Response to Anonymous Referee #1:

Based on comments from Referee #1 we plan to revise our manuscript as outlined below. All comments from the Referee are written with 'Italic' font. Our responses are written with 'normal' font.

This manuscript describes an important methodological advance for aquatic sciences by demonstrating that the eddy covariance method can be applied from the water-side of the air-water interface to measure oxygen and heat fluxes and to derive standard gas exchange coefficients. The method is used successfully at three shallow river sites where physical processes, especially heat exchange, are found to drive diurnal variations in gas exchange. The paper is well organized with careful, highly reasoned arguments for the approach and data treatments. The data examples are clear and mostly convincing.

We thank the Referee for the positive and constructive overall evaluation.

The only troubling part of the paper is sections that describe the possible methodological bias produced by temperature effects on the O2 sensor time series and how the authors have corrected their measurements for this bias. Although I agree this bias is likely and needs to be understood, I do not think the authors have shown they really know its magnitude. They estimate a ~3% change in the oxygen reading per 1 degree C, and even with relatively small temperature fluctuations (<0.1 oC) this creates a bias about 3 times the measured signal (Figure 6). What if the effect was 4% or 2% instead of 3%? How consistent is the effect between optode sensors and their films? Is the effect proportional to the oxygen concentration or independent of oxygen concentration? Since the response time of the thermistor is faster than the optode, does this alter the correction? In short, the authors need to independently measure the magnitude of the temperature effect before applying a correction. This might be done with experiments where the oxygen partial pressure is held constant but temperature varied. Otherwise, the applied corrections may be creating more bias than they are removing.

Temperature effects or biases were an unexpected ancillary finding – the main focus is on quantifying air-water gas exchange by aquatic eddy covariance – but we can see that some important details were missing, or not explained adequately. The following points were raised by the referee:

1) *They estimate a ~3% change in the O_2 reading per 1^{o}C.* – This is not an estimate, it is a calculation based on the signal conversion equations given by the dual O_2-temperature sensor's manufacture (JFE Advantech). According to the sensor manual we have (using manufacture's nomenclature):

$$P' = \frac{A}{1 + D(t-25) + F(t-25)^2} + \frac{B}{N\{1 + D(t-25) + F(t-25)^2\} + C}$$

$$P = G + H \times P'$$

where A, B, C, D, F, G, and H are fitting constants determined by the manufacture, t the temperature [°C], N the instrument output for O_2 [0 – 5 Volt], and P the dissolved O_2 concentration [%]. By varying t for fixed values of N, the sensor's temperature coefficient (% change in O_2 concentration reading caused by a temperature change of 1 °C) was calculated to range between 2.7 – 3.4%. We did the lab experiments suggested by the Referee and measured a temperature coefficient of 2.9. We can add this extra information in Section 3.3 describing temperature effects.

2) *What if the effect was 4% or 2% instead of 3%*? – Assuming that the temperature fluctuations are the same, the temperature bias (false flux) is proportional with the temperature coefficient. We can add this information as well

3) *How consistent is the effect between optode sensors and their films?* – Because the coefficients A, B, C, D, F, G, and H only change a small amount, if at all (< 0 – 10%), from sensor to sensor and from film to film, we conclude that the effect is well-described by our numbers provided above. We can, but suggest not to, add these very specific details to the text.

4) *Is the effect proportional to the O_2 concentration or independent of O_2 concentration?* – From the definition of the temperature coefficient (% change in O_2 concentration reading caused by a temperature change of 1 °C), the effect is proportional to the O_2 concentration. We can add this information to the text.

5) *Since the response time of the thermistor is faster than the optode, does this alter the correction?* – The response time curves presented in Berg et al. (2016) were determined separately for temperature and for O_2. In additional lab tests we inserted the dual O_2-temperature sensor from air into a water bath with both a significantly different temperature and O_2 concentration than the air. Since the response time curves looked similar to the ones shown in Berg et al. (2016), we assess that the minor difference in response times for temperature (0.34 s) and O_2 (0.51 s) does not affect the flux calculation or amplify the temperature bias. We can add this information , but suggest not to, as it appears to us to be too much detail. Again, the main focus of the manuscript is on quantifying air-water gas exchange by aquatic eddy covariance.

If the authors can address the above concern any remaining revisions to the paper will be minor. Below are listed areas by line number that might be clarified.

