WCRP Surface Flux Task Team

A white paper outlining the need for a coordinated high-level approach to improving our understanding of surface-atmosphere fluxes

Developed by
Carol Anne Clayson, Woods Hole Oceanographic Institution
Anna Rutgersson, Uppsala University
Brian Ward, University of Galway
Ronald Buss de Souze, National Institute for Space Research
James Edson, Woods Hole Oceanographic Institution
Peter Gleckler, LLNL
Petra Heil, Australian Antarctic Division and ACE CRC, University of Tasmania
Kazuhito Ichii, Chiba University and National Institute for Environmental Studies
Martin Jung, Max Planck Institute for Biogeochemistry
Diego Miralles, Ghent University
Pierre-Philippe Mathieu, ESRIN
Paul Stackhouse, NASA Langley

Contents

Summary

The exchanges of momentum, heat, water vapor and other gases across the surface-atmosphere interface drive atmospheric and oceanic circulations, and are fundamental to changes in vegetation, soil moisture and cryosphere. Global and local energy and water balances are highly dependent on these fluxes, and changes in atmospheric and surface chemical and biological compositions are greatly influenced by variability in these exchanges. Sea ice and glaciers are impacted by changes in fluxes and in turn changes in the cryosphere feedback to the atmosphere and ocean. While there are differing issues associated with atmosphere-surface fluxes over land, ice, and the ocean, there are also similarities that can be exploited, such as in methods for in situ and satellite measurements, in understanding of how to make the best use of sparsely spaced in situ observations, and in exploiting the complementary strengths of in situ and remotely sensed data and numerical modeling studies. A number of programs and projects under the WCRP umbrella have had surface fluxes as a component, such as the GEWEX Data and Assessments Panel (GDAP), CLIVAR, and CLiC, and there have been a few directed projects such as the WCRP/SCOR Working Group on AirSea Fluxes (WGASF). These efforts have all brought progress in measuring, understanding, and modeling the fluxes, but there is a growing recognition of not only the importance of air-surface fluxes in the climate system, but also that the rate of progress in this field has not matched the rate of progress in other components of the climate system, and that even when new observations, understanding, and models are developed, various communities remain unaware and lag considerably behind in making use of these assets. For that reason WCRP has requested the WCRP Data Advisory Council (WDAC) to establish a Surface Flux Task Team. This white paper briefly describes the current state of our surface flux knowledge, highlights some recommendations for progress over the next decade, and lists some specific ways in which this task team can move the science forward.

Background

The current inability to provide highly accurate global fields of air-surface fluxes is a critical source of uncertainty in closing the global energy, water, and carbon cycles (Figures 1 and 2; see discussions in Rodell et al. 2015, L'Eucyer et al. 2015, and Le Quéré et al. 2018 for discussions of the water budget, the energy budget, and the carbon budget, respectively). The interface between atmosphere-ocean, atmosphere-land, and ocean-ice systems represents the coupling of Earth System components operating physically on different timescales such that the interactions between them lead to variations and changes in the states of the climate system.

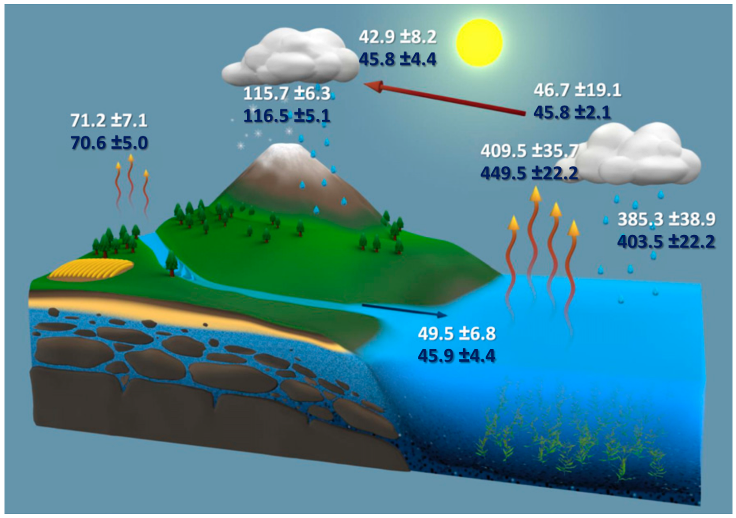

Figure 1. Mean annual fluxes (10^3 km^3 yr^{-1}) of the global water cycle, and associated uncertainties, during the first decade of the millennium. White numbers are based on observational products and data integrating models. Blue numbers are estimates that have been optimized by forcing water and energy budget closure, taking into account uncertainty in the original estimates. From Rodell et al. (2015).

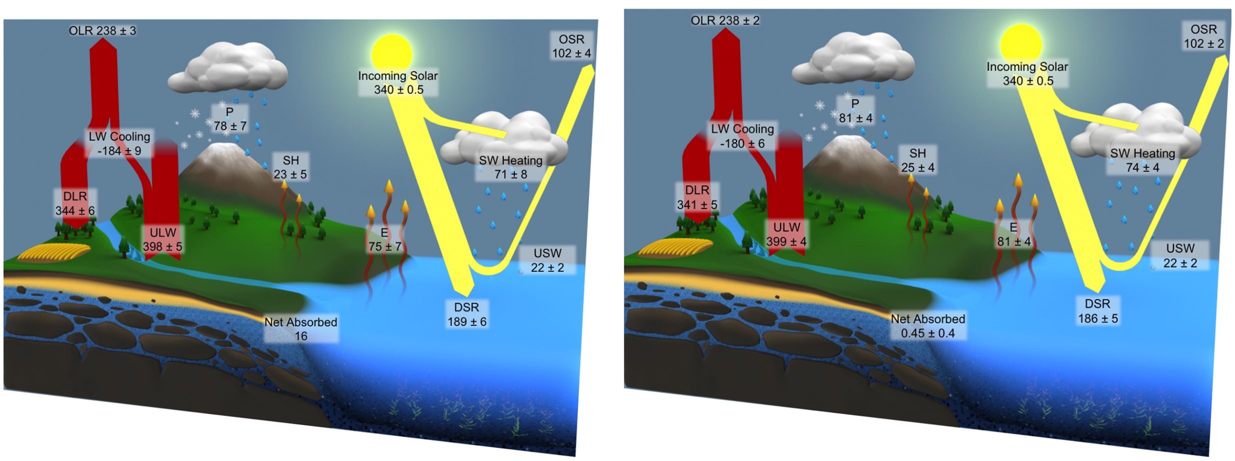

Figure 2. As in Figure 1, but for fluxes (W m^{-2}) of the global energy cycle. Panel on left is based on observational products and data integrating models; panel on right are optimized estimates. Adapted from L'Ecuyer et al. et al. (2015).

The exchange of heat, moisture, and momentum between the ocean and atmosphere helps drive the atmospheric circulation, contributes to precipitation variability, and modulates the heat storage of the ocean. The carbon exchange between the ocean and atmosphere represents one of the significant unknowns regarding uptake by the surface and the ability of the Earth system to store carbon outside of the atmosphere. Thus measuring these fluxes is imperative to understanding the processes leading to variability of the Earth system. For example, knowledge of how much heat and carbon the ocean absorbs is vital to understanding sea level rise and to predicting how much, how fast, and where the atmospheric temperature will change, as the ocean stores more than 90%

of the excess heat that has been added to the climate system over recent decades (e.g. Cheng et al. 2017) and roughly 30% of the excess carbon (Mikaloff-Fletcher et al., 2006; Le Quéré et al., 2010).

It is also essential to understand the partitioning of the global energy imbalances between the atmosphere and ocean, as this is critical to defining the climate response to anthropogenic forcings. Currently, although various analyses of the surface energy and water budgets close to within the uncertainties, these uncertainties are large enough to preclude the data being able to answer numerous scientific questions (e.g. Stephens et al. 2012; L'Ecuyer et al. 2015). Much of the uncertainty in these estimations are to a large extent due to imbalances between the radiative and turbulent heat fluxes and the evaporation and precipitation across the ocean surface. As can be seen in Figure 1, the evaporation from the ocean to the atmosphere represents the single largest term in the global water budget, and current estimates of the total magnitude require significant adjustment to bring balance to the budget. The warming of the upper ocean tends to be reflected in trends of increasing evaporation (i.e., water vapor flux) and precipitation in the global hydrologic cycle, all of which can be and have been observed by evaluating the upper ocean salinity over time (e.g. Durack and Wijffels, 2010). These variations in the global hydrologic cycle are not restricted to over-ocean locations but affect precipitation patterns across land surfaces as well. The movement of water from the ocean to the atmosphere, where it then becomes available to yield precipitation over both the ocean and land surfaces, is vital to life on land. However, our ability to predict the timing and magnitude of these variations is due in part to the uncertainties in the current global air-sea flux products which prevent them from being used to quantify the trends in either the heat or moisture budgets (IPCC 2013), as uncertainties are on the order of 10 to 20% (e.g. Gulev et al. 2010).

About half of the net radiation at the Earth's land surface is used as latent heat to evaporate water, while the remaining energy is returned as sensible heat to the atmosphere (L'Ecuyer et al. 2015, Jung et al. 2019). In water terms, while the flux of evaporation for the entire globe (oceans and continents) is expected to equal precipitation, over the continents evaporation still accounts for approximately two-thirds of the incoming precipitation (Gimeno et al. 2010). Therefore, continental evaporation is a key mechanism governing terrestrial hydrometeorological dynamics, from catchment to continental scales. A number of studies have highlighted its impact on climate trends (e.g., Douville et al. 2013, Sheffield et al. 2012) and the crucial role that the surface heat flux partitioning plays during meteorological extremes such as droughts or heatwaves (e.g., Teuling et al. 2013, Miralles et al. 2014a). The partitioning of net radiation between sensible and latent heat over the continents is still responsible for a large part of the uncertainty in global energy and water balances (Dolman et al. 2014). Therefore, the ability to monitor these turbulent fluxes over land is critical for climate research.

Despite this importance, rigorous and continuous *in situ* measurements of turbulent heat fluxes over land are mostly limited to those by the international network of eddy covariance sites (FLUXNET); their records are short and their coverage is insufficient for direct continental appraisals. Unfortunately, satellite sensors are unable to detect these fluxes over land in a direct manner. Nonetheless, pioneering efforts targeting the monitoring of continental sensible and latent heat have been proposed during the past decade (e.g., Fisher et al. 2008, Jung et al. 2010, Jung et al 2011, Mu et al. 2011, Miralles et al. 2011, Zhang et al 2016). Most methods build upon a long legacy of regional-scale studies aiming to combine remotely-sensed environmental and climatic drivers to indirectly derive sensible and latent heat fluxes (e.g., Price 1982, Anderson et al. 1997, Su 2002). Alternatively, machine learning methods are being used to upscale FLUXNET measurements to global gridded products using remote sensing as key input (Jung et al. 2010, Jung et al. 2011, Tramontana et al. 2016, Jung et al. 2019, Bodesheim et al. 2018). Multiple international activities have fostered the development of these remote sensing based algorithms, such as the European Union Water and global Change (WATCH) project, the LandFlux initiative of the Global Energy and Water cycle Exchanges (GEWEX) project, and the European Space Agency (ESA) Water Cycle Multi-mission Observation Strategy (WACMOS) ET project. Moreover, land evaporation and sensible heat fluxes have been included by the Global Climate Observing System (GCOS) within the next generation of Essential Climate Variables (ECVs).

