Biogeosciences Discuss., 9, C9345–C9357, 2013
www.biogeosciences-discuss.net/9/C9345/2013/

Biogeosciences Discussions

Interactive comment on "Remote sensing of LAI, chlorophyll and leaf nitrogen pools of crop- and grasslands in five European landscapes" *by* E. Boegh et al.

E. Boegh et al.

eboegh@ruc.dk

Received and published: 12 April 2013

We appreciate very much the thorough review given by the anonymous referee. It has been very valuable to clarify objectives and remove some misunderstandings and to improve the analysis and presentation of results. In particular, the paper has been restructured to clearly separate results and discussions related to the two objectives of the paper since we found that this provided the basis for some misunderstandings; more emphasis is given to the presentation and analysis of REGFLEC results, and uncertainty assessment of field data is now considered.

Please find detailed replies to the comments below.

REFEREE # 2

This paper sets out "to assess the utility of different remote sensing-based methods for regional mapping of CHLl, Nl and LAI in crop- and grasslands". To do so, they use SPOT satellite data to calculate various SVIs, and relate these to field measurements on the ground, made at the same time as the satellite data acquisitions. This is a very worthy aim, and the testing satellite-based estimates of vegetation properties is an area that needs much improving. This paper represents a reasonable attempt, and is useful in looking at a range of sites from across Europe, and collecting ground-based data at all of them. However, I feel there are various weaknesses in their approach and data analysis, and much more could be done. I would recommend some revisions before this is published.

1. The stated aim was to assess the capability of the selected remote sensing methods to quantify LAI, CHLl and Nl over a large range of environmental conditions in Europe. The crux of this is Figure 6, which shows the field measurements against the satellite estimates. However, the authors cherry-pick the methods which work best for each site, and combine these to give a misleading result. This is not a useful validation test. In any real application of the method, we would not know a priori which algorithm to use. Figure 6 needs to be expanded considerably to show the results for all data and each algorithm in a number of panels in the figure. Table 4 (rows for 'All') show the stats on the agreement for this comparison, but it is worth showing it graphically.

REPLY We agree that Figure 6 is not a useful validation test. The validation tests are shown in Tables 4 and 5. Rather than representing validation tests, the purpose of Figure 6 serves the second objective of the paper which is to "to assess the distribution and size of vegetation N pools in the five European agricultural landscapes". This was the main objective for using remote sensing data within the Nitroeurope project, and it is a contribution to landscape N budget evaluation. This is now clearly described in the introduction of the paper. Results in the paper showed that no single method worked equally well for a diverse set of landscapes. In order to improve the reliability of landscape based N estimates, the best remote sensing methods were therefore applied for separate each landscape, and the results from using such an approach are shown in Fig. 6.

a) In order to prevent that Fig. 6 is considered a validation test of the remote sensing methods, the paper has been restructured to clearly separate results and discussions related to the two objectives of the paper which are 1) “to assess the capability of selected remote sensing methods to quantify LAI, CHL and N over a large range of environmental conditions” (main results in Tables 4 and 5), and 2) “to assess the distribution and size of vegetation N pools in the five European agricultural landscapes” (results in Fig.'s 6, 7 and 8). This is accomplished by showing and discussing results in different sub-sections. In the sub-section on remote sensing capability, it is discussed that REGFLEC provides the best overall results when considering all landscapes (as already discussed and concluded in the paper). In sub-section 2, the reason for using different (best) methods for the individual landscapes is clearly explained by reference to objective 2 (landscape N assessment). Furthermore, the correlation coefficients expressing the agreement between field data and remote sensing based estimates for all landscapes are removed from Fig. 6, and the calculation of Model Efficiency (ME) has been removed. The purpose of providing the correlation coefficients and ME was to quantify the degree of confidence with which N could be mapped in the landscapes, but we realize that this presentation has been confused with a validation test. This would clearly be inappropriate since different methods were applied. We thank the referee for this important remark and the opportunity to correct this misunderstanding.