Line 25. I question whether it is known that physical controls are "prevalent in lotic systems". Perhaps it would be better to say "can be prevalent in lotic systems and adds uncertainty to assessments of biological activity for such systems that are based on water column O2 concentration changes".

We can change the text as recommended.

Line 30. What is meant by "erosion in the surface water"? Erosion of what?

We can clarify this by changing the sentence to: 'This was presumably caused by formation and erosion of vertical temperature-density gradients in the surface water driven by the heat flux into or out of the river that controlled the turbulent mixing'.

Lines 78-80. Awkward sentence. Please restructure.

We can change the sentence to: 'Turbulence, or turbulent-like motion, that affects or controls the thickness of the film on the water side, and thus the diffusive resistance to gas transport, can be driven by conditions both below and above the air-water interface'.

Line 87. Omit "but" in this sentence.

We can do so.

Line 93. Indicate where and how the tracer additions are made.

We can clarify this by changing the sentence to: "For smaller rivers and streams they include targeted parallel up-and across-stream additions of volatile tracers (e.g. propane) and hydrologic tracers (e.g. dissolved chloride), the latter is added to correct for dilution of propane due to hyporheic mixing (Genereux and Hemond 1992; Koopmans and Berg 2015)".

Lines 103-104. Change "studied" to "studies" and then clarify what is meant by "fitting measurements done in other aquatic systems".

We can clarify this by changing the sentence to: "Partly motivated by the substantial and often methodologically challenging effort required to measure *k* at specific sites with any of these approaches, many studies have simply relied on general empirical correlations for *k*

produced by fitting *k* values measured for other similar aquatic systems (Raymond and Cole 2001; Borges et al. 2004; Cole et al. 2010)".

Line 108. "many standard estimates" of what? Please clarify. Are you talking about carbon budgets?

We can add: "...such as gross primary production, respiration, and net ecosystem metabolism".

Line 125. Reword as: "we were able to derive parallel fluxes. . ."

We can do so.

Line 126. Reword as: "proof-of-concept tests that were up to 40 hours long at three river sites."

We can do so.

Line 131. Reword as: "All measurements were made from. . ."

We can do so.

Line 140. How was the measurement position ~5 cm below the interface determined?

We can clarify this by changing the sentence to: "This type of ADV allowed the sensor head to be positioned facing upwards (Fig. 1) while recording the velocity field right below the air-water interface, typically ~4 cm. This distance was determined post deployment from standard ADV output".

Later (line 222) can you indicate how sensitive the storage term correction is to changes in this measurement?

We can add this information in Section 3.1 where we present the O_2 fluxes shown in Fig. 2.

Line 144. Why is the senor not identified as manufactured by Rinko?

We can add this information.

Line 155. Separate into two sentences here. Indicate how reproducible the response times are with each fresh sensor film put on the optode tip.

We can split the sentence in two. We can add this information in Section 2.1 where we first present the sensor's response times for O_2 concentration measurements.

Line 158. Why reference Fig. 2a here?

This is a mistake. We can correct it to Fig. 1a.

Line 169. Change to "PAR sensor".

We can do so.

Line 183. "as level as possible" is vague. Can you indicate within a certain number of degrees from vertical? Please clarify how tilt changes were corrected for within the time span of a 15-min burst as the sensor must bob up and down some.

We can reword the sentence to: "Using a level and by placing dive weights on the platform (Fig. 1b) care was taken to ensure that it was horizontal within the tolerance of the level to minimize post-processing rotations of the velocity field to correct for sensor tilt". We don't think it is possible to specify a value for this tilt. However, for the deployment shown in Figs. 2 and 3, rotations to nullify the mean vertical and transverse velocity, resulted in an average rotation angle with horizontal of only 1.3 degrees, and did not affect the flux calculation flux. We can add this information in Section 3.1 presenting this deployment.

Also, as described in Section 2.2, all river sites used for our proof-of-concept tests were chosen to have "smooth and quietly flowing water without standing riffles or waves". As a result, our sensors did not bob up and down during measurements. Consequently, a correction for such complex sensor movements was deemed unnecessary. We can add a sentence stating that in Section 2.2.

Line 191. The key word here is "detectable". There may still be high frequency signals lost because they are not detectable by these sensors.

