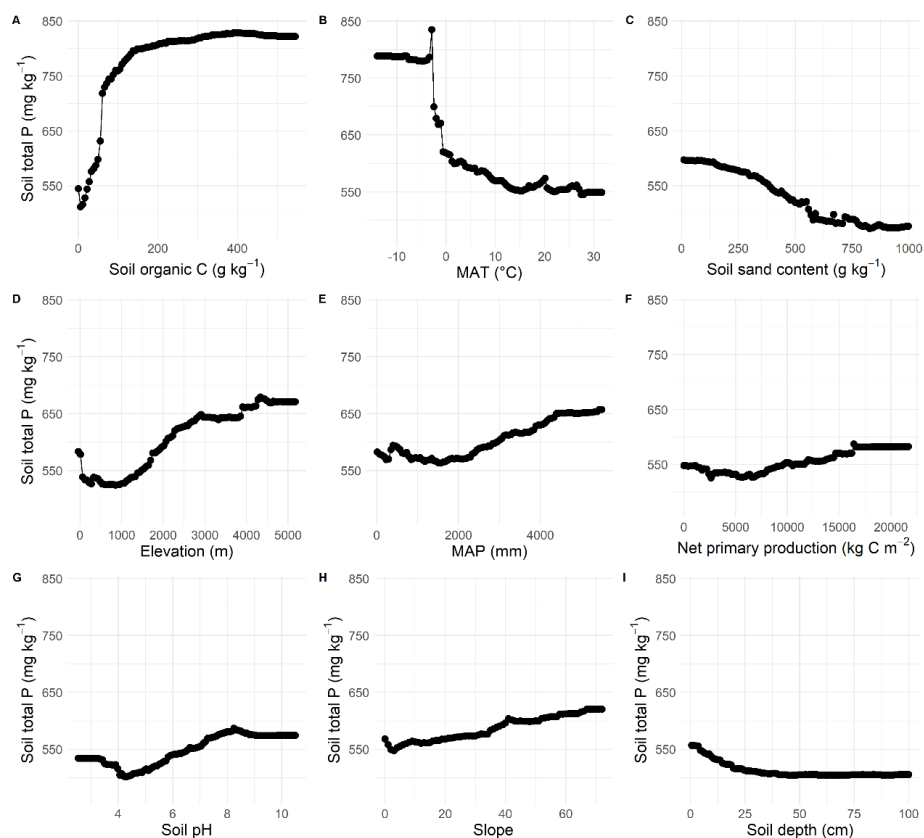



**Fig. 4 Partial dependence plots showing the dependence of soil total P concentration on predictors.** Soil total P concentration in relation
to SOC concentration, MAT, soil sand content, elevation, MAP, net primary production, soil pH, slope, and soil depth (A, B, C, D, E, F, G,
H, I, respectively).

### 3.3 Global patterns of soil total P

Our global map of predicted soil total P indicated that the total global P stocks in the topsoil (0-30 cm) and subsoil (30-100

cm) were 26.8 (standard deviation = 3.1) Pg and 62.2 (standard deviation = 8.9) Pg, respectively (excluding Antarctica; Table

4). Estimated area-weighted average soil total P concentrations in the topsoil and subsoil were 529.0 and 502.3 mg kg$^{-1}$,

respectively. Estimated area-weighted average soil total P content in the topsoil and subsoil were 209.7 and 487.0 g cm$^{-2}$,

respectively.

**Table 4 Analysis of the predicted global map of soil total P in natural ecosystems.** Area weighted average soil total P concentration was
calculated based on our predicted map. Converting soil total P concentration to soil total P content and stock used the soil bulk density (Hengl
et al., 2017) and land area.

| | 0-30 cm | | | 30-100 cm | | |
|---|---|---|---|---|---|---|
| Continent | Soil total P concentration (mg kg$^{-1}$) | Soil total P content (g m$^{-2}$) | Soil P stock (Pg) | Soil total P concentration (mg kg$^{-1}$) | Soil total P content (g m$^{-2}$) | Soil P stock (Pg) |

Earth System Science Data Discussions Open Access

| Africa | 390.2 | 164.1 | 4.5 | 360.7 | 362.8 | 10.1 |
|---|---|---|---|---|---|---|
| Asia | 603.0 | 238.7 | 10.3 | 576.3 | 565.2 | 24.3 |
| Europe | 632.4 | 240.9 | 2.4 | 601.1 | 581.7 | 5.7 |
| Oceania | 401.5 | 177.1 | 1.4 | 373.4 | 397.6 | 3.4 |
| South America | 411.5 | 158.1 | 2.8 | 392.0 | 358.6 | 6.3 |
| North America | 657.3 | 251.2 | 5.3 | 631.5 | 587.4 | 12.3 |
| *Global* | 529.0 | 209.7 | 26.8 | 502.3 | 487.0 | 62.2 |

The estimated global map of soil total P concentration revealed latitudinal patterns (Fig. 5), which were also be found from analysis of the site-level data (Fig. S4K). Soil total P concentration generally increased from the equator to high latitudes. The latitudinal pattern of soil total P concentration was not found in earlier work (Yang et al., 2013; Shangguan et al., 2014). Our

255    predicted soil total P concentrations were weakly correlated, though significantly, with earlier predicted maps, i.e., Yang et al. (2013) and Shangguan et al. (2014) (Fig. S6).

Highlands and mountains at low latitudes (e.g., the Tibetan plateaus, Andes, east Africa, west India etc.) had high soil total P concentrations. Our map also indicated some regional difference in soil total P, for example, central Australia was low in soil total P compared with east and west Australia. On a larger scale, South America, Oceania, and Africa had the lowest soil total

260    P concentration, while soil total P concentration was highest in Europe, North America, and Asia (Table 4). The estimated soil total P concentrations in the subsoil showed similar patterns to those found in the topsoil (Fig. 5A&C).