The empirical regression equations quantifying the relationships between field data and SVIs were not shown in the paper because we do not expect these to be generic (not transferable to other satellite images due to variability in environmental conditions, such as variable soil moisture which impacts background reflectance). We expect that field data will always be required to build empirical regression models for LAI, chlorophyll and leaf nitrogen estimation using SVIs, or alternatively, canopy radiative transfer

models can be used to build look-up tables relating spectral reflectance and SVIs to simulated LAI and chlorophyll (such an approach is included in REGFLEC and makes it independent on field data for calibration).

b) Based on the request from both Referee # 1 and Referee # 2, the empirical regression equations are included in the revised version of the paper. However, it is explained in the paper that we don't expect these equations to be generic, and we propose to include the equations in an appendix.

c) We agree that a graphical representation of Table 4 is worth showing graphically, and histograms were constructed for visualization of these results (please see also reply # 2 and # 7).

2. In both Fig 6 and Tables 4 & 5, only canopies without strong vertical profiles in chlorophyll are included. However, from optimisation theory, an exponential decline in N is expected, so this should be the norm rather than the exception. These points should be included, perhaps displayed with a different symbol in the graphs, and results of analyses shown with/without these points. Perhaps plot deviance between observations and SVI-estimates against the slope dChl/dh.

REPLY a) The data representing LAI, CHL and N of canopies with strong vertical profiles are now shown as smaller dots in Fig. 6. Since the paper is already very long, it was decided not to show regression results based on canopies with strong chlorophyll profiles for individual landscapes, but results for all landscapes (with and without non-uniform canopies) are shown graphically. Furthermore, all data will be uploaded as supplementary material, including measurements in non-uniform canopies.

b) Implementation of a criterium to remove canopies with strong chlorophyll gradient profiles (statistical significant regression slope) caused correlations between field data (LAI, CHL and N) and remote sensing estimates to increase, however a gradual transition from “uniform” chlorophyll profiles to canopies with “strong gradient” profiles is not easily observed or quantified. In this study, it was chosen to identify canopies with strong profiles using statistical significance as criteria. In fact, the chlorophyll structure within canopies is quite complex and very difficult to categorize. Even when regression slopes are insignificant and canopies therefore defined as "uniform" in this paper, canopies may still include large chlorophyll data variability (and structure) because many of the chlorophyll profiles are bell-shaped. Canopies with bell-shaped chlorophyll profiles do not have a strong chlorophyll gradient (slope) profile and are included in our reduced validation data sets as "uniform" canopies. This is clearly described in the revised manuscript.

It may be that other criteria or post-processing methods could also be useful, however many attempts (not shown in the paper) were already made to categorize the canopies and to average field data in various ways to evaluate the relationships between field data and remote sensing based results. The best results were obtained when disregarding canopies with strong CHL gradients. In order to remove these canopies, objective criteria was used. If others are interested in studying and working with the data (using the same or other criteria), they would be very welcome, and it has therefore been decided to upload data (incl extracted SPOT spectral reflectance data) as supplementary material.

3. Throughout the results, discussion & conclusions, the authors seem to have reversed the logic of the validation test. They say that by removing canopies with strong vertical gradients and horizontal clumping, or by focussing on a single land use type, they can improve the predictive capability. In any real application, we want a method that is generally applicable, without any a priori knowledge of the canopy/surface type. The key point here is that the predictive capability declines as the range of canopy/surface types increases. This is an interesting and publishable result, which merits analysis that is not presented here, and the authors seem compelled to put a very positive interpretation on the comparison, to an extent that could be misleading.

