Biogeosciences Discuss., 12, C60–C69, 2015
www.biogeosciences-discuss.net/12/C60/2015/
© Author(s) 2015. This work is distributed under
the Creative Commons Attribute 3.0 License.

Biogeosciences
Discussions
Open Access

Interactive comment on "Use of near-infrared spectroscopy to assess phosphorus fractions of different plant availability in forest soils" *by* B. Todt et al.

Anonymous Referee #1

Received and published: 3 February 2015

Todt et al. present one of the first comprehensive tests of the applicability of near infrared spectra to predict P fractions in soil. The authors used an extensive data set for which they measured both NIR spectra and P concentrations in different fractions by means of wet chemical extraction followed by photometric determination. They found that (i) the calibration and validation yielded a model quality mostly explaining more than 60% of the variation; (ii) model quality was higher for organic than for inorganic P fractions, and (iii) model quality was improved if heterogeneity in soil properties of the data set was reduced. Generally, the manuscript is well written and clearly structured. I feel that the topic is of broad interest to the readers of BioGeosciences. However, from my perspective, the manuscript has some restrictions and flaws. Therefore, I suggest major revision. The following remarks will illustrate my decision and might help to improve the manuscript.

General comments - While the authors clearly state that phosphate groups are not detectable by NIRS, it remains unclear whether or not spectra arise from ester bonds or other bonds associated with organic P fractions. Similarly, the reader is not able to follow which other soil properties might be linked to P fractions in terms of NIR spectra and how this could be explained in a more mechanistic way. Based on existing applications (that are able to focus on C-H vs. C-O or C-OH) one could at least come up with a very rough concept.

- As NIRS is intended to reduce the number of chemical analyses (still necessary for calibration), it would be highly useful to have an estimate of the mean error (in mass P/mass sample) associated with the predicted concentrations of P fractions of the validation subset (not included in model establishment). This would be comparable to common precision/trueness parameters used for quality assurance in wet chemistry analyses.

- Link between P compounds and spectra; standards to be analyzed (e.g. monoesters, diesters etc.)

Specific comments 1/23-25: Not only R^2 is relevant, but also whether or not the regressions were significant. I assume not all regressions were significant. If so, please state the proportion of significant regressions as well.

1/26: "homogeneity" in terms of? Range of soil properties? Range of P concentrations? Soil types? Specify!

2/13: This is controversially discussed, please add constraints of estimates and other views as well.

2/20: "diminish" because of? Timber harvest? Erosion? Be more specific and add evidence provided by other studies.

2/22: The initial idea of the fate of P during ecosystem development and pedogenesis dates back to 1976 (Walker, T.W., and J.K. Syers. 1976. Fate of phosphorus during pedogenesis. Geoderma 15: 1-19.). Should be acknowledged here as well.

4/32: As you state hypotheses, these will either be verified or falsified. This is not possible for Hypothesis 1 unless you define criteria associated with “sufficiently well”. Based on which criteria and thresholds do you rate a prediction as “good” or “not sufficient”?

5/1-3: Again more specific: “quality” in terms of?

5/21-35: I would like to see quantitative measures of the selection procedure. What criteria did you use to come up with “typical brown earths” as the final subset (apart from the fact that n = 84 is near to the 100 samples required for model development)? You state no correlation between total P and 25 individual P fractions or other soil properties such as total C, N and pH. But how could correlations aid in selecting subsets? Furthermore, your statements “less heterogeneous” (l. 28) and “still heterogeneous” lack a quantitative evaluation. What is the criterion for heterogeneous versus homogeneous data sets?

6/5: On the preceding page, please add approximate area covered by the BZE data set. Furthermore, add mean distance between two sites for the Chinese data set (maybe also for the BZE data set).

6/8-11: I do not understand the procedure here: the three (four) topmost diagnostic horizons were located deeper than 47 cm? Or did you select those diagnostic horizons only that did not duplicate the depth increments mentioned before? Please clarify.

6/11: The tree cluster samples were taken as replicate samples whereas (as far as I understood) all samples described before represent composite soil samples. Please add a critical remark concerning this difference (e.g. pseudoreplicates).