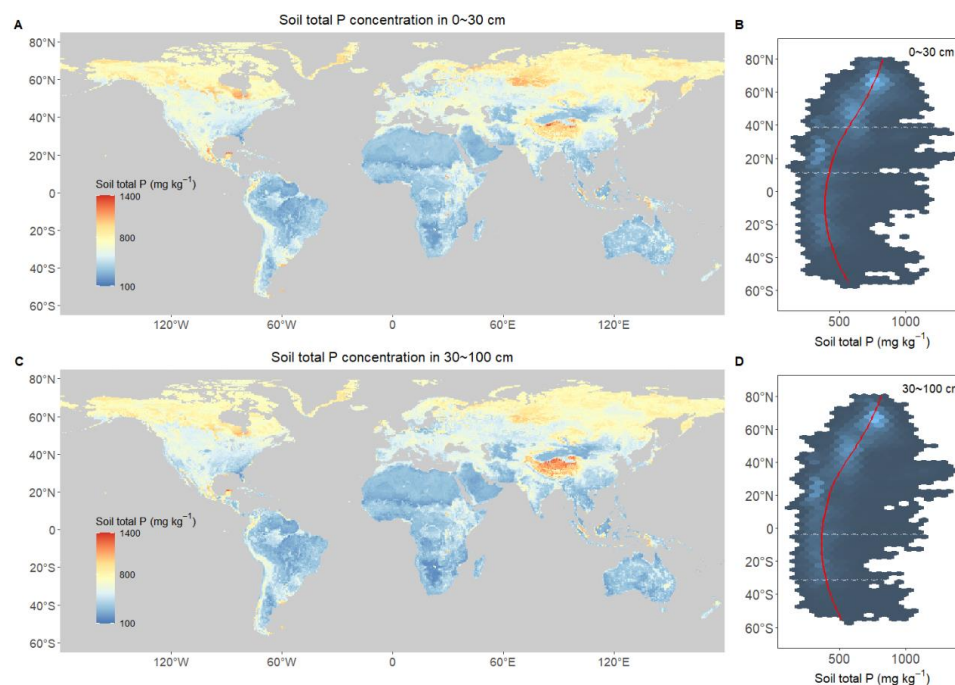



**Fig. 5 Global maps of total P concentration in the 0-30 cm and 30-100 cm of soils.** A and B are maps of topsoil (0-30 cm) total P concentration and the latitudinal patterns, respectively. C and D are maps of subsoil (30-100 cm) total P concentration and the latitudinal

265    patterns, respectively. Red lines in B and D indicate the locally weighted regressions between latitude and soil total P concentration in the

precited global map. Note that we did not mask out cropland or any other heavily influenced areas from the maps, so they can be used to represent soils without essential anthropogenic activities.

## 4 Discussion

270     With our soil total P concentration dataset, we quantified soil total P concentration in natural ecosystems, identified its key drivers, and predicted it for terrestrial ecosystems globally. Our work goes beyond previous studies (Delmas et al., 2015; Hengl et al., 2017; Shangguan et al., 2014; Viscarra Rossel and Bui, 2016; Yang et al., 2013; Cheng et al., 2016) which used limited data that did not represent the heterogeneous conditions found on Earth well, and did not separate natural soils from human-managed soils and therefore may not be able to distinguish natural drivers from anthropogenic factors (e.g. land use type, mineral fertilizer). In addition, we mapped soil total P concentration by considering more predictors, at multiple soil depths,
275     and at a higher resolution than previous studies.

### 4.1 Characteristics of soil total P concentration

     Given the larger number of measurements that we considered, the range of total P concentration in our study (1.4–9630.0 mg kg$^{-1}$) is wider than that reported in Cleveland and Liptzin (2007) (83.7–2746.6 mg kg$^{-1}$; n=186), Xu et al. (2013) (12.7–8400.1 mg kg$^{-1}$; n=536), Li et al. (2014) (30– 2744 mg kg$^{-1}$; n=178), and Hou et al. (2018) (4.8–2157.0 mg kg$^{-1}$; n=254). The
280     average soil total P concentration in our site-level database (570.0 mg kg$^{-1}$) was within the range of previous estimates by Cleveland and Liptzin (2007) (721.1 mg kg$^{-1}$), Xu et al. (2013) (756.4 mg kg$^{-1}$), Li et al. (2014) (463.6 mg kg$^{-1}$), and Hou et al. (2018) (471.9 mg kg$^{-1}$).

### 4.2 Soil total P concentration in relation to its predictors

     In agreement with previous studies, soil total P concentration was largely predicted by parent material type (Deiss et al.,
285     2018; Augusto et al., 2017; Porder and Ramachandran, 2013). This result supports the use of parent material to map soil total P concentration at the global scale (Yang et al., 2013). Parent material can affect soil total P concentration both directly and indirectly. Some parent materials tend to have higher P concentrations, which then translates into higher total soil P (Mage and Porder, 2013; Dieter et al., 2010; Kitayama et al., 2000). Additionally, parent material also affects soil total P indirectly via the influence of soil physiochemical properties such as soil texture, pH, and Al and Fe oxides (Siqueira et al., 2021; Mehmood et
290     al., 2018). For example, the retention of P in soil can be influenced by the soil content of clay, soluble calcium, and Fe oxyhydroxides (Delgado-Baquerizo et al., 2020; Mehmood et al., 2018; Achat et al., 2016). As such, parent material type is a critical predictor of soil total P from local to global scales.

     Interestingly, we found that SOC was one of the two most important predictors of soil total P concentration. The positive

relationship between soil total P and SOC has two possible explanations. First, this relationship may reflect the coupling

295 between P and C in soils (Hou et al., 2018a). Phosphorus couples with organic C in soil because soil P is composed largely of

organic P (typically in 30%–80%) which has a relatively fixed ratio to organic C (Achat et al., 2013a; Cleveland and Liptzin,

2007; Achat et al., 2013b). Second, P and organic C are stabilized and retained through similar processes in soil (Doetterl et

al., 2015). For example, reactive minerals can simultaneously stabilize both P and organic C in soil (Helfenstein et al., 2018).

As such, the strong relationship between SOC and total P at the global scale confirms that SOC is an integrated measure of

300 biotic (e.g. soil microbial activity) and abiotic (e.g. cation exchangeable capacity) factors that regulate soil total P (Spohn,

2020; Wang et al., 2020).

