Atmos. Meas. Tech. Discuss.,
doi:10.5194/amt-2017-176-RC3, 2017

Atmospheric Measurement Techniques Discussions

Interactive comment on "Uncertainty Characterization of HOAPS-3.3 Latent Heat Flux Related Parameters" *by* Julian Kinzel et al.

Anonymous Referee #3

Received and published: 8 November 2017

This paper develops a method to estimate errors of the HOAPS-3.3 ocean evaporation product. Ocean evaporation is a key component of the earth's energy and water balance, and error estimates are very much needed in budget studies and climate drift investigations. Too often conclusions are drawn on the basis of climatological data that does not have the accuracy that is needed. So this paper is very welcome as a complement to the HOAPS evaporation data. The methodology by itself is also of scientific interest because it can be applied to other data. The triple co-location method is used here in an interesting way as it allows to isolate the co-location error from the total error. Also random errors and systematic errors are distinguished, although I have questions about the analysis of systematic errors.

I have the feeling that the authors have done a lot of good work that is of interest to the

research community, and therefore I feel that the work should be published. However, the way the work is presented can and should be improved.

I have two major concerns:

1. The paper is hard to read. Often it requires re-reading a paragraph a number of times, to understand. It also has to do with the structure of the paper. It would help to define the main methodology of the data analysis and to have this as the main thread throughout the paper. I think I understand the methodology, but I am still not sure. Let me explain my interpretation of the method:

(i) Four dimensional look-up tables (LUT) are created of co-located data, so differences between data sets are stratified according to q_a, U, SST, and wvpa.

(ii) The mean difference between HOAPS and insitu data are interpreted as biases.

(iii) The variances of the difference are used for the triple co-location method, resulting in error estimates.

(iv) This results in LUT's of biases and random error estimates.

(v) In applications (e.g. global maps of mean and random error of q_a) the observations of q_a, U, SST and wvpa point to the table and provide errors of each observation. These can be averaged to obtain the desired map.

I feel that it would be helpful to describe upfront that this is the general methodology and follow it throughout the paper. So this would lead to 3 main sections in the paper: (i) Description of the methodology, (ii) Results of the methodology, i.e statistics on the LUT data, and (iii) Application to HOAPS evaporation. In case I am completely wrong on the interpretation of the paper, there is even more reason to be clear about the methodology.

Another question is: what is the main result of the paper? If my interpretation is correct, then the 4-dimensional table of error estimates is the main result, because it would allow a user to make estimates of anything he/she is interested in (e.g. monthly averages, daily averages, or El-Nino years). So it is worth thinking about communicating this 4D table to the users. Most of the current paper is about applying the methodology, but these are in fact just examples.

2. Estimation of biases is non-trivial. In fact this is very important because, as the authors point out, for long term averages the systematic errors dominate. My concern is two-fold:

(i) I have the feeling that it is assumed that DWD-ICOADS data is bias-free? If this is the basis for the bias estimates, then it deserves more discussion also in view of what has been published in literature.

(ii) Fig. 1 is used as an example to illustrate the estimation of biases. However, it is likely that artificial biases occur in binned scatter plots of noisy data if correlated variables are used on abscissa and ordinate. This applies to Fig. 1a where hair(HOAPS) is used on both vertical and horizontal axes. It also applies to hair versus wind because these variables are correlated due to the physics of the mixing (more wind brings hair closer to the surface value). To check, one could e.g. bin the differences of Fig. 1a in classes of hair(insitu). Also hair(insitu) is noisy because it has large representativeness errors (point observation, whereas HOAPS has a large footprint).

Finally, if one can be confident about the bias estimation, then it should also be trivial to apply a bias correction to HOAPS. This would just leave the uncertainty in C_E which is a parametrization constant used for satellite as well as in-situ data. Please discuss.

More detailed comments:

Section 1

Although well written, the introduction is rather long and contains sometimes fluffy language. For instance, the first 24 lines illustrate the importance LHF, which is well known and can be much shorter. The second sentence is another example. A side

issue is introduced by referring to turbulent heat fluxes (not latent heat fluxes). A good overview of the literature is given, but no reference is made to an earlier study by Kinzel et al. on the estimation of uncertainty of q_a. It is important to point out in the introduction what is new compared to the earlier work. My interpretation is that Kinzel et al. (2016) does an error analysis on q_a, and that the current paper extends it to U, q_s and LHF. It is important to clarify this.