Nowadays, we are able to map land heat and water fluxes remotely, and at multiple scales; these efforts support the current network of eddy covariance FLUXNET towers to enhance understanding on the continental variability of surface heat fluxes (Fisher et al. 2017; McCabe et al., 2017). Although many of these global data sets were originally intended for climatological-scale applications, some have evolved to provide estimates in operational mode, with ongoing efforts aiming to reduce product latency and improve spatial resolution (Ghilain et al. 2011, Anderson et al. 2011, Martens et al. 2018). The recent proliferation in the use of these datasets responds to a need for accurate land–atmosphere exchange data, not just within the climate and hydrology communities but also in the agricultural and water management sectors, and not just for scientific purposes, but also responding to government and commercial interests. In recent years, observation-based data of sensible and latent heat over land have been used for a wide variety of scientific explorations and societal applications, including, but not limited to (a) diagnosing the influence of the land surface on global warming and atmospheric CO_2 concentrations, (b) constraining model estimates of convection and cloud formation in the troposphere, (c) studying the relevance of water vapour, lapse rate and cloud feedbacks, (d) unravelling the two-way interaction between vegetation and climat e processes, (e) monitoring drought and heatwave occurrence and their impacts, (f) assessing crop water consumption and the efficiency of agricultural practices, (g) managing water resources at multiple scales, and (h) benchmarking climate model representation of these processes.

However, the WACMOS-ET and LandFlux projects also brought to light the large discrepancies among widely used, observation-based datasets of land turbulent heat fluxes, particularly in semiarid regions and tropical forests (e.g., Michel et al. 2016, Miralles et al. 2016, McCabe et al. 2016, Talsma et al., 2018). Consequently, further development and improvement retrieval algorithms appears crucial. This has been acknowledged by major national and international organizations, including the World Climate Research Programme (WCRP), the United Nations Food and Agriculture Organization (FAO), and the USA Global Change Research Program (USGCRP) and National Research Council (NRC) – see Fisher et al. (2017).

Sea ice and the associated snow cover significantly affect the ocean-air fluxes and exchanges despite their relatively small vertical extent within the ocean-ice-atmosphere system. Sea ice and the overlying snow are spatially and temporally highly variable and their effect on the surface fluxes may change depending on the timescale of interest. Together they engage in suite of complex processes with the other components of the polar climate and ecosystem, all of which modify the transfer of energy and materials from the ocean into the atmosphere and vice versa. The magnitude and persistence of the effect of sea ice on these exchanges depends on the structural characterisation of the ice and overlying snow. Due to the challenges in obtaining (accurate and) sustained measurements of the ocean-ice-atmosphere system, the current observational record is sparse in time and space, and are associated with a substantial uncertainty. In the heat-flux balance of sea ice, we discern the turbulent heat fluxes from the radiative fluxes, conductive flux and the oceanic flux. Even though the former, such as the turbulent flux, can be measured directly (e.g., Conway and Cullen, 2013), the need for specialized *in situ* measurements prohibits this at scales beyond the local scale. Therefore the turbulent fluxes are generally approximated using linear or quadratic bulk equations, which are driven by the vertical shear in observed parameters such as wind velocity, temperature, humidity, density or particle count. Bulk parameterisations are empirical and their results diverge with increasing magnitude of the underlying parameter.

The exchange of momentum in addition to heat and moisture through the ocean surface drives the ocean circulation, upper ocean mixing, surface wave fields, and provides a drag on the atmosphere. The overturning circulation in high latitudes, which help drive the larger-scale thermohaline circulation, is highly dependent on air-sea fluxes. In addition, the momentum flux to the ocean and resulting breaking wave fields provide a significant fraction of the aerosols that form the basis in the lower atmosphere for cloud formation, particularly in remote areas of the ocean. Uncertainties in how sea spray is generated by breaking waves limit our understanding of the transfer of heat and momentum between the ocean and atmosphere (e.g. Mueller and Veron, 2014), as well as limiting our ability to accurately reproduce the feedbacks between the upper ocean surface, aerosols, and low-level clouds. A number of analyses using combinations of satellite-based data sets, in situ observations, and reanalysis products have shown varying changes in wind-stress trends at decadal to centennial time scales; however, the variability between these analyses is large, and thus our knowledge of the actual change in time of these fields is low (IPCC 2013). Increases in wind-stress fields from the tropical Pacific and Southern Ocean regions have been noted (e.g. Swart and Fyfe, 2012; Merrifield et al. 2012), but variations due to interannual variability are large and poorly understood (IPCC 2013) and complicated by sun-synchronous observations of a variable with substantial diurnal and semidiurnal variability (Wentz et al. 2017).

Understanding the robustness of common approaches to derive flux components is critical, especially considering recent change in environmental conditions. For example, a recent study (Wanninkhof and Trinanes, 2017) mapped the trends of oceanic wind speed from 1988 to 2014. The high-latitude Southern Ocean (50° – 62°S) is marked as a circumpolar band of high wind speeds. Its Western Pacific sector holds the highest wind speeds measured with monthly means in excess of 12 m s^{-1}. In the northern hemisphere wind speeds peak in the North Atlantic at about 10 m s^{-1}. The 27 year trend to 2014 is dominated by increasing wind speed with few reductions. The largest increase at about 0.05 m s^{-1} y^{-1} have been observed in near equatorial regions and various near-coastal regions in both hemispheres. The observed change across the Southern Ocean has been in response to the predominantly positive SAM during this period. Such changes in the wind speed over the sea-ice zone and nearby ocean affect the ice in multiple ways, including due to changes of ice advection, which might result, for example, in increasing the Marginal Ice Zone, or the energy transfer due to turbulent fluxes with increasing winds.

Similarly, the scientific community lacks a robust understanding of how to estimate the exchange of CO_2 between the ocean and atmosphere given the few variables that can be measured and the significant uncertainties surrounding other key variables (e.g. wind speed and thermal stratification) that are needed to calculate this exchange. On centennial time scales, basic scientific principles dictate that the ocean CO_2 amount will equilibrate with atmosphere CO_2, but an open question remains as to the rate at which this ocean uptake will occur, and our understanding of some key processes that control the carbon distributions in the ocean is still limited. Due to the dearth of observations, it is still unclear whether the rate at which the ocean is taking up CO_2 is changing, with some analyses indicating a decline in the ocean uptake rate of total CO_2 (Le Quéré et al., 2007; Schuster and Watson, 2007; McKinley et al., 2011), and others finding a lack of evidence for a decrease (e.g., Knorr, 2009; Gloor et al., 2010; Sarmiento et al., 2010). Future estimates are even more uncertain, with recent studies suggesting that the Southern Ocean might increase carbon uptake, possibly enough to change the global net uptake to increasing rather than decreasing (Doney 2010). Current observation-based estimates of the climate response of the global air-sea CO_2 flux currently do not include feedbacks from ocean warming and circulation variability, which could make a difference of 20 to 30% in the ocean response (IPCC 2013). Given that the ocean stores roughly 50 times as much inorganic carbon as the atmosphere (Sabine et al., 2004), variations in the exchanges between the ocean and atmosphere can affect the atmospheric concentration of CO_2, while also impacting the rate and magnitude of the ocean acidification (Doney et al. 2009).

Terrestrial based measurements of the budgets of carbon and other important gases such as methane have significant global uncertainty (e.g. Ciais et al. 2013). Recent studies show both increased terrestrial sinks due to CO_2 fertilization (e.g. Keenan et al. 2016) and decreased terrestrial sinks due to droughts and other extreme events (e.g. Brienen et al. 2015). The fluxes of carbon between the land and atmosphere includes various processes, including photosynthesis, respiration, land use change emission, and biomass burning. In recent years, more and more attention is paid to the fluxes other than from photosynthesis and respiration to explain the differences in the estimated CO_2 budget among different methods. Direct estimates of the carbon flux between the land surface and atmosphere are being undertaken across a variety of regimes, including tidal wetlands and peatlands, and these are newly being augmented with remotely sensed techniques (e.g. Lees et al. 2018).

Processes associated with air-surface fluxes occur across a wide range of scales, from the microscale up to global, and from seconds to centuries. There is a growing recognition within the community of the importance of ever smaller spatial scales for driving many of the interactions between the atmosphere and the ocean, land, and ice systems. For instance, mesoscale ocean eddies which are particularly prevalent near the western boundary currents and the Southern Ocean, have been shown to influence the air-sea fluxes at these scales (e.g. Chelton et al. 2004; Small et al. 2008). Eddy footprints are seen in the wind stress curl residuals, and in the presence of strong SST gradients the coupling between the atmosphere and ocean varies from atmosphere leading to ocean driving SST variability (Bishop et al. 2017). How models resolve these eddies and their interactions with the marine atmospheric boundary layer can create strong differences in their climate outputs (Yang et al. 2018). Whether these mesoscale interactions are similar at the even smaller submesoscale range is an active research question. Similarly to the ocean, the impact of fluxes associated with local-scale features, such as surface roughness as well as opening-water areas within the ice pack, such as polynyas or leads, dominate that associated with a uniform ice cover. Vertical temperature gradients (i.e., in excess of 18 K) associated with openings in the ice give rise to large turbulent fluxes contributing excessively to the regional heat budget (Taylor et al., 2018). Furthermore turbulent ocean-atmosphere fluxes in the presence of sea ice are strongly episodic due to the dynamic and thermodynamic response of sea-ice properties to synoptic-scale atmospheric forcing.

Measurement/modeling techniques and uncertainties

In situ measurements

Vital to establishing long-term records of surface measurements of turbulent and radiative fluxes is the importance of measurement systems, maintenance, calibration and data quality. Another key feature is the number and types of locations sampled and the consistency of those measurements. In this section, we briefly summarize the major issues regarding the measurements of surface fluxes.

Radiative fluxes

The quality and numbers of measurements deployed to measure the surface radiative fluxes has improved through the years. The energy balance at the surface of the Earth is sensitive to net (down minus up) radiative fluxes for energy transmitted from the sun to the surface or shortwave (typically 0.25 um to 4 um) and the net thermal infrared emission from the surface or longwave (typically 4 um to 40 um). In practice, most measurement sites only provide measurements of the downwelling quantities. Instruments are designed to measure the surface solar irradiance and a thermal infrared measurement WCRP/GEWEX commissioned federated networks such as the Baseline Surface Radiation Network (BSRN) have standardized deployment, calibration and data quality assessment best practices (Ohmura et al., 1998) which continue to be updated and refined (Driemel et al. 2018). The standard measurements at BSRN locations is to measure shortwave radiation using the combination of a direct solar measurement (normal incidence pyrheliometer on a solar tracker) and/or and a diffuse measurements (shaded pyranometer and/or a black and white pyranometer). Using this method, BSRN procedures show that using instrument intercomparisons uncertainties of +/- 21 W m^{-2} 1-3 minute measurements, +/- 7 W m^{-2} monthly averaged under ideal conditions and 15.6 +/- 7.8 W m^{-2} for 15-min in operational conditions (Michalsky et al., 2011; WCRP-19/2012). Thermal infrared measurements made with pyrgeometers are calibrated relative to the entire thermal emission spectra generally show uncertainties of +/-6 W m^{-2} for 1-3 minute averages, +/-4 W m^{-2} for monthly averages and 5.8 +/- 2.0 W m^{-2} for 15-min averages in operational conditions. Upwelling measurements of the shortwave and longwave fluxes at BSRN sites are provided but their horizontal scale applicability is far more limited than the downwelling measurements. The BSRN handbook contains numerous other requirements and best practices and these set the standard ground-based measurements on stable platform (MacArthur, 2005). Measurements from other sites are available but the large majority of these sites aren't able to replicate the measurement and calibration. For instance, the WMO has also commissioned the World Radiation Data Centre (WRDC) to archive additional surface radiation measurements offered from nations throughout the world through an international agreement. Some of these measurements sites provide multi-decade records and these are made available via the Global Energy Balance Archive (GEBA; Wild et al., 2017). These measurements are of lesser quality relative to BSRN and mostly shortwave only that use an unshaded pyranometer for the measurements which are more subject to angular and other instrument effects that increase the uncertainties of the measurements.