REPLY What has been regarded as a positive interpretation, was meant to be an objective description of statistical results which includes explanations that the remote sens-

ing performance improves when separate landscapes or separate land use classes are studied. Clearly, from a remote sensing technical perspective (objective 1 on remote sensing capability), it is disappointing that none of the methods turned out to be best suited for all individual landscapes (REGFLEC did best when evaluated jointly for all landscapes), despite the detailed field sampling of ground-truth data which considered vertical CHL variability. However, from an ecological perspective (objective 2 on assessing spatial variations in leaf N pools in five landscapes), it was regarded as a positive result that the application of different remote sensing methods could explain 94 %, 70 % and 76 % of observed variations in LAI, CHL and N in five landscapes (using "uniform" canopies for validation). The remote sensing based estimates were averaged for the landscapes and found useful to "upscale" field measurements to landscape scale.

We do strongly agree with the viewpoint of the Referee that we want remote sensing methods to be generally applicable. In particular, we had very large expectations of the REGFLEC approach (we still have, but some limitations were identified for operational application, as discussed in the paper). In the revised version of the manuscript, results and discussions are clearly separated for objectives 1 and 2 to avoid any misleading over-optimistic interpretation on remote sensing capabilities.

Regarding the removal of canopies with non-uniform chlorophyll profiles, it is not suggested in the paper that remote sensing methods can only be applied to fields with uniform CHL profiles, but we find that it is important to consider vertical CHL variability when sampling data for evaluation of remote sensing methods. This viewpoint is clearly communicated in the revised paper.

4. Could the CORINE land cover data be used as an additional data input? The methods could be calibrated for each CORINE land cover type separately, which seems to work better. This is using prior knowledge that is easily available, but perhaps the spatial scales don't match well.

REPLY Yes, it might for some land cover types, however CORINE does not provide information on land use changes between years due to crop rotation. But in general, both land cover maps and soil maps would be useful to build regression models for different land cover classes (with possibly different soil backgrounds). This is also exploited within REGFLEC. REGFLEC estimates soil and vegetation model parameters for different land use classes by extracting spectral reflectance data for dense canopies within each land use class and for bare soil within each soil class. REGFLEC then uses a canopy radiative transfer model to build look-up tables representing the relationship between spectral vegetation indices and simulated LAI and leaf chlorophyll for different land cover classes. More information is included about REGFLEC in the revised version of the paper.

5. Something more imaginative might be done with the SVIs. Could they by combined in some multivariate way (PCA, EOFs, neural nets, etc.) to get the best from each algorithm. This could be applied to the raw data or or the SVIs themselves.

REPLY The main aim was originally to explore the utility of REGFLEC across a large range of environmental conditions, and a number of selected SVIs were included at a later stage for comparison with REGFLEC performance. In the revised version of the paper, more focus has been put on REGFLEC, as also suggested by Referee # 2 in Comment # 6 and by Referee # 1. However, many other methods could be explored, but the current paper is already very long. It is proposed to upload field data (and extracted remote sensing data) as supplementary material for free access, if others would like to study and/or use the data.

6. It is interesting that REGFLEC seems to perform the best when the data are pooled, but not when applied to a single land cover type. This is presumably because the algorithm can account for some of the complexities that different canopy structures and back ground surface introduce. Given that this algortihm has many extra parameters, it should of course perform better, but analysing the circumstances where it does might merit some more discussion.

REPLY Yes, REGFLEC is generally performing best when a large range of vegetation covers is present in the landscapes. The discussion on the performance of REGFLEC has been extended.

7. Some more thought might be given to evaluating the performance of the SVIs. Presenting Tables 4 & 5 graphically might help, as scatterplots or Taylor diagrams (see DOI: 10.1029/2000JD900719). What relative importance should be given to r2, RMSE, ME or other measures, in this context? It needs to be made absolutely clear whether any calibration was done on any of the SVIs or REGFLEC model - are all the comparisons tests on independent data, so true predictions?

