Van Dam et al. have put a lot of work into addressing the three very thorough reviews of their initial submission. In general, I think they did a good job addressing my constructive criticisms of the initial manuscript and it is much improved.

I appreciated Van Dam et al.'s response to my primary constructive criticism RE: missing advective terms in their metabolic rate models. They did a nice job diving back into their data to assess spatial gradients in TA and attempted to place them into context with their diel TA ranges to argue that advective fluxes can be safely ignored, after describing in the text. Below, I will show why I disagree with the assertion that these advective fluxes can be ignored. I will walk the authors through some examples and explain why I believe the authors should use these "missing" advective fluxes as error bounds on their metabolic rates.

In this example, we will continue to stick to NEC and TA fluxes, but the same concepts are analogous for DIC (and will need to be applied equally).

As the authors stated int their response, they observed spatial TA gradients of 75 umol/kg/km. Let's call this variable: dnTA/dx.

Now, observing NEC rates of +/- 15 mmol/m^2/hr (Fig. 6c), we can invert Eq. 1 to calculate the time rate of change of TA (dnTA/dt). Let's assume rho=1025 kg/m^3 and h = 2 meters. Furthermore, let's also do the calculation for NEC rates of 5, 10, and 15 mmol/m^2/hr (we will be sign-agnostic because it's the magnitudes we are most interested in).

So inverting Eq. 1 to solve for the dnTA/dt yields:

dnTA/dt = -2 * NEC / rho * h

And so for our three test cases of NEC = 5, 10, and 15 mmol/m^2/hr, dnTA/dt ~ 5, 10, and 15 umol/kg/hr (with some simplified rounding)

Now let's compare these estimates for the time-varying term against your spatial gradients. Your plots of TCMs suggest flow speeds below 1 cm/s (acknowledging that the limit of detection on the instrument is 2 cm/s). So let's consider two test cases of u= 0.5 cm/s and u = 1cm/s (I know the displayed values are even lower than this, but flow values of ~0.1 cm/s are likely to be too low, and the purpose of this analysis is to understand the limits to which you can state something accurately).

Advective flux = u * dnTA/dx, so at flow speeds of u = 0.5 and 1 cm/s, your advective flux = 1.35 and 2.7 umol/kg/hr

At the assumed low flow speed (0.5 cm/s), the "missing" advective flux is equivalent to ~ 27%, 13.5%, and 9% of your respective estimated dnTA/dt

At the assumed high flow speed (1 cm/s), the "missing" advective flux is equivalent to ~54%, 27%, and 18% of your respective estimated dnTA/dt

These values of u * dnTA/dx, relative to dnTA/dt, are sufficiently large such that they cannot be ignored (i.e. they are not ~1% or even 5% of your estimates; they may be as high at 50%).

As I described before, I believe your estimates of dnTA/dt, and hence NEC, are actually equal to dnTA/dt + u * dnTA/dx. If you knew the directionality of the flow (the sign of u), you calculate whether the term is equal to dnTA/dt + u * dnTA/dx (when u>0) or dnTA/dt - u * dnTA/dx (when u<0). In the absence of information on flow directionality, I think you have to treat the advective flux as an error on both sides of your NEC estimate (i.e. NEC +/- error). Practically speaking, this means that the lower bound on your NEC estimate becomes:

NEC_lower_bound = (rho* h) / -2 * (dnTA/dt - u * dnTA/dx)

NEC_mean = (rho * h) / -2 * dnTA/dt

NEC_upper_bound = (rho* h) / -2 * (dnTA/dt + u * dnTA/dx)

Since the spatial gradients may change throughout the day, your error bounds may as well. I will leave it up to you to choose an appropriately *conservative* value of u, recognizing that all of your recorded data are below the instrumental limits of detection.

The same set of calculations need to done for both NEP_DO and NEP_DIC. Given that TA and DIC are correlated, and similar in value, I think you could use the same value for dnDIC/dx and dnTA/dx.

And finally, the NEC error needs to be propagated through your DIC-based NEP calculations (Eq. 3), in addition to the error on the DIC fluxes. Then, the daily-integrated estimates need to have error estimates that propagate through the associated uncertainty for the hourly measurements.

I know this is difficult (none of us enter into environmental science in order to revolutionize error propagation :). I am not asking you to do this because I want to torture you. But I believe the assertions of net heterotrophy and dissolution are not very robust now, and I believe a more thorough treatment of all the errors in your calculations leading to your assertions will help you and readers assign appropriate confidence in your reported net heterotrophy and dissolution. It may even prevent someone else from rebutting your study since you are exposing all of your study's strengths and weaknesses.

Specific comments:

Eq. 1: This model still does not include a term for the production (consumption) of TA due to positive (negative) NEP, despite the negatively sloped line in Fig. 3 that acknowledges the relationship between DIC uptake and TA production. I had mentioned this is my previous review, but I think the comment was missed. I think the authors now

do a good job acknowledging that the simple TA and DIC models cannot resolve sulfate reduction and denitrification, but they are implicitly acknowledging the role of organic matter production in the TA budget in Fig. 3. I think Eq. 1 needs to be reformatted to include this term. Doing so, would mean that the TA budget might look something like this:

dnTA/dt = -2 * NEC / rho * h + 17/106 * (NEP / rho * h)

Thus, NEC would now look like this:

NEC = [(rho * h * dnTA/dt) - (16/107 * NEP)] / -2

Along this same line of logic, the statement that the delta_nTA/delta_nDIC slope is ~0 for ecosystem metabolism (p. 11, L7) is incorrect.

Fig. 2: I still find Fig. 2 difficult to follow. Panels a) and b) have primary y-axes that seem to indicate that PAR will be displayed in black, yet the legend in b) indicates that PAR will be shown in red and green (a color combination that is particularly problematic for colorblind individuals). Similarly, the secondary y-axis suggests that U_10 will be shown in red, but the legend indicates a times series displayed by the black lines. I still find the inclusion of four times series on a single plot, such as in panels c-h), difficult to read and to keep track of differences between the two sites. I recommend further revising this figure, possibly to display the different field sites as the different columns and to include the two sampling campaigns as different data series on each plot. After all, the one of the primary goals is to compare and contrast between the two field sites, right? Further, I recommend that the authors demo several versions of the figure with colleagues and get their feedback about the figure readability.