Point-by-point Responses to Reviewers

Note: text in black are the comments, and text in deep blue are our responses.

We appreciate reviewers' constructive comments on our manuscript. We carefully considered each comment and revised the manuscript accordingly.

Responses to Reviewer 1

Comment 1: The manuscript is very well written and presents very interesting and extremely useful data not only for a global soil science community but also for any related field interested in biogeochemical fluxes and pools. As the background, aim, methods and results are clearly and with high quality standards presented, I have only minor comments. One important request from my side would be, that the readme file of the presented data would be extended to make the data tables self-explainable. E.g., looking at the rf.dat.csv file, units are not clear and BIOMES, BEDROCK, SOIL TYPE is not clear what the categorical numbers given represent. Same holds true for the covstack.dat.csv and the raw.data.csv files. The readme.txt does not explain units or legends (e.g. what is a "1" for BIOME or BEDROCK). One option would be to include Table 1 somewhere in the readme file and add information needed on parent materials, vegetation types, bedrock and soil orders.

Response 1: We really appreciate your positive comments. As suggested, the revised readme file (a word file) has included Table 1 in our main text, which includes units of each numeric variables. And we have added three more tables in the readme file to explain the meaning of numbers in each categorical variable. We have uploaded the updated readme file to 10.6084/m9.figshare.14583375.

Comment 2: Another general comment would be that it is not clear how you excluded agricultural land and on which basis you separated between natural – semi natural and extensively used (grass-)land. It is a great idea, to not mask out cropland or other heavily influenced areas in your map and data tables but instead present the predicted "background" or "natural" values. However, I think this also needs to be described in the readme file very briefly to avoid misunderstanding and misuse of data. May be this should also be stated in the abstract, to make very clear, that you give "potential natural background" values of P in these areas.

Response 2: Many thanks for pointing this out. In the previous version, we have described how we excluded data from agricultural and other heavily influenced areas only in the method. In the revision, we have described our criteria to collect data in the introduction to make the criteria clear to readers when reading the introduction (Lines 78-79 in the revised manuscript). To avoid a misuse, in the Fig. 5 of main text, we have marked cropland areas in our predicted maps as a reminder that our predicted values in these grid-cells indicates "potential natural background" values. This has also been described in the introduction (Lines 80-82) and new readme file (Note 1).

Abstract

Comment 3: Very well written indeed!

Response 3: Thanks a lot.

Comment 4: 23: what do you mean "predictions increased"? The predicted amount increased? Or reliability of predictions increased?

Response 4: We meant the predicted concentration increased with latitude. In the revision, we have rewritten the sentence as "predicted soil total P concentration increased significantly with latitude" to avoid this misunderstanding (Line 24).

Introduction

Comment 5: Also very well written and interesting to read.

Response 5: Thanks a lot.

Comment 6: 78: do you mean to say that you explicitly exclude agricultural used soils? This is not clear (was only clear after reading the methods). And if so, on which basis did you do this separation between natural – semi-natural (extensively used grassland?) and agricultural (arable and intensive grassland?).

Response 6: Yes, we explicitly excluded agricultural used soils from our database. We collected data only from (semi-)natural ecosystems, based on sampling description in the original databases and literature. We defined (semi-)natural ecosystems as ecosystems without any documented significant anthropogenic activities such as tillage, fertilization, and heavy grazing. Forests with a stand age greater than 10 years were considered as (semi-)natural ecosystems. We have added these descriptions to the revised introduction (Lines 78-79).

Comment 7: 83: global total P stock including "background values" of agricultural soils? So this would be some kind value of potential P content with no human influence?

Response 7: Yes, we meant "potential background values" without direct human disturbance. We have described this clearly in the revised introduction (Lines 80-82).

Methods

Comment 8: 93-94: which efforts? What are the criteria?

