Biogeosciences Discuss., 7, C3431–C3441, 2010
www.biogeosciences-discuss.net/7/C3431/2010/

Biogeosciences
Discussions

Interactive comment on "Controls on winter ecosystem respiration at mid- and high-latitudes" *by* T. Wang et al.

W. Eugster (Referee)

werner.eugster@ipw.agrl.ethz.ch

Received and published: 22 October 2010

The authors use the La Tuille flux data set to extract 57 sites for which an investigation of winter respiration is presented. Besides purely descriptive sections, an attempt was made to look at the controlling factors of winter respiration, an aspect which I unfortunately found rather superficial and where I expected a much more systematic approach.

The whole paper reads like a manuscript that was started by a small group, then sent to a large set of coauthors, but the feedback from them appears to only have led to additional secondary statements and notes, such that at the end the overall text probably contains a short touching of any of the relevant aspects but not in an organized manner.

I unfortunately started to read the paper with a negative experience found in the abstract: these winter fluxes are so unrealistically high that I really wondered how the authors will justify so large numbers in the main text. They don't. It's just one of the plenty of technical issues or errors that I will mention at the bottom of my review.

The topic is certainly of interest.

Major Issues

1. Definition of winter. The authors use their own definition of winter as the primary approach, and the established climatological winter (December thru February) as the secondary, alternative approach, that is not consistently refered in all analyses. My view is that it should be the other way around, that is, the established convention is used as a reference and then a clear justification must be given for an alternative suggestion.

Here I even see a flaw: by adding the length of the winter period as a covariate to the definition, you get a slightly higher statistical correlation in a few analyses, but then you find that winter length does not have additional explanatory power on the residuals (page 7013/5–6). This is uncommmented but is quite clear: if you already used a variable in a statistical analysis, then of course if you use the same variable a second time for explaining the residuals you must find a zero correlation. Here I wonder whether the authors really have understood their statistical analysis.

2. It is less clear to me what the reason for the zero correlation with winter precipitation is on the same lines (page 7013/5–6). Here I see an interesting point which is however not commented at all. The text leaves the reader with the impression that winter precipitation is not relevant, but by rethinking what I find in the paper and what is not commented, I think this deserves a clearer interpretation on the functional level (these controls which are mentioned at the beginning of the title): if winter precipitation does not lead to significant variations among sites, then it may well be that thanks to the snow cover soil moisture is buffered and hence as long as there are no sites that never have a snow cover the importance of this variable is not seen in the variance, but there are good reasons to explain why such a variable is not necessarily irrelevant for the control of winter respiration. Here I would argue that this justifies why you do not need to use a soil moisture term in your temperature response model (because soil moisture is not limiting thanks to sufficient winter precipitation).

3. There is a very surprising hypothesis given in the introduction on page 7002: "Part of this soil organic carbon (SOC) mass could be decomposed more actively than fresh input (e.g. litterfall) in response to future warming." This is not elaborated in more detail and is not discussed at all later in the text. In fact, your statement on page 7009 directly contradicts your initial hypothesis (page 7009: "Across all the sites, R_{ref} significantly increases with ΔLAI (Fig. 2a). This indicates that substrate availability and quality exerts a significant control on the spatial variation of R_{ref} across sites, and thus supports the conclusions of Grogan and Jonasson (2005) who found that R_{ref} was significantly reduced after removing plant and litter in a birch and heath tundra."). Why are you not clearly structuring your paper in a way where you pick up your hypotheses presented in the introduction and then clearly show the evidence supporting vs. the evidence falsifying each hypothesis?

4. This brings me to another weak point that the clear structure is missing that includes establishing a testable hypothesis and then trying to reject the null-hypothesis with available data. Instead the authors use the "big is beautiful" approach, telling me about 256 site years (7004/25), but e.g. in Figure 2 I did not see the 256 points, and I did not see that the benefit of so many sites led to error bars for each site. Moreover I only count 42 points, but the authors claim they used 57 sites (7001/18). This is indicative of a strong subjective selection bias that really could skew the statistical interpretation. Why are you only using 42 sites out of your selection of 57 sites (which already is a subset of the 255 available sites with 965 site years)? At least an objective explanation

of how you removed even more sites must be given.

