Earth System Science Data

RE: Manuscript ID: essd-2020-398

Point-for-point responses to the comments from Reviewer 2

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

We appreciate your constructive comments on our manuscript. We carefully considered each of them and revised the manuscript accordingly. We hope that you will find the revisions satisfactory.

Zhang et al. mapped distributions of N and P pools in China terrestrial ecosystems, based on the most intensive field measurements in China ever, including all major (semi-)natural ecosystem types and ecosystem components. The study is generally well performed, and the manuscript is well written. I think the paper deserve a publication on Earth System Science Data and would be highly influential one after published. Before its publication, the authors may improve the manuscript by considering my comments and suggestions as follows.

Response: Thank you very much for reviewing our manuscript. We appreciate your helpful comments to improve this manuscript and revised it accordingly.

Major comments

• I think the authors should justify their use of artificial neural network for mapping. This method is a complex one but necessarily be the best one. Did the authors test or use other methods such as random forest?

Response: Thank you for your suggestion. According to your suggestion, we compared three different methods, the artificial neural network, support vector regression and random forest, and RF outperformed the other two. We therefore adopted random forest in this revised version for a better model performance (L178-L185).

"We used random forest to predict the nutrient densities and concentrations across China. The predictors included MAT, MAP, longitude, latitude, elevation, EVI and vegetation types (as dummy variables). We established one random forest model for N or P density in each component (in three plant organs, litter and five soil layers), respectively. In each model, six variables were randomly sampled at each split, and 500 trees were grown. Larger values of these parameters did not increase validation R^2 obviously. Model prediction were repeated for 100 times to obtain the average results..."

Furthermore, we fitted models for each soil layer, respectively, instead of the sum of all layers in the previous manuscript, and R^2 of these models were all around 0.5, much higher than the previous results. For details please see Fig 4-5 in the revised manuscript.

• Ideally, the authors may also show and discuss the relative importance of the predictors in predicting the nutrient densities. This will help readers to build a more mechanistic view of the patterns. Not sure whether neutral network can do this.

Response: Thank you for this comment. We analyzed the relative importance of variables in methods (L190-L192).

"We estimated the relative importance of predictors using the increase in node purity for the splitting variable, which was measured by the reduction in residual sum of squares."

The relative importance was discussion at L352–L365:

"These influences were reflected in our models (Fig. S8-S11). In the models for plant organs and litter, vegetation types and climate variables showed higher relative importance. Heat and water are usually limited in the plateau and desert regions in western China, where shrublands and grasslands are dominant vegetation type groups. More nutrients are allocated to root systems by dominant plants in such stressful habitats to acquire resources from soil (Eziz et al., 2017; Kramer-Walter and Laughlin, 2017). Spatial variables, longitude and latitude, also held high importance, especially in the models for soil nutrients. On the one hand, it may result from their tight links with climate conditions. On the other hand, it may imply the influence of spatial correlation on nutrient pools. The effects of elevation and spatial variables were obvious from the prediction maps. There were relatively larger values of soil nutrient densities in the plateau and mountainous area in western China, possibly because of the lower rates of decomposition, mineralization, and nutrient input as well as less leaching loss in high-altitude regions (Bonito et al., 2003; Vincent et al., 2014)." For detailed results please see Fig S8-S11 in the supplement of the revised manuscript.

• While I agree with the authors' argument that "the first time, we mapped N and P densities of leaves, woody stems, roots, litter and soil in forest, shrubland and grassland ecosystems across China", there are some previous estimates of nutrient stocks in China, maybe only for one ecosystem component or one nutrient. I think a comparison of the authors' estimates with previous estimates, e.g. Tian et al. (2010), would benefits the study. It will make the study well in context of previous studies, and will also show how the estimates are improved compared to previous estimates.

Tian, H., Chen, G., Zhang, C., Melillo, J.M. & Hall, C.A. (2010). Pattern and variation of C: N: P ratios in China's soils: a synthesis of observational data. *Biogeochemistry*, 98, 139-151.

