Response to reviewer comments

bg-2016-189: OzFlux Data: Network integration from collection to curation; Isaac et. al.

We would like to thank the 2 reviewers for the time they have taken to read the manuscript and for their comments, suggested changes and proposed additions.

We begin with the comments from Reviewer 2 and for each reviewer we deal with their general comments first and then their specific points. Original wording from the reviewers comments are given in plain text and our responses are given in *italics*.

Reviewer 2

General comments

Paragraph 1

"My only major comment is perhaps beyond the scope of the current dataset version, but I would like to emphasize it regardless in the hope that the authors will prioritize it in future efforts. Uncertainty."

The reviewer makes a very important point. The natural extension of the work presented here, and a natural extension to the OzFluxQC software, is the estimation of uncertainty in NEE, ER and GPP values. As the reviewer hints, we feel that this is beyond the scope of the present work. Estimation of uncertainty is discussed in Section 7 Future Directions as the third topic (fourth paragraph) after harmonising OzFlux data with FluxNet and resolving disparities in implementations of the same partitioning algorithms.
The estimation of uncertainty in this dataset will be the subject of future work across OzFlux. We have already developed code to estimate uncertainty due to random error, model error and u-threshold uncertainty and this has been tested at a single site in the last week. However, there is still significant work to be done to fully characterise the uncertainty such as resolving the differences between partitioning results from OzFluxQC, ReddyProc and FluxNet, estimating the effect of different gap-filling methods and different gap fractions and assessing the uncertainty and bias due to the lack of profile measurements of CO2 concentration within canopies (only 5 out of the 15 sites with substantial canopies currently have profile systems). We feel it is better to treat this topic comprehensively in a separate paper.*

Specific comments

1) Abstract, line 23-26: These lines could be removed from the abstract as telling the reader that processing used python, netCDF, OPeNDAP, etc is not really necessary at this stage.

   This is a good point that will result in a shorter and more concise Abstract. We will remove the sentences referred to in the above comment.

2) Page 2, line 18: Ameriflux was officially formed in 1996, though initial papers did not come out until 1999. I am not sure the Pryor et al. 1999 paper is the most appropriate, as it is just the first paper to come out that uses ameriflux data and has not been highly cited. Other initial papers from Ameriflux that had a larger impact were, for example:

1) Hollinger et al. 1999: Seasonal patterns and environmental control of carbon dioxide and water vapour exchange in an ecotonal boreal forest, GLOBAL CHANGE BIOLOGY, 5, 891-902
2) Wilson and Baldocchi, 2000: Seasonal and interannual variability of energy fluxes over a broadleaved temperate deciduous forest in North America, AGRICULTURAL AND FOREST METEOROLOGY, 100, 1-18
3) Schmid et al. 2000: Measurements of CO2 and energy fluxes over a mixed hardwood forest in the mid-western United States AGRICULTURAL AND FOREST METEOROLOGY 103, 357-374

Many thanks to the reviewer for correcting the historical inaccuracies in our Introduction. We will amend the text as suggested.

3) Page 2, line 20: It might be worth mentioning the FLUXNET network specifically here, and citing:
   1) Baldocchi et al., 2001 FLUXNET: A new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities, BULLETIN OF THE AMERICAN METEOROLOGICAL SOCIETY, 82, 2415-2434

   This is also a good point and would also serve to introduce the FluxNet syntheses mentioned on line 31 of the same page. We will amend the text as required.

4) Page 6, line 22: are also collected at all sites?

   This is correct, the additional quantities mentioned are collected at all OzFlux sites. We will amend the text as suggested.

5) Page 8, line 26: It is not clear why three different sources of alternative met data are required. Does one provide information the others do not? An opening sentence justifying the need to have different sources of met data would help the reader.

   The reviewer makes a valid point. Although the different quantities supplied by the 3 sources of alternate data are given in Sections 2.5.1, 2.5.2 and 2.5.3, it is somewhat buried among other detail. We will insert a sentence in Section 2.5 Ancillary Data to better explain the need for the 3 sources of alternate data.

6) Page 10, line 92: wide spread should be one word

   The text will be amended as suggested.

7) Page 11, Section 3.2 The difference between OzFluxQC and DINGO could be better articulated here. The text only highlights some differences in the plots generated and the format of output (csv vs netCDF) but these differences are only superficial. What are the differences in terms of the internal processing? Do both approaches use the same algorithms for gap-filling and partitioning, and if not, what are the fundamental differences?

   The Dynamic Integrated Gap-filling and partitioning tool for OzFlux (DINGO) is described in Beringer et al (2016), a paper in the same Special Issue as the current paper. As far as possible, we would like to avoid duplication of material between these 2 papers but we also acknowledge the frustration of having to flick between the 2 papers to compare methods. We will re-write the second paragraph of Section 3.2 The Roles of OzFluxQC and DINGO to briefly describe the main similarities and differences in the methods and algorithms used.

8) Page 13, Line 10 It appears that the uStar threshold identified can vary seasonally and from year to year. As the uStar threshold can have a large impact on the fluxes (particularly flux partitioning) it would seem important to highlight better whether uStar varied seasonally or not, and at what sites. Have the authors considered using also a fixed uStar threshold? This is also included for comparison in the Fluxnet 2015 data release.

   On re-reading Section 3.4, we feel there are 2 parts to this comment. The first is that the existing text does not include a description of how OzFluxQC uses the u threshold information from the CPD method. We will address this by splitting the second paragraph at the sentence beginning "At some OzFlux sites, ..." (page 13, line 27) and adding a sentence to the now truncated second paragraph that describes how OzFluxQC implements the u* thresholds found by the CPD method. This will include the information that, by default, OzFluxQC uses annual values of the u* threshold but that the user can also specify the threshold to be used as a function of time.*
   The second part, and this may have been the reviewers main point, concerns seasonal variability in the u threshold values detected by the CPD method. Our implementation of the Barr et al (2013) CPD technique provides the u* threshold for each 1000 point "season" but does not record the start and end dates for the season making it hard to relate the 1000 point "seasons" to any seasonal changes in ecosystem phenology.*
   As a compromise, we propose adding text and a table to Section 3.4 to describe the intra-annual and inter-annual variability in the u threshold values detected by the CPD method at the 9 OzFlux sites used in Figures 4, 5, 6 and 9 and testing any intra-annual variation for significance.*

9) Page 14, line 5: Usually -> often

   The text will be amended as suggested.

10) Page 21, line 10: The optimal window size likely changes depending on the time of year and the site. The authors claim that they have found an optimal size of 60 days seems somewhat ad-hoc. No evidence is presented nor any methodology for determining the optimal window size given. Perhaps rephrase. A more complex ANN design might not be necessary. Simply including day of year as one of the predictors should allow flexibility in the response.

    In the existing text, we state that the "best performance" is achieved with a window size of 60 days and that "performance degrades with increasing window length.". We did not intend to imply that 60 days was the "optimal window size" and accept that the wording in paragraph 5 of Section 6 Conclusions is ambiguous. We will amend the text along the lines suggested by the reviewer to clarify this point.

11) Page 21, line 19: As there are no measurements of ER, you cannot really claim that daytime partitioning methods overestimate ER. The true value is unknown.

    The reviewer is correct, the offending wording is sloppy. The sentence will be re-written to emphasise that the under-estimation is relative to the other methods.