

## ***Interactive comment on “Upscaling surface energy fluxes over the North Slope of Alaska using airborne eddy-covariance measurements and environmental response functions” by Andrei Serafimovich et al.***

**Andrei Serafimovich et al.**

andrei.serafimovich@gfz-potsdam.de

Received and published: 14 June 2018

This is a very promising synergy of airborne turbulence measurement, machine-learned environmental response functions, and mesoscale modeling (WRF) to project air-surface exchanges of water and heat out over large areas of the North Slope of Alaska. It looks to be a significant advance in addressing this important issue.

We thank referee #1 for the helpful comments, hints and suggestions how to improve

C1

our manuscript. We added these suggestions in the revised version of our manuscript. The details are documented below. We provide a supplement pdf-file, where we marked the respective changes (please see the link at the end).

### **Assets**

1. Airborne campaign was generally well designed including straight tracks with multiple passes over some tracks and frequent profile ascents to find the mixed-layer structure and depth. Instrumentation and sample rate (100/s at about 70 m/s true airspeed) are appropriate to the mission. One apparent deficiency was in radiation. Although the investigators measured shortwave insolation, they did not mention measuring reflected shortwave (they did determine the albedo), or upwelling and downwelling long-wave (infrared) radiation.

The 80% upwind footprint extent of our airborne flux measurements can vary from 800 m to several kilometers. In contrast, the 80% footprint of upwelling radiation measurements corresponds to a circle of ~350 m diameter with its center located approximately below the aircraft. In result, the flux and radiation observations represent entirely different surfaces, and cannot be directly used together. For this reason we decided to use the remote sensing data, which cover the actual flux footprint. It should be noted that the sole purpose of the remote sensing data is to explain spatial variation in the environmental response function. In contrast, temporal variation is explained through downward shortwave solar radiation and observation time. The albedo was determined from the MODIS data post-processed data by the National Research Council.

2. Pg 5 line 17: “Wavelet cross-scalograms were integrated in frequency over transport scales up to 20 km” This raises the question of the length of the flight tracks. From

C2

the map (Figure 1) they appear to be mostly about 100 km, which still provides an adequate 80 km of flux uncontaminated by edge effects at the largest width (20km?) of the cone of influence.

The transport scale of 20 kilometer for Wavelet integration was chosen based on spectral gap analysis, and exceeds the maximum observed boundary layer height by at least one order of magnitude. The eddy motions responsible for the net vertical transport are approximately confined to the scale of the boundary layer height. In result, further extending the transport scale beyond the spectral gap would effectively include compensatory flux contributions resulting from horizontal gradients, rather than actual vertical exchange. In theory, also these ultra-long scale contributions approach their ensemble value when averaged over space and time per ergodic hypothesis (Finnigan et al., 2003). However, they manifest themselves as erratic contributions to the individual flux observations, which cannot be related to surface properties in the turbulent flux footprint and thus complicate the extraction/projection of reliable relationships (Metzger, 2018).

3. Section 2.3: It's good to see a mesoscale model used in the study. In principle it can describe the mesoscale environment of the airborne campaign. It is difficult to relate model results quantitatively to what the aircraft and any fixed surface sites are reporting, especially in Alaska where the input data are relatively sparse. But for a "projection" (see item 6) it appears to work well.

Yes, the WRF model was very useful tool to improve our method to project turbulent fluxes.

4. Page 6, line 6: The verb "project" is a great word to describe the inference of

C3

air-surface exchange over large areas from measurements over small areas because it implies some sort of model which represents knowledge and draws on data both of which grow more sparse as the scale increases. "Upscaling" is commonly used but is troublesome because it implies that the system is (largely) scale-independent, like going from a model to a full-scale prototype.

It is a very good comment, this verb directly describes what we can do with ERFs and machine learning.

5. Page 8: MODIS has something like 500 m to 1000 m pixels, which appears a good match to the km-scale spatial averages of the fluxes derived from this campaign's measurements.

Yes, it is true. Moreover, for EVI we used composite products with 250 m pixels. They are based on the best available pixel value from all acquisitions from the observed period with low clouds and low view angle.

6. Figures 5 and 6, pages 9 and 10: The "rug" plots showing the distribution of the variables in deciles are appreciated.

We found "rug" plots very useful for the interpretation of response functions.

### Questions and Issues

1. Page 5 and thereabouts: Recognizing flight safety as the ultimate first priority, did they next prioritize near-constant height above ground or did they try to minimize

C4

pilot adjustments? These are often a trade-off because terrain is rarely really flat. Minimizing control pressures improves fidelity of the flux measurements by minimizing the flow distortion, but usually requires higher-altitude flight to clear the terrain safely. Higher altitude gives up the advantage of cleaner sampling of heterogeneous surface fluxes.

Actually, neither of both was really enforced. Occasionally pilot action was strong, but for most of the survey flights the measuring height was nearly constant. Additionally, the Alaskan North Slope is really flat. The median terrain height and its median absolute deviation along the flight lines was  $21 \text{ m} \pm 13 \text{ m}$  and allowed us to measure at the median height  $38 \text{ m}$  with median absolute deviation  $\pm 7 \text{ m}$ . We added this information to the manuscript (s. page 5).

2. Page 5, line 25 into Page 6: I presume that a sample of sensible or latent heat flux is one of the 1000 m mean fluxes computed every 100 m, perhaps further segregated by land-surface type. Because of the deep overlap, these are strongly autocorrelated implying fewer degrees of freedom (DF) than there are fluxes in the sample. There are ways to estimate this reduction in DF by determining the decorrelation length. A coarse guess would be to divide the sample size by 10 since the reporting interval (100 m) is one tenth of the averaging length (1 km). The loss is probably not that much. At the least the loss of DF by autocorrelation should be mentioned qualitatively.

