

## **Author's response to the general comments from referee II:**

Thank you for revealing your valuable criticism regarding the manuscript. Below, please find our responses to your specific comments, along with the implemented changes to our manuscript. All page and line numbers as well as figure numbering refer to the *revised* manuscript. Note specifically that the figure numbering has changed during the review process.

### **SPECIFIC COMMENTS:**

**1 ) Comment from referee:** *Page 5, Line 10, "... the pixel-level HOAPS-3.3 data in sensor resolution is used...". What are the spatial and temporal resolutions of the pixel-level HOAPS-3.3 data? Which nine sensors are used in the pixel-level HOAPS-3.3 climatology?*

**Author's reponse:** The spatial resolution of both data sources is channel-dependent. For SSM/I data (DMSP F08-F15), it varies from 69 km by 43 km (19 GHz channel) to 37 km by 28 km (37 GHz channel). Sampling frequencies take on a value of 25 km, corresponding to scan lines every few seconds. Regarding SSMIS (DMSP F16-F18): The spatial resolution varies from 74 km by 47 km (19 GHz channel) to 41 km by 31 km (37 GHz channel). As for SSM/I, sampling frequencies are given by 25 km. Overall, 9 different DMSP sensors contribute to HOAPS-3.3: F8, F10-F11, and F13-F18.

**Changes in the manuscript:** The DMSP satellite platforms have been included into the revised manuscript (P.5, L.10f). Furthermore, the spatial resolution has been implemented (P.5, L.11ff).

**2) Comment from referee:** *Page 5, Line 15: what is the temporal resolution of  $q_a$  retrievals? And at what height?*

**Author's reponse:** Unfortunately, no information is provided by Bentamy et al. (2003) as to the sensor heights of the (in situ)  $q_a$  retrievals. What is known is that their updated regression coefficients are derived using 1000 collocations between globally distributed ship data and validated DMSP satellite data (F10-F14) during 1996-1997. As the retrieval is based on these match ups, we believe that the expression of "temporal resolution" is somewhat misleading. The globally distributed match ups do not have a temporal resolution and are rather point measurements in time and space.

**3) Comment from referee:** *Page 5, Line 32: which surface pressure data are used in computing LHF?*

**Author's reponse:** The COARE-3.0 algorithm assumes a standard sea level pressure (SLP) of 1013.25 hPa when iteratively calculating LHF, which is also used for deriving HOAPS LHF. Brodeau et al. (2017) investigated the effects of this SLP approximation in bulk parameterizations of turbulent air-sea fluxes, amongst others. The authors conclude that errors of such an approximation remain well below discrepancies related to the computation of the transfer coefficients themselves. Their sensitivity experiments show that  $q_s$ - and  $\rho$ -induced errors range between merely  $\pm 5\%$  (given an SLP range from 950 hPa to 1040 hPa) with an opposite and therefore potentially compensating effect on LHF. Apart from this, both SSM/I and SSMIS are not capable of deriving SLP. Making use of auxiliary (e.g. reanalysis) data to implement SLP would violate HOAPS' unique feature of relying completely on satellite input.

**Changes in the manuscript:** A note has been added to the revised manuscript (P.6, L.9) that a constant SLP is presumed.

**4) Comment from referee:** *Page 5, Line 33: "...surface air temperature, which is estimated by assuming a constant relative humidity of 80 % (Liu et al., 1994) and air-sea temperature difference of 1K". How accurate is this assumption? During winter cold air outbreaks over the western boundary current regions, the air-sea temperature differences can exceed 10 K. In this case, the*

*assumption will lead to a bias in air temperature. How is surface air temperature compared to the in situ dataset?*

**Author's reponse:** Thank you for bringing this up. We did not investigate the uncertainty introduced by these two widely used assumptions, as it may be neglected for two reasons (for our purposes).