REPLY a) We thank the referee for the reference to the interesting paper by K.E. Taylor on “Summarizing multiple aspects of model performance in a single diagram”. The graphs shown in this paper are very illustrative, but unfortunately we do not have access to software for presenting such plots. However, Tables 4 and 5 can be visualized as histograms (please see figure below/attached).

b) The calculation and discussion of ME has been removed because the estimates are based on different methods (please see reply 1). Generally there is a near-linear (negative) relationship between r2 and rmse for individual vegetation types and landscapes which simplifies the selection of the best methods (however, these relationships are different for different land uses and landscapes, as illustrated in figure below). Due to the linear relationships between r2 and rmse for individual land use types and landscapes, only r2 results are shown graphically (because the paper is already very long), but all data and graphs are uploaded as supplementary material, and the rmse’s are also (still) shown in Tables 4 and 5.

c) The relationships between SVI’s and field data are based on calibration, and no independent validation was carried out. Many other studies gave empirical evidence for close relationships between SVI’s and vegetation properties. REGFLEC is not calibrated. In the revised version of the paper, this is clearly communicated. Based on recommendation from Referee # 1, more focus is on REGFLEC in the revised paper.

8. The authors mention the mismatch in spatial scales common when comparing satellite data with ground-based observations. The problem is less here because they have 10 or 20 m resiolution data, but still exists. At the UK site, they are comparing field measurements in a 0.25 m2 area with a 400 m2 pixel, with no obvious strategy to bridge that gap. The power analysis should demonstrate that they have enough field samples, such that the uncertainty around the mean for the 100 or 400 m2 pixel is less than the precision on the satellite-derived estimate. Can they show this? The power analysis as it is applied seems to be used for outlier detection, by comparing with "error margins", but this seems unfathomable to me.

REPLY a) Field measurements were generally done in fields which were deliberately selected due to large homogeneity. The variance of SPOT reflectance data has been calculated in 3x3 pixels centered at the geo-referenced field plot and for the complete field. These calculations have been included in the supplementary material and are reported in the paper to document the homogeneity of sampled fields.

b) The power analysis is applied only to test the adequacy of SPAD index sample sizes due to the very large variability of these data in some fields. Power analysis is often used to calculate the required sample size to represent data variability. The method was used in this study to objectively identify canopies where data variability is not sufficiently represented by the number of measurements. It was decided to use this criterium to remove (complete) field samples which were not adequately sampled instead of removing selected outliers. The problematic canopies turned out to represent mainly grass fields with small leaves. It is expected that the very high data variability in these fields may be caused by the SPAD sensor head not being completely covered by the small leaves (as explained in the paper). We believe that this processing step is necessary to ensure good data quality.

c) Replicate LAI measurements were made. Generally, within-field variability in LAI

measurements is low (though higher for UK grassland) while the variability in SPAD indices is quite high (as also shown in Table 3 in the paper). In the revised paper, uncertainty of LAI, canopy chlorophyll and N field data was evaluated based on available field data. For canopy chlorophyll and nitrogen, uncertainty is calculated in different measurement levels (strata) and propagated to canopy scale (thoroughly described in the reply letter to Referee # 1).

9. Can the location of the field measurements be shown explicitly in Fig 1? I get no sense of how large or widely distributed the field plots are in relation to the satellite pixels.

REPLY a) Farmers allowed us to work in their fields provided that the measurements could not be traced to their land. Therefore the locations are not shown. The variance of SPOT reflectance data was calculated to quantify field heterogeneity, and it is reported in the revised paper.

10. Precisely what is meant by the "soil line" needs explaining clearly, or removing. This journal has a general readership, and remote sensing jargon should not be used. Generally, the language needs tightening up in places, e.g. "Predictability" is used to mean "predictive capability".

REPLY A parenthesis “(the so-called “soil-line”)” has been inserted for clarification, so that the sentence now reads:

“Furthermore, the incorporation of local soil parameters, describing the linear relationship between red and nearinfrared (NIR) reflectance of bare soils (the so-called “soil-line”), improve the estimation of canopy “green-ness” (related to the product of LAI and CHLl) from SVIs (Broge and Leblanc, 2000). Since the empirical “soil-line” parameters depend on both the. . .”

We hope this improves understanding for the soil-line concept.