Consistent with a recent global synthesis that focused on soil P fractions (Hou et al., 2018), our result indicated that MAT

was a more important predictor of soil total P concentration than MAP. The negative relationship could be because soils under

low MAT are often found at high latitudes where soils were eroded during the last glaciation. These soils tend to be much

305 younger compared to soils at low latitudes with high MAT and thus have experienced less losses of P (Vitousek et al., 2010).

In addition, high MAT and MAP generally promote soil weathering as well as plant growth and P uptake, resulting in the

depletion of soil P (Huston and Wolverton, 2009; Arenberg and Arai, 2019; Huston, 2012).

Further, we provide two explanations for the negative relationship between soil total P concentration and sand content.

First, soil sand content is a surrogate for quartz content (Bui & Henderson, 2013), and the rock content in quartz is usually

310 negatively correlated with the total P content of siliceous rocks (Hahm et al., 2014; Vitousek et al., 2010). Second, soil sand is

worse at retaining nutrients including P than other soil fractions (Augusto et al., 2017). For example, loamy soils regularly lose

0.3–0.5 kg P ha$^{-1}$ yr$^{-1}$ by leaching, while coarse sandy soils lose up to 2.0 kg P ha$^{-1}$ yr$^{-1}$ (Amberger, 1996).

### 4.3 Global patterns of soil total P

Based on our predicted global map, we estimated that the area-weighted global average of soil total P concentration was

315 529.0 and 502.3 mg kg$^{-1}$ in the topsoil (0-30 cm) and subsoil (30-100 cm), respectively. Our estimated of the area-weighted

average soil total P in the topsoil was higher than previous estimates by Yang et al. (2013) (374.7 mg kg$^{-1}$) and Shangguan et

al. (2014) (484.7 mg kg$^{-1}$), but was very close to the estimate by Xu et al. (2013) (514.6 mg kg$^{-1}$). Our estimate of the global

soil P stock in the top 30 cm of soil (26.8 Pg; excluding Antarctica) was in line with the estimate of Shuanguan et al. (2013)

(26.7 Pg in the top 30 cm), but was higher than the estimates of Yang et al. (2013) (24.4 Pg in the top 30 cm), Wang et al.

320 (2010) (18.2 Pg in the top 30 cm), and Smil (2000) (24 Pg). Additionally, our estimate was much lower than a much earlier

estimate by Jahnke (1992) (about 120 Pg in the top 30 cm).

Our predicted soil total P concentrations decreased significantly with decreasing latitude. This result is consistent with our

theoretical understanding of the evolution of soils in soil chronosequences (Walker and Syres 1976) and the stark differences

in soil age and weathering intensity between low and high latitude regions. And this result is agreement with a recent meta-analysis that revealed P-limitation to plant growth decreased significantly with latitude (Hou et al., 2021). Soils at high latitudes are relatively young and P-enriched compared to soils at low latitudes due to erosion during the last glaciation (Vitousek et al., 2010; Reich and Oleksyn, 2004). Our result is consistent with Xu et al. (2013); by comparing soil total P concentration across the major biomes, the authors found the highest soil total P concentration in the tundra and the lowest in the tropical/subtropical forest. Previous global maps of soil total P concentration were not able to capture the latitudinal trend of soil total P concentrations (Yang et al., 2013; Shangguan et al., 2014), likely due to poorer spatial coverage of their measurements. For example, their measurements were mostly from the US and China, with a very small proportion of measurements from high latitudes.

While we found a latitudinal gradient in soil total P concentration, heterogeneity in soil total P concentration at the regional and local scales was large. For example, consistent with Brédoire et al. (2016), we found that the soil total P concentration was higher in Siberia than in northern Europe, both of which have similar latitudes. First, this difference may be due to the fact that glaciation was more regular and intense in Siberia than in northern Europe (Wassen et al., 2021), leading to a more intensive rejuvenation of soils. Second, the warmer and wetter climate in northern Europe may promote weathering which releases P from parent material (Goll et al., 2014) and makes it subject to loss (Fig. S7). Regional variation in soil total P concentration may also be attributable to regional variation in parent material. For example, higher soil total P concentration in eastern Australia than in central Australia was probably due to P-enriched basaltic lithologies in eastern Australia (6500–8700 mg kg$^{-1}$) (Viscarra Rossel and Bui, 2016). Moreover, regional differences in soil total P concentration may be related to topography conditions. For example, higher soil total P concentration in the Tibetan plateau than in eastern China may be the result of higher elevation and lower MAT in the Tibetan plateau (Zhang et al., 2005).

**4.4 Limitation and prediction uncertainty**

Despite our unprecedented effort to construct a database and perform global predictions, our study has some limitations. First, some regions, for example, northern Canada, Russia, middle Asia, and inner Australia, were still underrepresented, which may result in low accuracy of the predicted values in these regions (Ploton et al., 2020). Second, subsoils (> 30 cm depth) were not well represented in our dataset (14%) and therefore predicted P concentrations of subsoils may suffer from larger uncertainties than those of topsoils (< 30 cm depth). Third, some predictors were largely missing. Map-filled values suffer from large uncertainties, especially for the soil variables. This may cause some uncertainties in the predicted soil total P concentration. Finally, 36% of the variation in soil total P concentration was not explained, despite inclusion of 13 predictors using an advanced machine learning approach. This result may be because of measurement errors and/or methodological constraints. These limitations highlight the need for more measurements of subsoil total P concentration and closely associated

variables, especially from the underrepresented regions, as well as more advanced statistical methods for spatial predictions.

355 **5 Conclusion**

By constructing a database of total P concentration globally, we quantified the relative importance of multiple soil-forming variables for predicting soil total P concentration and further estimated it at the global scale. Our results indicated that no single variable can be used to predict soil total P concentration. Instead, it is a combination of variables that are needed to reliably predict soil total P concentration, among which SOC, parent material, MAT, and soil sand content are the most important

360 predictors. Soil total P concentration was positively correlated with SOC and negatively correlated with both MAT and soil sand content. Our predicted map captures the latitudinal gradient in soil total P concentration expected from our theoretical understanding. We estimated that P stocks in the topsoil (0-30 cm) and subsoil (30-100 cm) of soil of natural ecosystems (excluding Antarctica) were 26.8 and 62.2 Pg, respectively. Our improved global map of soil total P will be an important resource for future work which aims to tackle issues related to P cycling, including predicting future land carbon sink potential

365 and P losses to aquatic and marine ecosystems as well as modeling the P needs of crops to increase food security.

