## Earth System Science Data

## RE: Manuscript ID: essd-2020-398

# Point-for-point responses to the comments from Reviewer 1

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.

#### **General comment**

Yi-Wei Zhang et al. presented a data analysis study for terrestrial ecosystem N and P pools over China. The data collection, model fitting, regional and pft level aggregation and analysis are well done. The presentation is smooth.

Response: Thank you very much for reviewing our manuscript. We appreciate your kind praise for our study!

Below are my major suggestions and specific comments.

1 Root N, P and Soil N, P model fitting

Root and soil N, P models underperformed (e.g., R2 0.27~0.47), in comparison with models of other plant components (e.g., R2=0.56-0.81). I would suggest 1) trying more complex neural network models (more layers or more nodes within each layer) 2) trying different types of ML models (e.g., random forest, support vector regression) 3) including more explaining variables besides MAT, MAP, elevation, and PFT. For example, N/P deposition, land use history, soil order, GPP and so on.

Response: Thank you for this comment. After comparing several models, we have adopted the random forest method to reach a better model performance. Furthermore, we fitted models for each soil layer, respectively, instead of the sum of all layers in the previous manuscript. In the revised results,  $R^2$  of the models were higher than the previous version, with *ca* 0.6 for root N and P, and *ca* 0.5 for different layers of soil N and P concentration. We did not include other factors because of the lack of N/P deposition and soil order at such fine scale. As we focus on the natural ecosystems, land use history is not considered in this study. The process of random forest modelling was stated as follows (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 R2 obviously. Model prediction were repeated for 100 times to obtain the average results..." For detailed results please see Fig 4-7 in the revised manuscript.

#### 2 representativeness of data for regional extrapolation

It will be helpful to show 1) a map that includes the location of all data samples 2) MAT, MAP, elevation ranges for data samples, compared with those variables but

across China. The purposes are to reveal whether the data samples are spatially representative and whether the data reasonably cover the full range of T, P, Elevation so that the spatial extrapolation is reliable (for each vegetation cover).

Response: Thank you for mention this. We revised the manuscript according to this suggestion. Please see Fig S1 the supplement of the revised manuscript for the distribution map of sampling sites. The environmental variables of our sampling sites could generally cover the 99% ranges of the corresponding variables across China. We supplemented the relevant descriptions in the revised manuscript as follows (L164-L168):

"The ranges of these variables of our field sites (EVI:  $0.03 \sim 0.7$ ; elevation: -137 m~5797 m; MAP: 19.8 mm~2316.3 mm; MAT: -5.2 °C~ 26.0 °C) could generally cover the ranges of corresponding variables in the focused vegetation types across China (99% ranges of EVI:  $0.03 \sim 0.6$ ; of elevation: 24 m~5628 m; of MAP: 50.6 mm~2956.5 mm; of MAT: -6.6 °C~ 22.8 °C)."

#### 3. N, P mass concentration

This analysis focused on area-based N, P concentrations (g N/m2 of land surface), which do not directly link to ecosystem N/P limitations. And given that the vegetation is not evenly distributed, it will be helpful to also present the mass-based N, P concentrations (e.g., g N/g tissue biomass or soil) that could directly reveal the strength of plant and soil N, P limitation.

Response: Thank you for this suggestion. We revised the manuscript according to your suggestion and supplemented the results about mass-based concentrations in the revised manuscript.

In Methods, we mentioned that "The same procedures were repeated for the prediction of N and P concentrations in different components across China." (L192–193).

In Results, we supplemented that "The N and P concentrations in plant organs and litter were generally higher in northern and western mountain regions, but larger values of the former often occurs in northwestern part of China, while those of the latter often occurs in northeastern part of China (Fig. S6a–h). The spatial patterns of soil nutrient concentrations at different depths were consistent with those of soil nutrient densities (Fig. S6i–r)." (L274–L278). Please see Fig S6 for the detailed results.

#### 4. N:P stoichiometry

From an ecosystem N/P limitation perspective, the ratio of N and P within different plant tissues will be more informative than the individual concentrations. I would suggest also showing N:P stoichiometry, e.g., across pfts, leaf vs fine root.

Response: Thank you for this suggestion. We revised the manuscript according to your suggestion and supplemented the prediction maps of N:P in different plant tissues and soil in the revised manuscript.

In Methods, we mentioned that "The spatial pattern of N:P ratio was calculated from the predicted N and P density datasets of the corresponding component." (L194–195).

In Results, we supplemented that "N:P ratio of plant organs and litter showed similar distribution patterns, higher values occurring in southeastern and northwestern China

and Qinghai-Tibetan Plateau (Fig. S7a–d). Soil N:P ratio was higher in northeastern and southern China but lower in northwestern China (Fig. S7e)." (L279–L282). Please see Fig S7 for the detailed results.

#### **Specific comments:**

L54 independently or jointly L63 allocated to plant

Response: Thank you for the suggestions. We have corrected these in the revised manuscript.

L167 since the model uses re-scaled predictors (eq. 3), it is important the make sure the training data could represent the full climate envelopes over China.

Response: Thanks for this suggestion. We supplemented the descriptions about the ranges of environmental variables of the sampled data. Please see our answer to the major comment 2 above.

#### L226 what is site-averaged?

Response: Thank you for this comment. We corrected this description at L233 and other parts in the revised manuscript:

"The mean litter N densities for forest, shrubland and grassland sites were  $6.1 \pm 7.6 \times 10-2$  Mg N ha-1,  $3.8 \pm 4.6 \times 10-2$  Mg N ha-1 and  $5.5 \pm 9.3 \times 10-3$  Mg N ha-1, respectively..."

#### L238 density varied

#### Response: Thanks for this comment. We have made this correction.

L294 "soil N and P are stable" is not a convincing reason why soil models underperformed. In contrast, one would expect that stable N P pools shall be better modeled by long-term climatology, compared with e.g, seasonally changed leaf N/P concentrations.

Response: Thank you for the suggestion. Besides other reasons, we think the influence of soil properties (such geological conditions, soil age and parent material), which was not included our analysis, may weaken the soil models. This can be evidenced by the decreasing  $R^2$  of the models with soil depth. We have changed the text in this part of the manuscript (L310-L315).

"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 (Buol and Eswaran, 1999; Doetterl et al., 2015; Gray and Murphy, 2002), 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)." L309 this section needs more quantitative evidence for drivers that are included in this study (e.g., T, P, elevation) and should consider including potential drivers that are discussed if spatial data are available (e.g., soil age, soil order).

Response: Thank you for this comment. We supplemented the contributions of each variables in appendix fig S8-S11. Although we did not include soil age and soil order, we discussed the potential contribution of these variables in our models and possible drivers in this section

L352-L354: "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."

L358-L365: "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)."

L369-L372: "Additionally, such patterns reflect that the factors not investigated in this study, such as soil age and parent material, could contribute to the patterns of nutrient pools, which should be considered in future researches as potential drivers (Porder and Chadwick, 2009; Augusto et al., 2017)."

# 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 <sup>6</sup> 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