Biogeosciences Discuss., 7, C4042–C4048, 2010
www.biogeosciences-discuss.net/7/C4042/2010/

Biogeosciences Discussions

Interactive comment on "Role of de novo biosynthesis in ecosystem scale monoterpene emissions from a boreal Scots pine forest" *by* R. Taipale et al.

Anonymous Referee #1

Received and published: 23 November 2010

Taipale et al. present a well written report of the measurement and analysis of monoterpene emissions over a Scots pine forest in Finland. The canopy level measurements were taken every three hours from may to august in 2007. The authors analyze the monoterpene fluxes in terms of their temperature and light dependencies, as an analogy to emissions from storage or de novo bio-synthesis, making use of a hybrid emission algorithm. Based on their analysis, they conclude that there is a large de novo component to emissions, and suggest the use of the hybrid model in future modeling studies. The introduction and methodology used are quite clear, and the relevant literature is well cited. The importance of the topic has been recently discussed in the literature. I have some serious concerns about the analysis performed, however. In

particular, the statistical treatment of the raw data is very weak, and many of the main conclusions are either unfounded or based on questionable and untested assumptions. Given the potential for model error, the use of their model approach without any real statistical analysis of the data could be misleading. Hand waving interpretation of figures, 'Tentative indications', and 'qualitative hints' are not sufficient for a good scientific paper, especially if more detailed analysis is possible and has the potential to quantify them. I suspect the manuscript will need a major revision before becoming suitable for acceptance in Biogeosciences.

Concerns:

In order to justfiy the use of a hybrid emissions algorithm, the authors need to first show that there is a PAR dependency in the data (independently of model assumptions). This could be easily achieved applying an ANOVA test. Perhaps it might be better to do this on daily emissions values (excluding night time values?), given that the diurnal cycle of PAR and temperature are highly correlated. Whatever the approach taken, given the potential for model error (discussed below) the authors must provide statistical tests of the importance of light driven emissions (more on this below).

The authors apply the Guenther et al. (1991, 1993) and Guenther (1997) parameters. This is common, especially in large scale emissions studies. The problem is this: The inclusion of a de novo pool term (via Ghirado et al., 2010) essentially changes the temperature dependence of the emissions algorithm (by including a (weak, in this case, judging by figure 4) PAR dependency). So in effect, the inclusion of a de novo term acts to correct the original emissions algorithm for dynamics that it does not predict. But it is commonly observed that the original Guenther et al. parameters do not hold when working at the site and species specific level. In fact, there is not reason to assume that they would, given the large range of leaf structural traits and species specific levels of metabolic activity. So my question is, how can the authors be sure that the improvement gained by changing the temperature dependency of the original algorithm through introducing a temperature and PAR dependent de novo factor, is not just correcting for a misspecification of the temperature dependency of the original Guenther et al. model (when applied to this species and site). Could a reparameterised stored emissions only model do as well as the de novo emissions hybrid model? Different emissions models have remarkably different temperature and light sensitivities (see Arneth et al,. 2008). The Guenther et al. original algorithms were updated in the MEGAN version (Guenther et al., 2006) – including changes in some of the parameters and an improved dependence integrating recent days climate on emissions. Would it not have been more appropriate to use this improved version? I know it was developed for isoprene but if there is a light dependent component there is no reason to believe monoterpenes should behave differently (as is expressed in the light and temperature dependent version of the Guenther et al. monoterpene emission algorithm).

It is not quite clear to me how the de novo component was established (page 8024). The authors allude to a least squares approach to calculate Fsynth (Page 8024 line 25), but only for seasonal changes? No details as to the robustness of the calculations are given. In statistical terms, how successful was the separation? And given the comments above, how would uncertainty in the Guenther et al. original parameters affect this separation?

Page 8023, line 14-16: This is important and should be better explained. It is unclear to the reader how filtering suspect or data points with low confidence would decrease the data quality.

Page 8025, line 7: The statement regarding Canopy shadowing effects is not clear. This brings up the point of how the emission algorithms are related to the forest – it should be made clear that there is no consideration of sun or shade distinction for canopy emissions, or any incorporation of bark biomass (as a proxy for storage pools).

Page 8025, line 18-19: Please clarify what is meant by 'the sensitivity'.

Page 8025, line 20: This is a good example of poor analysis. If there is a correlation between emissions temperature and PAR then it should be quantified.

The statement: "the effects of cloudy periods are hard to discern from the figure" calls for either a better figure, or some quantification of these effects. The reader must be presented with a basis for such statements.

As mentioned before, an ANOVA might help, or a simple regression of Emissions vs. PAR and Emissions vs Temperature. Is there a significant correlation? Another approach could be to filter the data across a restricted range of air Temperature (say 20-25oC) and plot the emissions recorded within this range against PAR (thus removing the temperature effect, giving a clear picture of the PAR response), and vice-versa for PAR (filter emissions for a tight range of PAR, which would presumably include a wide temperature range), and plot emissions against Temperature. Without some more detailed analysis, the statements made here have little scientific basis.