**Data availability**

Raw datasets, R code, and global maps generated in this study are available at https://doi.org/10.6084/m9.figshare.14583375 (He et al., 2021).

370 **Author contributions**

X.H. and E.H designed this study. X.H., E.H, L.A., and Z.W. collected the data. X.H., E.H., L.A., D.S.G., B.R., Y.W., J.H., Y.H., and K.Y. discussed analyzing methods. X.H. conducted the analysis and drafted the manuscript. All authors discussed the results and contributed to the manuscript.

**Acknowledgements**

375 This research was funded by the National Natural Science Foundation of China (32071652, 31870464), the China Postdoctoral Science Foundation (2020M673123), Chongqing Technology Innovation and Application Demonstration Major Theme Special Project (cstc2018jszx-zdyfxmX0007), and the ANR CLAND Convergence Institute. We would like to thank Dr. Morgan Furze at Yale University for his assistance with English language and grammatical editing.

**Competing interests**

380 The authors declare that they have no conflict of interest.

**References**

Achat, D.L., Augusto, L., Gallet-Budynek, A. and Loustau, D.: Future challenges in coupled C-N-P cycle models for terrestrial ecosystems under global change: a review. Biogeochemistry, 131, 173-202, doi: 10.1007/s10533-016-0274-9, 2016.

385 Achat, D.L., Bakker, M.R., Augusto, L. and Morel, C.: Contributions of microbial and physical–chemical processes to phosphorus availability in Podzols and Arenosols under a temperate forest. Geoderma, 211-212, 18-27, doi: 10.1016/j.geoderma.2013.07.003, 2013a.

Achat, D.L., Bakker, M.R., Augusto, L., Derrien, D., Gallegos, N., Lashchinskiy, N., Milin, S., Nikitich, P., Raudina, T., Rusalimova, O., Zeller, B. and Barsukov, P.: Phosphorus status of soils from contrasting forested ecosystems in

390 southwestern Siberia: effects of microbiological and physicochemical properties. Biogeosciences, 10, 733-752, doi: 10.5194/bg-10-733-2013, 2013b.

Achat, D.L., Pousse, N., Nicolas, M., Brédoire, F. and Augusto, L.: Soil properties controlling inorganic phosphorus availability: general results from a national forest network and a global compilation of the literature. Biogeochemistry, 127, 255-272, doi: 10.1007/s10533-015-0178-0, 2016.

395 Adams, J., Tipping, E., Thacker, S.A. and Quinton, J.N.: Phosphorus, carbon and nitrogen concentrations in UK soil mineral fractions, 2013-2014. NERC Environmental Information Data Centre, doi: 10.5285/e6e9a85c-b537-4110-899f-2c1498bc826c, 2020.

Andersson, H., Bergström, L., Djodjic, F., Ulén, B. and Kirchmann, H.: Topsoil and Subsoil Properties Influence Phosphorus Leaching from Four Agricultural Soils. J. Environ. Qual., 42, 455-463, doi: 10.2134/jeq2012.0224, 2013.

400 Arenberg, M.R. and Arai, Y.: Uncertainties in soil physicochemical factors controlling phosphorus mineralization and immobilization processes. Adv. Agron., 154, 153-200, doi: 10.1016/bs.agron.2018.11.005, 2019.

Augusto, L., Achat, D.L., Jonard, M., Vidal, D. and Ringeval, B.: Soil parent material—A major driver of plant nutrient limitations in terrestrial ecosystems. Global Change Biol., 23, 3808-3824, doi: 10.1111/gcb.13691, 2017.

Augusto, L., Bakker, M.R., Morel, C., Meredieu, C., Trichet, P., Badeau, V., Arrouays, D., Plassard, C., Achat, D.L., Gallet-

405 Budynek, A., Merzeau, D., Canteloup, D., Najar, M. and Ranger, J.: Is 'grey literature' a reliable source of data to characterize soils at the scale of a region? A case study in a maritime pine forest in southwestern France. Eur. J. Soil Sci., 61, 807-822, doi: 10.1111/j.1365-2389.2010.01286.x, 2010.

Ballabio, C., Lugato, E., Fernández-Ugalde, O., Orgiazzi, A., Jones, A., Borrelli, P., Montanarella, L. and Panagos, P.: Mapping LUCAS topsoil chemical properties at European scale using Gaussian process regression. Geoderma, 355,

410 113912, doi: 10.1016/j.geoderma.2019.113912, 2019.

Baribault, T.W., Kobe, R.K. and Finley, A.O.: Data from: Tropical tree growth is correlated with soil phosphorus, potassium, and calcium, though not for legumes, Dryad, Dataset, doi: 10.5061/dryad.r9p70, 2012.

Brédoire, F., Bakker, M.R., Augusto, L., Barsukov, P.A., Derrien, D., Nikitich, P., Rusalimova, O., Zeller, B. and Achat, D.L.: What is the P value of Siberian soils? Soil phosphorus status in south-western Siberia and comparison with a global data

415 set. Biogeosciences, 13, 2493-2509, doi: 10.5194/bg-13-2493-2016, 2016.

Buendía, C., Kleidon, A. and Porporato, A.: The role of tectonic uplift, climate, and vegetation in the long-term terrestrial phosphorous cycle. Biogeosciences, 7, 2025-2038, doi: 10.5194/bg-7-2025-2010, 2010.

Carpenter, S.R. and Bennett, E.M.: Reconsideration of the planetary boundary for phosphorus. Environ. Res. Lett., 6, 014009, doi: 10.1088/1748-9326/6/1/014009, 2011.

420 Cheng, Y., Li, P., Xu, G., Li, Z., Cheng, S. and Gao, H.: Spatial distribution of soil total phosphorus in Yingwugou watershed of the Dan River, China. Catena, 136, 175-181, doi: 10.1016/j.catena.2015.02.015, 2016.