Response: Thank you for this suggestion. We compared the previous estimation of N and P pools with our results in the section *5.2 Nutrient pools in terrestrial ecosystems in China* (L327-L341):

"Previous researches have estimated N and P stocks in soil across China. For example, Shangguan et al (2013) estimated that the storage of soil total N and P in the upper 1m of soil in China were 6.6 and 4.5 Pg. Yang et al (2007) estimated China's average density of soil N at a depth of one meter which was 0.84kg m-2 and the soil N stock was 7.4 Pg. Zhang et al (2005) investigated soil total P pool at a depth of 50 cm in China and concluded that the soil stock was 3.5 Pg with the total P density of soil 8.3×102 g/m3. Our estimation of the soil N pool in China (6.6Pg) agreed with Shangguan et al (2013), but the estimated soil P pool (2.8Pg) was lower than the results of aforementioned studies. The mean soil N:P ratio in our study (2.5 of the predicted dataset and 2.1 of the training dataset) was lower than the result of Tian et al (2010), 5.2, while the spatial patterns in both studies are similar. Other than those researches focusing on soil, Xu et al (2020) estimated China's N storage by calculating the mean N densities of vegetation and soil from different ecoregions, and the reported that there were 10.43 Pg N in China's ecosystem, 10.14 Pg N in top 1 m soil and 0.29 Pg N in vegetation, both higher than our results (6.6 Pg N in soil and 0.16 Pg N in vegetation)."

Minor comments

L18-19: "the limitation of these two" may be changed to "their limitations". L26-31: the numbers are unreadable. Mg is million gram? Given the use of 10⁶, you may use bigger units (e.g., Tg).

L49: here you may also cite Sun, Y., Peng, S., Goll, D.S., Ciais, P., Guenet, B., Guimberteau, M. *et al.* (2017). Diagnosing phosphorus limitations in natural terrestrial ecosystems in carbon cycle models. *Earth's Future*, 5, 730-749. L91-92: Not very clear. Du et al. (2020) showed either N or P limitation. If you mean ubiquitous limitation by N and P, you may refer to Elser et al. (2007), LeBauer and Treseder, K.K. (2008), Augusto et al. (2017), and more recently Hou et al. (2020). Similarly, L46-60 may cite more recent papers on the topic to reflect recent progresses in the field.

LeBauer, D.S. & Treseder, K.K. (2008). Nitrogen limitation of net primary productivity in terrestrial ecosystems is globally distributed. *Ecology*, 89, 371-379. Hou, E., Luo, Y., Kuang, Y., Chen, C., Lu, X., Jiang, L. *et al.* (2020). Global meta-analysis shows pervasive phosphorus limitation of aboveground plant production in natural terrestrial ecosystems. *Nature Communications*, 11, 637.

Augusto, L., Achat, D.L., Jonard, M., Vidal, D. & Ringeval, B. (2017). Soil parent material $\hat{a} \in \bullet$ a major driver of plant nutrient limitations in terrestrial ecosystems. *Global Change Biology*, 23, 3808-3824.

L100: "a high-resolution map" to "high-resolution maps"

Response: Thank you for these comments. We have corrected these inappropriate descriptions in the text and cited the recommended papers (L93-94).

L111-112: are all plots the same size for forests, shrublands, and grasslands?

Response: Thanks for the suggestion. We stated the plot sized in different vegetation types. Please see 2.1(L113-L114).

"At each site, one 20×50 m² plot was set for forests, three replicated 5×5 m² plots were set for shrublands, and ten 1×1 m² plots were established for grasslands."

L123-126: you may give references for the methods here. Equation 1: should the sum symbol with "i = 0" to "n" added? n is the total number of plant species. Similar for Equation 2.

Response: Thank you for the suggestion. We made this correction at L135 and L147 and corrected the description in text.

