Biogeosciences Discuss., 10, C2259–C2261, 2013
www.biogeosciences-discuss.net/10/C2259/2013/

Biogeosciences Discussions
Open Access

Interactive comment on "Testing the applicability of neural networks as a gap-filling method using CH_4 flux data from high latitude wetlands" *by* S. Dengel et al.

Anonymous Referee #2

Received and published: 24 May 2013

The manuscript presents a challenging study on gap-filling of incomplete time series of methane flux measurements. Gap-filling of flux time series is of practical and scientific importance in understanding the gas exchange balance between the atmosphere and ecosystems, in particular for methane. Since methane is a compound which exchange cannot be easily predicted from process-based knowledge, statistical models have been frequently used to establish functional relationship between the controlling environmental variables and methane exchange. Therefore, neural networks serve as potential and novel approach to gap-fill the methane time series at about hourly time scale by using information on environmental variables.

The manuscript provides overview about the current knowledge on methane emission

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

drivers and studies as well as basis knowledge on neural network model and data processing. Although neural networks perform as “black box model”, the authors make also good effort to interpret the results in terms of functional dependencies. Neural networks show high performance in predicting the missing flux data from high latitude wetlands based on basic environmental variables. Surprisingly, also high variability and extreme flux events are predicted by the model. Such events are not characteristic to carbon dioxide and water exchange, which makes methane flux time series gap-filling much more challenging.

As demonstrated in the manuscript, the neural network approach shows high potential in becoming widely accepted tool in gap-filling of methane flux time series and other flux series with similar sporadic nature. I can recommend it for publication in ACP after considering the items below.

Specific comments and questions

1. The environmental variables as well as 4 fuzzy sets representing the time of day were used as drivers for predicting CH4 emissions for six sites. It is evident that many environmental variables are highly correlated with time of day, which by itself should not be a direct driver for CH4 fluxes. What is the main reasoning including time of day via those 4 fuzzy sets? The issue is explained in section 2.3 (that it is useful to include fuzzy sets to represent diurnal and seasonal variation to reduce cumulative weight of time) but it does not become clear why it is important to include time of day and if and how much it improves predictability.

2. P.5, l.23 explains that eddy covariance data from 3 sites were filtered according to u*. Please be more specific about the criteria of filtering because Fig. 3b GAM plots seems to indicate wide range of friction velocity values for these 3 sites. Also because at many sites turbulence conditions indicate the state of coupling of atmospheric layers with surface emissions.

3. Consider moving section 2.4 after 2.5 because readers not closely familiar with NN

BGD
10, C2259–C2261, 2013

method might get misleading impression that the statistical analysis described in that section is part of NN model.

4. P.10, l.2, please make easier to understand for the reader the meaning of scenario with short explanation. Text refers to Figures 4 and 5, which do not directly reveal what are the 3 scenarios per gap length, i.e. in total 3 x 5 gap length scenarios = 15 scenarios in total? What is the meaning of "several neural network iterations" in the figure captions?

5. P.12, l. 26-27, what do the referred ranges mean for the confidentiality of the prediction, is it the 95% confidentiality range?

6. General question: are all plots in Figures 4 and 5 informative? Consider reducing and keeping only a few plots as the best examples. In turn it would be interesting to summarize and possibly discuss more different functional dependencies of fluxes at 6 different sites as revealed by the statistical analysis and GAM (as presented by Fig. 3).

Technical comment

I wonder if equation (1) is strictly correct because xi represent the 10 input variables and wi the 4 weights, both carrying the same index i. Should there be double summation? Correct also notations/subscripts following the equation.

Interactive comment on Biogeosciences Discuss., 10, 7727, 2013.