5. In my view the US-Atq site that you still show in Fig. 2b but then removed from the regression analysis (page 7009) would be of more interest. Before having read your manuscript I was expecting that you will present a nice and clear assessment on how permafrost-controlled sites differ from sites without permafrost. That's what is the suggestive meaning of "mid- and high-latitudes" in the title (in my perception). But then you screen out possible indications of effects of permafrost from your analysis. Here I am really concerned if we really do not address more rigorously such key issues which are however heavily underrepresented in a data set like the La Thuile data se.

6. Overall the site selection bias is extremely strong in this paper and the interpretation further increases this skewness, such that in fact the statements relate mostly to forests. For example, 7007/11–12 states that wetlands are essentially wet tundra sites in the Arctic. But Table 1 shows that there are only 9 tundra site years vs. 7 non-tundra site years included in the analysis. Since Table 2 does not separate between different types of wetlands I do not agree with such statements. In my view "essentially" cannot possibly be synonymous to 56% (that is, $9/(7+9)\cdot 100\%$). Another critical point is 7005/13–16: "The seasonal amplitude (ΔLAI) is defined as the difference between maximum and minimum of LAI and can be considered as a proxy for recent carbon inputs to soil, i.e. substrate available for sustaining winter respiration." – do the cropland and grassland co-authors really agree with such a statement? If they do then some specific explanation should be given with respect to croplands (which are anyway heavily underrepresented) and grasslands. Currently my interpretation is that this definition automatically leads to the selection bias for forests, since ΔLAI might be a weak indicator for winter respiration conditions in croplands and grasslands.

7. The main conclusion ("First, winter RECO temperature sensitivity obtained on space and temporal scales should be treated differently, since the RECO sensitivity to warming obtained from spatial gradients will definitely be exaggerated when extrapolated to future warming") is based on very weak evidence presented in the paper. One problem

I have is that the authors do not really present a clear analysis of spatial and temporal scales, they refer to these terms implicitly by assuming that the reader e.g. automatically sorts the sites according to their latitude etc. I am strongly objecting to such implicit argumentation: if this is the main conclusion we need to see serious and careful analysis of (a) the spatial gradient and (here I have one more question mark) (b) temporal gradient.

8. Your conclusions are further weakened by the fact that you consider the two parameters in the Lloyd and Taylor model as independent from each other and hence discuss them seperately without considering the probably high correlation between them (7009). This appears quite misleading unless you can document a low correlation between the two parameters (which you most likely can not, given Fig. 2). In my view the overall effect with respect to anticipated climate change is not at all as clear as you try to phrase your conclusions you draw from this analysis.

Other Important Issues

6998: affiliation 16 is next door from my own office, but the associated co-author has not been seen here in the past few years. Maybe there are other such errors, please check your names and affiliation

Abstract, 7001: after having read the whole paper I saw that you multiplied your numbers by ten (never ever do something like this!) in Table 2 and that's why these winter fluxes are way too high, in fact one order of magnitude too high.

Abstract, 7001:25–28: what do you want to express with a phrase like "The increase in winter RECO with a 1°C warming based calculated from the spatial analysis was almost that double that calculated from the temporal analysis."? This is not intelligible to the general reader (see guidelines).

Abstract, 7001/8: ratios never have units; here you wanted to write "rates".

Throughout the paper: use SI units (e.g. http://physics.nist.gov/cuu/Units, link SI units and link prefixes. See also the special note on degree Celsius).

That means that K must be used at many places where you use °C (e.g. 7003/16)

That also means that at places where you use K but wanted to use the prefix k for kilo you need to correct the error (e.g. 7001/21, 7005/22 and more).

