Biogeosciences Discuss., 6, C1902–C1907, 2009
www.biogeosciences-discuss.net/6/C1902/2009/

Biogeosciences
Discussions

BGD

6, C1902–C1907, 2009

Interactive
Comment

Interactive comment on "Towards global empirical upscaling of FLUXNET eddy covariance observations: validation of a model tree ensemble approach using a biosphere model" *by* M. Jung et al.

M. Jung et al.

mjung@bgc-jena.mpg.de

Received and published: 9 September 2009

We appreciate the positive and constructive comments by the two anonymous reviewers that help to improve the manuscript. Here is our response:

The main concern of reviewer I is how we have dealt with the concept of scale and scaling, emphasising the role of different processes of different scales. Reviewer 1 also misses mechanistic explanations and the role of feedbacks while reviewer II disagrees with reviewer I here. We think that there is some misunderstanding here regarding the

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

meaning of ‘upscaling’ that needs to be clarified in the paper. We regard upscaling as the process of generating spatial fields from point data, and injected this into the introduction. Reviewer I criticism regarding the lack of explicit treatment of processes and feedbacks misses a bit the point that machine learning tools are empirical methods that generate functions from the data without injecting explicit process knowledge. Thus, these tools are independent of theoretical assumptions and thus also fully independent to land surface process models.

Reviewer I also made the point that we emphasise the case of extrapolation which in his opinion is only part of the upscaling. Reviewer II suggested clarifying the terms extrapolation, interpolation etc. We think that this clarification is crucial since the criticism of reviewer one seems to originate from a different understanding of ‘upscaling’ and ‘extrapolation’, and so we introduced a paragraph in the introduction:

“Upscaling exercises of eddy covariance based carbon fluxes to large regions has been conducted for the US (Xiao et al., 2008, Yang et al., 2007) and Europe (Jung et al., 2008; Papale and Valentini, 2003; Vetter et al., 2008), which are both characterized by a comparatively dense network of towers. The upscaling principle generally employs the training of a machine learning algorithm to predict carbon flux estimates based on measured meteorological data, remotely sensed vegetation properties, and vegetation type. The trained model can then be applied spatially using grids of the respective input data. Upscaling generally involves both, interpolation and extrapolation. We refer to interpolation when fluxes are predicted at locations whose environmental characteristics are captured by the training data set. Extrapolation occurs if fluxes are predicted for environments, which are not present in the training data set. It is important to note that it is not necessarily the geographical space which determines if inter- or extrapolation takes place but the environmental space. In our sense, an example for interpolation would be where ecosystems from the northern hemisphere may be used to predict carbon fluxes of structurally similar ecosystems in the southern hemisphere. Extrapolation may happen, if for example data from temperate coniferous forests are used

BGD

6, C1902–C1907, 2009

to predict the response of temperate grasslands which are geographically nearby but structurally different. In practise, the distinction between inter- and extrapolation can be more fuzzy if it is not exactly known what determines structural similarity or if important characteristics are not known. For example, let us assume that data points from forests on shallow and acidic soils are used to estimate the behaviour of forests on deep and fertile soils which is different. If the soil information is present and if we know that soil is important then we would call it extrapolation, if not we would think of interpolation. This is an example of 'hidden extrapolation', i.e. where predictions are made for conditions that are not sampled by the training data (here different soils) although the measured characteristics are captured by the training data (e.g. same climate and vegetation type)."

Reviewer I asked for pseudo code of our model tree algorithm and was concerned if it could be reproduced from the paper by interested readers. We had delivered pseudo code in the supplementary material which reviewer one apparently missed. We have now inserted some more references to the pseudo code in the supplementary material.

Reviewer I made a very good point regarding interannual variability and lag effects and how they could be better represented. We have inserted a discussion on that in section 3.1:

"However, if the interannual variability is of particular interest it may be possible to further improve the performance of MTE for interannual variability by training directly on the anomalies, i.e. using temperature, precipitation, and FAPAR anomalies etc. to predict carbon flux anomalies. This approach would solve the first two problems that the factors controlling interannual variability can differ from those determining the spatial and seasonal gradients, and that only little emphasis is normally given to reproducing the interannual variability due to its small contribution to the total variance. Further improvements may include the addition of variables with lag (e.g. precipitation anomaly of the previous one, two, or three months) or cumulated variables such as temperature sums or cumulative water balance indicators as proxy for soil moisture as additional

BGD

6, C1902–C1907, 2009

explanatory variables. Such additions would enable to describe memory effects, i.e. effects of past conditions on the current fluxes."