Measurements from the BSRN networks and other networks implementing BSRN standards have been and are being used for large numbers of studies and are invaluable for the validation of satellite and model based data sets. However, issues still remain in processing and utilization of the measurements. For instance, BSRN archives 1-3 minute average data products, but most users require longer temporal averages. Thus, this requires the generation of those averages. Many of these time series contain time periods of missing data and thus users must devise strategies for determining those averages despite those missing data periods. Missing data periods can increase to the uncertainties of the direct solar irradiance to near 10 Wm^{-2} but sensitive to various filling methods were mostly are in the 1-2 W m^{-2} range (Roesch et al., 2011). There is a lack of community standards in regards to the generation of time series averages.

Perhaps, the largest is with the surface measurements is simply due to the underrepresentation of the large regions of land surfaces in different climate regimes. Large areas of Asia and Africa have no BSRN quality measurements. There are currently about 66 BSRN measurements sites distribution worldwide, but only a few of these span longer than 20 years at this time. Some of the sites have stopped measurements and submitting data to the archives. There is a need to encourage researchers and governments to establish new radiative measurement sites and at least maintain the sites that are already existing.

In situ ocean measurements of radiative fluxes are made are a wide range of platforms including ocean buoy, platforms and ships. Ocean platforms provide the opportunity to deploy BSRN type quality measurements due to the stability present, but there are a very limited number of these locations. Ocean buoy measurements provide the opportunity to increase the spatial coverage of measuring systems and the uncertainties of those basic systems are summarized by Calbo and Weller (2009) under ideal conditions. However, buoy systems have the issue of sway depending upon the ambient wind conditions and that effect increases the uncertainties of the measurements depending upon location. Additionally, maintenance in relatively isolate regions can be an issue for certain buoys where aerosol deposition can bias the measurements in time (Foltz et al., 2013). Despite this ocean buoy measurements provide measurements of surface radiation where little other observations exist. The issues of the persistence of the buoys at a fixed location, regular maintenance and calibration, and data consistency and available do impact the usage of these more than at the BSRN land based networks. However, the largest issue in regards to the sampling of ocean radiative fluxes is vast areas particularly outside of the tropics where buoys are not typically deployed and no observations are made. Some of these gaps in data records are compensated by measurements from ship based measurement systems for which standard procedures have been established (Bradley and Fairall, 2006) and subsequently deployed Fairall et al., (2008). The difficulty of using ship data for model and satellite data validation is in movement of vessel. Thus, long-term accurate in situ surface flux measurements over the ocean are subject to higher uncertainties and even greater undersampling issues than the land-based surface measurements networks.

It is difficult to obtain high accuracy for radiative flux measurements on sea ice. Contributing factors are the icing over and/or condensation. Furthermore radiation measurements depend on the view angle. Instrument deployment on sea ice, which drifts in response to wind and ocean forcing and is subject to shear and deformation, is challenging. Consequently any deviation from level seating of the radiometer will degrade the accuracy of derived radiative fluxes. Typically radiation measurements on sea ice or snow include downward and reflected shortwave radiation as well as downward and upward longwave radiation. Standard sensors may be upward- and downward looking pyranometers for the shortwave and upward- and downward looking pyrgeometers for the longwave radiation. Sensors are mounted off a mast assembly about 1 to 1.5m above the surface. It is important to avoid sensor shading or sensor heating in the mount design (Vihma et al., 2009). Artificial temperature increase may be corrected empirically based on diffuse and direct shortwave radiation fluxes from the same site. Sea ice based measurements are mostly limited to various field campaigns such as SHEBA (Uttal et al., 2002) and thus long-term surface in-situ measurements of radiative fluxes are not available.

Turbulent fluxes

Turbulence is an efficient transport mechanism for both physical quantities such as heat and momentum as well as mass transport i.e. gases and particles. The vertical transfer or turbulent flux of these quantities can be directly obtained from the eddy-correlation (EC) or direct covariance (DC) technique. The EC method computes the fluxes by correlating the turbulent fluctuations of the vertical velocity with fluctuations of the air constituent in question (e.g., horizontal velocity for momentum, temperature for sensible heat, and specific humidity for latent heat). The instrumentation is based on a short-path length (in the order of 0.1m) 3-axis sonic anemometer/thermometer complemented with additional high-frequency sensors, e.g., infrared gas analyzers (IRGAs) to measure water vapor and carbon dioxide.

Surface fluxes measured with the EC method over the ocean are mainly taken at ships (Fairall et al. 1997; Pedreros et al. 2003: Moum et al. 2014), drifting spar buoys (Drennan et al. 2003; Sahlée et al. 2012; Edson et al. 2013), stationary moorings (Weller et al. 2012; Bigorre et al. 2013; Farrar et al. 2015; Clayson et al. 2019); over-ocean towers including the R/P FLIP (Smith et al. 1992; Mahrt et al. 1998, 2001; Edson et al 2007; Grare et al. 2013; 2018); land-based towers (Rutgersson et al., 2008) and occasionally low-flying aircraft (Mahrt et al. 2001). These different settings are complementary and have different advantages and disadvantages. Ships and buoys are frequently used to estimate buoyancy and momentum fluxes from sonic anemometers. However, the velocity measurements on these platform have to be corrected for motion contamination before the EC method is applied (e.g. Anctil et al. 1994, Edson et al. 1998, Pedreros et al. 2003, Miller et al. 2008; Landwehr et al., 2014; Flügge et al. 2016) making the method less direct than can be accomplished on fixed towers. Fixed towers over the open ocean, however, are few and far between. For land-based towers attempting for measure oceanic fluxes from shore it is crucial to make a careful analysis of the site and the data to investigate potential disturbances of limited water depth or land areas in the flux footprint (Rutgersson et. al, 2019).

The inertial dissipation (ID) method was widely used by marine researchers to reduce the impact of motion contamination of velocity measurements. The ID method relied on the insensitivity of the inertial subrange to platform motion (i.e., wave-induced motion is generally seen at lower frequencies in the energy containing subrange). Accurate estimates of the dissipation rate of TKE from the inertial subrange are then combined with the law-of-the-wall for dissipation to estimate the momentum flux (e.g., Fairall and Larsen, 1986; Edson et al. 1991). The ID method relies on the assumption that all of the energy flux into the surface layer is ultimately dissipated within that layer. Over the ocean, however, a significant fraction of that energy is transferred to ocean waves and currents. This leads to less dissipation than predicted by the law-of-the-wall and an underestimation of the momentum flux (Janssen 1999; Cifuentes et al. 2018) even if the dissipation itself is accurately estimated from the inertial subrange. Similar effects are seen over forest canopies and other surfaces. This dissipation deficit scales with the height of above the ocean surface, so the use of the ID method should be avoided near the ocean surface and care must be taken when the method is used with ship-based measurements at higher elevations (e.g., Edson and Fairall 1998). Additionally, accurate measurement of the dissipation rate from sonic anemometers are hindered by path averaging and aliasing (e.g., Henjes et al. 1999), which limits their frequency/wavenumber resolution at low heights and/or high winds.

EC systems in marine conditions therefore need to consider a number of concerns including sensor separation, flow distortion, motion contamination, salt contamination, limited frequency response and corrections due to density differences. Several studies have been discussing uncertainties in EC flux measurements including those found in Fairall et al. (1996, 2000), Vickers et al. (2010), Cronin et al. (2019) and references therein. The uncertainty is highest for large moving platforms such as ships with significant flow distortion (O'Sullivan et al. 2015). The uncertainty is reduced for small moving platforms such as discus moorings and drifting spar buoys due to reduced flow distortion and higher signal-to-noise ratios. These measurements work best when the sensors are in the surface layer but far enough away from the surface to ignore pressure-related terms associated with wind-wave coupling in the wave boundary layer (Hara and Sullivan, 2015).

Carbon dioxide fluxes are inherently difficult to measure over the ocean due to the very small CO_2 gradient (and low signal-to-noise ratio), and several corrections usually need to be applied due to the limitations of the presently available sensors and sampling systems (McGillis et al. 2001; Edson et al. 2011; Blomquist et al. 2014; Nilsson et al., 2018). The relative instrumental uncertainty in tower-based estimate of the CO_2 flux is estimated to be 17–20 % (Rutgersson et al. 2008; Vickers et al. 2010). The uncertainty in ship-based estimates of CO_2 are higher, but have been significantly reduced in recent field programs using improved methodologies such as drying the air sample prior to measuring the CO_2 with closed path IRGAs (e.g., Blomquist et al. 2017). Marine researchers benefit from large single-to-noise ratios for high frequency water vapor (H_2O) measurements form IRGAs. However, open-path IRGAs do not operate in rain and fog and suffer from contamination of their optics by sea-salt and dust. The use of closed-path IRGAs to measure water vapor fluctuations in high humidity marine environment is problematic due to signal attenuation within the sampling tube. As a result, marine researchers are moving towards the use of closed path IRGAs with dried air-samples for CO2 fluxes and frequently cleaned open path IRGAs to measure evaporation and latent heat fluxes.

Additionally, one needs to consider difficulties in severe environmental conditions, such as in the sea-ice zone, where the EC approach and required sensors are not suitable for seasonal or longer deployments. There have been few deployments of autonomous EC sensors on sea ice; most of which were short deployments during field campaigns (e.g., in the East Antarctic pack ice during October-November 2012). Measurements with sonic anemometers may be contaminated by icing or corrupted by reflections off mounts, the deployment tower or other nearby sensors or hardware. In general sonic anemometers are pointed into the direction of the prevailing wind direction. However, when deployed on pack ice, an anemometer moves and rotates within the framework of the drifting sea ice. This leads to a lower ratio of valid EC measurements obtained on sea ice than at (near-) stationary sites.

As a result, the deployment of EC systems over the ocean requires significant effort to deploy and analyze the required measurements. Instead, it is far more common for marine researchers to rely on the less direct bulk aerodynamic (BA) method (Liu et al. 1979; Smith 1988; Clayson et al. 1996; Fairall et al. 1996, 2003; Bourassa et al. 1999; Fairall et al. 2011; Edson et al. 2013). The BA method requires sea-surface differences of the velocity, temperature, humidity and gas concentrations depending on the desired flux. The differences are combined with the appropriate transfer coefficients to estimate the momentum, heat and mass fluxes. When using the BA method, Fairall et al. (1996) showed that the radiative fluxes; air-sea temperature, specific humidity and velocity differences and their related transfer coefficients need to be estimated with high accuracy to measure the surface heat budget with an uncertainty of less than 10 W/m^2. There exist numerous studies on bulk coefficients parameterised using wind speed, atmospheric stability, wave information, boundary layer height and other environmental conditions. The COARE-algorithm presently represents state-of-the art knowledge on the bulk coefficients (Fairall et al. 1996, 2003; Edson et al. 2013).