Please also check the definition of temperatures in the SI document referred to above; you seem to mix up the freezing point of water (273.15 K) with the tripple point of water (273.16 K). This must be corrected (e.g. 7005/22).

Use the approximate sign ($\approx$) in place of the proportionality sign ($\sim$)

I am not happy with the mixture of symbolic writing with abbreviations. The two estabilshed ways to refer to ecosystem respiration is either as a symbol (with subscript), R_{eco} or as an abbreviation TER where each letter represents a word (Terrestrial Ecosystem Respiration). Your RECO is a mixture of both which I am not in favor of. Use one of the established versions instead.

The p-ratio is an utterly confusing concept because it is easily mixed up with the statistical p-value. This is most pronounced when you sloppily write of the p-ratio value (7007/22) and moreover express it a percentage instead of ratios. A better and non-confusing symbol and wording must be found.

The same applies to the p'-ratio (7007). Here I was even unable to really grasp its definition. Should it be **mean** winter R_{eco} rates to annual mean R_{eco}? The confusion exists, since your p-ratio is defined based on cumulative winter R_{eco}. Please clarify this. In my view the two ratios only differ by the inclusion of the winter duration variable, hence it is associated with your problematic definition of winter D1.

With respect to the winter definition I of course agree that there is need for critical consideration of how to lump data for the analysis. The climatological definition has the advantage that it has a clear duration and hence the variable of true length of winter would be a second (independent) variable. But here I do not see the value of your definition because in reality for plants and ecosystem the length for each specific year appears to be important, not the mean duration or mean temperature (7006/13) of a few to a few more years (max. 10 in your study). Here I see a serious flaw in your concept that in my view does not really advance our understanding.

Moreover the definition of D1 is entirely unclear with respect to which data granulation was used. 7004/7–8 mentions that you used half-hourly and daily values; does this mean that you have two different definitions of D1, one derived from half-hourly data (where a single 30-minute value above your threshold might exclude a full 10-day period from being considered winter) and one for daily aggregated data (where a short warm period during peak daytime may still not elevate the daily mean above your threshold)?

Equations should be presented in journal style mode, not Fortran source code, same for numbers (e.g. 7011/9–11). Moreover, the use of brackets in equations must be homogenized and corrected. Numbers should be rounded to their significant digits. In this example 6E-18 is probably just the internal digital resolution of the variable type used in the analysis. Standard computers typically use 16 significant digits for floating point variables, so such a number in a statistical analysis is rather unlikely to be significantly different from zero. Specifying the standard error of parameter estimates would probably directly have helped to see the uncertainty of these estimates. My suggestion is to add the standard errors to all parameter estimates and more carefully inspect the statistical output before copying the information to a manuscript.

Language, 7008: "It should be noted that the use of the open-path gas analyzers for eddy covariance estimates of small fluxes relaying on the WPL correction (Webb et al., 1998) can introduce the errors (e.g. Kondo and Tsukamoto, 2007), and CO2 releases can be systematic underestimated" – please improve

7009/18: what if US-Atq is included? Would it still be significant or would it tell you that

you have to consider permafrost vs. non-permafrost sites as one of the key controls for winter respiration? On 7012/4 you even exclude yet another elucidating site, US-Ivo – would you not agree with me that these sites clearly indicate that permafrost vs. non-permafrost conditions would be a key issue to address in a paper that claims to focus on controls on winter respiration?

7010/4–7: "Extrapolating the relationship in Fig. 2 from space to time would imply that the future warming trends reduce the activation energy of winter soil C decomposition, hence dampening the potential increase of RECO with temperature." This would be important to do if you want to support your conclusion, but why do you not show this analysis/extrapolation to convince the reader that there is evidence in your data set to accept (not reject) your hypothesis?