Response 8: We have added one sentence to describe how we excluded soils from anthropogenic disturbance (Lines 94-95). **"**We carefully checked description of soil sampling in every cited paper to ascertain if the soil sample experienced any documented significant anthropogenic activities such as tillage, fertilization, and heavy grazing." Forests with a stand age greater than 10 years were considered as (semi-)natural ecosystems and included in our database.

Comment 9: 97 why web of science and not google scholar? Web of science often seems exclusive of some journals or data sets which are still peer reviewed.

Response 9: Data in papers included in the databases of Web of Science have been generally strictly peer reviewed, and thus should be generally reliable. Some data or journals that are not included in Web of Science but included in Google Scholar may be still peer reviewed, but could be difficult for us to distinguish them from those not peer reviewed, and may be difficult for us to judge the quality of the data (e.g., while without clear land-use history description). Meanwhile, our aim is to collect a spatially representative database, not a comprehensive one (which is very unlikely to be achieved, given the huge amount of data in literature). There is a trade-off between selectivity (only relevant references are selected) and sensitivity (all the selected references are relevant) here. Google Scholar is excellent for sensitivity, but very inefficient for selectivity. Therefore, we have used the Web of Science to search and collect the data.

Comment 10: Figure 1 and lines 140-141: would it be possible to also give WRB soil types? This might increase understanding (and citation) in the whole of Europe soil science community.

Response 10: Many thanks for this advice. We have added WRB soil type as a variable in the updated rf.dat data frame. Please check the updated rf.dat data frame stored in Figshare (https://figshare.com/s/cbd0840d9d32da1a7f81). We also summarized the soil total P concentration in relation to different WRB soil types and added these results in a table in the supplementary file (Table S6). We extracted WRB soil type of each site from a global WRB soil type map (Hengl et al., 2017) based on the geographical coordinates. And we mentioned the method to extract WRB soil type of each site (Lines 148-150) and the result in main text (Lines 227-229).

Reference:

Hengl, T., et al.: SoilGrids250m: Global gridded soil information based on machine learning. PLoS One, 12, e0169748, 2017.

Table S6 Soil total P concentration (mg kg^{-1}) in WRB soil types at 0-100 cm depth. Results based on our database. P10, P25, P75, and P 90 indicate the percentile rank of 10%, 25%, 75%, and 90%. Only WRB soil types with more than 10 observations are shown here.

WRB soil type	Count	Min.	P10	P25	Median	Mean	P75	P90	Max.
Cryosols	59	36.3	282.6	646.5	1078.0	1152.2	1525.0	1968.4	3470.0
Phaeozems	153	43.1	96.3	205.1	641.0	1122.6	987.7	3192.0	9630.0
Leptosols	170	35.0	253.1	390.6	592.9	1114.1	957.3	3155.0	9020.0
Nitisols	24	99.1	398.8	578.3	742.8	766.5	1012.9	1164.7	1367.0
Andosols	258	11.0	111.7	240.9	578.1	656.5	881.6	1362.7	2850.0
Albeluvisols	69	124.8	264.4	401.1	585.1	655.3	808.7	1039.8	2374.8
Cambisols	1010	9.8	171.8	358.5	581.5	650.7	805.0	1110.3	4433.0
Vertisols	56	14.1	175.0	247.5	415.5	634.0	723.8	1259.0	2900.0
Histosols	21	90.6	167.2	184.4	305.3	631.7	1370.7	1450.6	1505.2
Calcisols	107	17.7	88.9	232.7	450.0	630.4	658.5	1317.4	4243.0
Luvisols	534	3.3	99.9	239.1	489.4	585.6	799.5	1187.0	4800.0
Alisols	223	34.0	190.0	319.1	476.0	578.5	665.5	1010.0	3680.0
Chernozems	122	37.0	107.2	238.2	470.9	573.0	654.4	1327.4	3480.0
Podzols	185	14.5	104.3	203.9	327.5	546.7	732.0	1160.5	3444.2
Kastanozems	341	20.3	38.7	222.1	411.9	513.0	604.0	760.0	5520.0
Fluvisols	82	83.5	154.4	245.0	331.5	477.8	516.3	785.0	3320.0
Solonchaks	50	16.7	254.8	329.5	518.0	466.7	641.7	674.6	685.7
Acrisols	916	3.0	105.8	200.0	364.4	443.6	575.4	856.9	3898.0
Gypsisols	34	63.0	176.6	289.3	410.9	384.2	472.3	573.3	664.1
Gleysols	46	58.6	72.4	99.7	147.5	373.6	400.3	900.0	3200.0
Ferralsols	257	16.0	86.9	148.0	254.4	307.7	380.0	537.3	1997.0
Regosols	23	100.0	116.0	155.0	250.0	277.7	315.0	356.0	820.8
Lixisols	106	1.4	11.0	21.3	106.7	258.7	326.2	653.2	3090.0
Arenosols	167	24.3	31.8	38.3	52.9	171.1	267.8	543.5	1355.0