First, air temperature only has a secondary effect on LHF (in contrast to SHF) through the stability of the atmospheric column. The assumption of 1 K temperature difference with respect to SST is a good approximation for vast regions over the global oceans. However, we agree that during cold air outbreaks over the WBCs or in upwelling regimes, which are very confined compared to the global oceanic area, this approximation is violated. Compare conclusion section of Wells and King-Hele (1990).

Second, our uncertainty estimation procedure described in Sect. 3 is exclusively based on high-quality match-ups of HOAPS and in situ measurements. The data density of both ship and buoy records is comparably low in the regions addressed above, which further reduces the impact of our two assumptions. Due to the comparatively small amount of reference data, we presumably underestimate resulting uncertainties in these regions. Using for example ancillary reanalysis-based data would violate our ambition to create a completely remotely-sensed data record, which is a key feature of HOAPS.

No SSM/I or SSMIS retrievals exist that are capable of accurately retrieving oceanic surface air temperature (SAT) from space. This implies that SAT is not available as an official HOAPS product and has thus not been compared to the in situ reference. Future efforts will take on this challenge.

**5) Comment from referee:** *Page 6, Line 1: Provide a map showing the spatial distribution of in situ (ship and buoy) reference data density over the global domain.*

**Author's reponse:** We agree that providing such a map is useful to the reader. We therefore implemented a map showing the spatial distribution of match ups (ship/buoy vs. satellite) over the global oceans, exemplarily for  $q_a$ . It shows all collocated match ups between 2001-2008 that contribute to Fig. 2 ( $\approx 13.8$  million match ups per subplot in total). Match ups for U and  $q_s$  occur even more frequently, but are not shown in the revised manuscript.

**Changes in the manuscript:** A map showing the distribution of  $q_a$  collocations between 2001-2008 has been implemented into the revised manuscript (Fig 1, left panel). It is briefly described in terms of density distributions (P.8, L.29ff).

**6) Comment from referee:** *Page 6, Lines 4-5: Does the reference dataset include the 1996-97 period that is used in training  $q_a$  algorithm?*

**Author's reponse:** We are not able to answer this question, as Bentamy et al. (2003) does not provide any information as to which ship records were used to train their  $q_a$  retrieval. Yet, the multi-dimensional bias analyses are restricted to match ups between 1998 and 2008 (depending on the parameter, see P.10, L.25f). This implies that no temporal overlap between the reference data archive and the ship records used for training purposes exists.

**7) Comment from referee:** *Page 7, Lines 28-29: The “instantaneous and climatological uncertainties” are not explained. How are they related to systematic, random, and sampling uncertainties?*

**Author's reponse:** Sorry for not being precise enough here; we agree that this needs clarification. “Instantaneous“ uncertainties are pixel-level uncertainties. These uncertainties can either be systematic (compare Fig. 4 over revised manuscript) or random (see Fig. 3 of revised manuscript). On an instantaneous basis, sampling uncertainties do not exist.

By contrast, we define “climatological“ uncertainties as *total* uncertainties averaged over the time period 1988-2012 (as illustrated in Figs. 4 and 5 of revised manuscript). That is,  $E_{\text{clim}}$  is formally the mean root mean squared sum of  $E_{\text{sys}}$ ,  $E_{\text{retr,ran}}$ , and  $E_{\text{smp}}$  averaged over 1988-2012. As  $E_{\text{retr,ran}}$  scales with  $1/N$ , with  $N$  being the amount of observations per grid box (see Eq. 3), it becomes virtually

zero when averaging over long time periods. Likewise, monthly mean  $E_{\text{smp}}$ , which applies even more so to multi-annual averages. On climatological time scales,  $E_{\text{clim}}$  and  $E_{\text{sys}}$  therefore hardly differ. This is why Fig. 4 of the revised manuscript can be treated as both „systematic“ and „climatological“ uncertainty.

