Biogeosciences Discuss.,
https://doi.org/10.5194/bg-2020-172-RC2, 2020

Interactive comment on "Technical note: Measurements and data analysis of sediment-water oxygen flux using a new dual-optode eddy covariance instrument" *by* Markus Huettel et al.

Karl Attard (Referee)

karl.attard@biology.sdu.dk

Received and published: 8 June 2020

General comments

Huettel et al. present a technical study describing a new dual-optode eddy covariance system. The authors integrate two independent O2 sensors within a standard eddy covariance setup to cross-check fluxes extracted using two independent O2 sensor output streams, and to identify any biases in the measurements which are most likely caused by sensor fouling. Dual O2 sensor eddy systems are not new per se (e.g.

McGinnis et al. 2011, Attard et al. 2014), but it is the first time that the two sensor signals have been compared in the level of detail provided in this study. The authors also perform chamber incubator measurements in parallel with eddy covariance to resolve O2 fluxes using two different state-of-the-art methods. Finally, the authors also provide a comparison between the three most popular O2 sensor systems for eddy covariance measurements.

The paper by Huettel et al will find broad interest among the growing community of aquatic eddy covariance users. The length of the paper and the angle of the study make it appropriate to be published as a Technical Note in Biogeosciences. The scientific methods are clearly outlined, language is fluent and precise, referencing is appropriate and up-to-date, and the overall presentation is well-structured and clear. I have one main comment and several smaller comments that the authors may wish to address. My main comment concerns how flux quality is evaluated. Currently, the authors determine quality based on (a) diel dynamics of O2 fluxes in relation to PAR, and (b) by comparison to chamber incubator measurements. If the quality-checking aspects could be expanded to include other metrics, then I foresee that the dual sensor approach would be useful in a broader range of settings.

Specific comments

Introduction

L20-67: It would be fair to mention that dual O2 sensor eddy systems have been in use for years (e.g. McGinnis et al. 2011 L&O Methods, Attard et al. 2014 L&O) but that so far, no detailed comparison between sensor signal output has been presented.

Methods

L70: It is worth mentioning that these meters only use half the potential voltage range of the Vector analog channels (0-2.5V), but developments are in place to increase this to the full range (0-5V).

L71: I cannot find this model on the Pyroscience website. Do you mean the FSO2-SUBPORT? https://www.pyroscience.com/en/products/all-meters/fso2-subport

L86-87: It would be useful to specify whether you powered the analog channels through the Vector

L95: Firesting O2Mini: Again here, please check that this is the right model, or specify whether this was an older model that has since been replaced by the FSO2-SUBPORT.

L97: 30mm is quite a large distance. Any reason for not moving closer to the measurement volume (e.g. 1 cm)?

L143-145: This reads like Results.

L148-161: It would be useful to describe the chamber measurements in more detail. How were the chamber O2 fluxes calculated? Did you have an optode inside the chamber measuring O2 concentration continuously? How long did the deployments last? Ultimately, what are we comparing in Fig. 7?

L156: Do the chambers attenuate PAR?

Results

L164: The first sentence seems out of place here.

L171-172: What deployment hours did you use for this analysis? Was it all of daylight hours i.e. until approx. 20:00? If so, does linear regression adequately represent these dynamics?

L179-183: I would expect that both optodes located at 35 cm above the seafloor and 1 cm apart would capture these variations, though?

L231: Fig 6E: I suppose that the jump at hour 22 in the cumulative flux for sensor P is not real, but it was offset in post processing to indicate that the two sensors match one another very well beyond this point. I understand the wish to illustrate this, but I think it

is confusing, because it suggests that despite the fluxes from both sensors being very different prior to hour 22, the daily integrated flux is very similar, which cannot be the case.

L233-234: I generally agree with this interpretation, it makes intuitive sense. One concern I have is that identifying what sensor works best at what time seems somewhat subjective. For instance, in Fig 6D hour 15, PAR drops from 150 to below 50 umol m-2 s-1, and sensor P registers a concurrent decrease in flux, but sensor Q does not. After hour 16 the fluxes from sensor P are clearly 'compromised', but then again, this assessment is based upon what we'd typically expect to see. I would otherwise be tempted to interpret the drop in fluxes at hour 15 in sensor P as 'real', unless there is some other metric we could use to establish flux quality.

Overall, I fear that if we do not adopt some quantitative metrics for establishing flux quality beyond what we expect to see (e.g. diel dynamics in relation to PAR), then we might miss out on something new and interesting. This is especially true during the nighttime or in non-photic habitats. In the absence of light, would we be able to say with the same certainty what flux dynamics is 'true' and what isn't?

We've been using a two-sensor setup since we started using eddy covariance in 2010, and I fully agree that this setup drastically increases the chances of obtaining good data. I typically evaluate the two sensor signals for their performance throughout the deployment by (a) comparing the mean O2 microsensor concentration to the O2 optode, (b) point-to-point noise in the 8 Hz data streams, and (c) linearity of the instantaneous cumulative fluxes for each 15 min flux period (Attard et al MEPS in press https://doi.org/10.3354/meps13372). Yamamoto et al (2015) L&O (https://doi.org/10.1002/lno.10018, Fig. 3) adopt a similar approach. Would fitting linear regressions to the cumulative instantaneous fluxes for each 15 min flux for sensors P and Q, and evaluating the coefficient of determination (R2 value), help to shed light on this? An additional analysis could be to fit P-I relationships to the data and see which sensor produces the best R2 value, like the approach described in Attard & Glud

(2020) Biogeosciences Discussions (https://doi.org/10.5194/bg-2020-140, Fig. 2).

L243-244: Also here, it would be good to mention what part of the integrated curve was used for this analysis.

Discussion

L222-271: A two sensor setup provides redundancy and cross comparison, no question about that. However, it is also twice the cost in hardware, and twice the amount of work in postprocessing. If fouling seems to be such an issue, wouldn't the right approach be to try to eliminate fouling, rather than to add more sensors? I believe there is scope in the Discussion and in the Conclusion to comment on what future modifications might be valuable. For instance, should we install a pump and back-flush the sensors before each measurement burst? Can we monitor buildup of sensor fouling in some other way?

Technical corrections

L78: Remove extra 'relative'

L81: . . .established using the jet-nozzle method. . .

L86: Analogue should read 'analog'

L127: Apostrophes should be replaced with primes

L137: should read 'products of instantaneous. . .'

L156: 'permitting' rather than 'facilitating'

L174: Should read 'The close agreement. . .'

Please also note the supplement to this comment:
https://www.biogeosciences-discuss.net/bg-2020-172/bg-2020-172-RC2-supplement.pdf

Interactive comment on Biogeosciences Discuss., https://doi.org/10.5194/bg-2020-172, 2020.