Vaquer-Sunyer et al BC 2012 211

The paper reports measurements of GPP, NCP and CR rates in the European Arctic – the work itself must have been a major undertaking and the data is potentially very valuable. The paper makes the point in the introduction that whereas there is work on the rates in the Antarctic (authors please note there is also UK and US work, as well as that from your laboratory) there is a paucity of work in the Arctic.

Its presentation leaves a lot to be desired – it contains more than its share of often quite silly mistakes - for example it's glaringly obvious that the line "fitted" to the data in Figure 5 is wrong, it certainly does not match the equation given in the text.

There are quite a number of matters the authors need to give attention to – I have numbered the important specific points to make them clear.

The paper is accompanied with Supplementary material in the form of a Table containing the data in the rates and standard errors. This is valuable to the reader but a mystery to me is that in the GPP and CR column of rates there are blank values with errors. I fail to understand how non-existent value can have an error.

In a paragraph (p.7710, l. 21 et seq) they discuss their errors. Most of the numbers in this short paragraph appear to be incorrectly reported:

1) The report that "*The experimental standard errors (SE) among replicate samples varied between 0.04 and 6.27 mmol O_2 m^{-3}, with a mean of 0.66 ±0.03 mmol O_2 m^{-3}.*" From the supplementary material, the smallest SE value of the rate measurements I can find is 0.12, the largest is 54.85 and the mean is 1.3 mmol O_2 m^{-3}, all very different to what they report.

2) They note that: "*These errors represent a mean of 0.19 % of the total value of the measurement"*. This is incorrect; it must be 19%.

3) They then go on to claim "*These errors are very close to the limit of analytical detection, reported to vary between 0.06 and 1 mmol O_2 m^3 (Robinson and Williams, 2005)"*. Robinson and Williams (Table 9.1) reported a range of 0.06 to 0.1 mmol O_2 m^3 – regardless of which the means you adopt (0.66 or 1.3) the upper end of the range of the author's values is 6-10 times greater that reported by Robinson and Williams, not by my measure "*very close"*.

They then go on to discuss their findings, which is useful. Their rates are summarised in Table 3 and Figure 3. I have two questions relating to Figure 3.

4) Why, if it's an analysis of seasonality, present the data in chronological order and not seasonal

5) NCP must be presented as a linear plot and there is a logic in plotting GPP/CR logarithmically, but why plot GPP and CR on a log basis – doing so gains nothing and prevents easy comparison with NCP.

Table 3 reports volumetric and integrated rates. The depth-integrated value for the ARTICOS

study in based on a single point. The depth of integration used is 20m – whereas the reported mixed layer is given in the text as 67.7m (p. 7710, l. 15) during the dark period

> 6) What are the grounds for using 20m - the justification is given (p. 7707, l. 13 onwards) is not convincing. Author note: that if the euphotic zone is defined by light penetration this is independent of the irradiance level, i.e. it is the same in the dark as the light.
>
> 7) The standard error reported in the table I calculate to be the derived from the variance of the values used for the mean. This is quite in order but it should be made clear in the Table caption.

The authors then (p. 7711 & 7712) engage in a useful discussion of their work. They note correctly that although NCP is related to the f-ratio they cannot derive a value from their work – that is a sound conclusion but to be expected – f-ratio is in effect a time-averaged value.

They then (p.7713, l. 8 onwards) discuss the relationship between GPP and CR, as log(GPP/CR) v log (GPP) plots. I have noted that the line shown is incorrect and not consistent with the relationship (p.7713, l.19). This caused me to re-run the regression and was unable to obtain the same value are reported in the text. The difference was quite significant when it came to calculating the "threshold" value (see discussion bottom of p.7717) – I obtained a value of 1.7 against their value of 3.8. I spent some time trying to locate the cause and eventually found that data appear to have been omitted from the plot given in Figure 5, notably the "fliers" such as the 549 value for ATP 2010 6, 15m, but also several others.