**Changes in the manuscript:** The explanation of the methodology has been extended (P.8, L.12ff). This includes a link from instantaneous and climatological uncertainties to systematic, random, and sampling uncertainties. A mathematical description of  $E_{\text{clim}}$  is furthermore provided (P.15, L.12f).

**8) Comment from referee:** *Page 8, Line 10: Definition of water vapour path?*

**Author's reponse:** The water vapour path (“wvpa”) refers to the vertically integrated water vapour and is therefore a measure of humidity contents in the atmospheric column. It is thus suitable to use as an indicator of the ambient atmospheric conditions. For more information regarding the HOAPS-3.3 wvpa retrieval, please refer to Schlüssel and Emery (1990).

**Changes in the manuscript:** The term „water vapour path“ has been replaced by “vertically integrated water vapour“ (P.9, L.2).

**9) Comment from referee:** *Page 8, Lines 11-14: It seems that HOAPS  $q_a$  is wet biased in the tropical wet zone and dry biased in the subtropical dry zone. The bias pattern seems to be similar to GSSFT v3  $q_a$  product (Prytherch et al. 2014, Int. J. Climatol.; Jin et al. 2015, J. Atmos. Ocean. Technol.).*

**Author's reponse:** Thank you for pointing this out. Indeed, Figure 4c in Prytherch et al. (2014) shows a strong resemblance between HOAPS-3.2 and GSSTF3. Both data records are based on the same algorithm and follow an inter-satellite calibration procedure. The minor differences in the tropics are thought to be related to either different quality control standards or differing Earth incidence angles. Given the close resemblance of GSSTF3 and HOAPS-3.2 shown in Prytherch et al. (2014), the difference pattern (GSSTF minus buoys and OAFlux) shown in Jin et al. (2015) was to be expected. The distribution is closely related to the  $q_a$ -dependent bias pattern shown in our manuscript (Fig. 2a).

**Changes in the manuscript:** Prytherch et al. (2014) is cited in this context (P. 9, L.9f).

**10) Comment from referee:** *Page 8, Lines 21-22: Indeed, the 1-D bias analysis is not sufficient. Please provide a figure showing the global pattern of the mean differences between HOAPS and the reference data. Need to discuss the uncertainty pattern in terms of humidity regimes.*

**Author's reponse:** Thank you for your suggestion. Originally we thought the reader would be distracted by such a difference map, as we would like to emphasize the importance of considering *multiple* atmospheric state parameters, i.e., the multi-dimensional bias analysis. However, we agree that the manuscript improves when including such a difference map (HOAPS minus in situ  $q_a$ ).

**Changes in the manuscript:** The difference map has been included into the revised manuscript (Fig 1, right panel). It is briefly described in Section 3.1 (P.9, L.10f,L.25f), where a connection to Fig. 2a (of revised manuscript) is established.

**11) Comment from referee:** *Page 9, Line 24: “Recall that the aim is to characterize uncertainty and not bias patterns”. The sentence is confusing. Bias is one kind of uncertainties.*

**Author's reponse:** We disagree with this statement. According to the International Vocabulary of Metrology (VIM, 2012), the (measurement) uncertainty is a *non-negative* parameter characterizing the *dispersion* of the quantity values being attributed to a measurand, based on the information used (VIM, 2.26). By contrast, a (measurement) bias (VIM, 2.18), which corresponds to an estimate of a systematic measurement error (VIM, 2.17), may be either positive or negative and, if known, can be corrected for. Keeping these two definitions in mind, a bias, which is a signed value, is strictly speaking not a kind of uncertainty. In order to turn the bias into an uncertainty estimate, we use the absolute systematic difference as an upper boundary of the (more simple) bias distribution.

**Changes in the manuscript:** The wording in the revised manuscript has been modified and moved

further up in Sect. 3.2 (P.10, L.17-22).

