

In this manuscript, Hoshina et al. describe the construction, installation and performance of a fuel-cell/NDIR-based instrument for measuring O₂ and CO₂ aboard ships of opportunity. They also present a year of data collected while underway across the North Pacific. Overall, this is a very straightforward presentation of careful work. Nothing about this project is revolutionary or ground-breaking, but the continued development of O₂-CO₂ systems is important for the advancement of atmospheric potential oxygen (APO) as a useful oceanographic tool, and the data they have collected fill an significant gap in the APO community's records. This is valuable work and definitely deserves to be published. I have very few scientific questions or concerns about the content. With many years of excellent work in this field, the Tohjima lab knows what it is doing, and this paper reflects that expertise. It is also quite well written.

We would like to thank you the anonymous referee for his/ her careful reading and for constructive feedbacks. We have revised the manuscript and describe the changes in the following. The referee's comments are in *blue italics*, our responses to the specific comments below.

Questions/comments on substance

P3 line 3: Change to "The mole fraction is not used as a measure of O₂ abundance because, for example, a change" Furthermore, it would be helpful if the authors gave an explanation of why mole fraction changes when per meg doesn't. This is alluded to on page 7 (around eq. 4), but it should be stated more explicitly at the outset.

In accordance with the reviewers' comments, we have changed the relevant sentence to "The mole fraction is not used as a measure of O₂ abundance because the changes in the mole fraction of major atmospheric constituents like O₂ are sometimes very confusing. For example, adding 1 μmol of O₂ to an air parcel containing 1 mol of dry air results in a 0.79-ppm increase in the O₂ mole fraction and adding 1 μmol of CO₂ results in not only a 1-ppm increase in the CO₂ mole fraction but also a 0.21-ppm decrease in the O₂ mole fraction. These confusing results are attributed to influences of the changes in the total number of moles in the air parcel on the mole fractions or dilution effect (e.g., Keeling et al., 1998; Tohjima 2000). However, adding 1 μmol of O₂ to 1 mol of dry air, which contains 0.2094 mol of O₂ (Tohjima et al., 2005), always results in a 4.77-per meg change in the δ(O₂/N₂) value." (P3 line3-10).

P5 line 5-6: I believe you are trying to say that the outlet pressures (and by extension, the pressures in the analysis cells), are kept at the same absolute value at all times by actively matching them to a reference volume. If this is what you do in fact mean to say, you should probably make it a bit more explicit.

Yes, your understanding is correct. In accordance with your suggestion, we have changes the relevant sentence to “The outlet pressures of the analyzers are kept at the same absolute value at all times by actively matching them to a reference pressure using the piezo actuator valve and a differential pressure sensor (Model 204, Setra Systems, USA).” (P5 line 12-14).

P5 line 11: Is the “mechanical mass flow controller” the same as the one mentioned on P4 line 25? If so please make it clearer. If not, specify this one more completely.

The “mechanical mass flow controller” is not same as the “mass flow controller” on P4 line 25, but a kind of needle valve. To mention it clearly, we have changed the wording “a mechanical mass flow controller” to “a needle valve (2204, KOFLOC, Japan)” (P5 line 18). According to this change, we have also changed the labels “Flow controller” to “Needle valve” and changed the symbol of the valve in Fig. 1.

P5 line 23: Is there more than one SUS tube? If so, explain how many and why.

Yes, one SUS tube is used for the cold trap. To mention it clearly, we have changed the relevant part to “Figure 2a shows ... two disk-shaped aluminum blocks, a 1/8-inch SUS tube, and a drum-shaped glass vessel with a volume of about $1.0 \times 10^3 \text{ cm}^3$. The aluminum blocks had four grooves for the 1/8-inch SUS tubes, including spare tubes to address clogging.” (P5 line 28-P6 line 3).

P6 line 10: To what do you attribute the improved performance of the 2nd trap design? Is it simply larger glass volume, more complete chilling of the glass due to the insulation, the coaxial design of the glass trap. . .? Please give a few words of explanation.

