Biogeosciences Discuss., 6, S873–S876, 2009
www.biogeosciences-discuss.net/6/S873/2009/

Biogeosciences Discussions

Interactive comment on "Influence of landscape heterogeneity on spatial patterns of wood productivity, wood specific density and above ground biomass in Amazonia" *by* L. O. Anderson et al.

F. Gerard (Referee)

ffg@ceh.ac.uk

Received and published: 20 April 2009

The authors present work which tests the representativeness of field plots established in Amazonia, or in other words whether the location of the plots has introduced bias in the estimates of regional forest parameters. The test basically consists of producing an EO derived forest type map from ETM imagery which include the field plots and use this map to produce a weighted regional average of the estimates of interest. Little or no difference between the weighted regional averages and the original regional averages are the authors' measure of representativeness or absence of bias.

I am not trained as a statistician, however I have issues with the manner in which 'absence of bias'; and 'representative samples'; are being tested and discussed. I suggest the authors ask for advice from a statistician.

Comment 1, bias: From a team of experts who are currently developing a sampling strategy for biodiversity monitoring in Europe I learned that the most popular approach for ensuring 'representative' sampling is by means of stratified random sampling. The stratification is often introduced to reduce within stratum variability and strata are also used as the reporting unit.

It is clear that the RAINFOR sites and the sample plots were not positioned following a stratified random sampling scheme. So as the authors correctly observe, there is a good chance there may be bias. From wikipedia: 'A biased sample is a statistical sample of a population in which some members of the population are less likely to be included than others. If entire segments of the population are excluded from a sample, then there are no adjustments that can produce estimates that are representative of the entire population. But if some groups are underrepresented and the degree of underrepresentation can be quantified, then sample weights can correct the bias.'

So basically to test bias the authors, should check (1) if some members of the population are less likely to be included then others and (2) if entire segments of the population are excluded. The next step, if no segments of the population have been excluded, could then involve a correction of the bias by quantifying the degree of underrepresentation and weighting the samples accordingly.

The population here are the forests of Amazonia. The reporting units (for which estimates are being produced) which could also be interpreted as one possible stratification of Amazonia are: eastern, middle and western Amazonia. No longitudinal coordinates are given to identify the boundaries between these regions but the ownership of the sample plots to the regions is given. Based on that information (Table2): 25 plots (3 sites) in the western, 16 plots (1 site) in the middle and 4 plots (1 site) in the eastern

BGD

6, S873–S876, 2009

region, I would say the sample is biased unless the heterogeneity of the regions and thus the variability of the reported parameters (biomas, wood productivity and wood density) increases from east to west. Also as a direct consequence of using one 10km x 10km EO window per site the number of windows used to represent these regions (3, 1, 1) are introducing a bias! This bias could be corrected by weighing the EO windows with the area covered by the three regions they represent but no boundaries are given for these regions.

Comment 2, stratification: How are the boundaries of eastern, middle and western Amazonia actually being defined. I suspect in the mind of the authors the distinction is more complex than a longitudinal coordinate. So which other climatic, geographical, topographical etc... characteristics are being used ? The authors mention soil type, inundation patterns, and elevation. Why not produce such a stratification (however crudely it may be) to revisit the bias question?

Comment 3, forest types and EO: If forest type has an impact on the variability of the reported parameters, then ideally the EO windows which are used to map the forests should be randomly placed within this biogeographical/environmental stratification to correctly capture the degree of forest type heterogeneity within a stratum. I realise this is a though proposition for the Amazonian area if the aim is to accurately map forest type classes identified from in situ knowledge. However if the only aim was to capture the spatial variability in forest reflectances using the hypothesis that in (i) tropical environments the spectral characteristics of a vegetation type is strongly correlated to the physiognomic characteristics of that vegetation type and (ii) differences in physiognomies have a strong relationship with the reported parameters, then an intelligent unsupervised classification should be enough to capture within stratum variability and assess bias. It is not clear which of the two options was implemented.

Comment 4, Two main forest types ... and the rest?: The authors only focus on the mapping and distribution of two main forest types and the variability (spectral, elevation, soil) within these types. What are the statistical implications of not considering the

BGD

6, S873–S876, 2009

Interactive Comment

whole 'forest' population and how is that incorporated in the weighted average calculations?

Comment 5, class definitions and classification systems: The authors propose to focus on the spatial distribution of two forest types: Terra Firme and Alluvial Terrian forests. The criteria used to define such forest types can vary depending on which classification system is used (floristic, physiognomic, ecological). It is really important to specify clearly the class criteria used in this work to ensure the readers and authors hold the same forest image in their minds. Labeling spectral classes using forest classes which are not clearly defined leads to confusion and misunderstandings. Table 3 confirms this, listing the worst type of class definitions I have seen for some time: 'alluvial forest physiognomy with the spectral property... however without inundation periods also located in Holocene/Pleistocene alluvial formations'. For example, what is meant by 'alluvial forest physiognomy'?

Interactive comment on Biogeosciences Discuss., 6, 2039, 2009.