Errors in BA estimates of the surface fluxes arise from 1) errors in the state variables required to compute the sea-air temperature, specific humidity and velocity differences and 2) errors in the bulk parameterization used to relate these differences to the turbulent fluxes. Errors in the bulk parameterizations are generally attributed to the uncertainty in the transfer coefficients for momentum, sensible heat and latent heat. The transfer coefficients are generally parameterized as a function of atmospheric stability and surface roughness. The impact of atmospheric stability on the air-sea differences (i.e., the temperature, specific humidity and velocity profiles) are typically accounted for using Monin-Obukhov Similarity theory. Decades of research has shown that the

MOS theory is valid as long as the assumptions required by the theory are satisfied. Several of the key assumptions include stationarity, horizontal homogeneity and a constant turbulent flux layer.

For field programs designed to largely satisfy these assumptions, researchers have developed stability functions that work as well over a Kansas wheat field as they do over the open ocean. Departures from these assumptions, however, are relatively easy to find. For example, land surface are often characterized by horizontal inhomogeneity within the flux footprint. Although to a lesser extend, significant horizontal variability is observed in coastal ocean regions, near western boundary currents (e.g., the Gulf Stream) and over marginal ice zones and leads. The resulting horizontal variability in the fluxes drives variability in the remotely sensed surface characteristics that is difficult to quantify. Additionally, diurnal variability, frontal passages, squall lines and numerous other phenomena can violate the assumption of stationarity. This requires consideration of appropriate averaging times in the field and what defines turbulent versus mesoscale transport. The depth of the constant flux layer is clearly a function of atmospheric stability. However, numerical models tend to have insufficient vertical resolution and static grids, which often violates the use of BA flux algorithms based on MOS theory. Additionally, wave-induced surface fluxes and their modulation of the near-surface profiles are amongst many processes that lead to lead to violation of the constant turbulent flux layer assumption.

The second major component of a BA flux parameterization is surface roughness, which models the surface drag or friction. Typically, the transfer coefficients are first adjusted to neutral conditions using the stability functions. Neutral values of the transfer coefficients are then used to estimate the aerodynamic roughness lengths using the additional assumption of semi-logarithmic velocity, temperature and specific humidity profiles. Over land, static or slowing evolving (e.g., seasonal) roughness length can be used. Over the ocean, the surface roughness is a dynamic variable that tends to increase with increasing atmospheric forcing over short time scales. This forcing drives surface waves that provide the roughness elements. As such, the surface roughness has been parameterized as a function of surface forcing, sea-state (e.g., wave steepness and wave height) and wave-age. This has resulted in a wide variety of algorithms and significant uncertainty amongst the parameterizations. In particular, the coupled wind-wave processes that are thought to be important under extreme wind conditions (i.e., wind speeds greater than 25 m/s) are poorly understood due to the scarcity of measurements under these conditions. These processes include wave breaking, flow separation, bubble production and generation and transport of evaporating sea-spray, all of which impact the momentum, heat and energy exchange under these conditions. Innovative means to investigate these processes must be encouraged and funded.

Lastly, different methodology used to develop BA flux algorithms can drive errors and uncertainty in estimates of the fluxes. For example, modern BA algorithm require measurements of the skin temperature to provide the appropriate air-land, air-sea and air-ice temperature difference. However, the skin temperature can be difficult to measure in the field and provide for model initialization and updates. Over the ocean, the sea temperature is most likely measured (or provided to models) at depth and should be adjusted to account for the diurnal warm layer and a cool skin correction. This is not common practice and results in additional uncertainty in the fluxes. Additionally, BA algorithms should be developed using velocity measurements that are relative to the ocean surface, which requires estimates of the surface currents. Again, inclusion of surface currents to estimate the surface fluxes is uncommon both in the field and in numerical models.

For land-based EC fluxes, the comparatively large random errors (20-50%) of measured half-hourly latent and sensible heat fluxes by EC are not critical if the sample size is large enough but systematic errors propagate to the gridded flux products and cause biases. Perhaps the largest concern about land-based data arises from the so called energy balance closure gap. On average across all FLUXNET sites the sum of measured latent and sensible heat flux account for only roughly 80 % of measured net radiation (Wilson 2002, Stoy 2013), which implies that the latent, the sensible, or both fluxes are systematically biased low. The reasons for this problem remain debated and unresolved. Many different factors can contribute to an energy balance closure gap and can be broadly grouped into (a) issues of instrumental set-up and methodologies, and (b) flux advection. Some instrumental issues (e.g. calibration issues, footprint mismatch between net radiometer and turbulent fluxes, influence of topography or tower structure on net radiation measurements) can cause sizeable energy balance closure gaps but are unlikely to cause a systematic bias across the entire network of sites. Other instrumental problems can generate a systematic fractional loss of the signal, e.g. when averaging periods are not adequate to capture low frequency contributions (Charuchittipan et al., 2014; Finnigan et al., 2003), due to structural elements of the ultrasonic anemometer ('angle of attack' (Nakai et al., 2006) or loss of high frequency covariance due to line-averaging, instrument separation, and tube attenuation (Leuning et al., 2012a; Mammarella et al., 2009)). In sites equipped with closed path gas analyzers, the attenuation of the high-frequency fluctuations of measured water vapor related to the tube that connect the inlet to the analyzer is known to affect the LE measurements. Mammarella et al., (2009) show a systematic underestimation of LE due to tube length, and the underestimation increases with tube length and age, and for increasing relative humidity (RH). Flux advection occurs during conditions of insufficient turbulence (peres-priego) but also arises from landscape scale circulation patterns due to differential heating of heterogeneous surfaces and topographic variations (Stoy et al. 2013). The uncertainty arising from the energy balance closure gap at FLUXNET sites currently appears to be the most limiting factor for reducing the uncertainty of global latent and sensible heat flux estimates over land based on remote sensing and machine learning methods (Jung et al. 2019).

Direct flux measurements on sea ice are sparse, as instrument deployment and maintenance are an issue. Generally such measurements are limited to the duration of manned field campaigns. For example in early summer 2004/05 Ice Station Polarstern (ISPOL) was deployed in the western Weddell Sea hosting a variety of oceanographic, sea ice/snow and atmospheric observation campaigns. As part of ISPOL a 1.5km x 1.5km ice floe was instrumented with flux equipment at multiple sites. Instruments were deployed on the ice in late November 2004 and most equipment as removed in early January 2005. The few automated sensors which remained on the ice, were integrated in drifting sea-ice buoys with autonomous data relay (Heil et al., 2008. There have been few deployment of drifting ice buoys with AWS-type instrumentation to support *in situ* measurements of ice-atmosphere fluxes, This is mainly due to the difficulties to maintain autonomous ice-tethered atmospheric measurement suites (Lee et al., 2017).

While the lack of observations over sea ice and the ocean is particularly acute, there are a number of regions across land surfaces which are quite undersampled. While the ~200 FLUXNET sites with available data span large gradients from arctic to tropical climates, the limited number of sites and the uneven distribution is not ideal. More sites, in particular in not well represented regions like the tropics, the sub-tropics and the tundra would be needed to better capture the existing diversity of environmental conditions and ecosystem functional properties, in particular because these regions are subject to particularly large climate variability and change. Compared to the rest of the global oceans, the tropics are relatively well-sampled with long-term stations, in part due to the long term TAO, RAMA, and PIRATA arrays and other OceanSITES assets. Relatively fewer long-term measurements related to surface fluxes are available in the mid and high-latitude oceans; notable exceptions being the highly instrumented buoy deployed and maintained by the NSF Ocean Observing Initiative (OOI), the Kuroshio Extensive Observatory (KEO) and the Southern Ocean Flux Station (SOFS). Overall, however, the lack of data in mid to high latitudes and in key regions like the Gulf Stream has a significant negative impact on our process understanding for improved modeling and validation efforts for gridded products.

Some alleviation of the observation issue may be possible from the many new platforms and sensors that are being developed and deployed for the atmospheric and oceanic boundary layers and land surfaces. However, best practices and accepted theory need to be considered when developing methods to compute the fluxes. For example, the sensor response should be adequate to resolve the turbulent motions of interest at a given height. Validation of the fluxes against accepted standards should be undertaken as part of the development. The scientific community should encourage these efforts by developing test facilities that receive funding to support these activities.

Satellite-derived measurements

Unlike the fluxes at the top of the atmosphere, satellites cannot directly observe surface fluxes. Over the most recent decades satellite missions have gathered data on the energy fluxes from Sun and space to the Earth, measuring components of the Earth's radiation budget (Wielicki et al., 1996) or exploring the Sun's radiation budget (Anderson and Cahalan, 2005). A necessary component of the global observing system for surface fluxes are satellite-based products. Methodologies for retrieval of the radiative fluxes are similar across all surfaces, while there are significant differences in the estimation of the turbulent fluxes across land, ice, and ocean surfaces. Multiple international projects, including the GEWEX LandFlux and SeaFlux initiatives, are working to understand/reduce the uncertainties and improve the retrievals. The importance of improving these estimations was highlighted recently by the 2017 ESAS Decadal Survey, where every panel (except the Earth Surfaces and Interior panel) had improved measurements of boundary layer structure and/or surface heat fluxes and the heat and water cycles as components of at least one of the science/society questions, and of the 18 Science Questions with "Most Important" or "Very Important" objectives 11 called out the need for these observations (NAS 2018). What follows is a brief description of the current issues associated with retrievals of these fluxes, and estimates on the uncertainties.

Turbulent fluxes over the ocean

Over the ocean, satellite-based estimation of the surface turbulent sensible and latent heat fluxes requires use of a bulk aerodynamic flux parameterization and measurements of multiple parameters over the ocean surface, including winds, temperature, and humidity. At the ocean surface, measurements of sea surface temperature are needed, and depending on the wind measurement used, currents could be a necessary component (see above for discussion on this topic). Satellite-based SST measurements has a large community of developers and a well-organized international presence (the Group for High Resolution Sea Surface Temperature, GHRSST) designed to improve products, as well as sensors that can produce direct retrievals of SST. The International Ocean Vector Winds Science Team (IOVWST) is similarly robust, and based on sensors that have a relatively direct retrieval of winds. There are no such sensors nor large-scale groups focused on near-surface temperature and humidity, and these are the variables most in need of improvements. Issues still associated with SST and satellite flux measurements are related to the need for diurnal skin SST as opposed to bulk SST (Clayson and Bogdanoff, 2013, Cronin et al. 2019, and discussion above). Few products are available with diurnally-varying skin, and these typically use a modeled diurnal warming on top of the bulk SST (e.g. Clayson and Brown, 2016), or are based on new geostationary products such as the Himiwari-8 and MTSAT-2 SSTs (Ditri et al. 2018).

Although there is no direct sensitivity to the near-surface thermodynamic state, regression-based approaches have been used since the work by Liu (1986) that highlighted the connection between columnar water vapor and surface humidity on monthly time scales, and has since been followed by a number of researchers (Schulz et al. 1993; Roberts et al., 2010; Bentamy et al., 2013; Tomita et al. 2018). Near-surface air temperature is an even more challenging retrieval. Typically, air temperature is either based on reanalysis data, or as part of a retrieval that attempts to retrieve atmospheric surface humidity, temperature, winds, and SST concurrently (Roberts et al. 2010; Bentamy et al., 2013; Tomita et al. 2019). Despite this progress in retrieving the near-surface temperature and humidity fields from microwave imagers, significant regional biases and uncertainties remain (e.g. Bentamy et al. 2017; Roberts et al. 2019; Cronin et al. 2019). The additional use of microwave sounder information is also used for producing fields of near-surface humidity and temperature, sometimes with additional inputs from buoys and reanalysis fields (e.g. Jackson et al. 2009; Jackson and Wick 2010; Jin et al., 2015).