L135:

" $N(P) = \sum_{i=0}^{n} B_i \times \theta_i$

N(P) represents the community-level N or P density (Mg ha-1); n is the total number of plant species in one site..."

L147:

 $"SND(SPD) = \sum_{i=0}^{n} (1 - \delta_i) \times \rho_i \times C_i \times T_i / 10$

where *SND* (*SPD*) is the total N or P density of the soil within top 1 m (Mg ha-1); *n* is the total number of soil layers (ranging from one to five) in one site..."

L259, the unit of 5?

L269-281: one digit after decimal is enough and would be easier to read.

Response: Thank you for this comment. We made the corrections to the description of results according to your suggestions.

L295: I can't understand the reason. The reason may be expanded to be clear. Response: Thank you for the suggestion. We have changed the text in this part of the manuscript(L323-L324).

"Models for soil showed relatively poorer accuracy than models for plant organs and litter (Fig. 4 & 5), partly because that soil N and P were largely influenced by geological conditions, soil age and parent material (Gray and Murphy, 2002; Buol and Eswaran, 1999) (Doetterl et al., 2015), which were not included in our analysis because of the limited data availability. The can be evidenced by the decreasing validation R2 of the models for soil N densities and P densities and concentrations with soil depths (Fig. 5 and S3)."

L303: "the predicted SDs" is confusing. You may mean "SDs of the predictions" L313: remove "the"

L330: You may also cite the classic paper on this topic: Walker, T.W. & Syers, J.K. (1976). The fate of phosphorus during pedogenesis. *Geoderma*, 15, 1-19. L346: not necessarily more accurate predictions, depends on whether the models are informed by measurements such as those used in this study. "could" may be changed to "may".

Response: Thank you for the suggestions. We made these corrections and cited this paper at L370.

Fig. 3 color legend in panel (a) may include colors only for leaf/stem/root, with colors for vegetation/soil moved to panel (c), because panel (a) and (b) do not have vegetation vs. soil.

Response: Thank you for the suggestion. We moved the legend for vegetation/soil to panel (c). Please see Fig. 1.

Fig. 4: is there a reason for the slopes to be consistently higher than 1.0 across ecosystem components and nutrients? It seems to be a systematic bias in the models: overestimate when observed values are low and underestimate when observed values are high.

Response: Thank you for the suggestion. We changed the prediction method, and the slopes and intercepts are close to 1 and 0, respectively. Please see Fig 4 and 5.

Tables and figures in the revised manuscript

2	Table.1. N and P	stocks of vegetation,	litter, soil and total	l ecosystem in forests,	shrublands and	grasslands in China.
---	------------------	-----------------------	------------------------	-------------------------	----------------	----------------------

Vegetation	Vegetation	Area	N pool (Ta)				D pool (Ta)			
type group	type	$(10^{6} ha)$	N pool (1g)				P pool (1g)			
			Vegetation	Soil	Litter	Ecosystem	Vegetation	Soil	Litter	Ecosystem
Forest	EBF	40.6	18.0	476.4	1.7	496.1	1.7	154.8	0.1	156.6
	DBF	66.3	43.1	811.3	3.7	858.1	6.9	346.5	0.4	353.8
	ENF	83.8	28.4	952.8	2.8	984.0	3.7	349.2	0.2	353.1
	DNF	11.5	5.6	177.7	0.5	183.8	1.5	73.6	0.1	75.2
	MF	9.6	4.6	107.6	0.5	112.8	0.9	41.5	0.1	42.4
	subtotal	211.9	<i>99</i> .8	2525.8	<i>9.3</i>	2634.9	14.6	965.6	0.9	981.1
Shrubland	EBS	18.7	2.1	213.6	0.5	216.2	0.2	80.9	< 0.1	81.1
	DBS	48.7	5.5	570.9	1.2	577.6	0.5	233.6	0.1	234.2
	ENS	1.0	0.1	12.4	< 0.1	12.5	< 0.1	4.9	< 0.1	4.9
	SS	11.9	0.5	66.1	0.1	66.7	< 0.1	61.6	< 0.1	61.6
	subtotal	80.3	8.1	863.0	1.8	873.0	0.7	381.0	0.1	381.8
Grassland	ME	44.2	11.6	806.9	0.1	818.5	0.9	247.2	< 0.1	248.0
	ST	137.4	21.3	1348.5	0.3	1370.1	1.5	573.1	< 0.1	574.6
	TU	22.8	2.3	230.4	0.1	232.8	0.2	112.9	< 0.1	113.2
	SG	103.8	13.6	860.6	0.1	874.4	0.9	506.3	< 0.1	507.2
	subtotal	308.2	48.8	3246.4	0.6	3295.8	3.5	1439.5	< 0.1	1443.0
Total		600.4	156.7	6635.2	11.7	6793.1	18.8	2786.1	1.0	2806.0

