

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 #2

Received and published: 26 February 2015

Todt et al. prepared an interesting paper about the use of near infrared spectroscopy for an estimation of P fractions in forest soils. I found many ideas in the paper exciting. However, the manuscript has some flaws and I suggest that a major revision is required.

Comments 1. The selection subsample sets and the procedures used in the calibration/validation or cross-validation need to be much better explained and justified. The authors described four different subsample sets used for calibration (p. 568). Here much more information and justification is required. Important issues are for each of the four subsample sets - what were the sample numbers? - which depth ranges were considered? - the samples are considered to be representative for which population? - did the authors make sure that no pseudoreplicates (i.e. in case of calibration/validation:

C343

samples from one site were NOT in the calibration and validation data set or i.e. in case of cross-validation: the authors did NOT carry out a leave-one-out-cross-validation and made sure that samples from one site were not in different groups) were present and thus no overoptimistic results? The authors are urged to follow the recommendations by Brown et al. (2005, Validation requirements for diffuse reflectance soil characterization models with a case study of VNIR soil C prediction in Montana. *Geoderma* 129, 251–267).

2. The manuscript has some peculiar statements. The authors wrote: "Since there was no indication of autocorrelation between samples of different depth, we included all samples in our calibration and validation step". I strongly disagree with that statement. Firstly, the authors should study the paper by Brown et al. (2005). Secondly, the authors should give their scale of interest for each data set and should avoid pseudo-replication. I do not see the need for a test of autocorrelation in this study, since the mineralogical background does affect the spectra. The presence of the same mineralogical background reduces the noise and increased accuracies for the estimations can be expected.

The authors wrote: "Development of robust NIRS-models requires sample populations that cover the whole calibration range with an approximately even distribution of samples across the range of the variable to be predicted. In contrast, populations with normally distributed samples tend to overestimate low values and underestimate high values in model calibration (Williams, 2001)". This may be ok, but the authors still have to give essential information: whenever they present r^2 and RPD values (which are calculated from SD and SECV values), they rely on a normal distribution. Thus, skewness and kurtosis should be given for all data sets and constituents, where RPD and r^2 are presented and the interpretations are dependent on that additional information.

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