**12) Comment from referee:** Page 10, Eqs (2)-(3): Which figures are produced from Eqs.(2)-(3)?

**Author's reponse:** Figs. 3-6 are based on Eqs. 2 and 3. Details are provided in the following.

Whereas Eq. 3 merely expresses that the total instantaneous LHF uncertainty consists of a systematic and a random component, Eq. 2 forms the basis of LHF pixel-level uncertainties using uncertainty propagation. That is, applying Eq. 2 equips each LHF pixel with a *total*, that is systematic plus random uncertainty contribution. In consequence, Figure 4d directly results from Eq. 2, that is the systematic uncertainty contribution (the random component converges to zero, due to averaging over long time period). Likewise, the systematic uncertainty contributions by  $U$ ,  $q_s$ , and  $q_a$ , which contribute to Eq. 2, are illustrated in Figs. 4a-c.

Note that the random uncertainty measures resulting from Eq. 2 still incorporates random uncertainty contributions of the collocated in situ data ( $E_{ins}$ ) as well as the collocation procedure itself ( $E_c$ ). Each random uncertainty contribution resulting from Eq. 2 needs to therefore be corrected to isolate the random *retrieval* uncertainty. This random retrieval uncertainty is what we would like to characterize in the HOAPS climatology. The random LHF uncertainty resulting from Eq. 2 is therefore corrected pixelwise, using the results of the random uncertainty decomposition (see Sect. 3.4 and e.g. Figure 2 in Kinzel et al. (2016) for  $q_a$ ). The average field of these instantaneous, corrected random retrieval uncertainties is shown in Fig. 3d. Respective random retrieval uncertainty components contributed by  $U$ ,  $q_s$ , and  $q_a$ , are shown in Figs. 3a-c, respectively. As noted in the manuscript, Fig. 3 shows the *instantaneous* point of view, that is  $N=1$ . Likewise, Fig. 5 shows both systematic (rectangles) and instantaneous random retrieval (bars) uncertainties. It therefore shows the maximum uncertainty one can expect for a single pixel for different geographical regimes. Figure 5 is therefore based on both Eqs. 2 and 3. The same accounts for Fig. 6. The technical aspects are described in Sect. 3.4-3.5 .

**13) Comment from referee:** Page 10, Line 10: Why only random satellite retrieval component, not the total random uncertainty, is computed?

**Author's reponse:** The purpose of our uncertainty characterization is to assign systematic, random, and sampling uncertainties to all *satellite-related* LHF parameters. This approach is unique and important, as simply assigning total random uncertainties does not allow the user to understand to what extent they are associated with the retrieval itself or other uncertainty sources. This implies that contributions by collocation ( $E_c$ ) and in-situ data ( $E_{ins}$ ) need to be corrected for (i.e., removed) by applying the random uncertainty decomposition (Sect. 3.3). What remains is the random retrieval uncertainty, which consists of both random model uncertainty ( $E_M$ ) and sensor noise ( $E_N$ ) (see Kinzel et al. (2016), their Eq. 5).

Immler et al. (2010) formulate an implication of such an approach for consistencies like this: „Roughly speaking, consistency is achieved when the independent measurements agree within their individual uncertainties“ (their Sect. 2.5, Eq. 6). In other words, the decomposition of uncertainties allows for comparing two independent measurements with *own* (that is, independent) uncertainties, which makes conclusions regarding consistency more meaningful. The decomposition and contributing random uncertainties are thoroughly explained in Kinzel et al. (2016), their Sect. 2c.

**Changes in the manuscript:** Immler et al. (2010) has been added to Sect. 1 for a clearer motivation of our uncertainty decomposition approach (P.4, L.1f).

**14) Comment from referee:** Pages 10-11, sections 3.4-3.5: The two sections are not directly related to any figures. Suggest to revise and combine.

