Biogeosciences Discuss., 7, C1335–C1337, 2010
www.biogeosciences-discuss.net/7/C1335/2010/

Biogeosciences Discussions

Interactive comment on "The use of machine learning algorithms to design a generalized simplified denitrification model" *by* F. Oehler et al.

Anonymous Referee #2

Received and published: 8 June 2010

The presented manuscript "The use of machine learning algorithms to design a generalized simplified denitrification model" describes the use of two machine learning algorithms to design a simplified denitrification model, which is a new and interesting approach in this field.

However, to use the ANN model not only as a black box but as a generalized simplified denitrification model and interpret its response shapes, requires further work in proving the generality of the obtained model. Therefore, I would like to ask the authors to address the following three comments:

- Inconsistent choice of methods: Artificial neural networks (ANNs) and boosted regression trees (BRT) were used to determine the configuration of the final ANN model. These two machine learning algorithms present very different mathematical

Full Screen / Esc
Printer-friendly Version
Interactive Discussion
Discussion Paper

approaches. An influencing variable in a BRT is one that minimizes the variance of the output and fulfils a splitting task. This variable might not have a similar importance to present a (non-linear) response in an ANN with a minimized root mean square error. There are several methods to determine the input relevance of variables directly for ANNs (Gevrey, M., et al. (2003). "Review and comparison of methods to study the contribution of variables in artificial neural network models." Ecological Modelling 160(3): 249-264.). Please discuss this inconsistency.

- Missing evidence of the generality of univariate response functions: The univariate response shapes presented in Figure 8 are only ONE very special case of the hyperspace ANN(T,WFPS,NO3,OM,pH), with the values of the other variables fixed. For example, if the value of WFPS would have been fixed not to 100% but to 60%, the response of OM in Figure 8 D would have shown a very flat response instead of a steep increase according to Figure 9 D. Therefore these results should be discussed as a very specific cases of the fixed values, or evidence should be provided that the found univariate responses bear any generality beyond the presented fixed values.

- Missing evidence of the generality of the coupled response functions: The same argument is true for the different couples of response in Figure 9. Would these surfaces still bear the same features, if for example the pH would have been fixed at 5.5 or the WFPS at 40%? Which features would hold, which features change? Again, please provide evidence that the response surfaces presented in Figure 9 have relevance beyond the fixed values of the input variables before discussing their implication.

- Missing evidence of the generality of the simplified model: The more input variables are used, the more freedom is in the ANN mapping and the more likely will the ANN model have ecologically implausible mapping. Does an ANN with only the three major variables as inputs ANN(T,WFPS,NO3) exhibit the same univariate and coupled responses? Does an ANN with a different set of five input variables (e.g. ANN(T,WFPS,NO3,DEA,pH)) still shows the same features of the Da rates as in Figure 9? How general and plausible is this simplified model?

BGD

7, C1335–C1337, 2010

Furthermore, the writing style of the manuscript could be improved for clarity and consistency. For example, rather than jumping back and forth between the three methods, the calibration section could follow the description of each technique, and the headers in the results section could be more meaningful, e.g. almost the whole chapter describes "ANN results", not only 3.2. The abstract is difficult to read, again jumping back and forth between ML, BRT, ANN, and Nemis. The first two sentences of the abstract have a repetitive start of “we designed”, rather than the overall focus and goal of this paper. Please quote the list of abbreviations.

Overall, the presented paper has the potential to provide a new alternative to design a simplified denitrification model. However, the generality of the presented simplified model remains to be proven and I am looking forward to the revised manuscript.

Interactive comment on Biogeosciences Discuss., 7, 2313, 2010.