Comment 11: 179-180 this is a great idea, to not mask out cropland or other heavily influenced areas. However, I think this also needs to be described in the readme file very briefly to avoid misunderstanding and misuse of data. May be this should also be stated in the abstract, to make

very clear, that you give “natural background” values of P in these areas.

Response 11: We appreciate this advice. As described in Response 2, to avoid a misuse, we have marked cropland areas in our predicted maps (Fig. 5 of main text) as a reminder that our predicted values in these grid-cells indicates “potential natural background” values. This has also been described in the introduction (Lines 80-82) and new readme file (Note 1).

Results

Comment 12: I think it would be very interesting, if you would extend Tables 2 and 3 to the 0-30 cm layer, so give 0-30 and 0-1m, separately.

Response 12: Many thanks for this advice. In the revision, we have extended Table 2 and 3 to give 0-30 cm and 0-100 cm separately (Table 2 and 3 in revision).

Comment 13: 214 – 216 this is surprising as we would expect strongly weathered soils to be significantly lower than intermediate weathered soils, and intermediate soils higher than young, low weathered soils. Any explanation?

Response 13: To our knowledge, as Walker and Syers (1976) conceptual model predicts (Figure 1): soil total P concentration decreases during soil development. At the beginning of soil development, all soil P is in the primary mineral form. With time, P in primary minerals decrease, while organic P and occluded P accumulate, total P declines due to losses in dissolved or particulate form. The decreasing trend of soil total P with soil development (or weathered extent) is well supported by our results and previous studies (i.e., Cross and Schlesinger, 1995; Yang and Post, 2011; Yang et al., 2013) with uplift modulating the evolution of total soil P (Buendía et al., 2011). We didn’t discuss the relationship between soil total P and soil order, because a similar pattern has been discussed in these studies, and soil order is less important than four other predictors we discussed in the manuscript, i.e., parental material, SOC, soil sand content and MAT.

Indeed, as Walker and Syers model predicted, soil available P may peak at intermediate weathering stage, but not total P.

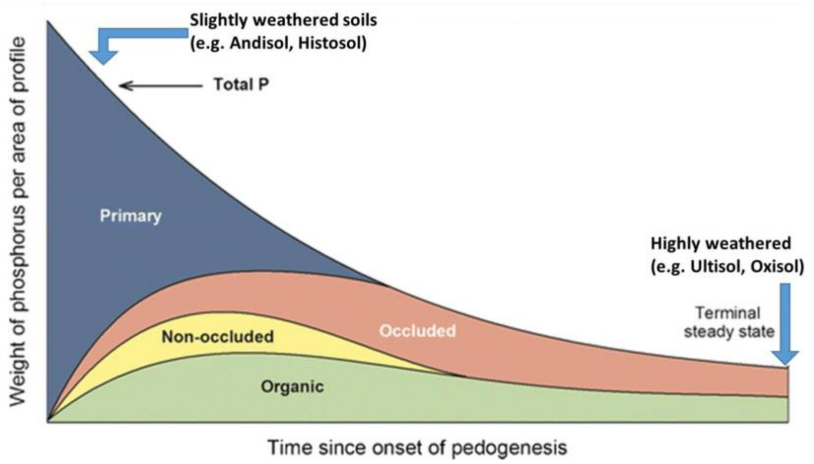

Figure 1. The Walker and Syers (1976) conceptual model of phosphorus dynamics during long-term ecosystem development. Figure adapted from Walker and Syers (1976), Cross and Schlesinger (1995), and Yang and Post (2011).