6/28-7/2: Add a critical remark on how different sample preparation procedures might affect the relationship between spectra and wet chemical extraction procedures.

9/6-7: You state functional groups, but show bonds: O-H no functional group (-OH); C-H/-CH3 or –COOH or….; N-H/-NH2. Would you like to refer to the bonds? If so, would this include C for N and O as well (C-N-H; C-O-)? Without such information it is difficult to guess how NIRS could be adapted for P fractions.

10/15-17: Contradiction to pre-selection of typical brown earth (5/21-35). You state that you tested different groupings including soil type. This would not be possible if you pre-selected “typical brown earth” only!? Finally, after the confusing statements on inclusion or exclusion of data subsets (starting five pages before!), the reader is relieved to find the reasoning… (10/23-30). These should precede any statements on in-/exclusion of data to ease readability. Please restructure this section accordingly and rephrase if necessary.

10/17-20: Above you stated that NIRS measurements of P are possible BECAUSE of correlations with soil organic matter properties. As organic P forms part of SOM, I do not understand what is meant by “original properties of soil P”. Please clarify.

11/9: For readers not familiar with model evaluation please explain how to interpret the RPD. 11/12-13 implies that high RPDs are desirable but why should one aim at high standard errors of prediction used as numerator in the ratio calculation?

11/28-12/11: Comparison with variables (pH, C, N) used to classify the data sets as heterogeneous/homogeneous?

12/7-9: Please add the proportion of the organic NaOH P fraction relative to total P to enable the reader to judge the relevance of these high Po concentrations.

12/16: State range of R^2 and RPD for models of the fractions the at least.

13/26-14/7: You did not state it explicitly in the methods (add information 10/19), but here as well as in Fig. 7 you mention the Spearman Rank correlation coefficient as independent variable. For continuous and metric C or N and P concentrations the Pearson’s correlation coefficient is to be preferred. What was the reasoning for choosing a non-parametric coefficient? Irrespectively, the two variables used for the regression are differently detailed: i) the goodness of fit represents the percentage of data of data that can be predicted by the calibration model; ii) any correlation coefficient will yield the “strength” of the relationship between two variables be it an approximation of the slope of a regression (Pearson) or the relative position if ordering the data from low to high values (Spearman). However, a correlation coefficient of 1 does not mean that the data can be predicted well, because these coefficients are not necessarily related to the scatter in the data. For example, a correlation coefficient of 1 could arise despite the that fact that data points scatter greatly along the 1:1 line. Therefore, no meaningful interpretation can be derived from Figure 7. If the authors used a regression between concentrations of Ct or Nt and P concentrations, the resulting R^2 might be used as an independent variable in Figure 7. Delete this paragraph and rewrite it according to the new results. There are already six figures in this manuscript, therefore, a list of results without a figure is sufficient.

14/24-26: Stated at this prominent position (concluding sentence of a paragraph) I would like to see some details of this quality check (coefficient of variation or mean difference between repeatedly analyzed samples or similar) without displaying all the data.

15/28-30: Given the fact that the preceding sentences repeat information provided in the introduction already and thus, do not lead to an increased knowledge before and after conducting the measurements, I would like to read an educated guess how the different P compounds could influence the spectra. Why should monoesters result in spectra different from that of diesters?

16/1-6: I do not agree that this conclusion can be derived from the results because Fig. 7 does not allow for a meaningful interpretation (see comment on Fig. 7).

16/6-9: I am lost now: at several places throughout the manuscript, it is stated that the P-O bond cannot be characterized by NIRS and that P compounds must be detected

indirectly based on other soil properties with organic matter being the most promising proxy because of the influence of functional groups/bonds in organic molecules (e.g., 9/10-11). If this is true, what did lead to “sufficiently good” predictions of P fractions and pools in your study?

16/23-25: Not all studies listed above stated increasing prediction quality with increasing heterogeneity. Order preceding list of studies accordingly and evaluate which studies agree/disagree with your findings and, most importantly, why there are similarities/differences.

16/29-17/2: The BZE brown earth model deviates only slightly from the BZE model. Did this improvement lead to a higher class assigned to model quality in any case? If not, please tune down the statement on improvement of the model.

