Here we listed our responses to the comments of reviewer 1 in tabular form. The page and line numbers of the referee's comments refer to the original manuscript: soil-2018-40, (https://doi.org/10.5194/soil-2018-40).
Page and line numbers of the Author's reply refer to the revised manuscript.
We want to thank the anonymous reviewer for the valuable input to improve the manuscript.
In behalf of all authors, Jörg Niederberger

Index	Referee's comment	Author's reply
1	The paper is rather long and sometimes cumbersome to read. It should be shortened by 20-30% to attract more readers and focused on the most relevant issues.	All reviewers, including Reviewer 1 suggested a variety of improvements, which we implemented in the revised manuscript. These revisions should also improve the readability of the manuscript. At the same time, they did not allow for major reductions in the length of the paper. Furthermore, the other two reviewers did not criticize the length of the manuscript. Thus we believe that the size of our manuscript is appropriate to address all relevant issues of our study.
2	Soil texture and SOC at least in a given depth increment most often are strongly correlated with each other, because low sand and high (silt)/clay contents favor SOC accumulation by formation of organo-mineral associations and aggregates which impede SOC mineralization. This multicollinearity effect could be used either form an amalgamated predictor or to remove one of the two variables in order to shorten the paper.	We are aware of the problem of multicollinearity among predictor variables in soil. Therefore, we checked our predictor set before modeling for correlations, multicollinearity, as well as for autocorrelations. We could not observe any of the above mentioned effects. Additionally, the model output parameter provided by SPSS, as described in the material and method section (Chapter 2.4. P 5 L 24 ff.), provided no indication for autocorrelation (Durbin-Watson) or multicollinearity (VIF). Therefore, we believe that all predictors that we considered describe important and also different properties.
3	The authors should test whether splitting the sample collective into 2 sub-collectives (non-calcareous soils [pH < 6.5] vs. calcareous soils [pH > 6.5]) may improve the predictive power of the Hedley fractions for characterizing the P nutritional status of the trees.	Although we have a large number of sites included in our survey, there were only 8 out of 143 sites with a soil pH above 6.5. This number is too low to develop robust statistical models for these calcareous soils. However, we checked also models excluding these 8 calcareous sites and compared them with models including all sites. Model results for the group of soil samples with pH < 6.5 (non-calcareous soils) did not change substantially when compared to models including all sites. We found only some minor improvements as well as some minor deterioration of model quality. Nevertheless, we could not observe changes in the selected predictor variables or in the dominant predictor

		variable for non-calcareous soils. See Methods section chapter 2.4, P5 L 13 ff. However, it would be very interesting to address the issue of calcareous soils in a future study with a different collective of soil samples.
4	Moreover, I suggest to test whether the calculated total topsoil stocks of the different Hedley fractions (the latter should be available according to the statement made in page 12, line 19/20 of the manuscript) may improve the predictive power of the Hedley fractions for characterizing the P nutritional status of the trees.	We followed this suggestion and calculated also models with P stocks (stocks of total Hedley P and P Hedley P pools) as predictor variables to explain foliar P concentrations. However, we did not find any improvements in model quality, on the contrary, models were of consistently lower quality. One reason might be that we needed to make very broad assumptions for soil bulk density and stone content to calculate P stocks. . See Method section, 2.4 P5 L22ff.
5	There are several papers dealing with the relation between (operationally-defined) soil P fractions and the P nutritional status of German forest ecosystems, whose results could be compared with the results of the Hedley procedure. For example the papers of Prietzel and Stetter (2011), Prietzel et al. (2013; DOI 10.1007/s11104-014-2248-9); and the recent paper of Manghabati et al 2018 (JPNSS; DOI: 10.1002/jpln.201700536) all found a good predictive power of citric-acid extractable soil P on tree P nutrition, whereas HCO3 was suitable only under particular conditions (Manghabati et al.).	We included these studies in the discussion section (Chapter 4.2, P13 L10 ff., P14 L4 ff.) of the revised manuscript.
6	At least in some pages, the paper contains a lot of typos/spelling mistakes and sloppy grammar. For example, page 10 reads. The English should be brushed up before resubmission	The paper has undergone a thorough language revision.
P2 L34	Reference De Schrijver et al 2012 is missing in the Reference Section	The reference De Schrijver et al. 2012 was accidentally listed under „S“. We corrected this in the revised manuscript.
P3 L10:	C content. Is this total C or organic C? Should be clarified. The text on P4 L 34 suggests that “C content” means SOC, whereas the large Max C/N values (437; 61) presented in Table 1 indicate that at in the calcareous soils C content included inorganic C in addition to SOC.	This was indeed SOC and therefore changed to soil organic carbon. Unfortunately, the table in the manuscript was not the final version and included still samples from two sites with peaty soils. Here we found of course very high SOS values. These two sites were removed from our study.
P4 L9:	In deciduous tree stands, leaves are always from the current year	We rephrased this sentence. Now it becomes clear that the leaves resp. needles were sampled

		in the same year as the soil samples were taken.
P4 L22:	Nitric acid digestion does not completely retrieve total P, because Si-bound P is only partially mobilized. Underestimation is between 15 and 37% (Schwartz & Kölbel, 1992; Z Pflanzenernähr Bodenkd 155: 281–284; Hornburg & Lüer 1999; J Plant Nutr Soil Sci 162:131–137.	Indeed, that is the case; therefore, we used in the text usually the phrase "sum of all Hedley fractions". We think, that it is appropriate to use the same digestion technique like the strongest Hedley fractionation step to assess if the sum of the Hedley fractions corresponds with the independent P sum. If we had used the HF digestion instead, we were not able to compare the sum of fractions with the independently measured total P content. Nevertheless, we clarified this in the text, see Method section chapter 2.3, P4 L26ff.
P5 L3:	Has the MWU-Test been corrected for multiple comparisons? Please indicate!	A correction of MWU-Test for multiple comparisons is necessary if multiple questions are tested simultaneously with one dataset. In our case, we tested our target soil variables independently with - according to the target variable - newly arranged datasets. Therefore, a Bonferroni correction was not necessary.
P13 L3:	forest floor mass and Corg/Porg ratio or P , Porg content should be used	We included this addition into the discussion section at this point
Figure 2:	see comment to P4 L22	See response to Comment P4 L22
Figure 3:	Nice figure.	Thank you!
Table 2:	Pi residual: Is this really Pi or may it also include Porg which is liberated and converted into Pinorg by nitric acid/H2O2 digestion?	Indeed, digestions with strong acids could liberate organic P and convert it into inorganic P. We observed this problem in the fractionation step with concentrated HCl, where we found very inconsistent Po and Pi relations for repeatedly analysed samples, but very consistent results for the sum of Pi and Po. For that reason, we refrained from using Pi and Po for that fractionation step and instead, we used the sum of both. This issue was addressed in Niederberger et al. 2015 and 2016 and cited in our method section. Additionally, we didn't observe this phenomenon for the residual fraction in our preliminary studies, where we tried to address such problems. It might be that the organic P is already digested by the preceding HCl_{conc} step which is also a strong acidic digestion. Nevertheless, we changed Pi residual into P residual to avoid any misunderstandings.
Table 2: Table 5:	As the data are non-normally distributed would it make sense to describe the variation in box plots rather than by standard deviation, which requires normal distribution?	We replaced the tables by box plots