We agree, and note that we did use the word "detectable" here. We doubt though, that the "undetectable" part of the flux signal has any significant magnitude given the steep drop-off of the flux contribution at the high-frequency end of the co-spectrum (Fig. 4) combined with the sensors response time ($t_{90\%}$: 0.51 s for O_2 and 0.34 s for temperature). We mention this at the very end of section 3.1 and suggest that we not add a more detailed discussion of this complex question.

Line 227. It would be helpful here if the authors gave more information about how the "Spectra version 1.2" code treats the data. Also, what is meant by "several consecutive data segments"? How does this relate to what is shown in Fig. 4?

We can add the requested information so that the paragraph reads: "To examine the eddy frequencies that carried the flux signal, cumulative co-spectra of the O_2 concentration and the vertical velocity were calculated for representative periods in each deployment with minimally varying fluxes using the software package Spectra version 1.2 (P. Berg unpubl.). This software performs essentially the identically flux calculation in the frequency domain after fast Fourier transforming the de-trended data as EddyFlux does in the time domain. Both software packages rely on the same means of de-trending and time shifting data".

Also, in the presentation of Fig. 4, the specific time interval behind the two co-spectra is mentioned specifically.

Line 237. It would be helpful for the authors to present the relationships for calculating k600 even though they are in the papers cited.

To do this in a meaningful way would add two equations and a separate paragraph to explain this calculation well. Since it is a standard conversion in the gas exchange literature we suggest that we do not add this. However, because the conversion is outlined best and most straight forward in the referenced Cole et al. 2010 paper, we suggest to remove the two other papers cited.

Line 256. Since the data is presented as hourly fluxes, why not change the units in the figures to per hour rather than per day?

We prefer to use the unit mmol m^{-2} day^{-1}, in part because this unit is often used for measures such as net ecosystem metabolism.

Line 269. Suggest reword as: "controlled by a driver apart from the river current velocity or winds. . .".

We can do so.

Line 272. Unclear what results are being referred to here.

These are data from a stable independent dual O_2-temperature sensor. It is defined specifically in line 167, and we can specifically add/state that we refer to this sensor as "the independent sensor" throughout the manuscript.

Line 317. I do not see why the authors reference Gundersen et al. 1998 here? This paper discusses the temperature sensitivity of oxygen microelectrodes that operate by different principles than optodes. The microelectrode temperature effect is usually related to the gas solubility in the membrane and changes in the diffusion rate.

This referenced paper does indeed focus on microelectrodes. It is the only reference we have been able to find that gives information on the temperature coefficient for any type of fast-responding O_2 sensor. It is relevant because microelectrodes that apparently suffer from the same temperature dependency as optical sensors are still by far the most common sensor type used for aquatic eddy covariance. However, we acknowledge that this was not explained well, and we can elaborate on this as suggested in our response to the Referee's main comment (see above).

Lines 352-355. Good argument here. Correct the spelling of "concentration" in line 355.

We can correct this.

Line 371. Reword as: "This, in turn, changed the. . .."

We can correct this.

Line 380-383. Can the authors take this argument further perhaps with an illustrative example?

We do not have a measure for the average water depth, but we can assess a value based on visual inspection of the site. This would allow us to use the "open water" technique (Odum 1956) to estimate the benthic flux for comparison.

Line 399. Do the authors have any temperature profiles from their sites that may illustrate temperature stratification during the day?

No unfortunately not.

Line 410. Change to: "was first developed"

We can correct this.

Lines 416-417. It is unsatisfying that the authors call for more studies of the temperature bias. As noted above, they need to include more concrete studies in the context of this paper.

In our detailed response above to the Referee's main comment we suggest that we add more information along these lines. However, the main focus of our manuscript is the new approach for determining air-water gas exchange rates and coefficients, whereas the temperature bias is an ancillary finding. Conversely, we find it acceptable to suggest that more work is needed to further document this bias.

Line 436-437. It would be helpful if earlier in the paper the authors indicated the magnitude of the O2 storage term relative to Jeddy (Equation 3).

We can add this information in Section 3.1 where we present the O_2 fluxes shown in Fig. 2.

Table 1. Add standard deviations to the parameters in the right three columns.

We have reported SEs throughout the paper and can add these to Table 1.

Figure 1c. Add arrows to indicate each identified item and indicate that the "independent dual O2-temperature sensors" are the miniDOT sensors and the sensor used for EC is a Rinko sensor.

We can add this information.