Studies such as Prytherch et al. (2014) that have intercompared these approaches find annual mean differences exceeding 1 g kg^{-1} and regional monthly mean differences of 2 g kg^{-1}. Further, they find monthly mean biases against surface observations that exhibit strong regional coherence; these regional variations were also noted in direct comparisons to research vessel observations in Brunke et al. (2011). Typical biases reported by the producers of the datasets range from 0.8 - 1.6 g kg^{-1} for humidity and 0.2 - 1.5K for temperature (Jackson and Wick 2010; Roberts et al. 2010; Tomita et al. 2018; Jin et al. 2015). However, there is little consistency in methods used, or comparison datasets, and getting to an uncertainty applicable to the global ocean state that can directly compare methodologies is still an open question.

Currently there are few groups working internationally to provide latent and sensible heat fluxes from satellite over the oceans, including the SeaFlux-CDR group (e.g. Clayson and Brown, 2016); the OAFlux group (e.g. Yu and Weller 2007); the J-OFURO group (e.g. Tomita et al. 2019); the IFREMER group (e.g. Bentamy et al., 2013) and the HOAPS group (e.g. Andersson et al., 2011). These groups mainly rely on the long time series of passive microwave imagers to provide combinations of SST, near-surface temperature, humidity, and wind speed, as well as columnar vapor and liquid water, which are inherently underconstrained by the imager observations. Mitigating these correlated errors will require providing additional constraints. Future satellite missions which are designed specifically to enhance our ability to determine near-surface temperature and humidity profiles will provide much-needed improvements to the measurements. In addition, these products have spatial resolutions at best at the 25 km (or 0.25o) scale, which does not resolve the mesoscale or submesoscale fields, and may lead to unknown additional errors. Finally, the use of the microwave imagers also precludes estimations of coastal fluxes, which will also lead to unknown additional errors in regional or global budgets.

Although there have been a number of comparisons of these latent and sensible heat flux datasets, there is no clear community consensus for how best to determine the global uncertainties of each of these products. Methodologies of the comparisons range from the use of buoys and research vessel data as a main *in situ* comparison (high quality data, but problematic given the non-uniform distribution of buoys and R/V ships across the global ocean), to comparisons of the datasets to each other and reanalyses, or to the use of the more globally-distributed ships of opportunity such as through the ICOADS dataset (Brunke et al. 2011; Bentamy et al. 2017; Roberts et al. 2019; Liman et al. 2018). Sampling issues related to *in situ* observations have been discussed by Barry and Kent (2016), where they note that there has been a 20% reduction in global coverage since the early 1990s. Thus the problem is not improving with time. Recommendations for improvements to the in situ coverage are described in Cronin et al. (2019). To carefully evaluate satellite-derived fluxes at a wide range of atmospheric conditions, more observations in mid- and high-latitude regions are needed.

Oceanic momentum fluxes and transfer transfer velocity

Scatterometers and microwave radiometers have been deployed on orbiting satellites for decades to estimate near surface wind speed from the backscattered signal from their emitted radiation. The rationale behind this approach is simple; i.e., the greater the wind speed, the rougher the surface; and the rougher the surface, the greater the return. Scatterometer systems include QuikSCAT, ERS-1, ASCAT-A & B, OSCAT, ISS-RapidScat, and SCATSAT-1. Satellite radiometer systems include SSM/I, AMSR-2 & E, GMI, and WindSAT. The principle difference between these system is the wavelength used to make the measurements, which determines the primary scatterers that provide the roughness seen by each system. For example, Ku-band QuikSCAT and C-band ASCAT reflect off gravity-capillary and short gravity waves, respectively, to provide estimates of wind speed and direction (Bourassa et al. 2010). The brightness temperature of the sea surface measured by microwave radiometers is also related to the ocean surface roughness (Meissner and Wentz, 2012). Radiometers only provide wind speed with the exception of the polarimetric radiometer, WindSat, which also provides the wind direction.

Geophysical model functions (GMF) have been developed for each of these systems to relate the returned signal to wind speed and, when possible, wind direction using data from buoys and numerical models. Since these systems measure surface roughness, researchers have long postulated that the returned signal is more closely related to wind stress than wind speed. This is particularly true of scatterometers, which reflect off the short wind waves locally generated by wind stress (Bourassa et al. 2010). Wind stress is proportional to the equivalent neutral (i.e., stability adjusted) wind speed relative to the sea surface. For this reason, the equivalent neutral wind speed adjusted to some reference height has been used to develop the GMFs. The reference height is traditionally chosen to be 10-m such that scatterometer wind retrievals are usually defined as the 10-m equivalent neutral wind relative to the sea surface, U_{r10N}.

The retrieved neutral winds can then be used to estimate the momentum flux at the ocean surface (i.e., the surface stress) using a neutral drag coefficient from bulk parameterizations. However, there are a number of issues associated with these estimates that could be addressed to improve surface stress retrievals. One basic issue involves the measurements used to compute U_{r10N}, which requires wind speed and direction, air and sea surface temperature, pressure, relative humidity and surface currents. However, surface currents are rarely available and RH measurements are often missing or suspect. Therefore, most of the U_{r10N} estimates used to train the GMF are only corrected for temperature stratification and are measured relative to earth. Additionally, the adjustment of the measured winds to neutral conditions at 10-m for GMF development requires the use of a bulk formula, which may not be the preferred formulation. In fact, it may be more appropriate to use the drag coefficient used to adjust the winds even if it is demonstrably less accurate than the state-of-the-art bulk parameterization. Lastly, the effect of non-locally generated wind-waves and swell are incorporated into the scatterometer GMFs, which increases the uncertainty in these functions.

To combat these issues, members of the remote sensing communities have stated the need to develop GMFs based directly on estimates of the surface stress. The stress estimates would come from the EC measurements being made from an increasing array of surface moorings, towers and mobile platforms. These would be supplemented with bulk estimates of the stress using state-of-art bulk parameterizations such as COARE with buoys and mobile platforms that measure all of the required variables needed to estimate the flux. The retrieved stress could then provide estimates of U_{r10N} by inverting the bulk parameterization. A number of the issues stated above would still apply, but these may be easier to quantify using stress as the dependent variable.

The flux of carbon between the ocean and atmosphere from satellite methods is difficult to quantify. Jackson et al. (2012) use satellite observations of near-surface temperature, humidity, and wind speed as well as observations of SST and longwave and shortwave radiative fluxes in combination with a gas version of the COARE flux algorithm to determine global estimates of CO_2 transfer velocities. These could be used with climatologies of pCO_2 in the atmosphere and ocean, to produce a flux; however, no current method exists for retrieving pCO_2 from satellites.

Continental turbulent fluxes

Satellite-derived continental turbulent fluxes require different inputs and methods than those for the ocean surface. While space-based Earth observing systems have provided an unprecedented ability to measure processes at the Earth's surface, sensible and latent heat fluxes cannot be directly observed from space, whether at local or continental scales. Current methodologies for large-scale retrieval concentrate on the derivation of latent heat flux (*λE)* by combining satellite observable variables that are linked to the evaporative process using process-based and empirical formulations. In all instances, there is an assumption that the models developed at the local scale are equally applicable at the field and larger scales: an occasionally heroic assumption, particularly where issues of strong land surface heterogeneity and paucity of meteorological forcing are prevalent (McCabe *et al.* 2015).

As over the ocean, continental sensible heat flux can be calculated from satellite-derived values of air temperature, surface temperature, and the judicious use of a bulk aerodynamic parameterization for the heat transfer coefficient, which requires information on winds and specific humidity. Using this direct approach, Siemann et al. (2018) combined satellite air temperature data with reanalysis-based air temperature, wind, and humidity products combined with land cover information to calculate the sensible heat flux over the global continents. Biases in the air temperature products compared to NOAA ground stations ranged from 0.2K to 1.2K. The resulting set of sensible heat flux products had biases from roughly -5 to 15 W m^{-2} when compared with FLUXNET towers. Other approaches use an energy balance model to estimate the terrestrial surface sensible heat flux (e.g. Timmermans et al., 2007) or a simpler model using estimates of radiometric temperature at two times per day, thereby constraining the diurnal cycle and deriving the sensible heat flux (e.g. Norman et al., 2000); however these two approaches appear to be limited in regional scope due to the need for local input data. A comparison of a number of differing approaches with various levels of utilization of reanalysis data and off-line surface models is shown by Jimenez et al. (2011); these global means vary by 39 W $m^{-2.}$ As noted by Stephens et al. (2012), no definitive measure of uncertainty of these flux estimates exists.

For terrestrial latent heat flux, the retrievals of solar radiation, humidity, air temperature, wind speed, and soil moisture are crucial, as well as information on phenology, vegetation cover and vegetation stress. At large spatial and temporal scales, net radiation is the most important driver of *λE*, with the exception of water-limited ecosystems. Global scale *λE* models are highly reliant on accurate net radiation, with this forcing explaining up to 80% of *λE* variability (Fisher *et al.* 2008, Miralles *et al.* 2011). Second to radiation, the availability of moisture at the land surface is expected to be the most dominant driver of global *λE* (Miralles *et al.* 2014). However, the incorporation of precipitation or surface soil moisture as inputs in *λE* retrieval models requires substantial amounts of modelling and remains very uncertain, largely due to the existing large, but not understood, diversity of drought stress responses by different ecosystems (Miralles *et al.*, 2011; Martens *et al*., 2017). Some approaches use vapour pressure deficit to estimate the effect of dryness instead (Fisher *et al*., 2008; Mu *et al*., 2011), usually taken from reanalysis data. To detect the effect of different stressors on vegetation and the changes induced by vegetation phenology on the surface energy partitioning, information on the state of vegetation becomes critical, either in the visible and near infrared – where measurements can be attained at high spatial and temporal resolutions (10–100 m, daily–weekly) – or in the microwave part of the spectrum – at lower resolutions but with all-weather capability. Observable constraints such as microwave vegetation optical depth or solar induced fluorescence can be further exploited to capture the λE response and to better discriminate coupled water and carbon dynamics (Alemohammad *et al.* 2016; Pagán *et al.*, 2019).

Currently, most global λE retrieval approaches are based on different modifications of traditional local-scale physical parameterizations of λE, such as those by Monteith (1965) or Priestley and Taylor (1972), driven by satellite observations (see e.g. Fisher *et al.* 2008, Mu *et al.* 2011, Miralles *et al.* 2011, Zhang *et al.* 2016). Conversely, others have applied satellite data within statistical frameworks (Jimenez *et al.* 2009), sometimes in combination with ground meteorological measurements of evaporation (Jung *et al.* 2009), and more recently machine-learning algorithms (Tramontana et al. 2016; Jung et al. 2010, Jung et. al 2011, Jung et al. 2019, Bodesheim et al. 2018). As a response to these developments, the LandFlux initiative from the GEWEX Data and Assessments Panel (Jimenez *et al.* 2011, Muller *et al.* 2011; McCabe *et al.* 2016) emerged almost a decade ago, with the aim of organizing these efforts towards the creation of a merged datasets of continental λE, targeting the long-term goal of achieving global closure of surface water and energy budgets. More recently, the European Space Agency (ESA) WACMOS-ET project was initiated in response to the need for a thorough and consistent model inter-comparison across a range of spatial and temporal scales to understand the uncertainties in λE data sets (Michel *et al.* 2016, Miralles *et al.* 2016). A group of scientists interested in the machine learning based assessment of global land-atmosphere fluxes gathered as FLUXCOM (www.fluxcom.org) and produced a large ensemble of global energy flux products (Tramontana et al. 2016, Jung et al.2019). The ultimate goal of these initiatives is to enhance our understanding of the global energy and water cycles, and to provide high quality benchmark datasets for global climate model developers to improve predictions of future climate.