Page 8026 line 10 Most bark is not in the canopy, but on the bole. I suspect you mean above ground biomass?

Page 8026 line 11 This treatment of the data in a statistical fashion leaves a lot to be desired. The authors are mixing diurnal variations with monthly variations. It is therefore not surprising that the resulting picture is not very clear, and no statistical significance could be found. I would suggest using only midday data (say from 1200 to 1500) in order to reduce the standard deviation contribution by diurnal variation. Another approach would be to analyze the data in terms of paired observations – data from days in different months which experience similar climatic conditions.

Page 2026, line 17 Were the monthly medians corrected in any way for the distribution of data gaps? For example, if May had more data gaps during night, whilst July had more data gaps at peak emission times, then your approach would report that the median emissions for July was less than that of May. Again, it is not surprising that there is no statistical difference – noise from other uncontrolled aspects is muddying the picture.

Page 2026, line 24 Are the results of this paragraph worth reporting?

Page 8027, line 18-24 Why did PAR nearly double in June? Reduced cloudiness? This is direct PAR then, what about diffuse? An increase in direct PAR would also be correlated with an increase in leaf temperature (independent of air temperature). These are aspects which should be discussed.

The data treatment here is again somewhat vague. What does "a similar occurrence was observed in the monoterpene emissions" mean. Or " temperature variations had a divergent pattern".? The paragraph gives the reader the impression that emissions were independent of temperature during June, but is not substantiated by any real statistical analysis. See my earlier comment on restricted range analysis to separate temperature and PAR effects. Again, the authors need to quantify the temperature vs. PAR weights independently of their model framework. Qualitative hints are very weak and could be misleading if not backed up by real scientific approaches.

Page 8028, line 20 It is not clear to me what structural changes in bark could explain the decrease in emissions potentials during the summer.

Page 8028, line 27 Please give the confidence intervals along with these values. I.e. 30% (+/- . . .)

Page 8029, line 19-21 This statement is unsubstantiated – the authors really have not shown that the algorithm should incorporate emissions from storage pools. In my opinion they have done quite the opposite, as no real improvement was observed in using the hybrid algorithm (see comments below). In any case, such unsubstantiated text should be made clear as opinion.

Section 3.3 (pages 8029, 8030)

My interpretation of the results are quite different from that derived by the authors in this section. This is in part due to the fact that again no statistical evidence is reported. The authors claim that the hybrid model performs better than the original model, but only a slight difference is observable between the two (Fig. 4) and no assessment seems to

have been done other than visual inspection. True, the hybrid models seems to give somewhat lower median emissions at night, though a reparameterisation of the original model might also reproduce such dynamics. What about June? The authors report that PAR in June was almost double that of other months, but yet the hybrid model gives exactly the same mid-day emissions as the pool algorithm. So the hybrid model is insensitive to variation in daily PAR? There must be something wrong here. Rather than the contribution of de novo emissions on total emissions, a more interesting question is why the June emissions are so high, and neither model gets them?

Given the large standard deviation inherent in the data presented in Fig. 4 (if the authors feel strongly about presenting the data in this format then error bars (model and data) must be included in the figure please), it is unlikely that the hybrid model is statistically better when analyzed in this manner. If the error bars obscure the signal, then the authors must find a better way to analyze the data. How about the r2 or RMSE of the whole time series? Or the independent correspondence of each model and the data with climate variability.

From figure 4 it seems that the hybrid model improved night time emissions simulation, but gave the exact same values for emissions during the day as the pool model. I do not understand how the authors can claim a significant improvement in emission without even mentioning the night-day distinction in model performances. Regardless of all the above, I find the authors conclusion that the hybrid model be adapted questionable. There are rigorous methods for comparing models of differing complexity (e.g. Akaikes information criteria). Such assessments will always invoke Occams razor, selecting the simpler of two models which perform in a statistically similar manner. There is no reason to add biological realism if it does not significantly improve results. Until it can be shown statistically to improve results such suggestions are purely the beliefs of the authors, and should be stated as such. In any case, judging from Figure 4 the inclusion of a PAR dependent de novo component only seems to improve night-time emissions estimates, which i find somewhat suspect.

Conclusions section: (1) Again, I would suggest using mid-day values to get a better handle on trends of PAR and temperature

Page 8031, line 1-2 I would argue that the determination of monthly changes should be done in this study (not future studies) – the large uncertainties which authors claim mask the signal could be easily reduced by careful treatment of the data.

All Figures – please increase the font size of the axis labels.

Figure 1: Please remove data gaps. The authors report ~20% data removed, and some gaps which span a few days, yet the Emissions graph shows a continuous time series. Why include the monthly average temperature for the past 30 years? It does not seem to have been used anywhere in the text or data interpretation.

Figure 4. Please add error bars (standard deviation for each data point) for models and data.

Interactive comment on Biogeosciences Discuss., 7, 8019, 2010.