Cheng, Y., Li, P., Xu, G., Li, Z., Yu, K., Cheng, S., Zhao, B. and Wang, F.: Factors that influence soil total phosphorus sources on dam fields that are part of ecological construction programs on the Loess Plateau, China. Catena, 171, 107-114, doi: 10.1016/j.catena.2018.07.006, 2018.

425 Cleveland, C.C. and Liptzin, D.: C:N:P stoichiometry in soil: is there a "Redfield Ratio" for the microbial biomass? Biogeochemistry, 85, 235-252, doi: 10.1007/s10533-007-9132-0, 2007.

Cross, A.: Phosphorus fractions in grassland and shrubland soils at the Sevilleta NWR sev.149.2 (https://knb.ecoinformatics.org/knb/metacat/sev.149.2/default).,1989.

De Schrijver, A., Vesterdal, L., Hansen, K., De Frenne, P., Augusto, L., Achat, D.L., Staelens, J., Baeten, L., De Keersmaeker,

430 L., De Neve, S. and Verheyen, K.: Four decades of post-agricultural forest development have caused major redistributions of soil phosphorus fractions. Oecologia, 169, 221-234, doi: 10.1007/s00442-011-2185-8, 2012.

Deiss, L., de Moraes, A. and Maire, V.: Environmental drivers of soil phosphorus composition in natural ecosystems. Biogeosciences, 15, 4575-4592, doi: 10.5194/bg-15-4575-2018, 2018.

435 Delgado-Baquerizo, M., Reich, P.B., Bardgett, R.D., Eldridge, D.J., Lambers, H., Wardle, D.A., Reed, S.C., Plaza, C., Png, G.K., Neuhauser, S., Berhe, A.A., Hart, S.C., Hu, H., He, J., Bastida, F., Abades, S., Alfaro, F.D., Cutler, N.A., Gallardo, A., García-Velázquez, L., Hayes, P.E., Hseu, Z., Pérez, C.A., Santos, F., Siebe, C., Trivedi, P., Sullivan, B.W., Weber-Grullon, L., Williams, M.A. and Fierer, N.: The influence of soil age on ecosystem structure and function across biomes. Nat. Commun., 11, 4721-4721, doi: 10.1038/s41467-020-18451-3, 2020.

Delmas, M., Saby, N., Arrouays, D., Dupas, R., Lemercier, B., Pellerin, S. and Gascuel-Odoux, C.: Explaining and mapping
440     total phosphorus content in French topsoils. Soil Use Manage., 31, 259-269, doi: 10.1111/sum.12192, 2015.

Deng, Q., McMahon, D.E., Xiang, Y., Yu, C.L., Jackson, R.B. and Hui, D.: A global meta-analysis of soil phosphorus dynamics after afforestation. New Phytol, 213, 181-192, doi: 10.1111/nph.14119, 2017.

Dieter, D., Elsenbeer, H. and Turner, B.L.: Phosphorus fractionation in lowland tropical rainforest soils in central Panama. Catena, 82, 118-125, doi: 10.1016/j.catena.2010.05.010, 2010.

445 Doetterl, S., Stevens, A., Six, J., Merckx, R., Van Oost, K., Casanova Pinto, M., Casanova-Katny, A., Muñoz, C., Boudin, M., Zagal Venegas, E. and Boeckx, P.: Soil carbon storage controlled by interactions between geochemistry and climate. Nat. Geosci., 8, 780-783, doi: 10.1038/ngeo2516, 2015.

Dokuchaev, V.V., The Russian chernozem. St. Petersburg, Russia: Imperial Univ. of St. Petersburg. Report to the Free Economic Society. [in Russian],1883.

450 Elser, J.J., Bracken, M.E.S., Cleland, E.E., Gruner, D.S., Harpole, W.S., Hillebrand, H., Ngai, J.T., Seabloom, E.W., Shurin, J.B. and Smith, J.E.: Global analysis of nitrogen and phosphorus limitation of primary producers in freshwater, marine and terrestrial ecosystems. Ecol. Lett., 10, 1135-1142, doi: 10.1111/j.1461-0248.2007.01113.x, 2007.

Fleischer, K., Rammig, A., Kauwe, D.M.G., Walker, A.P., Domingues, T.F., Fuchslueger, L., Garcia, S., Goll, D.S., Grandis, A., Jiang, M., Haverd, V., Hofhansl, F., Holm, J.A., Kruijt, B., Leung, F., Medlyn, B.E., Mercado, L.M., Norby, R.J., Pak,
455     B., Randow, V.C., Quesada, C.A., Schaap, K.J., Valverde-Barrantes, O.J., Wang, Y.P., Yang, X., Zaehle, S., Zhu, Q., Lapola, D.M. and Oak Ridge National Lab. ORNL, O.R.T.U.: Amazon forest response to CO2 fertilization dependent on plant phosphorus acquisition. Nat. Geosci., 12, 736-741, doi: 10.1038/s41561-019-0404-9, 2019.

Gama-Rodrigues, A.C., Sales, M.V.S., Silva, P.S.D., Comerford, N.B., Cropper, W.P. and Gama-Rodrigues, E.F.: An exploratory analysis of phosphorus transformations in tropical soils using structural equation modeling. Biogeochemistry,
460     118, 453-469, doi: 10.1007/s10533-013-9946-x, 2014.

Garrett, R.G.: Relative spatial soil geochemical variability along two transects across the United States and Canada. Appl. Geochem., 24, 1405-1415, doi: 10.1016/j.apgeochem.2009.04.011, 2009.

Goll, D.S., Brovkin, V., Parida, B.R., Reick, C.H., Kattge, J., Reich, P.B., van Bodegom, P.M. and Niinemets, Ü.: Nutrient limitation reduces land carbon uptake in simulations with a model of combined carbon, nitrogen and phosphorus cycling.
465     Biogeosciences, 9, 3547-3569, doi: 10.5194/bg-9-3547-2012, 2012.