Within the framework of the WACMOS-ET project, four commonly-used satellite-based λE data sets were evaluated: the Surface Energy Balance Model, SEBS (Su 2002); the Penman–Monteith approach that forms the basis for the official Moderate Resolution Imaging Spectroradiometer (MODIS) evaporation product, PM-MOD (Mu et al. 2007); the Global Land Evaporation Amsterdam Model, GLEAM (Miralles et al. 2011); and the Priestley and Taylor model from the Jet Propulsion Laboratory, PT-JPL (Fisher et al. 2008). These data sets against 24 in situ stations from the FLUXNET archive (Baldocchi et al. 2001). The algorithms were forced using common satellite observations and in situ meteorological data from the period 2005–2007. The four models indicated robust performances in terms of changes in forcing types and temporal resolutions, with rms differences across the products ranging from 0.08 to 0.12 mm hr^{-1}. Regional differences between the various products are evident, and differences between the products are more than 400 mm yr^{-1} near the tropics but much less in the northern mid latitudes (Miralles et al. 2016). Total continental estimates of evaporation in these products range from 54.9 to 72.9 (x 10^3 km^3), with ERA-Interim values at 84.4 x 10^3 km^3 (Miralles et al. 2016). Findings from both this project and the GEWEX LandFlux effort highlighted the need to improve the parameterization of λE under water stressed conditions and the partitioning of λE into different sources for all algorithms (Michel et al. 2016, Miralles et al. 2016).

The machine learning based approaches are showing very good skill in reproducing the between site variability of mean energy fluxes, the seasonal cycles and even the diurnal cycles in dedicated cross-validation analysis but also highlighted challenges with respect to capturing drought stress (Jung et al, 2011, Tramontana et al 2016, Bodesheim et al. 2018). This methodology has been used for sensible and latent heat flux as well as carbon flux. Good consistency of corresponding global products with independent estimates (e.g. GLEAM) was shown for spatial and seasonal patterns while sizeable discrepancies also remain between different products, for example in tropical and subtropical regions (Jung et al. 2019). However, a consistent evaluation of machine learning based and remote sensing based methods at FLUXNET level is still missing. Based on the FLUXCOM ensemble of machine learning and remote sensing based estimates the relative uncertainty of global mean annual latent and sensible heat fluxes was estimated to be 12% but can be much larger regionally (Jung et al 2019). In this example uncertainty due to FLUXNET energy balance issues dominated over uncertainty associated to the choice of different machine learning algorithms. Recent analyses of climatological trends in λE have also highlighted the potential of these data sets for climate research, and at the same time the discrepancies existing in these data sets at annual to multi-decadal scales (Jung et al. 2010, Miralles et al. 2014, Mao et al. 2015, Zhang et al. 2016).

Other methods

Other indirect methods for estimating the net global or ocean surface fluxes make use of balances between the atmospheric radiation and water budgets, ocean heat content, and ocean salinity. Trenberth and Fasullo (2017) combined top-of-atmosphere radiation estimates with atmospheric reanalyses of vertically integrated atmospheric total heating to estimate annual net upward ocean surface flux. A CLIVAR Research Focus (CONCEPT-HEAT; Consistency between planetary energy balance and ocean heat storage; von Schuckmann et al. 2016) compares estimates from atmospheric radiation data, surface air-sea fluxes, the ocean heat content, global sea level to provide estimates of the planetary heat balance for comparison with climate models. Similar projects could include ocean salinity and the water cycle. These methods can provide useful constraints on the estimation of the net surface heat budget on a mean basis. What they do not provide is information about the individual components of the heat or water cycle, which is crucial for guiding improvements to satellite/gridded products, and forcing of atmosphere/ocean/land/coupled models.

NWP analyses and re-analyses

Atmospheric analysis is used by Numerical Weather Prediction (NWP) centres to create an atmospheric initial state from which a forecast can be made. It is rather obvious that a good initial condition is a prerequisite for a good forecast. This is the reason that NWP centres pay a lot of attention to the processing, quality control, bias correction and analysis of as many observations and observation types as possible. Highly sophisticated statistical techniques have been developed to make optimal use of a wide variety of observations that are irregular in time and space (e.g. Kalnay 2003).

A high quality forecast model is a key component of any atmospheric data assimilation system, because it is needed to propagate the atmospheric state in time, to support simulation of the model equivalent of satellite radiances, and to provide consistency between variables. Most analysis systems work with data slots of 6 or 12 hours and use the observations over this time interval to correct a background field (first guess), which is the forecast from the previous analysis. In fact the first guess field has more weight than the observations, i.e. the changes to background due to the observations tend to be small. The analysis combines background, and observations in an "optimal" way by giving weights according to error estimates. The analyzed state is often obtained through 3D (space) or 4D (space and time) variational techniques, in which a cost function is minimized that gauges the distance to the background and the observations (conventional observations and satellite radiances), according to their error estimates.

The result of continuous data assimilation cycling, is a consistent representation of the atmosphere in space and time. The data are available on the model grid and there is no missing data. Theoretically, the result is better than for instance retrievals from a particular satellite only, because NWP analysis uses many different data sources and reduces uncertainty be combining all pieces of information. In practice, the analysis may not be optimal for a variety of reasons, e.g. the estimation of observation and background errors is not perfect, the model has deficiencies, and observations may have biases. Another weak point of most current analysis systems is that they do not provide an estimate of the analysis error. Development of flow dependent error estimates is a very active research area of research (e.g. Houtekamer and Zhang 2016; Bonavita et al. 2016; Liu et al. 2015) and the recent ERA5 reanalysis has a measure of uncertainty that is available to the user (Hersbach et al. 2018).

Surface fluxes over the ocean

Through their short range forecasts, atmospheric data assimilation systems also provide surface fluxes over land and ocean. Over the ocean the SST is constrained through an independent SST analysis, and the turbulent fluxes are derived from standard bulk formulations relating fluxes to wind, temperature and moisture at the lowest model level of the assimilation model. There is currently a variety of bulk formulations used in NWP models, with some being more state-of-the-art than others. The main error in the fluxes is related to errors in the difference between wind, temperature and moisture at the lowest model level and the surface. Although it is impossible to draw firm conclusions about accuracy, it is possible to get a feel by considering statistics of first guess departures from observations. Such departures include first guess errors, observation errors and observation representativeness errors. For example, the standard deviation of the difference between scatterometer winds and first guess is typically 0.5 to 1 m/s dependent on the climatological regime. The first guess departures with respect to buoy and ship winds are typically 1.5 to 1.7 m/s. The higher values of the latter are due to the point character of buoys and ships. One could argue that the buoys and ships are less important, but it should be realized that they play an important role in the control of biases because they come with absolute calibration. Temperature is constrained by satellite observations within a few tenths of a degree over deep layers, but realizing such an accuracy near the ocean surface is a challenge. Total column water vapor is also very well controlled in data assimilation by microwave observations, but the vertical distribution is less certain and rather strongly affected by the assimilation model. NWP models also produce radiative fluxes and precipitation at the surface. These fluxes are the result of a radiation model, and the precipitation formulations respectively. The observational constraint on these parameters is rather indirect. Clear sky radiation is fairly accurate, but the final surface flux carries the uncertainty of its input, particularly clouds and aerosols. Precipitation can also show large errors particularly at the smallest scales and in convective regions. The accuracy of surface fluxes from NWP analyses can be evaluated through comparison with independent observations (e.g. Brunke et al. 2011; Renfrew et al. 2002). Such studies show that analysis systems benefit from the wide variety of observations that are used and from the excellent representation of synoptic variability. However, fluxes are not independent of the assimilation model and can have biases. NWP assimilation does also not apply large scale budget constraints, and the implied ocean heat and water budgets may not close (e.g. Berrisford et al., 2011).

Surface fluxes over land

The character of the surface boundary condition over land is completely different from the one over the ocean. To simulate surface fluxes, NWP centres use a comprehensive land surface model to describe soil moisture, soil temperature, snow and vegetation. Indirect data assimilation is used to control soil moisture and sometimes soil temperature (Bouttier et al. 1993; De Rosnay et al. 2014). These models close the surface energy balance and the main aim is to simulate a realistic energy partitioning between sensible and latent heat flux, which both have a pronounced diurnal cycle. Again, these systems benefit from a good atmospheric analysis and show good synoptic variability. However, errors can be large and systematic. Evaluation relies heavily on flux towers, which also have limitations e.g. lack of energy closure and representativeness (Jimenez et al. 2011; Decker et al. 2012).

Model re-analyses

NWP analysis depends on the forecast model, data assimilation method, and data handling (quality control and bias correction). All these components benefit from ongoing research and increasing computer power. It is therefore desirable to go back in time to re-analyze historic observations with a frozen system that uses the most up-to-date forecasting model and data assimilation method. The advantage of a frozen system is also that no discontinuities can occur due to system changes. The only aspect that can change the quality of the result, is the volume and accuracy of the observations.

Re-analyses are very popular in the atmospheric research community, because they produce data on a regular grid, have consistency between parameters, do not have gaps, and are easy to use (for an overview of reanalysis projects see Dee et al. 2016). The big advantage is that they incorporate millions of observations that are irregular in space and time. They include conventional data from SYNOP stations, ships and buoys, radiosondes, aircraft, satellite radiances from passive and active instruments e.g. infrared, microwave, scatterometer, radio occultation. Satellite platforms are both geostationary and polar orbiting.

As mentioned above, re-analysis has limitations, which should be considered particularly for "derived" quantities like surface fluxes as they are not directly constrained by observations. The best constrained are the turbulent fluxes over the ocean through wind speed and atmospheric temperature and moisture. These fields show very good synoptic variability, because the air carries the history of the upstream flow. However, turbulent fluxes are not bias-free and no attempt is made to satisfy global constraints. Precipitation and radiation fields also show good synoptic variability, but are generally less accurate because they rely heavily on vertical velocity and the model formulation of cloud and precipitation processes. So far reanalysis systems were basically duplicated operational Numerical Weather Prediction (NWP) systems with very few adaptations for climate applications. The consequence is that for instance biases exist in the net ocean fluxes which is less relevant for NWP (Trenberth et al. 2009; Stephens et al. 2012). More research is needed to consider conservation, ocean and top of the atmosphere constraints, and use of top of the atmosphere radiation budget observations.

Recommendations for the future

In order to make advances in estimates of the surface fluxes useful for progress in weather and climate modeling predictions and for diagnosing climate variability and its causes, uncertainties in our regional to global and short-term to long-term estimates of heat, water, momentum, and gas fluxes need to be reduced. Current constraints on the accuracies of the surface fluxes arise from errors in input parameters; sampling errors resulting from regions, regimes, or time or space scales that are not observed regularly (or at all); and a lack of physical knowledge of the importance of various phenomena contributing to the fluxes. No *in situ* or satellite system will, in the near future, provide a global observational network of the fluxes at high temporal and spatial scaling to catch all of the sub-hourly and submesoscale/sub-basin scale variability that contributes to air-surface flux variability and means. Thus, improvements in the observational systems must be commensurate with model improvements to improve realistic atmospheric and surface information (i.e., ocean boundary layer, soil moisture, vegetation, etc.). Improved scientific understanding of the role of smaller-scale variability (such as submesoscale eddies) towards local and global balances must also be a priority.