References:

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, https://doi.org/10.5194/bg-7-2025-2010, 2010.

Cross, A.F. and Schlesinger, W.H.: A literature review and evaluation of the. Hedley fractionation: Applications to the biogeochemical cycle of soil phosphorus in natural ecosystems. Geoderma, 64, 197-214, https://doi.org/10.1016/0016-7061(94)00023-4, 1995.

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

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

Yang, X. and Post, W.M.: Phosphorus transformations as a function of pedogenesis: A synthesis of soil phosphorus data using Hedley fractionation method. Biogeosciences, 8, 2907-2916, https://doi.org/10.5194/bg-8-2907-2011, 2011.

Comment 14: 252 increase only from equator to high northern latitudes? Any explanation why this increase is seen?

Response 14: We found soil total P concentration increased with decreasing latitude in both northern and southern hemispheres (Fig. 5 and Fig S3K). Explanation for this result has been added to the revised discussion (Lines 341-343), as follows: "Lowland tropical soils are relatively more weathered compared to soils at high latitudes due to warmer climate (Hou et al., 2018). Moreover, the last glaciation could have eroded soils at northern higher latitude and have caused relatively young and P-enriched soils (Vitousek et al., 2010; Reich and Oleksyn, 2004)".

References:

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, 2018.

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.

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.

Comment 15: 257 African highlands do not show this?

Response 15: Yes, African highlands also show higher soil total P concentration than low African lands in our predicted maps (Fig. 5). We stated African highlands as east Africa in the previous version. To avoid this misunderstanding, we have rewritten east Africa to African highlands in the revision (Line 272).

Discussion

Comment 16: Generally, a very clear and good discussion. Only the above-mentioned points, why younger soils are clearly higher in P than intermediate soils and intermediates soils not higher than strongly weathered (old) soils are missing. This kind of contradicts our text book knowledge of young soils being low in P, then slowly accumulating P to a certain peak (intermediate weathered soils) and finally loosing soil again.

Response 16: Please see our Response 13. In revision, we have added two sentences to discuss this pattern shortly (Lines 223-227). The declining trend of soil total P during soil development supports the Walker and Syers (1976) conceptual model of phosphorus dynamics during long-term ecosystem development. And this pattern is consistent with previous studies (i.e., Cross and Schlesinger, 1995; Yang et al., 2011; Yang et al., 2013).

Comment 17: And may be you should briefly discuss high southern latitudes (no glaciations, so strongly weathered?)

Response 17: Many thanks for this advice. In the revision, we have re-written the discussion of this issue (Lines 341-343). Also see Response 14 above.

Conclusions:

Comment 18: 361: this is not totally correct, as we would have expected highest P contents in medium aged, intermediate weathered soils (e.g. peak of weathering, not so much lost yet).

Response 18: Please check our response 13 above.

Responses to Reviewer 2

Comment 19: Overall this is a well-written paper describing a dataset that may be useful for many purposes. While it is clear that the data were derived from semi-natural sites, the information may be relevant to agricultural issues as well, provided there is clarity on what is represented by the global totals presented. I offer the following comments and suggestions for improvement of the clarity and interpretation of the paper and dataset.

Response 19: We appreciate these positive comments. Below we respond to your comments point-by-point.

Comment 20: Line 15 - The term “reserves” has a specific definition applying to the mining of geological phosphate rock resources. I would suggest the word “stocks” be substituted for the word “reserves” as is done in line 26.