We thank the referee for this valuable suggestion and mentioned the loss of DF in the manuscript (s. page 6).

3. Page 8, Fig 4: in the second-row left node the dashed arrow intrudes into the ellipse partially obscuring the criterion for insolation making it hard to read ( $S > 380 \text{ W/m}^2$ ?).

C5

Corrected.

4. Pg 11 Figure 8 Caption: Is it proper to call the quality parameter a standard error, or is it more an uncertainty? If an error, what constitutes "truth"?

We do not mean the quality parameter of flux measurements. This figure shows the median flux maps and the "standard error" is a statistical parameter of this average equal to median absolute deviation divided by the square root from the size of the sample. We replaced the words "standard error" by "standard error of the median value".

5. Figure 7 is interesting. There are two slopes of BRT predicted vs aircraft measured. The majority of the aircraft data match the BRT predicted quite well, but the majority of the spread of BRT vs measured has a shallower slope. Away from the training data, the BRT underpredicts the strong fluxes and overpredicts the weak fluxes. Were measures taken to avoid overfitting?

At this time we did not use any methods to avoid under- or overprediction of BRT. We will take this suggestion into account for future investigations. However, we have to mention that only few data are located in the range of overfitting. For the sensible heat flux only 10% of the data are less than  $-5$  and more than  $80 \text{ W/m}^2$  and located outside of the black cloud. For the latent heat flux only 6% are less than 0 and more than  $110 \text{ W/m}^2$ . We added this remark to the manuscript (s. page 9-10).

6. Page 17, line 7ff: It is indeed likely that under free-convection conditions (light winds

C6

and strong insolation) that fixed sites do not experience the larger patterns that aircraft can readily penetrate. If these structures are in fact turbulence (hence “random”) they can be treated by integral techniques such as time/frequency-domain methods or time/space-domain averages. If they are not random it is not clear that an individual aircraft track in space and time will capture the relevant (nonrandom) structure unless it can be designed to do so.

Yes, they are non-random. For example, surface heterogeneity and as you mention light wind and strong isolation could result in such non-random structures, which can be captured by randomness of flight tracks. We learned it also from our machine learning method: the more data used for training, the better the convergence of the algorithm. Therefore, we always tried to extend the coverage of our airborne measurements during the reported campaign and campaigns performed in following years.

7. Page 19, lines 2 and 3: The data set is no longer unique. Sayres et al. (2017), cited in this manuscript, also used an aircraft to measure fluxes (primarily latent heat and methane) over heterogeneous Arctic landscapes. Exclusive use of airborne data can be considered a liability because of the inherently greater uncertainty in airborne measurements. Having a surface reference, as in this manuscript with Atqasuk and Happy Valley, is important both for temporal continuity and for sanity check. Data are, of course, the more readily acquired from aircraft than from fixed sites over remote areas difficult of access as on Alaska’s North Slope.

We agree with this comment, but we would like to point out that this data set is unique in its spatial extent and clarified the related sentence (s. page 19).

C7

**Copy-editing items** (a few were found and noted, not guaranteed to be exhaustive list)

1. Page 4 line 16: “reference period”

Corrected.

2. Page 14 line 15: Likely: “periods that were”

Corrected.

3. Page 14 line 20: Sentence starts with Greek letter (beta for Bowen Ratio). Recasting of the sentence is recommended.

Corrected. The sentences before and after were also corrected for the same reason.

4. Page 15 line 19 Remove “An”; insert “an”: “measure water vapor <an> LI-7200 gas analyzer...”; remove “for water vapor measurements.”

Corrected.

5. Page 17 line 25: Turbulent fluxes over water surfaces are more likely to be suppressed (due to lack of both mechanical and buoyant generation) than to be directed to the surface.

Yes, it is correct. We added this remark to the discussion. However, we also observed

C8

negative sensible heat fluxes during similar campaigns in the Mackenzie River Delta, Canada, and they are also reported by Overland et al. (2000), who observed flux compensation during the SHEBA experiment.

6. Page 17 line 30: “Project” (verb) would be a really great word here instead of “upscale.”

Corrected.

**Referencies:**

Finnigan, J. J., Clement, R., Malhi, Y., Leuning, R., and Cleugh, H. A.: A re-evaluation of long-term flux measurement techniques. Part 1: Averaging and coordinate rotation, *Boundary Layer Meteorol.*, 107, 1-48, doi:10.1023/A:1021554900225, 2003.

Metzger, S.: Surface-atmosphere exchange in a box: Making the control volume a suitable representation for in-situ observations, *Agric. For. Meteorol.*, 255, 68-80, 2018.

Overland, J. E., McNutt, S. L., Groves, J., Salo, S., Andreas, E. L., and Persso, P. O. G.: Regional sensible and radiative heat flux estimates for the winter Arctic during the Surface Heat Budget of the Arctic Ocean (SHEBA) experiment, *J. Geophys. Res.*, 105, 14 093 – 14 102, 2000.

Please also note the supplement to this comment:

C9

<https://www.atmos-chem-phys-discuss.net/acp-2017-1166/acp-2017-1166-AC1-supplement.pdf>

---

Interactive comment on *Atmos. Chem. Phys. Discuss.*, <https://doi.org/10.5194/acp-2017-1166>, 2018.