- 3 EBF, evergreen broadleaf forest; DBF, deciduous broadleaf forest; ENF, evergreen needle-leaf forest; DNF, deciduous needle-
- 4 leaf forest; MF, broadleaf and needle-leaf forest; EBS, evergreen broadleaf shrub; DBS, deciduous broadleaf shrub; ENS,
- 5 evergreen needle-leaf shrub; SS, sparse shrub; ME, meadow; ST, steppe; TU, tussock; and SG, sparse grassland.

Vegetation type group	Vegetation type	Area (10 ⁶ ha)	N pool (Tg)		P pool (Tg)			
			Leaf	Stem	Root	Leaf	Stem	Root
Forest	EBF	40.6	3.9	10.1	4.0	0.3	1.0	0.3
	DBF	66.3	6.1	26.6	10.5	0.6	4.6	1.6
	ENF	83.8	8.6	13.4	6.4	0.9	2.0	0.8
	DNF	11.5	1.3	2.9	1.4	0.2	0.9	0.3
	MF	9.6	1.0	2.6	1.0	0.1	0.7	0.2
	subtotal	211.9	21.0	55.5	23.4	2.1	9.2	3.3
Shrubland	EBS	18.7	0.6	0.7	0.7	< 0.1	0.1	0.1
	DBS	48.7	1.4	1.4	2.7	0.1	0.1	0.2
	ENS	1.0	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1
	SS	11.9	0.1	0.1	0.3	< 0.1	< 0.1	< 0.1
	subtotal	80.3	2.1	2.3	3.8	0.2	0.2	0.2
Grassland	ME	44.2	0.9	0.0	10.7	0.1	0.0	0.8
	ST	137.4	2.2	0.0	19.2	0.2	0.0	1.3
	TU	22.8	0.5	0.0	1.7	0.1	0.0	0.2
	SG	103.8	1.1	0.0	12.5	0.1	0.0	0.8
	subtotal	308.2	4.7	0.0	44.1	0.4	0.0	3.1
Total		600.4	27.7	57.8	71.2	2.7	9.4	6.7

6 **Table.2.** N and P stocks of plant organs (leaf, stem and root) in forests, shrublands and grasslands in China.

7 See table 1 for abbreviations.



9 Fig. 1. Frequency distributions of N densities in soil, roots, leaves, litter and woody stems in

10 forests (a-e), shrublands (f-j) and grasslands (k-n) in China.



12 Fig. 2. Frequency distributions of P densities in soil, roots, leaves, litter and woody stems in

- 13 forests (a–e), shrublands (f–j) and grasslands (k–n) in China.
- 14



15

Fig. 3. N and P density allocations among leaf, stem and root (a & b) and between vegetation and soil (c & d) in 13 Vegetation types. See table 1 for abbreviations. The error bar represents standard error. Notice that the y axes above and below zero are disproportionate.