**Author's reponse:** We disagree that these two sections are not directly related to any figures/tables in the manuscript. For transparency, we believe a clear separation of all HOAPS-related uncertainties, that is systematic and random retrieval uncertainty (Sect. 3.3-3.4) and sampling uncertainty (Sect. 3.5), is appreciated. Sect. 3.3 is a main prerequisite for what is shown in Figs. 3 and 5, respectively. Sect. 4.3 (and Table 2 therein) is dedicated to only  $E_{smp}$ , which is first picked up

in Sect. 3.5.

**15) Comment from referee:** *Page 13, Line 10: Fig.2 is regarded as a 2-D representation of the error bar magnitude of Fig.1a. A figure showing the global pattern of HOAPS3.3 - minus - in situ needs to be provided to help interpret Fig.2.*

**Author's reponse:** The differences map points at biases, which are not linked to the random retrieval uncertainties shown in Fig 3a. Yet, the differences map (HOAPS minus in situ) has been added to the revised manuscript, where it is also commented on (P.9, L.10f,L.24f). This is already picked up in a different context (see comment #10 on this). As noted in the manuscript, the quoted passage is meant to qualitatively link the error bars in Fig. 2a to the four-dimensional (Fig. 3a) *random retrieval uncertainty* representation. Differences in their magnitudes were to be expected, as the bars in Fig. 2a include both  $E_C$  and  $E_{ins}$ , which have been corrected for in Fig. 3a. However, the  $q_a$ -dependent distribution of error bar magnitudes (Fig. 2a) are very closely related to the  $E_{retr,ran}$  pattern (Fig. 3a) . That is, random retrieval uncertainties are largest for subtropical ranges of  $q_a$  (11-17  $g\ kg^{-1}$ , Fig. 3a), which is mirrored in largest uncertainty bars in Fig. 2a. Likewise, these magnitudes reduce for tropical  $q_a$  ranges of roughly 20  $g\ kg^{-1}$ . Smallest magnitudes are generally found in high latitudes, where  $q_a$  is smallest (below 7  $g\ kg^{-1}$ , see Fig. 2a). The intention was to show the spatial distribution of random uncertainty in HOAPS-3.3  $q_a$ . As mentioned later on, this random uncertainty can be neglected if monthly to multi-annual averages are considered, while systematic components become the dominating source of uncertainty. Spatial maps of these long-term means of systematic uncertainties are provided in Fig. 4.

**Changes in the manuscript:** See comment #10.

**16) Comment from referee:** *Page 13, Fig. 2: The instantaneous random uncertainty map of  $q_a$  (Fig.2a) has a pattern similar to the uncertainty map of  $q_a$  produced by OAFflux (Yu et al. 2008, OAFflux technical report), though HOAPS3.3 has a much larger magnitude.*

**Author's reponse:** Thank you for bringing up this comparison. We agree that the error distribution shown in Yu et al. (2008) resembles our instantaneous random uncertainty distribution. Regarding uncertainty magnitudes: Yu et al. (2008) declare “mean errors“ as monthly mean standard deviations (std) (time period: 1958-2006). This definition considerably differs from our approach. Furthermore, it remains unclear as to how this std is derived. Apparently, several data sets contribute to its estimation (NCEP1, NCEP2, ERA40, satellites), which may be the cause for lower magnitudes shown in their Fig. 21. Whereas our uncertainty estimates are exclusively HOAPS-related (that is, related to only one data record), the error estimation presented in Yu et al. (2008) does not clarify as to how the global error distribution includes contributions by the individual data sets.

**17) Comment from referee:** *Page 14, Line 3: In addition to Table 2, please add a zonal-mean average of the monthly mean sampling uncertainties to show the latitudinal distribution of the uncertainties.*

**Author's reponse:** We investigated the latitudinal dependency of all sampling uncertainties. Due to the large averaging time period (monthly means), there is hardly any zonal dependency evident in any of the parameters (not shown). This was to be expected, as a differentiation between tropical and extratropical buoys for quantifying monthly mean sampling uncertainties did not reveal differences in uncertainty magnitudes (see end of Sect. 3.5). As indicated in Table 2, sampling uncertainties averaged over such long time scales only show a dependency on the amount of orbiting platforms. However, this effect is not seen in the zonal means, as at least three instruments were in operational mode between 1995-2008.