The term predictability is no longer used. Instead we use the term “remote sensing capability". The reason for not using the term "predictability" is that the methods are not applicable for future forecasts.

11. p 10169: "s" is not defined I think. Slope of dChl/dh I think.

REPLY "s" is defined in the beginning of section "4.3 Within-canopy variations" (p 10169, l. 4)

12. The "statistical significance" of relationships is repeatedly quoted, with p values. This implies a test of the null hypothesis "no relationship between SVIs and the surface properties", which seems irrelevant to me. I'd suggest removing all references to null hypothesis testing.

REPLY We find that it is important to indicate when correlations are statistical significant (in particular due to different sample sizes), and that it is normally done by using the expression "($p<0.05$)" which then also includes information about the confidence level. In the revised paper, it was decided to test all correlations at 5 % confidence level, as we find that the reference to different confidence levels in the submitted paper may have triggered this comment from the Referee.

13. Table 3 caption is very confusingly worded. I think they just mean they calculate the variance at different strata, in ANOVA terminology.

REPLY The text, "Standard deviations of leaf scale measurements are shown to represent between-field variability (sd1) which is the "standard deviation of the mean CHLl or Nl of different fields" and within-field variability (sd2) which is the averaged "CHLlor Nl standard deviation of data measured in individual fields"

has been reformulated to

"Standard deviations of leaf scale measurements are shown to represent variability between field canopies (sd1) and the mean variability within field canopies (sd2)".

We hope this improves readability.

14. Fig 4 - the z-axis needs to be explained clearly. I think this is a histogram in two dimensions, with colour scale showing the number of pixels. Axes need units.

REPLY Axes are unitless (reflectance). This is clearly indicated on axes in the revised version of the paper. The figure shows a "density scatter plot" of near-infrared reflectance (NIR) versus red spectral reflectance (red). It is like a normal scatter plot, but with colors indicating data density.

15. Fig 6 might have error bars in the x- (and y?) dimension.

REPLY The Referee also requested to show data from canopies with non-uniform chlorophyll profiles in Fig. 6, and it was decided to include these data in the figure. In our opinion, it is not possible to display more information in this figure. Uncertainty assessment of the field data have however been assessed, reported and discussed in the revised paper. The standard deviations of the field data are (still) shown in Table 5.

16. Fig 8. Why not show the distribution of the field observations as well, rather than just a single point?

REPLY At first, it was included but difficult to see due to overlap with the remote sensing standard deviation bar. Because both the mean and the standard deviation of the observations are already shown in Table 3, only the mean was included in Fig. 8. However, based on the question, it has been included in the figure again.

17. To be clear, conversions between leaf-scale and canopy-scale nitrogen and chlorophyll contents should be made explicit. Is it simply Nc = Nl x LAI, or are different canopy strata integrated separately?

REPLY It is simply Nc=Nl x LAI. This is now clearly explained in section

Interactive comment on Biogeosciences Discuss., 9, 10149, 2012.

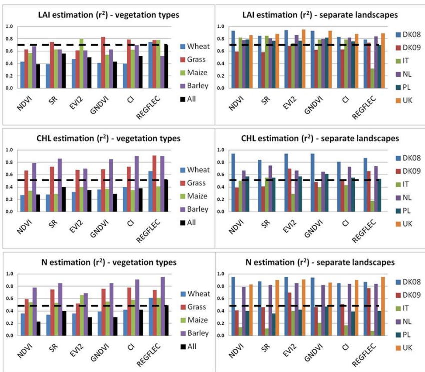

Figure1. Squared correlation coefficients of relationships between remote sensing methods (NDVI, SR, EVI2, GNDVI, CI and REGFLEC) and field measurements (LAI, total chlorophyll, total leaf nitrogen) for different vegetation/land use classes (left) and different landscapes (right). The broken horizontal lines show the performance when evaluated jointly for all landscapes.

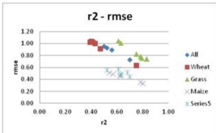

Fig. 1.