Response 20: We appreciate that you point this out. In the revised manuscript, we have replaced “reserves” with “stocks” in lines 15 in the revised manuscript.

Comment 21: Line 15 & 32 - Apparently, P in geological formations and rock below the land surface are not considered. The reference Zhang et al. 2021 in line 32 pertains only to China, and compares soil P only to that in leaves, woody stems, roots, and litter in forest, shrubland and grassland ecosystems. It does not assess geological P. Suggested alternate wording: “In terrestrial ecosystems to a depth of one meter from the land surface, most of the P is found in the soil.”

Response 21: Really appreciate pointing this out. In the revision, we have revised "on land" to "in terrestrial ecosystems" (Lines 15). Ecosystem only includes land surface interacting with living organisms, while excluding deep geological formations and rock. As for the first sentence in the introduction, we have re-written it as you suggested, i.e., "In terrestrial ecosystems to a depth of one meter from the land surface, most of the P is found in the soil." (Line 32).

Comment 22: Line 26 (and lines 243-244) - Clarity needs to be provided for these figures on global soil P stocks. Do they include agricultural lands that were not included in the study? If so, do they represent and estimate of the soil total P content prior to land conversion? What percentage of the total would be comprised of such lands? From global fertilizer consumption figures, it can be estimated that ~0.9 Pg of mined P has been added to soils since the early days of fertilizer manufacturing around, and similar figures for the removal of P in crop harvests could be calculated. Such figures would be useful for those concerned about the depletion of soil P relative to the global reserves of phosphate rock (estimated at ~7 Pg P in USGS Mineral Surveys).

Response 22: Many thanks for pointing this out. Yes, we didn't mask out the cropland or other heavy influenced areas from the predicted maps. And our predicted maps and estimates represent "potential natural background values", i.e., background soil P without direct human influence. To avoid misunderstanding and misuse, in the revision, we have marked cropland areas in our predicted maps (Fig. 5 of main text) as a reminder that our predicted values in these grid-cells indicates "potential natural background" values. This has also been described in the introduction (Lines 80-82) and new readme file (Note 1).

According to an empirical model, about 20% fertilizer P is fixed in soil as stable soil P (Sattari et al., 2012). As a simple estimation, the soil P stock estimation gap between our prediction and current state in global cropland areas could be 0.9 Pg * 20% = 0.18 Pg. It is an interesting and great idea to estimate the cropland soil P depletion relative to global phosphate rock reserves, in view of the increasing food demand accelerated flow of phosphate rock from Earth's crust for fertilizer (Cordell and White, 2014). And our predicted maps could be used as an improved natural background map in future studies.

References:

Cordell, D. and White, S.: Life's Bottleneck: Sustaining the World's Phosphorus for a Food Secure Future. Annu. Rev. Env. Resour., 39, 161-188, https://doi.org/10.1146/annurev-environ-010213-113300, 2014.

Sattari, S. Z., Bouwman, A. F., Giller, K. E. & van Ittersum, M. K. (2012) Residual soil phosphorus as the missing piece in the global phosphorus crisis puzzle. Proceedings of the National Academy of Sciences, 109, 6348-6353.

Comment 23: Line 34 - Regarding "soil P form depends on the amount or total concentration of P in soils" — This dependence does not seem clear or intuitive, and is not clearly supported by the three references. In fact Turner and Engelbrecht conclude "We conclude that soil properties exert a strong control on the amounts and forms of soil organic phosphorus in tropical rain forests, but that the proportion of the total phosphorus in organic forms is relatively insensitive to variation in climate and soil properties."

Response 23: Many thanks for pointing this out. In the revision, we have re-written this sentence as: "Moreover, the amount or total concentration of P in soils determines P concentration in all major forms in soils (Hou et al., 2018a; Turner and Engelbrecht, 2011).” (Lines 35-36).