Goll, D.S., Moosdorf, N., Hartmann, J. and Brovkin, V.: Climate-driven changes in chemical weathering and associated phosphorus release since 1850: Implications for the land carbon balance. Geophys. Res. Lett., 41, 3553-3558, doi: 10.1002/2014GL059471, 2014.

Goll, D.S., Vuichard, N., Maignan, F., Jornet-Puig, A., Sardans, J., Violette, A., Peng, S., Sun, Y., Kvakic, M., Guimberteau,
470     M., Guenet, B., Zaehle, S., Penuelas, J., Janssens, I. and Ciais, P.: A representation of the phosphorus cycle for ORCHIDEE (revision 4520). Geosci. Model Dev., 10, 3745-3770, doi: 10.5194/gmd-10-3745-2017, 2017.

He, X., Hou, E., Liu, Y. and Wen, D.: Altitudinal patterns and controls of plant and soil nutrient concentrations and stoichiometry in subtropical China. Sci. Rep.-UK, 6, 24261, doi: 10.1038/srep24261, 2016.

He, X., Augusto L., Goll D.S., Ringeval B., Wang Y., Helfenstein J., Huang Y., Yu K., Wang Z., Yang Y. and Hou E.: Global
475     patterns and drivers of soil total phosphorus concentration, doi: 10.6084/m9.figshare.14583375, 2021.

Helfenstein, J., Tamburini, F., von Sperber, C., Massey, M.S., Pistocchi, C., Chadwick, O.A., Vitousek, P.M., Kretzschmar, R. and Frossard, E.: Combining spectroscopic and isotopic techniques gives a dynamic view of phosphorus cycling in soil. Nat. Commun., 9, doi: 10.1038/s41467-018-05731-2, 2018.

480    Hengl, T., Leenaars, J.G.B., Shepherd, K.D., Walsh, M.G., Heuvelink, G.B.M., Mamo, T., Tilahun, H., Berkhout, E., Cooper, M., Fegraus, E., Wheeler, I. and Kwabena, N.A.: Soil nutrient maps of Sub-Saharan Africa: assessment of soil nutrient content at 250 m spatial resolution using machine learning. Nutr. Cycl. Agroecosys., 109, 77-102, doi: 10.1007/s10705-017-9870-x, 2017.

Hengl, T., Mendes, D.J.J., Heuvelink, G.B., Ruiperez, G.M., Kilibarda, M., Blagotic, A., Shangguan, W., Wright, M.N., Geng, X., Bauer-Marschallinger, B., Guevara, M.A., Vargas, R., MacMillan, R.A., Batjes, N.H., Leenaars, J.G., Ribeiro, E., 485    Wheeler, I., Mantel, S. and Kempen, B.: SoilGrids250m: Global gridded soil information based on machine learning. PLoS One, 12, e0169748, doi: 10.1371/journal.pone.0169748, 2017.

Hou, E., Chen, C., Luo, Y., Zhou, G., Kuang, Y., Zhang, Y., Heenan, M., Lu, X. and Wen, D.: Effects of climate on soil phosphorus cycle and availability in natural terrestrial ecosystems. Global Change Biol., 24, 3344-3356, doi: 10.1111/gcb.14093, 2018a.

490    Hou, E., Luo, Y., Kuang, Y., Chen, C., Lu, X., Jiang, L., Luo, X. and Wen, D.: Global meta-analysis shows pervasive phosphorus limitation of aboveground plant production in natural terrestrial ecosystems. Nat. Commun., 11, 637, doi: 10.1038/s41467-020-14492-w, 2020.

Hou, E., Tan, X., Heenan, M. and Wen, D.: A global dataset of plant available and unavailable phosphorus in natural soils derived by Hedley method. Scientific Data, 5, doi: 10.1038/sdata.2018.166, 2018b.

495    Hou, E., Wen, D., Jiang, L., Luo, X., Kuang, Y., Lu, X., Chen, C., Allen, K.T., He, X., Huang, X. and Luo, Y.: Latitudinal patterns of terrestrial phosphorus limitation over the globe. Ecol. Lett., 00, 1-12, doi: 10.1111/ele.13761, 2021.

Huston, M.A.: Precipitation, soils, NPP, and biodiversity: resurrection of Albrecht's curve. Ecol. Monogr., 82, 277-296, doi: 10.1890/11-1927.1, 2012.

Huston, M.A. and Wolverton, S.: The global distribution of net primary production: resolving the paradox. Ecol. Monogr., 79, 500    343-377, doi: 10.1890/08-0588.1, 2009.

Jenny, H., Factors of soil formation; a system of quantitative pedology. McGraw-Hill, New York,1941.

Kaňa, J. and Kopáček, J.: Impact of Soil Sorption Characteristics and Bedrock Composition on Phosphorus Concentrations in two Bohemian Forest Lakes. Water, Air, and Soil Pollution, 173, 243-259, doi: 10.1007/s11270-005-9065-y, 2006.

Kitayama, K., Majalap-Lee, N. and Aiba, S.: Soil Phosphorus Fractionation and Phosphorus-Use Efficiencies of Tropical 505    Rainforests along Altitudinal Gradients of Mount Kinabalu, Borneo. Oecologia, 123, 342-349, doi: 10.1007/s004420051020, 2000.

Kuhn, M.: caret: Classification and Regression Training. R package version 6.0-86., https://github.com/topepo/caret/, 2020.

Lang, F., Bauhus, J., Frossard, E., George, E., Kaiser, K., Kaupenjohann, M., Krüger, J., Matzner, E., Polle, A., Prietzel, J., Rennenberg, H. and Wellbrock, N.: Phosphorus in forest ecosystems: New insights from an ecosystem nutrition 510    perspective. J. Plant Nutr. Soil Sc., 179, 129-135, doi: 10.1002/jpln.201500541, 2016.

Li, P., Yang, Y., Han, W. and Fang, J.: Global patterns of soil microbial nitrogen and phosphorus stoichiometry in forest ecosystems. Global Ecol. Biogeogr., 23, 979-987, doi: 10.1111/geb.12190, 2014.

Li, X., Li, Y., Peng, S., Chen, Y. and Cao, Y.: Changes in soil phosphorus and its influencing factors following afforestation in Northern China. Land Degrad. Dev., 30, 1655-1666, doi: 10.1002/ldr.3345, 2019.