**Changes in the manuscript:** A comment has been included into the revised manuscript (P.12, L.26f) that no latitudinal dependency of the sampling uncertainties exists on the monthly mean basis.

**18) Comment from referee:** Page 14, Line 13: How is  $E_{clim}$  defined? Please provide a mathematical expression of  $E_{clim}$ .

**Author's reponse:** Please refer to comment #7 on this.

**Changes in the manuscript:** Please refer to comment #7 on this.

**19) Comment from referee:** Page 14, Line 15: “Figures 3a-e can also be treated as the systematic uncertainty distribution”. What is the relation between Figures 3a-e and the mean difference map of HOAPS-3.3 minus in situ? See comment Page 13, Line 10. The maps shown in Figures 3a-e are not bias patterns, as bias has both positive and negative signs. What is the meaning of the systematic uncertainty?

**Author's reponse:** We apologize that the current formulation may be confusing. Regarding the phrase you quoted: When averaging over 25 years (1988-2012), random and sampling uncertainties become virtually zero. This implies, given our definition of  $E_{clim}$  (see comment #7 on this), that  $E_{clim}$  is practically equal to the systematic uncertainty ( $E_{sys}$ ), which in turn is the absolute representation of the bias (see Sect. 3.2). Throughout our manuscript, we do not speak of „bias patterns“, as we are characterizing *uncertainties*, which are per definition non-negative. The average of an array of biases with respect to a reference can be zero, while none of the individual match ups are actually equal. This automatically points at a non-zero uncertainty. In this regard, we agree that Fig. 4 (of revised manuscript) does not show bias patterns (unlike Fig. 1 (right) in the revised manuscript), but rather patterns of  $E_{sys}$ .  $E_{sys}$  is therefore the *upper boundary* of the (more simple) bias distribution (see Sect. 3.2).

**Changes in the manuscript:** The  $q_a$  difference map (HOAPS minus in situ) has been included into the revised manuscript (Fig. 1 (right), see comment #10 and #15 on this). It is briefly described and related to Figs. 2a (P.9, L.25). Also, the composition of Sect. 3.2 has been changed.

**20) Comment from referee:** Page 18, Line 5: “On average, it increases by roughly  $4.5 \text{ W m}^{-2}$  (4.7%) per decade...”. Which term gives rise to this large increase,  $q_a - q_s$  or  $U$ ? The continuing increase in LHF during the “hiatus” period in the 2000s does not seem realistic from the perspective of the global water budget balance (see Robertson et al. 2014, J.Clim).

**Author's reponse:** Thank you for bringing this up. As mentioned in Sect. 4.7, this linear LHF increase over time is picked up by numerous studies and is resolved in several climatologies. Yu et al. (2007), for example, point at an OA Flux LHF increase of  $9 \text{ W m}^{-2}$  over a time period of 22 years (1981-2002), which closely resembles our linear trend estimate. Our trend analysis includes a strong negative offset in HOAPS LHF during 1991. As pointed out in the manuscript (P.18, L.25ff), this is associated with retrieval issues related to the Mount Pinatubo eruption and is therefore an artificial signal. If this is solved, as has been done for the latest HOAPS version, HOAPS 4.0 (Andersson et al., 2017), the offset is smaller, which ultimately reduces the linear trend. Also, possibly related to the hiatus, global mean HOAPS LHF slightly decrease after 2008. GSSTF3 also exhibits an LHF increase up to 2007/8 (which is even stronger than that of HOAPS) and a subsequent decrease (see Robertson et al. (2014), their Fig.2b and Fig.8). Regarding the increase up to 2008, the same conclusion may be drawn for SeaFlux (Robertson et al. (2014), their Fig. 2c). As to the cause of the LHF increase: Q-term analysis indicates that linear trends of both  $U$  and  $q_s$  are positive, whereas that of  $q_a$  is negative. In consequence, both  $U$  and  $(q_s - q_a)$  give rise to the observed LHF increase. For the time period of 1988-2005, this also becomes evident in Iwasaki et al. (2014), their Fig. 9.