Fig. 4. Fitting performance of random forest models for nutrient densities of leaves (a & b), woody stems (c & d), roots (e & f) and litter (g & h) of terrestrial ecosystems in China based on 100 times of replications with the 10% validation data. Solid lines represent all the fitting lines, and the displayed parameters stand for the average conditions. The dashed line denotes the 1:1 line.



Fig. 5. Fitting performance of random forest models for nutrient densities of 0–10 cm (a & b),

- 29 10-20 cm (c & d), 20-30 cm (e & f), 30-50 cm (g & h) and 50-100 cm (i & j) soil layers of
- 30 terrestrial ecosystems in China based on 100 times of replications with the 10% validation data.
- 31 Solid lines represent all the fitting lines, and the displayed parameters stand for the average
- 32 conditions. The dashed line denotes the 1:1 line.
- 33
- 34



- 36 Fig. 6. Predicted spatial patterns of N and P densities with a resolution of 1 km (a–j) in leaves
- 37 (a & b), woody stems (c & d), roots (e & f) and litter (g & h) of terrestrial ecosystems in China.



- 39 Fig. 7. Predicted spatial patterns of N and P densities with a resolution of 1 km in vegetation (a
- 40 & b, the sum of leaves, stems and roots), soil (c & d, the sum of five layers) and ecosystems (e
- 41 & f, the sum of vegetation, litter and soil) of terrestrial ecosystems in China.
- 42

43 Supplement



Fig. S1. The spatial distributions of sampling sites (a) and the topographic map of China (b).



Fig. S2. Fitting performance of random forest models for nutrient concentrations of leaves (a & b), woody stems (c & d), roots (e & f) and litter (g & h) of terrestrial ecosystems in China based on 100 times of replications with the 10% validation data. Solid lines represent all the fitting lines, and the displayed parameters stand for the average conditions. The dashed line denotes the 1:1 line.



55 **Fig. S3.** Fitting performance of random forest models for nutrient concentrations of 0–10 cm

- 56 (a & b), 10–20 cm (c & d), 20–30 cm (e & f), 30–50 cm (g & h) and 50–100 cm (i & j) soil
- 57 layers of terrestrial ecosystems in China based on 100 times of replications with the 10%
- validation data. Solid lines represent all the fitting lines, and the displayed parameters stand for
- 59 the average conditions. The dashed line denotes the 1:1 line.
- 60



62 Fig. S4. Frequency distributions of standard deviations of the predictions in models for N and

- 63 P densities in different components.
- 64



65

66 Fig. S5. Frequency distributions of standard deviations of the predictions in models for N and

- 67 P concentrations in different components.
- 68
- 69



- Fig. S6. Predicted spatial patterns of N and P concentrations with a resolution of 1 km (a–j) in
- 72 plant organs (a–f), litter (g & h), and soil layers (i–r) of terrestrial ecosystems in China.



Fig. S7. Predicted spatial patterns of N:P ratios with a resolution of 1 km (a-j) in leaves (a),

76 woody stems (b), roots (c), litter (d) and soil (e) of terrestrial ecosystems in China.

77



- 79 Fig. S8. The relative importance of variables in random forest models of N and P densities for
- 80 leaf (a & b), stem (c & d), root (e & f) and litter (g & h).



- 83 Fig. S9. The relative importance of variables in random forest models of N and P densities for
- 84 0-10 cm (a & b), 10-20 cm (c & d),20-30 cm (e & f) 30-50 cm (g & h) and 50-100 cm (i & j)
- soil layers.
- 86



- 88 Fig. S10. The relative importance of variables in random forest models of N and P
- 89 concentrations for leaf (a & b), stem (c & d), root (e & f) and litter (g & h).



- 92 Fig. S11. The relative importance of variables in random forest models of N and P
- 93 concentrations for 0-10 cm (a & b), 10-20 cm (c & d),20-30 cm (e & f) 30-50 cm (g & h) and
- 94 50-100 cm (i & j) soil layers.
- 95