515    Liaw, A. and Wiener, M.: Classification and Regression by randomForest. R News, 2, 18-22,2002.

Mage, S.M. and Porder, S.: Parent Material and Topography Determine Soil Phosphorus Status in the Luquillo Mountains of Puerto Rico. Ecosystems, 16, 284-294, doi: 10.1007/s10021-012-9612-5, 2013.

McGroddy, M.E.: LBA-ECO TG-07 Forest Soil P, C, and N Pools, km 83 Site, Tapajos National Forest. Data set. Available on-line [http://daac.ornl.gov] from Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, 520    Tennessee, U.S.A .http://dx.doi.org/10.3334/ORNLDAAC/1085.,2012b.

Mehmood, A., Akhtar, M.S., Imran, M. and Rukh, S.: Soil apatite loss rate across different parent materials. Geoderma, 310, 218-229, doi: 10.1016/j.geoderma.2017.09.036, 2018.

Meinshausen, N.: Quantile Regression Forests. J. Mach. Learn. Res., 7, 983-999, doi: 10.5555/1248547.1248582, 2006.

Meinshausen, N.: quantregForest: Quantile Regression Forests, https://CRAN.R-project.org/package=quantregForest, 2017.

525    Minasny, B.M., Alex, B.M.M., Malone, P., B., B, A. and Budiman, Using R for Digital Soil Mapping. Springer, Cham, doi:

10.1007/978-3-319-44327-0, 2017.

Ploton, P., Mortier, F., Réjou-Méchain, M., Barbier, N., Picard, N., Rossi, V., Dormann, C., Cornu, G., Viennois, G., Bayol, N., Lyapustin, A., Gourlet-Fleury, S. and Pélissier, R.: Spatial validation reveals poor predictive performance of large-scale ecological mapping models. Nat. Commun., 11, 4540, doi: 10.1038/s41467-020-18321-y, 2020.

530    Porder, S. and Chadwick, O.A.: Climate and Soil-Age Constraints on Nutrient Uplift and Retention by Plants. Ecology, 90, 623-636, doi: 10.1890/07-1739.1, 2009.

Porder, S. and Ramachandran, S.: The phosphorus concentration of common rocks—a potential driver of ecosystem P status. Plant Soil, 367, 41-55, doi: 10.1007/s11104-012-1490-2, 2013.

R Core Team: R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna,
535    Austria, https://www.R-project.org/, 2018.

Reed, S.C., Yang, X. and Thornton, P.E.: Incorporating phosphorus cycling into global modeling efforts: a worthwhile, tractable endeavor. New Phytologist, 208, 324-329, doi: 10.1111/nph.13521, 2015.

Reich, P.B. and Oleksyn, J.: Global patterns of plant leaf N and P in relation to temperature and latitude. Proceedings of the National Academy of Sciences, 101, 11001-11006, doi: 10.1073/pnas.0403588101, 2004.

540    Ringeval, B., Augusto, L., Monod, H., van Apeldoorn, D., Bouwman, L., Yang, X., Achat, D.L., Chini, L.P., Van Oost, K., Guenet, B., Wang, R., Decharme, B., Nesme, T. and Pellerin, S.: Phosphorus in agricultural soils: drivers of its distribution at the global scale. Global Change Biol., 23, 3418-3432, doi: 10.1111/gcb.13618, 2017.

Rodionov, A., Bauke, S.L., von Sperber, C., Hoeschen, C., Kandeler, E., Kruse, J., Lewandowski, H., Marhan, S., Mueller, C.W., Simon, M., Tamburini, F., Uhlig, D., von Blanckenburg, F., Lang, F. and Amelung, W.: Biogeochemical cycling of
545    phosphorus in subsoils of temperate forest ecosystems. Biogeochemistry, 150, 313-328, doi: 10.1007/s10533-020-00700-8, 2020.

Shangguan, W., Dai, Y., Duan, Q., Liu, B. and Yuan, H.: A global soil data set for earth system modeling. J. Adv. Model. Earth Sy., 6, 249-263, doi: 10.1002/2013MS000293, 2014.

Shangguan, W., Hengl, T., Mendes De Jesus, J., Yuan, H. and Dai, Y.: Mapping the global depth to bedrock for land surface
550    modeling. J. Adv. Model. Earth Sy., 9, 65-88, doi: 10.1002/2016MS000686, 2017.

Siqueira, R.G., Schaefer, C.E.G.R., Fernandes Filho, E.I., Corrêa, G.R., Francelino, M.R., Souza, J.J.L.L. and Rocha, P.D.A.: Weathering and pedogenesis of sediments and basaltic rocks on Vega Island, Antarctic Peninsula. Geoderma, 382, 114707, doi: 10.1016/j.geoderma.2020.114707, 2021.

Smeck, N.E.: Phosphorus dynamics in soils and landscapes. Geoderma, 36, 185-199, doi: 10.1016/0016-7061(85)90001-1,
555    1985.

Spohn, M.: Increasing the organic carbon stocks in mineral soils sequesters large amounts of phosphorus. Global Change Biol., 26, 4169-4177, doi: 10.1111/gcb.15154, 2020.

Steffen, W., Richardson, K., Rockstrom, J., Cornell, S.E., Fetzer, I, Bennett, E.M., Biggs, R., Carpenter, S.R., de Vries, W., de Wit, C.A., Folke, C., Gerten, D., Heinke, J., Mace, G.M., Persson, L.M., Ramanathan, V., Reyers, B. and Sorlin, S.:
560    Planetary boundaries: Guiding human development on a changing planet. Science, 347, 1259855, doi: 10.1126/science.1259855, 2015.

Strobl, C., Boulesteix, A., Kneib, T., Augustin, T. and Zeileis, A.: Conditional variable importance for random forests. BMC Bioinformatics, 9, 307, doi: 10.1186/1471-2105-9-307, 2008.