17/3-5: Without any chemical information on the link between P compounds and NIR spectra, the reader is not able to follow this paragraph. How might spectra be related to P compounds? See comments on chemical structures above.

Technical corrections

1/1: “near-infrared”; “phosphorus fractions”

1/15: “fractionation of . . . into fractions”; awkward phrasing, please rephrase

1/27: “Meaningful models”

2/2: “useful” might depend on the view point. Please phrase more specifically what is meant by “useful” (e.g. match between NIRS data and results of chemical extraction).

2/25: “monitoring the” (delete “of”)

2/30: hyphen in “plant-available P”; check throughout manuscript

3/6: “dynamic”

3/8: (relevance. . .) “has been”

3/14: “Hedley fractionation” (without hyphen)

3/16: red dot?

3/16: “Hedley P” (without hyphen); check throughput manuscript (e.g. 4/29)

3/17: “less expensive” or “cheaper”

3/25: “2010).”

3/26: “Furthermore,”

3/27: “which commonly constitute the major portion”

3/28: As it is phrased now, the first part of the sentence is contrary to the second part. The spectral information cannot be complex/heterogeneous and uniform at the same time. What does “its” refer to? Better state an own subject for the first part of the sentence.

4/7: “<2mm”

4/17: It would be logical if high variation in chemical composition was associated with high spectral variation. If this was the case, please rephrase (“chemical composition associated with high spectral variation”).

4/26: “soil P” (no hyphen)

4/28-30: awkward phrasing; merge to one sentence.

5/16: “grouped by soil type”

5/26: “and, “

6/1: “research project”

6/25: “measured in”

6/31: “< 2mm”

6/32: “the determination of P fractionation in soil.”

7/8: “authors discussed”

7/17: tense: “considered”, “used”

7/27: “2008).”

7/29: consistent hyphens

7/31: insert Po in parantheses

8/2: “the resin”

8/22-23: “the Hedley fractionation method”

8/33: “did not”

9/3: “bending, and”

10/6: “This was carried”

10/6-10: I do not understand the last part of the sentence (“second to min and max values”)? Split sentence and rephrase.

10/11: “optimize”? Be consistent throughout manuscript (e.g. characterize 1/16)

10/31: “Set 3” (incl. space); 11/2: “Set 2”, “Sets 1 and 2”; 11/4: “Set 4”; 11/7: “Set 4”

11/11: “was discussed”

12/11: Accuracy by definition includes precision and trueness of measurements (the opposite for inaccuracy, of course). I cannot see how low concentrations fit in either of these meanings. Maybe you refer to the limit of detection or similar? Rephrase.

13/13: “(Fig. 5)” (space)

13/29: “(Fig. 7)”

14/21-22: awkward wording (“can make it difficult”), rephrase.

15/2-16: Pure description without interpretation, move to (method)/results section.

17/10: Add references on knowledge vs. knowledge gaps concerning inorganic P.

17/12-14: Awkward sentence, rephrase.

Table 1: Replace comma by dots (2 times); reduce decimal places (one) for skewness and curtosis.

Table 2: “carbon”; “nitrogen”; Be consistent with Table 1: one decimal place only.

Table 3: “parameters”

Figure 1: too many figures in manuscript; this procedure is described well and easy to understand in the method section. Delete Figure 1.

Figure 3: This might represent one basis for a quantitatively-driven data subset selection. However, neither the method details nor results were described (Which variables are included in the PCA?; Which procedure was used to create the n-dimensional space [e.g. varimax rotation]?; How many principal components were derived?; Which proportion of variance was explained by the two components displayed in Figure 3). Why should this principAL (please change spelling in Figure 3 and caption accordingly) be preliminary as stated in the methods section?

Figure 4 and 5: How could negative concentrations be predicted? The model should set these to zero!?

Fig. 4: I find it strange that the calibration and validation figures both use measured P concentrations as independent variables. For the validation data set, the modelled values were not derived from P concentrations of the wet chemistry protocol (“measured P”) but directly from the NIR spectra. Therefore, the modelled P concentrations should represent independent values plotted at the x axis.

Fig. 4/Table 3: Redundant data display; either as a figure or a table, but not both.

Interactive comment on Biogeosciences Discuss., 12, 555, 2015.