In order to bring the various communities together, there needs to be a high-level group set up to interact with CLIVAR, GEWEX, GSOP, and all of the other communities who work on atmosphere-surface fluxes. This group could be the current WCRP WDAC Surface Flux Task Team, or a new configuration placed within the WCRP framework. Members of the modeling, observational, and remote sensing communities should be involved.

Below are some recommendations that our task team has determined to be of high value in achieving significant gains in understanding, observing, and modeling the surface flux system. Some of these recommendations would require significant resources; others are suggestions that could be implemented simply through workshops or other community-gathering activities.

Major Recommendations

- Consistent collaboration between *in situ* and satellite communities that focus on aspects of surface fluxes, to improve understanding on what *in situ* data is useful for satellite product improvements, and to produce error statistics from both types of datasets that are of more value to the other community.

- Land, ice, and ocean observationalists should be encouraged to collaborate more consistently. For instance, BSRN has established a best-practices document for land-based radiation measurements, while similar guidance for over ocean sites is not available. Both communities work with sparse and irregularly-spaced data and have developed methods for upscaling/gridding that could inform the other.

- There should be an Increased focus on co-location of various *in situ* flux networks and observations. In corollary, we need to expand the number of sites that are measuring not only the surface fluxes, or parameters for calculating them, but also observations that will improve our understanding of the physics for modeling and satellite algorithm development. For example, improved models and satellite retrievals of radiative and turbulent fluxes require improved atmospheric boundary layer profiles of water vapor and temperature. Over the ocean, enhanced wave measurements and ocean mixing processes are also needed. For all locations, data throughout the coupled boundary layers at semi-permanent sites need to be established in key locations (continental tropics, semi-arid regions, arctic, coastal, mid- and high-latitude oceans) to provide the required parameters for model and process improvements.

- Enhancement of current and development of new *in situ* platforms and sensors needs to continue. As more groups make measurements of fluxes and flux parameters, with current or new technology, clear guidelines need to be established for assessment and quality-control of that data. Flux measurements are particularly difficult to make and interpret accurately and, especially with the number of autonomous vehicles that are making measurements of the boundary layers, some community guidance in assessing and evaluating these new methodologies is required.

- Closer integration of the observational and modeling communities is required to convert our *in situ* measurements into more useful and more easily accessible products for numerical modelers. The use of measurements will provide insight into what observations would be most useful for modelers to improve boundary layer/surface models. This will enhance our understanding of the needs for resolution and accuracy from the models for producing flux products, which will improve the accuracy of latent heat flux over land surfaces through improved land models.

- Uncertainty estimates for data, either *in situ,* gridded, or satellite, remain highly variable depending on the temporal and spatial scale in question and the method used to quantify the uncertainty. Clear guidelines need to be established to estimate the uncertainty and to quality control in situ, gridded products and satellite data. Assessments of products with realistic uncertainties, including temporal stability and reliability particularly in the case of changing satellite data streams, is necessary to help both drive continued progress and to inform users about which products are best for which types of research.

- Improvements in the following aspects of re-analysis systems are of high priority: (i) constrain top of atmosphere radiation by satellite observations, (ii) develop coupled land/atmosphere and coupled sea-ice/atmosphere data assimilation systems to obtain more satellite constraint on surface fluxes, and (iii) to include global constraints on energy fluxes in the variational algorithms.

- Foster the use of high resolution models to evaluate the importance of temporal and spatial variability of turbulence and radiative fluxes, and to identify key locations and parameters to focus on when suggesting additional measurements for improved climate simulations.

- Establish a forum to report results of comparisons, possible model errors and improvements for communities who use models for both predictive purposes (such as the NWP community) and for producing global gridded products (such as the continental satellite flux community).

Other Recommendations

- Observations need to be freely available and easy to find for researchers not directly associated with the gathering of this data.

- An inventory of additional complementary data necessary to understand the processes controlling the fluxes is needed.

- The community needs a clearer understanding of the spatial resolution needed in fluxes and their associated bulk parameters to accurately both capture their variability in mean analyses but also elucidate the processes required for model resolution.

- Innovative means to investigate coupled wind-waves processes that are thought to be important on measurements and estimations of fluxes under extreme (greater than 25 m s^{-1}) winds must be encouraged and funded. These processes include the role of wave breaking, flow separation, bubble production and the generation and transport of evaporating sea-spray on momentum, heat and energy exchange under high winds.

- The energy balance closure gap that occurs from land-based EC measurements of latent and sensible heat fluxes needs to be understood and overcome, whether by improved instrumentation in the future or by developing appropriate correction methods.

- The modeling and satellite communities should be encouraged to work together to make more use of techniques such as data assimilation and machine learning. From NWP results, it is clear that modern data assimilation techniques are very powerful to define the state of the atmosphere. Variational techniques for optimization of, e.g., land surface state variables can even be extended to poorly known land parameters (augmented control variable approach). Research in this area should be encouraged, to ensure that optimal use is made of observations.

Task Team Structure and Mission

In 2015 the WCRP Data Advisory Council (WDAC) determined that a Task Team with a specific focus on surface fluxes was needed to help provide coordination for addressing the many remaining uncertainties in observing, understanding, and modeling the surface fluxes across the weather and climate continuum. To that end, a Task Team was set up and the following Terms of Reference were produced:

1. Provide a single point-of-contact for surface flux observations and analysis in the WCRP. Communicate with other relevant entities regarding WCRP surface flux activities through work on committees, a website, and other published articles and information.

2. Establish and encourage the publication and use of data, metadata, and documentation standards for global surface flux (ocean, land, or ice and atmosphere) data sets that are consistent with standards and infrastructure used in major climate model intercomparison efforts (e.g., CMIP, ESGF, and Obs4MIPs), thereby facilitating intercomparison of the data sets and their use in evaluation of Earth System models and their components.

3. Establish conventions for intercomparisons of global datasets, and for assessment of the global datasets with available in situ data, making use of established assessments for other components of the Earth system from GEWEX and other WCRP entities.

4. Report to the WDAC and WCRP Core Projects (*e.g.,* GEWEX/GDAP and CLIVAR) on progress, status, and plans for activities overseen by the Task Team.

Task Team Activities

In order to achieve the goals of the WDAC, the Surface Flux Task Team has identified several activities to focus on in the near-term:

- Encourage continued acquisition of eddy-covariance flux measurements, particularly in extreme locations and conditions that are currently not well represented. In addition, campaigns with mobile towers to help in better capturing the large diversity of ecosystems on land. All efforts with respect to improved processing, standardization, and quality control of EC data need to be fostered. The expansion of flux observations to include all of the heat (turbulent and radiative), momentum, gas, aerosols, and freshwater fluxes is also encouraged. Support continued acquisition of key input parameters for bulk flux estimation and radiative fluxes at existing and planned long-term sites.

- Support increased collaborations between *in situ* measurement community, satellite users, and modeling community. Hosting of workshops fostering collaboration between the communities. We have also assisted with multiple recent recommendation reports and urged the inclusion of all three communities.

- Initiate a working group consisting of *in situ* and satellite experts in radiation from both the land and ocean communities in order to homogenize information about best practices and uncertainty analysis.

- Oversee a website containing available direct and indirect flux estimates in a standardized format. In addition, provide hosting for *in situ*, satellite, and reanalysis flux data sets in agreed standards, a repository of results of comparisons, and an inventory of the literature published using these data sets for the benefit of the larger community.

- Collaborate with other data science communities to develop a benchmark for uncertainty analysis that can be used by the flux community.

- Support activities from non-surface flux communities, for example the ocean heat content and the atmospheric radiation communities, to work with constraining and understanding uncertainties in the air—surface fluxes.

- Encourage research into and support for satellite missions that provide enhanced air-surface flux capabilities.

Links to the larger community

There are a number of international projects that relate to the goals of this Task Team. Over the ocean, the International SOLAS (Surface Ocean - Lower Atmosphere Study) project has a primary objective "to achieve quantitative understanding of the key biogeochemical-physical interactions and feedbacks between the ocean and atmosphere, and of how this coupled system affects and is affected by climate and environmental change" (http://www.solas-int.org/about/solas.html). The Global Climate Observing System (GCOS) maintains the definitions of the Essential Climate Variables (ECVs) and work towards sustaining, coordinating and improving physical, chemical, and biological observations. Observing networks include the Global Ocean Observing System (GOOS, http://www.goosocean.org), which comprises the oceanographic component of the Global Earth Observing System of Systems, and overseas such assets as the Argo program, drifting and moored buoys (including the OceanSITES program), the GO-SHIP program, and other observational networks. There are regional flux networks including ICOS and Ameriflux. The Integrated Carbon Observation System (ICOS, https://www.icos-ri.eu/home) is a pan-European research infrastructure which provides harmonised and high-precision scientific data on carbon cycle and greenhouse gas budget and perturbations, with stations located across Europe and covering the North Atlantic and European marginal seas including Voluntary Observatory Ships, fixed stations and research vessels. AmeriFlux is a network of PI-managed sites measuring ecosystem CO_2, water, and energy fluxes in North, Central and South America. It was established to connect research on field sites representing major climate and ecological biomes, including tundra, grasslands, savanna, crops, and conifer, deciduous, and tropical forests. FLUXNET is a global network of micrometeorological tower sites that use eddy covariance methods to measure the exchanges of carbon dioxide, water vapor, and energy between terrestrial ecosystems and the atmosphere. The Baseline Surface Radiation Network (BSRN; https://bsrn.awi.de) maintains a network of high quality radiation land-based sites for collecting surface radiation data.

References

Alemohammad, S. H., Fang, B., Konings, A. G., Green, J. K., Kolassa, J., Prigent, C., Aires, F., Miralles, D. G., and Gentine, P. (2016). Water, Energy, and Carbon with Artificial Neural Networks (WECANN): A statistically-based estimate of global surface turbulent fluxes using solar-induced fluorescence. *Biogeosciences Discussion*, doi: 10.5194/bg-2016-495.

Anctil, F., M.A. Donelan, W.M. Drennan, and H.C. Graber, 1994: Eddy-Correlation Measurements of Air-Sea Fluxes from a Discus Buoy. *J. Atmos. Oceanic Technol.*, **11**,1144–1150.

Anderson, M., Norman, J., Diak, G., Kustas, W., and Mecikalski, J. (1997). A two-source time-integrated model for estimating surface fluxes using thermal infrared remote sensing. *Remote Sensing of Environment*, 60(2), 195–216.

Anderson, D. E., and Cahalan, R. F. (2005). The solar radiation and climate experiment (SORCE) mission for the NASA earth observing system (eos). *Solar Physics*, 230(1–2), 3–6, doi:10.1007/S11207- 005-1592-6.

Anderson, M. C., Kustas, W. P., Norman, J. M., Hain, C. R., Mecikalski, J. R., Schultz, L., Gonzalez-Dugo, M. P., Cammalleri, C., Urso, G., Pimstein, A., and Gao, F. (2011). Mapping daily evapotranspiration at field to continental scales using geostationary and polar orbiting satellite imagery. *Hydrology and Earth System Sciences*, 15(1), 223–239. http://doi.org/10.5194/hess- 15-223-2011.

Andersson, A., Klepp, C., Fennig, K., Bakan, S., Grassl, H., and Schulz, J. (2011). Evaluation of HOAPS-3 ocean surface freshwater flux components. *Journal of Applied Meteorology and Climatology*, 50, 379–398, doi:10.1175/2010JAMC2341.1.