Comment 24: Lines 45-48 - the importance of bioavailable soil P relative to total soil P to ecosystem functioning should be discussed.

Response 24: Many thanks for this advice. This manuscript didn’t include data or results on soil bioavailable P, so we didn’t discuss the importance of bioavailable soil P relative to total soil P either in introduction or in discussion. But yes, this is a very important issue. We have complied a database of Hedley P fractions concentration in global natural soil, which including more than 1800 site-level measurements. Now we are drafting a manuscript, in which we will discuss the importance and drivers of soil bioavailable P relative to total P across terrestrial ecosystems.

Comment 25: Line 63 - the two references, Carpenter and Bennett 2011 and Steffen et al 2015 do not support the reliance on accurate soil P maps.

Response 25: Many thanks. We have replaced these two references with four other references to support the reliance on accurate soil P maps, i.e., (Alewell et al., 2020; Ringeval et al., 2017; Beusen et al., 2015; Wang et al., 2010) (Lines 64-65).

Reference:

Alewell, C., Ringeval, B., Ballabio, C., Robinson, D.A., Panagos, P. and Borrelli, P.: Global phosphorus shortage will be aggravated by soil erosion. Nat. Commun., 11, https://doi.or/10.1038/s41467-020-18326-7, 2020.

Beusen, A.H.W., Van Beek, L.P.H., Bouwman, A.F., Mogollón, J.M. and Middelburg, J.J.: Coupling global models for hydrology and nutrient loading to simulate nitrogen and phosphorus retention in surface water – description of IMAGE – GNM and analysis of performance. Geosci. Model Dev., 8, 4045-4067, https://doi.org/10.5194/gmd-8-4045-2015, 2015.

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, https://doi.or/10.1111/gcb.13618, 2017.

Wang, Y.P., Law, R.M. and Pak, B.: A global model of carbon, nitrogen and phosphorus cycles for the terrestrial biosphere. Biogeosciences, 7, 2261-2282, https://doi.or/10.5194/bg-7-2261-2010, 2010.

Comment 26: Line 92 - “heavily” should be “heavy”

Response 26: Done (line 93).

Comment 27: Line 99 - “potentially useful” should be better defined. What factors decided the exclusion of 77% of the papers?

Response 27: In the revision, we have described in more details how we screened the literature in this step (Lines 102-103). As we want to search existing global or regional databases that may include soil total P concentration measurements in (semi-)natural ecosystems, by looking at title and abstract, we excluded studies at site level (or local scale) or with artificial treatment (e.g., fertilizer treatment, elevated temperature, or elevated CO_2 etc.). And 163 papers were removed at this step.

Comment 28: Line 121-124 - For the under-represented regions, no exclusion of agricultural land was described. Are these regions then more likely to include land that is not semi-natural?
Response 28: For the under-represented regions, we have also excluded agricultural soils or other human influenced soils at this step. We have stated (Lines 89-93) we collected soil total P concentration measurements in (semi-)natural terrestrial ecosystems. To avoid any misunderstanding, in revision, we have clarified this again here (Lines 126-127).

Comment 29: Line 127-128 - What were the data sources for climate, vegetation, etc?
Response 29: We collected climate variables (i.e., MAT and MAP), vegetation type, soil physiochemical properties (e.g., SOC, soil clay and sand contents, soil pH) of corresponding soil total P measures from source paper when they are reported. In revision, we have clarified this (Line 133). And "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." was stated later (Lines 158-159).

Comment 30: Line 179-180 - need to state the assumption that cropland in its native state had the same set of relationships as for semi-natural land. The possibility that land with different total P levels, or different total P relationships, should be discussed explicitly.
Response 30: We really appreciate this suggestion. We have stated this assumption in the revised manuscript "Here we assume that cropland and other heavy influenced areas in its native state had the same set of relationships as for (semi-)natural land." (Lines 188-189). And yes, there is a possibility that in these underrepresented areas different soil P levels and different total P relationships could exist. So, we have pointed this out as a data limitation in our revised manuscript (Lines 363-365), which is "First, there are some regions were still underrepresented, e.g., northern Canada, Russia, middle Asia, and inner Australia, which may result in a low accuracy of the predicted values in these regions (Ploton et al., 2020). Further, our assumption that soils which are or have been in agricultural use can be characterized in their native state by the same relationships as semi(natural) soils might not hold true. For example, as fertile soils are preferred in agriculture and forestry.".