We change the trap design for more complete chilling. To clarify it, we have changed the relevant sentence to “due to the more complete chilling of the glass vessel” to the end of the relevant. (P6 lines 16-17).

P6 line 17: At what interval did you operate your changeover valve? “1-5 minutes” is

not specific enough, as the frequency of these switches can have a significant impact on the precision of your results (see, for example, Keeling et al, 2004). How did you settle on this interval?

To respond the reviewers' comments, we have changed the ambiguous descriptions to "In previous studies (e.g., Stephens et al., 2007; Thompson et al., 2007; van der Laan-Luijkx et al., 2010; Goto et al., 2013), the sample and reference air are alternately introduced into each fuel cell sensor by switching the 4-way 2-position valve at 1- to 5-min intervals. In this study, we adopted 2 min for the valve-switching intervals in light of the responses of the O₂ and CO₂ analyzer after valve switching, as described below." (P6 lines 22-25).

P6 line 20-21: It's not at all clear from the plot that the CO2 ever reaches a plateau before the changeover valve changes state again. My concern is that the CO2 values you infer will be consistently biased toward the value of whatever gas was previously in the analysis chamber. Perhaps a better approach is to fit each transition with function like (1 – exponential), with 2 free parameters (rise time and asymptote) and use the asymptote as the true mixing ratio, even this value is never measured in the instrument.

As the reviewer suggested, Figure 3 didn't clearly show whether the CO₂ signal reaches a plateau during the 2-min intervals. To respond the reviewer's concern, we have redrawn Figure 3 to show more details of the changes in the CO₂ signals. We hope the added Figure 3d convince the readers that the CO₂ signal reaches a plateau after 90 s after the valve switching. In the original manuscript, we mentioned that the last 10 s the signals were averaged to calculate ΔCO_2 , but actually we used the last 20 s data. Consequently, the sentences "The signal plateau ... last 10 s of each 2-min interval" (P6 line20-P7 line3) have been changed to "The signal plateaus at least 1 min after the valve switching, and the output signal is averaged from the second minute of the cycle (Fig. 3c). The deviation of the O₂ mole fraction in the sample gas from that of the working reference gas for the i -th 2-min interval, $\Delta O_{2,i}$, is computed based on the 1-min average according to the following equation:

$$\Delta O_{2,i} = (-1)^{(i-1)} [v_i - (v_{i-1} + v_{i+1})/2]/2 \quad (1)$$

where v_i represents the average of the output signal for the second minute of the i -th 2-min interval and the output signal represents the difference of the sample gas minus working reference gas when i is an odd number greater than 1.

In contrast to the O₂ analyzer, the CO₂ analyzer alternately measures the sample and

working reference gases. The temporal variation of the output signal of the CO₂ analyzer is depicted in Figs. 3d and 3d. As shown in the figures, the output signal does not plateau until after more than 90 s because of the relatively low flow rate in comparison with the volume of the optical cell of the LI-840A analyzer (14.5 cm³). Therefore, we compute an average of the output signal for the last 20 s of each 2-min interval.” (P6 lines 29-P7 line 12).

P6 lines 26: What happens when $i=1$? I recognize that this formula represents the difference between a block and the average of the blocks before and after; but the formula doesn't work for the first block after a calibration run.

When $i=1$, we did not calculate for ΔO_2 and ΔCO_2 . We have added “greater than 1” in the manuscript. (P7 line 6-7 and line 16).

P7 line 24-25: I'm a bit confused by your calibration procedure. I think you're saying that you used two span tanks for O₂ and calculated a linear response function for the Oxzilla every 25 hours. How did you interpolate the response of the instrument for the many observations that were made during the intervening 23.5 hours? Similarly, it seems like you had a 3-point calibration for CO₂ made just once at the beginning of the week of observations. If you have a 3-point calibration, did you assume a linear response function, or did you allow a 3-parameter form with some curvature? Am I correct that you did not repeat the CO₂ calibration at any time during or after the 7-day period of observations? Please clarify these points.