> 8) If I am correct, the point one would make to the authors is that, whereas there may be good reason to omit particular values, this is a dangerous area in science and one needs to tread carefully. The justification needs to be spelt out - and an objective procedure needs to be adopted and declared. The convention is the deemed "filers" are shown on the plots, singled out in parenthesis (or some form of identification).

The research group lays importance on the production rate at which P=R – viewed as a "threshold" value, which they seem to regard as potentially a universal biological ecological property – akin to (but obviously not the same) as the compensation point. This might be the case in a closed ecosystem, however in a different context - when endeavouring to account for net heterotrophy – the group argues that the system is receiving significant external subsidies. This organic import will be embedded in the calculated "threshold", so the resultant value will have a "geochemical" component (net organic import, which supplements in situ photosynthesis) as well as a biological component. An in depth discussion of this is probably beyond the present paper, but there are a couple of elementary practical problems that need addressing in the present paper. The principle issue is what form of regression analysis to be used – an ordinary least squares or a Model II regression. The present paper uses an OLS regression (as far as I can determine – it is not specified), whereas the two Regaudie-de-Goux papers from the same group used Model II regressions. If I make an OLS and a Model II (MRA) analysis of a single dataset, I get a 2 to 3-fold spread of values, in some cases much greater spreads. Which one should we use?

> 9) The question one asks is, if it is possible to get a range of values from such simple alternative forms of data processing, what, if any meaning, do you give to the numbers you obtain. I think an in depth discussion of these and other matters (e.g. what, other than some empirical property, are we measuring) relating to the "threshold" notion is long overdue – without clarification we are just generating numbers with no idea what they mean, if anything.

10) Authors please make clear the form of regression used if OLS, why not the same as in the two Regaudie-de-Goux papers, and please check the consequence on the obtained "threshold' value.

A second issue is whether you should include non-significant values in regression analyses – they account for quite a significant fraction (30%) of the values in the present work. Whereas I can see that you can use non-significant values to derive a mean (as the SE can be incorporated into the overall SE) this is not available in the case of regression. The matter is made worse if log-log plot are used, as the non-significant values are characteristically the lower values in the data set and as log values they pull the line with as strong a moment as the higher, significant values. This needs some consideration.

In the section on Metabolic rates they raise the very pertinent problem of the supply of organic material to the heterotrophic population during the dark period. I think it should be pointed out that of the 7 reported values, only 2 are significant (>2*SE), so we're dealing with a very thin data base. As noted earlier the SE they report in Table 3 appears to be derived from variance of the mean, if only 2 out of 7 values are significant then it is a pretty dubious value. More appropriate in this case is the error of this mean value – derived from the errors of the individual measurements – this by my calculation is ±1.12 mmol/m^3 d, greater than the absolute mean rate (0.84 mmol/m^3 d).

They argue that the winter respiration is sustained by the DOC prior to the onset of the dark period, i.e. that the calculated requirement (75.3 mmol/m^3 d - presumably 90 dark days - authors please give us the basic details/justification of the calculation) is less than the mean DOC for the area (89 mmol/m^3) thus there is sufficient DOC to sustain this demand. The implication is that there would be 89-75 = 14 mmol/m^3 remaining by the end of the winter. This is wholly at variance with our understanding of the biogeochemistry DOC: there are several thousand DOC analyses and you simply do not encounter values below 35 mmol/m^3 in the oceans – even in the deep ocean where there have been c.1,500 years for decomposition to occur. So, the proposition does not stand up to the most elementary analysis.

11) The authors need to give more careful consideration to this analysis, as I feel it is flawed.

In summary, the authors present a useful set of data for a region of the oceans where next to nothing exists; in that respect the work is potentially valuable and welcome. The presentation and aspects of the analysis need attention. The authors need to go though the paper checking the details and correcting the errors. Clearly at this stage the paper is nowhere near suitable for publication, but the authors should be given the opportunity to resubmit as the data it contains is useful.

Peter Williams
August 7, 2012