7010/20: Here comes snow cover as a quick side aspect. But why did you not use the albedo measurements in the La Thuile data set for sites that have them, and combine this with albedo values from a similar data product as you used for ΔLAI? In my view this would have strengthened your interpretation.

The structure of the manscript needs improvement. As an example I put a big question mark on 7009/12, where you very surprisingly write "This indicates that substrate availability and quality exerts a significant control". This statement was purely based on the fact that Lloyd and Taylor mentioned R_{ref} to be a function of substrate availability and quality. But you cannot draw the inverse conclusion that if you see an effect in R_{ref}, then this must be both factors. So far you have only provided a variable on substrate availability (ΔLAI) but none on quality. This is not the correct deductive approach for interpreting statistical results.

In general your wording should be more careful with statistical analyses. If you make a variance analysis and you do not find variance, then this does not mean that a variable (like snow cover) is not important, it just means that it does not express its importance in the variance among sites. There is no causal relationship in statistical analysis, the causial relationships ("controls") must be established on physical and biological considerations. In your paper I have the impression that you tried to deduce all controls from statistical considerations alone.

Another statistical problem that you do not address with one single word is the significance of ratios. A simple t-test is not a correct measure for rations, unless you make a couple of assumptions. Since it has been shown that the t-test is rather robust against many violations about its assumptions, this is not a big deal, but as a reviewer I would have gotten more confidence that you actually understood your statistical analyses if you had critically discussed the limitations of your approach or warned the reader about the shortcomings that such an analysis has.

7012/15: typo on units of $\triangle$LAI

7012: use subscripts for indices of parameters both in equation and text (no source code!)

Tables

Table 1: inconsistent rounding of numbers. Recall that the convention is that the mean and SD or SE must have the same number of significant digits, and if only the mean is given, then the last shown digit must be significant (and all significant digits must be shown!)

Table 1: why is there a footnote sign (1) for Type, but the footnote sign is not found in the footnotes?

Table 1: what does the asterisk denote?

Table 2: you surprised yourself with multiplying numbers by 10. This must be avoided. If necessary it would be smart to use an SI prefix (but in this case it would be quite exotic to use deca-grams)

Table 2: since I did not grasp your definition of p'-ratios I wonder how the **big** differences between p-ratios and p'-ratios should be interpreted. An assessment would only be possible with a clear definition.

Figures

Figure 1: misleading display of histograms:
(i) histogram bars have gaps between them
(ii) it is not clear whether the histogram bars are placed in the correct interval
(iii) the x-axis range is 0–30 for D1 but 0–20 for D2 which does not allow a comparison
(iv) in the same panels the intervals are chosen differently for D1 and D2 which also does not allow for a comparison.

Figure 2: leading zeros before decimal points missing; panel lables at unorthodox locations; caption unclear, not standalone (points seem to be a further selection and an average of years without range bars)

Final Remark

Given all the issues that I struggled with, I rejected this paper, but think that the topic and wealth of information hidden in the La Thuile dataset would encourage a resubmission as a new paper that really tries to bypass all technical issues on first attempt. It is a well-known fact that reviewers tend to be harsh and critical if so many (so many!) technical issues exist in a paper. Hence I want to place a critical note to the editors of this journal (or at least the co-author who is also an editor of this journal): For us as reviewers it still remains unclear how step 2 "Access Review" under www.biogeosciences.net works. If this should lead to technical revisions **before** the paper is published in BGD (that's how I interpret the guidelines and in particular the associated scheme), then there must be a clearer procedure to screen out such manuscripts, have them technically rectified, and then present them in the BGD publication. My expectation is that we should expect a technically OK paper when we agree to review it, and then can focus on its contents. Here I had to tell the respected group of authors that (a) they should check their affiliation, (b) present equations and numbers in a technical correct way (not as Fortran code) and so on. This does not really allow for an unbiased assessment of the contents of the paper.

Interactive comment on Biogeosciences Discuss., 7, 6997, 2010.