We used linear response functions both for the O₂ and CO₂ analyzers. And we determined a single calibration line for the O₂ values from the repeated measurements of the two O₂ reference gases during the 7-day observations. We didn't repeat the CO₂ calibration procedure during and after the preliminary measurements. To mention above clearly, we have changed the sentences to “We determined a single calibration line of linear response function for $\delta(O_2/N_2)$ values from all the measurements of the two standard gases during the observation. As for the CO₂ mole fraction, a single calibration line of linear function was determined from measurements of three standard gases with 387, 406, and 434 ppm only before the observation.” (P8 line 6-9).

P8 line 19: I think you mean “During every 24hr period, these three reference gases were measured for 32min each.” If that's right, please reword. If not, clarify.

Your understanding is correct. We have changed to “During every 24 h period, these three standard gases were measured for 32 min each.” (P9 line 4).

P8 line 28: There is no way for us to tell (from Fig. 6b) that there are 20s cycles. It would be very helpful if there were an inset figure that zoomed in on a subset of the data with a much-expanded time-scale.

To respond the reviewer’s comment, we have added an inset figure in Fig. 6b, which shows closeup variations for 60 seconds period.

P9 line 15-16: Are the “hourly averages” mentioned on line 16 tank values, or atmospheric data? If the latter, wouldn’t natural variability in the atmosphere lead you to expect more scatter? See also my comments on Fig. 7 below.

The “hourly averages” are tank values. We have changed “the expected standard deviation of hourly average of the $\delta(\text{O}_2/\text{N}_2)$ value” to “the expected standard deviation of the hourly $\delta(\text{O}_2/\text{N}_2)$ value for the standard gases” (P10 line 2).

P10 lines 10-14: I am puzzled about why the authors didn’t just deconvolve the flask errors to get a separate value for the uncertainty in the shipboard data. If the (shipboard data – flask values) have a scatter of XX per meg and the flasks themselves have an uncertainty of YY, then isn’t the uncertainty in shipboard data alone just given by $ZZ = \sqrt{XX^2 - YY^2}$? This way one can report a meaningful error for the system, rather than an upper limit.

In accordance with the reviewer’s suggestion, we have reported the estimated uncertainty not the upper limit. Additionally, since we have updated the $\delta(\text{O}_2/\text{N}_2)$ values of the flask samples in association with an update of a working reference gas used for the flask measurements in our laboratory, which was a normal process, we have recalculated the shipboard-flask differences and redrawn Fig. 10. Consequently, the relevant two sentences have been changed to “The averaged differences with standard deviations were -2.8 ± 9.4 per meg of $\delta(\text{O}_2/\text{N}_2)$, -0.02 ± 0.33 ppm of CO_2 , and -2.9 ± 9.5 per meg of APO. Taking into account the uncertainties of the flask measurements (5 per meg for $\delta(\text{O}_2/\text{N}_2)$ and 0.05 ppm for CO_2 measurements, Tohjima et al., 2003), we conclude that the uncertainties of the in situ measurements aboard NC2 were 8.0 per meg for $\delta(\text{O}_2/\text{N}_2)$ and

0.33 ppm for CO₂.” (P10 line 27-P11 line 4).

P10 line 20: I cannot see the “noticeable decrease” in Figure 9. Perhaps a zoomed-in inset would show it. In fact, when I look at Figure 8, I think I seen an _increase_ in APO during the eastbound leg of NC2-125.

We are afraid that our ambiguous description, probably wordings of “decrease” and “increase”, confused the reviewer. Here, we mentioned the low APO values (< -180 per meg) during NC2-125 and high APO values (> -100 per meg) during NC2-127. To mention them clearly, we have changed the sentence, “For example, the APO values show ... in late May”, to “For example, the APO values show relatively low values (< -180 per meg) during the eastbound voyage NC2-125 in early March and high values (> -100 per meg) during the eastbound voyage NC2-127 in late May.” (P11 line 9-11).

P11 line 5: The period 2014-2016 is not 4 years in length. Please correct either the dates or the duration.

“2-yr” is correct. We have changed “during the 4-yr period 2014-2016” to “during the 2-yr period 2014-2016” (P11 line 24).