**21) Comment from referee:** Page 19, Line 14: Remove the sentence. Aren't the uncertainty estimates supposed to be a common practice for all gridded products?

**Author's reponse:** We think that it is appropriate to include this sentence in our manuscript, as we are not aware of any other satellite climate data set with such an (extensive) uncertainty characterization. We certainly agree that this should be a common practice in the future. It seems, however, that HOAPS-3.3 (and HOAPS 4.0, by now) leads the way.

### **Cited studies:**

**Andersson, A.** Graw, K., Schröder, M., Fennig, K., Liman, J., Bakan, S., Hollmann, R. and Klepp, C.: Hamburg Ocean Atmosphere Parameters and Fluxes from Satellite Data - HOAPS 4.0, Satellite Application Facility on Climate Monitoring (CM SAF), doi:10.5676/EUM\_SAF\_CM/HOAPS/V002, 2017.

**Bentamy, A.,** Katsaros, K. B., Mestas-Nuñez, A. M., Drennan, W. M., Forde, E. B., and Roquet, H.: Satellite Estimates of Wind Speed and Latent Heat Flux over the Global Oceans, *J. Climate*, 16, 637–656, doi:10.1175/1520-0442(2003)016<0637:SEOWSA>2.0.CO;2, 2003.

**Brodeau, L.,** Barnier, B., Gulev, S. K., and Woods, C.: Climatologically Significant Effects of Some Approximations in the Bulk Parameterizations of Turbulent Air-Sea Fluxes, *Journal of Physical Oceanography*, 47, 5–28, doi:10.1175/JPO-D-16-0169.1, 2017.

**Immler, F. J.,** Dykema, J., Gardiner, T., Whiteman, D. N., Thorne, P. W., and Vömel, H.: Reference Quality Upper-Air Measurements: guidance for developing GRUAN data products, *Atmospheric Measurement Techniques*, 3, 1217-1231, doi: 10.5194/amt-3-1217-2010, 2010

**International Vocabulary of Metrology** – Basic and General Concepts and Associated Terms (VIM 3rd edition), JCGM, 200, 2012

**Iwasaki, S.,** Kubota, M., and Watabe, T.: Assessment of various global freshwater flux products for the global ice-free oceans, *Remote Sensing of Environment*, 140, 549-561, doi: 10.1016/j.rse.2013.09.026, 2014.

**Robertson, F. R.,** Bosilovich, M. G., Roberts, J. B., Reichle, R. H., Adler, R., Ricciardulli, L., Berg, W., and Huffman, G. J: Consistency of Estimates Global Water Cycle Variations over the Satellite Era, *Journal of Climate*, 27, 6135-6154, doi: 10.1175/JCLI-D-13-00384.1, 2014

**Schlüssel, P.:** Satellite Remote Sensing of Evaporation over Sea, in: *Radiation and Water in the Climate System: Remote measurements*, Vol. 45, NATO ASI Series, pp. 431–461, Springer-Verlag, Berlin, Germany, 1996.

**Wells, N.** And King-Hele, S.: Parameterization of tropical ocean heat flux, *Quart. J. Roy. Meteor. Soc.*, 116, 1213–1224, doi: 10.1002/qj.49711649511, 1990

**Yu, L.,** Xiangze, J., and Weller, R.A.: Multidecade Global Flux Datasets from the Objectively Analyzed Air-sea Fluxes (OAFlux) Project: Latent and Sensible Heat Fluxes, Ocean Evaporation, and Related Surface Meteorological Variables, OAFlux Project Technical Report, OA-2008-01, WHOI, 2008.