Sun, Y., Peng, S., Goll, D.S., Ciais, P., Guenet, B., Guimberteau, M., Hinsinger, P., Janssens, I.A., Peñuelas, J., Piao, S.,
565    Poulter, B., Violette, A., Yang, X., Yin, Y. and Zeng, H.: Diagnosing phosphorus limitations in natural terrestrial ecosystems in carbon cycle models. Earth's Future, 5, 730-749, doi: 10.1002/2016EF000472, 2017.

Tipping, E., Somerville, C.J. and Luster, J.: The C:N:P:S stoichiometry of soil organic matter. Biogeochemistry, 130, 117-131, doi: 10.1007/s10533-016-0247-z, 2016.

Turner, B.L. and Engelbrecht, B.M.J.: Soil organic phosphorus in lowland tropical rain forests. Biogeochemistry, 103, 297-
570    315, doi: 10.1007/s10533-010-9466-x, 2011.

Viscarra Rossel, R.A. and Bui, E.N.: A new detailed map of total phosphorus stocks in Australian soil. Sci. Total Environ., 542, 1040-1049, doi: 10.1016/j.scitotenv.2015.09.119, 2016.

Vitousek, P.M. and Chadwick, O.A.: Pedogenic Thresholds and Soil Process Domains in Basalt-Derived Soils. Ecosystems, 16, 1379-1395, doi: 10.1007/s10021-013-9690-z, 2013.

575 Vitousek, P.M., Porder, S., Houlton, B.Z. and Chadwick, O.A.: Terrestrial phosphorus limitation: mechanisms, implications, and nitrogen–phosphorus interactions. Ecol. Appl., 20, 5-15, doi: 10.1890/08-0127.1, 2010.

Walker, T.W. and Syers, J.K.: The fate of phosphorus during pedogenesis. Geoderma, 15, 1-19, doi: 10.1016/0016-7061(76)90066-5, 1976.

Wang, Y., Zhang, Q., Pitman, A.J. and Dai, Y.: Nitrogen and phosphorous limitation reduces the effects of land use change
580 on land carbon uptake or emission. Environ. Res. Lett., 10, 14001, doi: 10.1088/1748-9326/10/1/014001, 2015.

Wang, Y., Zhang, X. and Huang, C.: Spatial variability of soil total nitrogen and soil total phosphorus under different land uses in a small watershed on the Loess Plateau, China. Geoderma, 150, 141-149, doi: 10.1016/j.geoderma.2009.01.021, 2009.

Wang, Z., Tian, H., Yang, J., Shi, H., Pan, S., Yao, Y., Banger, K. and Yang, Q.: Coupling of Phosphorus Processes With
585 Carbon and Nitrogen Cycles in the Dynamic Land Ecosystem Model: Model Structure, Parameterization, and Evaluation in Tropical Forests. J. Adv. Model. Earth Sy., 12, e2020MS002123, doi: 10.1029/2020MS002123, 2020.

Wang, Z., Wang, M., Yu, K., Hu, H., Yang, Y., Ciais, P., Ballantyne, A.P., Niklas, K.J., Huang, H., Yao, B. and Wright, S.J.: Global synthesis for the scaling of soil microbial nitrogen to phosphorus in terrestrial ecosystems. Environ. Res. Lett., 16, 044034, doi: 10.1088/1748-9326/abed78, 2021.

590 Wardle, D.A., Walker, L.R., Bardgett, R.D. and Sveriges, L.: Ecosystem Properties and Forest Decline in Contrasting Long-Term Chronosequences. Science, 305, 509-513, doi: 10.1126/science.1098778, 2004.

Wassen, M.J., Schrader, J., van Dijk, J. and Eppinga, M.B.: Phosphorus fertilization is eradicating the niche of northern Eurasia's threatened plant species. Nature Ecology & Evolution, 5, 67-73, doi: 10.1038/s41559-020-01323-w, 2021.

Wieder, W.R., Cleveland, C.C., Smith, W.K. and Todd-Brown, K.: Future productivity and carbon storage limited by terrestrial
595 nutrient availability. Nat. Geosci., 8, 441-444, doi: 10.1038/ngeo2413, 2015.

Xu, X., Thornton, P.E. and Post, W.M.: A global analysis of soil microbial biomass carbon, nitrogen and phosphorus in terrestrial ecosystems. Global Ecol. Biogeogr., 22, 737-749, doi: 10.1111/geb.12029, 2013.

Yan, T., Zhu, J. and Yang, K.: Leaf nitrogen and phosphorus resorption of woody species in response to climatic conditions and soil nutrients: a meta-analysis. Journal of forestry research, 29, 905-913, doi: 10.1007/s11676-017-0519-z, 2018.

600 Yanai, R.D.: The effect of whole-tree harvest on phosphorus cycling in a northern hardwood forest. Forest Ecol. Manag., 104, 281-295, doi: 10.1016/S0378-1127(97)00256-9, 1998.

Yang, X., Post, W.M., Thornton, P.E. and Jain, A.: The distribution of soil phosphorus for global biogeochemical modeling. Biogeosciences, 10, 2525-2537, doi: 10.5194/bg-10-2525-2013, 2013.

Yang, X., Thornton, P.E., Ricciuto, D.M. and Post, W.M.: The role of phosphorus dynamics in tropical forests - a modeling
605 study using CLM-CNP. Biogeosciences, 11, 1667-1681, doi: 10.5194/bg-11-1667-2014, 2014.

Zhang, C., Tian, H., Liu, J., Wang, S., Liu, M., Pan, S. and Shi, X.: Pools and distributions of soil phosphorus in China. Global Biogeochem. Cy., 19, GB1020, doi: 10.1029/2004GB002296, 2005.

Zhang, Q., Wang, Y.P., Pitman, A.J. and Dai, Y.J.: Limitations of nitrogen and phosphorous on the terrestrial carbon uptake in the 20th century. Geophys. Res. Lett., 38, L22701, doi: 10.1029/2011GL049244, 2011.

610 Zhang, Y., Guo, Y., Tang, Z., Feng, Y., Zhu, X., Xu, W., Bai, Y., Zhou, G., Xie, Z. and Fang, J.: Patterns of nitrogen and phosphorus pools in terrestrial ecosystems in China [preprint]. Earth Syst. Sci. Data Discuss, doi: 10.5194/essd-2020-398, 2021.