P11 line 7: Presumably the figure referenced here is the one labelled “A12”. It’s odd to me that this one figure would comprise an entire appendix. I would suggest simply making it a part of the normal figures (since it is a valuable one). At the very least, make sure it is correctly referenced in the text.

In accordance with the reviewer’s suggestion, we have changed the appendix figures (Fig. A1) to the normal figures, which are referred to as Figure 11, and “Figure 11” of the original manuscript to “Figure 12” in the revised manuscript.

Figure 3: This would be much more valuable if there were fewer cycles shown with more detail. Perhaps an inset with one representative cycle (for each of plots “a” and “b”) would make the true stability of the measurements more apparent.

As is mentioned in the responses to the reviewer’s comments (P6 line 20-21), we have redrawn Figure 3 to show fewer cycles and detail variations near the end of the individual cycles.

Figure 7: I am puzzled that the error bars on the black circles (when the ship is in port) are not apparently any smaller than the scatter in the blue points (taken at sea). I'm just going by eye, but it certainly looks like the error bars in port capture more than 68% of the blue points. Wouldn't the rocking of the ship and the resulting response in the Oxzilla II make the scatter a bit larger when at sea? In lines 19-20 you say the scatter is actually bigger in the blue points, but I just can't see that. Also, on lines 20-21, you say you are only using the in-port measurements for calibration of the Oxzilla. Really? What if the instrument response varies during the time between ports? Why not use the information from the tank runs taken at sea (the blue points in Figure 7) to address this possibility.

We are afraid that the rather large error bars on the black circles, which sometimes exceed 10 per meg, confuse the reviewer. Unfortunately, the Oxzilla-II at port was not so stable as it was in our laboratory. However, we continued the measurements of the standard gases at port for more than 5 hours for the individual standard gases. So, the standard errors were smaller than 2 per meg. Therefore, we calibrated the instrument by using the calibration lines based on the standard gas measurements before and after each round-trip voyage as is mentioned in the manuscript. Consequently, we have changed the last four sentences in Section 3.1 to “In Fig. 7, the average $\delta(\text{O}_2/\text{N}_2)$ values of the standard gases determined when NC2 was berthed at the port of Tahara are also plotted as black circles with error bars showing the standard deviations. Unfortunately, the standard deviations for the standard gases were larger than the expected values obtained in our laboratory, as discussed in Section 2.3. However, the standard errors were lower than 2 per meg because the measurements were continued for more than 5 h for the individual standard gases. Therefore, we calibrated the O₂ analyzer using calibration lines based on the results at the port just before and after each round-trip voyage.” (P10 line 3-9).

Strictly editorial comments:

In accordance with the reviewer's comments, the manuscript has been changed except the following comments:

P3 line 17: Change to “transects of the data in the western Pacific region revealed that variation in the magnitude of the bulge in annual mean APO”

We have changed to “transects of the data in the western Pacific region revealed that

variation in the magnitude of the bulge in annual mean APO”.

P6 line 25: Change to “signal for the second minute of the i-th 2-minute interval”

We have changed to “signal for the second minute of the *i*-th 2-min interval”.

P7 line 6: Change to “the average of the last 10s of data for the i-th 2-minute interval.”

We have changed to “the average of the last 20 s of data for the *i*-th 2-min interval”. Because we mistook the averaged span, “20 s” is corrected.

P8 line 1: Change to “O₂/N₂) shown in Figure 5, reveals a diurnal cycle”

In accordance with the comments from both reviewers #1 and #2, “The observed $\delta(\text{O}_2/\text{N}_2)$ showed a diurnal cycle ...” has been changed to “As shown in Fig. 5, the observed $\delta(\text{O}_2/\text{N}_2)$ revealed a diurnal cycle ...”.

P8 line 16: Change to “-579 mer peg (tank #CPD-00010)”

“(CPD-00010)” has been changed to “(tank #CPD-00010)”, and “(CPD-00011)” has also changed to “(tank #CPD-00011)”, and (CPB-17350)” has also changed to “(tank #CPB-17350)”.

P8 line 19: Change to “During the shipboard measurements”

As the above your suggestion, we changed to “During every 24 h period, these”.