Comment 31: Line 201 - the reported mean and median values then represent a mix of topsoil and subsoil values. Since subsoil comprised only 15% of the samples, it has little effect, but the numbers for a single specified depth (topsoil) would be more useful and relevant.
Response 31: Many thanks for this suggestion. In revision, we have re-organized table 2 and 3 in the main text. Now these two tables show soil total P in different biomes and soil orders at 0-30 cm and 0-100 cm, respectively.

Comment 32: line 239 - in Figure 4 it is clear that many of the predictors have non-linear relationships to soil total P. Does the random forest method account for non-linearity?
Response 32: Yes, the random forest method can deal with non-linear relationships (Breiman, 2001). Here we used partial dependence plots to illustrate the relationships between one predictor and the predictions of the trained model. This is a common diagnostic to illustrate the dependence

of ‘black box’ machine learning predictions to potential drivers (Heffelfinger et al., 2020; von Fromm et al., 2021; Berkström et al., 2020). Partial dependence plots look at the variable of interest across its corresponding range in training data. At each value of the variable, the model is evaluated for all observations of the other model inputs, and the output is then averaged. Then, a partial dependence plot can show different types of relationship between predictor and prediction, such as a step function, curvilinear, linear, and so on.

Reference:

Breiman L: Random Forests. Machine Learning 2001, 45:5-32.

Berkström, C., Eggertsen, L., Goodell, W., Cordeiro, C.A.M.M., Lucena, M.B., Gustafsson, R., Bandeira, S., Jiddawi, N. and Ferreira, C.E.L.: Thresholds in seascape connectivity: the spatial arrangement of nursery habitats structure fish communities on nearby reefs. Ecography, 43, 882-896, https://doi.org/10.1111/ecog.04868, 2020.

Heffelfinger, L.J., Stewart, K.M., Shoemaker, K.T., Darby, N.W. and Bleich, V.C.: Balancing Current and Future Reproductive Investment: Variation in Resource Selection During Stages of Reproduction in a Long-Lived Herbivore. Frontiers in Ecology and Evolution, 8, https://doi.org/10.3389/fevo.2020.00163, 2020.

von Fromm, S.F., Hoyt, A.M., Lange, M., Acquah, G.E., Aynekulu, E., Berhe, A.A., Haefele, S.M., McGrath, S.P., Shepherd, K.D., Sila, A.M., Six, J., Towett, E.K., Trumbore, S.E., Vågen, T., Weullow, E., Winowiecki, L.A. and Doetterl, S.: Continental-scale controls on soil organic carbon across sub-Saharan Africa. SOIL, 7, 305-332, https://doi.org/10.5194/soil-7-305-2021, 2021.

Comment 33: Line 295 - “soil P is largely composed of organic P” is contradicted by Turner and Englebrecht 2011 who reported organic was 26% of total P for lowland tropical rain forests. Exceptions include the tundra and boreal sites included in this study. These sites were likely the drivers of the SOC-TP relationship.

Response 33: We appreciate this comment. In the revision we have removed this description. Now the first explanation why significantly positive correlation was found between soil total P and SOC across terrestrial ecosystems only emphasizing the coupling between P and C in soils. “Phosphorus couples with organic C in soil because soil P has a relatively fixed ratio to organic C (Spohn, 2020; Cleveland and Liptzin, 2007).” (Lines 310-312).

Comment 34: Line 312 - Amberger reference missing.

Response 34: It has been added in the revised manuscript (Lines 413-414).