

## ***Interactive comment on “Spatial distribution of soil organic carbon stocks in France” by M. P. Martin et al.***

**Anonymous Referee #2**

Received and published: 20 December 2010

This paper is an analysis of the French soil monitoring data to predict soil carbon stocks on a 12km grid – rather than the 16km original sampling. I am concerned that by relating each point to a number of soil, landuse and climate variables in space and then treating this dataset as non spatial they are throwing away a lot of information. They have analysed this dataset using Boosted Regression Trees. One of the drawbacks of regression trees is that the dependant variable, in this case organic carbon stocks, will be put in bins so the size and range of those bins will influence how good the predictions will be. I suspect the authors only used 10% of their data to validate the models as using more would have shown how poor the model is. I am concerned that they appear to be overfitting – something that boosted regression trees are well known for. Had they used some of the variables to fit a General linear model and then used some kind of spatial residuals they would have been able to use the spatial structure

C4450

of the data to predict and use the actual data points to control that prediction as well as the underlying variables on the smaller grid. I would have thought an approach such as Marchant et al 2010 – which applied a robust spatial prediction method to different soil properties measured in the same survey would have generated a more robust methodology. The actual variables that are included in the BRT are not well described and I find it hard to understand how they generated all these variables across the whole of France. In particular what is the difference between wregime and wlogging? The general standard of English needs to be improved throughout the paper. I think the authors have a really good dataset from which it should be able to predict the SOC stocks for France but I think the paper needs major revision before it can be published

---

Interactive comment on Biogeosciences Discuss., 7, 8409, 2010.