Reviewer I made an interesting question how Canadian and European towers help constrain Siberian fluxes and if this is a feature of the model tree approach or LPJ. We inserted a brief discussion on that:

"Here the strength of the model tree approach is illustrated because data from ecosystems that function similarly are identified via the stratification and can be used to predict similar ecosystems that are geographically far away. However, the good performance in Siberia will be also related to simplifications made in LPJmL to some extent, such as one general parameter set for boreal forests."

Reviewer II made a very good suggestion to introduce an improved discussion on how the conclusions of the synthetic test case can be transferred to the actual FLUXNET upscaling. We inserted a new section (3.3.) on that:

"3.3 Remarks regarding FLUXNET upscaling

In this section we discuss briefly the meaning of our synthetic test case for real FLUXNET upscaling projects and propose additional steps that can be taken to further study and improve FLUXNET upscaling. The primary objective of this paper was to introduce the method of model tree ensembles and an evaluation of its efficiency to derive spatial and temporal fields from highly clumped and irregularly spaced data like FLUXNET. The presented test case using a biosphere model as surrogate truth is a necessary first step to gain confidence in the technique. If MTE would have failed to adequately reproduce the LPJmL simulations from the flux tower locations, it would not be worth applying it using real world flux data. We suggest that our approach of testing the method of empirical upscaling flux tower data to continents and the globe should become a required standard.

The fact that MTE could reproduce the global LPJmL simulations very well does not

BGD

6, C1902–C1907, 2009

Interactive Comment

proof that the real FLUXNET upscaling products using MTE will generate global carbon flux fields of comparable accuracy. Most importantly, the biosphere is more complex than LPJmL, which uses a relatively small number of plant functional types associated with constant sets of parameters to discretize global ecosystems, uses necessary simplifications of physical and physiological processes, and lacks other processes that can be important such as nutrient cycles. The artificial experiment using the biosphere model also represents a 'perfect world' without any uncertainties in the data while data uncertainties are present (and an issue) in both, the flux tower data, and driver data such as grids of meteorological fields and satellite based estimates of fAPAR. The current study is also simpler in the sense that we use training data at 0.5° resolution to predict at 0.5° resolution globally. In reality, the training data (meteorology and fluxes) are measured at the towers plus time series of satellite fAPAR products of 1 or 2 km resolution. On top of the uncertainty of all these measurements there is additional uncertainty originating from a mismatch between the footprints of the individual instruments at the tower and of the satellite. This footprint mismatch introduces additional noise. The role of the data uncertainties could be assessed for example by adding noise and bias to the (simulated) training data that are comparable to those inferred from studying the uncertainties of eddy covariance data (Lasslop et al., 2008; Richardson et al., 2006). This approach would be effective to evaluate how well the machine learning tools are capable of extracting general relationships from the noisy data or tend to overfit. If successful, machine learning tools could be used to assess the information content and signal to noise ratios of real world data. Clearly, additional confidence of FLUXNET derived upscaling products is required by corroboration against independent data."

We also introduced some more information on how TRIAL works in section 2 as suggested by reviewer II:

"The key question of model tree algorithms is how the best split is found for a given node. Since an exhaustive test of all possible splits is often computationally impractical, a smart subset of possible splits is evaluated by computing multiple linear regressions for the left and right child for each tested split location. Subsequently, the location where the joint error of the model in the right and left child is minimal is chosen as the best split of a node. The search for the best spilt of a node is necessarily different for continuous and categorical variables and the next two sections describe the individual strategies in more detail. We refer the interested reader also to the pseudo code in the supplementary material."

Reviewer II also made a valid point by noting that some more information regarding the cross-validation is needed in section 2 and we added this accordingly:

"The cross-validation operates in the leaves of the tree and thus provides an assessment of the robustness of the multiple linear regressions with selected regression variables (see below). The MSE of the tree is calculated by adding up the sum of squared errors (SSE) from the cross-validation of all leaves and then dividing by the total number of data points."

We followed the call by reviewer II to better clarify terminology such as extrapolation as this can be interpreted differently. We introduced that in the introduction. See our response to reviewer I also.

We also corrected the trivial changes indicated by reviewer I and II.

Interactive comment on Biogeosciences Discuss., 6, 5271, 2009.

BGD

6, C1902–C1907, 2009

Interactive Comment