The introduction makes references to other data sets and to studies that provide error estimates. However, nothing is said about published error estimation methods. Since it is the topic of the current paper, it is necessary to explain what is different about the own method compared to others.

Page 5, Line 32

The sentence with "The latter depends" suggests that it refers to q_a in the sentence before, but what it intends to say is that the COARE algorithm needs stability and that specific assumptions are made. Please rephrase.

Page 6, Line 20-24

The non-correction of q_a for measuring height is confusing. Why not using the real measuring height in the bulk formula? Perhaps it is possible to say in one sentence what the results are of the height difference effects as estimated by Kent et al. (2014)

Page 7, Line 13

Cool skin corrections are applied to in situ observation but not to HOAPS-3.3 SST (AVHRR based). This makes sense in priciple because AVHRR measures the skin temperature. However, there must be a calibration procedure of AVHRR, which is probably against bulk SST data. So, what does calibrated AVHRR data represent, bulk or skin SST?

Pages 4-5 section 2.1

It would be informative to mention pixel size of the microwave sensors.

Page 8, Line 11

The sentence "Figure 1a overestimates ...." is confusing. Formally it is correct, but, after reading the first time it suggests that the biases range from 7-12 g/kg and that the plot is for the inner tropics.

Page 8, Line 17

The expression "over-(under-)estimated" is perhaps better than "over-(under-)represented"

Scatter plots in Fig.1

In all the plots except (c) the variables on the vertical axis are correlated with the variable of the horizontal axis. This is most obvious for Fig. (a) where hair-HOAPS is used in both abscissa and ordinate. In such cases the binning according to one axis can show biases that are not necessarily real. Whether this is really the case can be easily demonstrated by making the same plot but now with hair-insitu on the horizontal axis. Similarly unrealistic bias may be seen in (b) and (d) because wind and wvpa are derived with from the same satellite channels and therefore correlate with hair-HOAPS. Please discuss.

Page 9, Line 21

Please specify what "even stronger winds" are.

Page 9, Lines 24-26

This paragraph is hard to read. After reading, a number of times times, I think I understand. Is it not better to say: "Our goal is to document the upper bound of the bias and therefore we take the absolute value of the possible systematic error in CE"?

Page 10, line 15 and page 11, line 7

I suggest to replace "Next to" by "In addition to"

Page 11, section 3.5

This section is hard to read. If I understand correctly, it addresses the question: Does it matter for the averages that the satellites sample the ocean at particular times of the day only, given that a diurnal cycle may be present? The authors investigate by looking at buoy data and by comparing averages that cover the full diurnal cycle with samples at satellite overpass times only. Part of the confusion is because it mentions spatial sampling, but I don't think this section covers that? Please simplify for clarity.

page 13, Lines 9-11

I am not sure that it is helpful here to refer to Fig. 1a, because it is showing the combination of E_ins(qa_a) and E_retr(qa), which is different from Fig2_a. The authors point this out but instead of clarifying something it confuses.

Page 13, 23

Suggestion: replace "merely" by "only"

Page 13, Line 24

What is meant by "local minimum in that region for q_a"? E_retr(q_a) has a maximum over the warm pool.

Page 13, Line 29

In the sentence "Respective values partly exceed 50 W/m2", what is meant by "respective" and "partly"? Do the authors mean: "In these areas, values are found in excess of 50 W/m2"?

Page 14, Line 33

"direct eddy covariance" is not wind speed.

Page 16, Lines 1-2

This is an interesting example, where it is explained that q_a retrievals may be in error because of dry air advection. However, it is not clear how the systematic error analysis picks up the area of the Agulhas current. The systematic error estimation is entirely driven by U, q_a and SST and wvpa (if I understand correctly).

Section 4.6 and Fig. 4

Here both systematic and random errors are discussed region by region and climatologically versus January/July. Earlier in the paper it was concluded that the random errors were small compared to the systematic errors. However in Fig. 4 the random errors are larger than the systematic errors. Furthermore I would expect that the climatological data (I assume averaged over the entire period) has much more data than the January or July data and therefore much smaller random errors.

Page 18, Line 31

Please replace "outperforms" by "exceeds"

Conclusion:

The material is well worth publishing, but major re-structuring and revision of the paper will improve its quality and accesibility.

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2017-176, 2017.