Anderson, M. C., Allen, R. G., Morse, A., and Kustas, W. P. (2012). Use of Landsat thermal imagery in monitoring evapotranspiration and managing water resources. *Remote Sensing of Environment*, 122(0), 50-65.

Andersson, A., Klepp, C., Fennig, K., Bakan, S., Grassl, H., and Schulz, J. (2011). Evaluation of HOAPS-3 ocean surface freshwater flux components. *Journal of Applied Meteorology and Climatology*, 50, 379–398, doi:10.1175/2010JAMC2341.1.

Baldocchi, D., Falge, E., Gu, L., Olson, R., Hollinger, D., Running, S., Anthoni, P., Bernhofer, C., Davis, K., Evans, R., Fuentes, J., Goldstein, A., Katul, G., Law, B., Lee, X., Malhi, Y., Meyers, T., Munger, W., Oechel, W., Paw, K. T., Pilegaard, K., Schmid, H. P., Valentini, R., Verma, S., Vesala, T., Wilson, K., and Wofsy, S. (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 American Meteorological Society*, 82, 2415– 2434.

Berry, D. I., and E. C. Kent (2016), Assessing the health of the in situ global surface marine climate observing system, *Int. J. Climatol.*, *2259*, 2248–2259, doi:10.1002/joc.4914.

Bentamy, A., Piolle, J. F., Grouazel, A., Danielson, R., Gulev, S., Paul F., Azelmat, H., Mathieu, P. P., Von Schuckmann, K., Sathyendranath, S., Evers-King, H., Esau, I., Johannessen, J. A., Clayson, C.A.; Pinker, R.T.; Grodsky, S.A.; Bourassa, M.; Smith, S.R.; Haines, K.; Valdivieso, M.; Merchant, C.J.; Chapron, B.; Anderson, A.; Hollmann, R.; Simon, J. (2017). Review and assessment of latent and sensible heat flux accuracy over the global oceans. *Remote Sensing and Environment*, 201, 196-218.

Bentamy, A., Grodsky, S. A., Katsaros, K., Mestas-Nunez, A. M., Blanke, B., and Desbiolles, F. (2013). Improvement in air-sea flux estimates derived from satellite observations. *International Journal of Remote Sensing*, 34(14), 5243–5261, doi:10.1613/jair.301.

Berrisford, P., Kållberg, P., Kobayashi, S., Dee, D., Uppala, S., Simmons, A.J., Poli, P. and Sato, H. (2011). Atmospheric conservation properties in ERA-Interim. *Quarterly Journal of the Royal Meteorological Society*, 137(659), 1381-1399.

Bigorre, S, Weller, R. A., Lord, J., Edson, J. B., and Ware, J. D. (2013). A surface mooring for air-sea interaction research in the Gulf Stream. Part 2: Analysis of the observations and their accuracies. *Journal of Atmospheric and Oceanic Technology*, 30, 450–469.

Bishop, S. P., Small, R. J., Bryan, F. O., and Tomas, R. A. (2017). Scale dependence of midlatitude air-sea interaction. *Journal of Climate*, 30(20), 8207– 8221.

Blomquist, B. W., Huebert, B. J., Fairall, C. W., Ludovic, B., Edson, J. B., Hare, J. E., and McGillis, W. R. (2014). Advances in air-Sea CO2 flux measurement by eddy correlation. *Boundary-Layer Meteorology*, 152, 245–276. doi: 10.1007/s10546-014-9926-2.

Bodesheim, P., Jung, M., Gans, F., Mahecha, M. D., and Reichstein, M. (2018). Upscaled diurnal cycles of land–atmosphere fluxes: a new global half-hourly data product. *Earth Systems Science Data*, 10(3), 1327-1365.

Bonavita, M., Hólm, E., Isaksen, L. and Fisher, M. (2016). The evolution of the ECMWF hybrid data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 142(694), 287-303.

Bourassa, M. A., Vincent, D. G., and Wood, W. L. (1999) A flux parameterization including the effects of capillary waves and sea state. *Journal of Atmospheric Sciences*, 56, 1123–1139.

Bourassa, M. A., S. T. Gille, D. L. Jackson, J. B. Roberts, and G. A. Wick (2010). Ocean Winds and Turbulent Air-Sea Fluxes Inferred From Remote Sensing, *Oceanography* 23(4):36–51, https://doi.org/10.5670/oceanog.2010.04.

Bouttier, F., Mahfouf, J. F. and Noilhan, J. (1993). Sequential assimilation of soil moisture from atmospheric low-level parameters. Part I: Sensitivity and calibration studies. *Journal of Applied Meteorology*, 32(8), 1335-1351.

Bradley, F. and C. Fairall, A Guide To Making Climate Quality Meteorological And Flux Measurements At Sea. NOAA Technical Memorandum OAR PSD-311. 2006.

Brunke, M. A., Wang, Z., Zeng, X., Bosilovich, M., and Shie, C. L. (2011). An Assessment of the Uncertainties in Ocean Surface Turbulent Fluxes in 11 Reanalysis, Satellite-Derived, and Combined Global Datasets, *Journal of Climate*, 24(21), 5469–5493, doi:10.1175/2011JCLI4223.1.

Charuchittipan, D., Babel, W., Mauder, M., Leps, J. P. and Foken, T. (2014). Extension of the averaging time in eddy-covariance measurements and its effect on the energy balance closure. *Boundary-Layer Meteorology*, 152(3): 303-327.

Chelton, D. B., Schlax, M. G., Freilich, M. H. and Milliff, R. F. (2004). Satellite measurements reveal persistent small-scale features in ocean winds. *Science*, 303, 978–983.

Cheng, L., Trenberth, K. E., Fasullo, J., Boyert, T., Abraham, J. and J. Zhu, 2017: Improved estimates of ocean heat content from 1960 to 2015. *Science Advances*, 3 (3), doi:10.1126/sciadv.1601545.

Ciais, P., C. Sabine, G. Bala, L. Bopp, V. Brovkin, J. Canadell, A. Chhabra, et al. 2013. Carbon and other biogeochemical cycles. Pp. 465-570 in *Climate Change 2013: The Physical Science Basis* (eds. T.F. Stocker, D. Qin, G.K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P.M. Midgley). Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge: Cambridge University Press.

Cifuentes-Lorenzen, A., J. B. Edson, and C. J. Zappa (2018). Air–sea interaction in the Southern Ocean: Exploring the height of the wave boundary layer at the air–sea interface. *Bound.-Layer Meteor.,* doi.org/10.1007/s10546-018-0376-0.

Clayson, C. A., Fairall, C. W., and Curry, J. A. (1996). Evaluation of turbulent fluxes at the ocean surface using surface renewal theory. *Journal of Geophysical Research*, 101(C12), 28503–28513.

Clayson, C. A., Brown, J., and NOAA CDR Program (2016): NOAA Climate Data Record (CDR) of Ocean Heat Fluxes, Version 2. NOAA National Centers for Environmental Information. doi: http://doi.org/10.7289/V59K4885.

Clayson, C.A., and A.S. Bogdanoff (2013): The effect of diurnal sea surface temperature warming on climatological air-sea fluxes. *Journal of Climate* 26(8):2546-2556.

Clayson, C. A., Pinker, R. T., Grodsky, S. A., Bourassa, M., Smith, S. R., Haines, K., Valdivieso, M., Merchant, C. J., Chapron, B., Anderson, A., Hollmann, R., and Josey, S. A. (2017). Review and assessment of latent and sensible heat flux accuracy over the global oceans. *Remote Sensing of Environment*, 201, 196-218.

Clayson, C. A., J. B. Edson, A. Paget, R. Graham, and B. Greenwood, 2019: The effects of rainfall on the atmosphere and ocean during SPURS-2. *Oceanography*, accepted.

Conway, J. and Cullen, N. (2013). Constraining turbulent heat flux parameterization over a temperate maritime glacier in New Zealand. *Annals of Glaciology*, 54, 41–51, doi:10.3189/2013AoG63A604, 2013.

Cronin, M. F., C. L. Gentemann, J. B. Edson, I. Ueki, M. Bourassa, S. Brown, C. A. Clayson, C. Fairall, T. Farrar, S. T. Gille, S. Gulev, S. Josey, S. Kato, M. Katsumata, E. Kent, M. Krug, P. Minnett, R. Parfitt, R. T. Pinker, P. W. Stackhouse, Jr., S. Swart, H. Tomita, D. Vandemark, R. A. Weller, K. Yoneyama, L. Yu, D. Zhang, 2019: Air-sea fluxes with a focus on heat and moment. *Frontiers in Marine Science*(for OceanObs '19), revised.

De Rosnay, P., Balsamo, G., Albergel, C., Muñoz-Sabater, J. and Isaksen, L., (2014). Initialisation of land surface variables for numerical weather prediction. *Surveys in Geophysics*, 35(3), 607-621.

Decker, M., Brunke, M. A., Wang, Z., Sakaguchi, K., Zeng, X. and Bosilovich, M.G. (2012). Evaluation of the reanalysis products from GSFC, NCEP, and ECMWF using flux tower observations. *Journal of Climate*, 25(6), 1916-1944.

Dee, D., Fasullo, J., Shea, D., Walsh, J., and NCAR, S. (2016). The climate data guide: atmospheric reanalysis: overview & comparison tables. National Center for Atmospheric Research, Boulder, CO. Available at https://climatedataguide.ucar.edu/climate-data/atmospheric-reanalysis-overview-comparison-tables. Accessed June, 1, 2017.

Ditri, A. L., Minnett, P. J., Liu, Y., Kilpatrick, K., and Kumar, A. (2018). The accuracies of Himawari-8 and MTSAT-2 sea-surface temperatures in the tropical western Pacific Ocean. *Remote Sensing*, 10, doi:10.3390/rs10020212.

Dolman, A. J., Miralles, D. G., and De Jeu, R. A. M. (2014). Fifty years since Monteith's 1965 seminal paper: The emergence of global ecohydrology. *Ecohydrology and Hydrobiology*, 7, 897–902.

Doney, S. C., V. J. Fabry, R. A. Feely, and J. A. Kleypas, 2009: Ocean acidification: The other CO2 problem. *Ann. Rev. Mar. Sci.*, 1:1, 169-192.

Doney, S. C., 2010: The growing human footprint on coastal and open-ocean biogeochemistry. *Science*, 328, 1512-1516.

Douville, H., Ribes, A., Decharme, B., Alkama, R., and Sheffield, J. (2013), Anthropogenic influence on multidecadal changes in reconstructed global evapotranspiration. *Nature Climate Change*, 2, 1–4.

Drennan, W. M., Graber, H. C., Hauser, D., and Quentin C. (2003). On the wave age dependence of wind stress over pure wind seas. *Journal of Geophysical Research*, 108(C3), 8062, doi:10.1029/2000JC000715.

Driemel, A., Augustine, J., Behrens, K., Colle, S., Cox, C., Cuevas-Agulló, E., Denn, F. M., Duprat, T., Fukuda, M., Grobe, H., Haeffelin, M., Hodges, G., Hyett, N., Ijima, O., Kallis, A., Knap, W., Kustov, V., Long, C. N., Longenecker, D., Lupi, A., Maturilli, M., Mimouni, M., Ntsangwane, L., Ogihara, H., Olano, X., Olefs, M., Omori, M., Passamani, L., Pereira, E. B., Schmithüsen, H., Schumacher, S., Sieger, R., Tamlyn, J., Vogt, R., Vuilleumier, L., Xia, X., Ohmura, A., and König-Langlo, G. (2018). Baseline Surface Radiation Network (BSRN): structure and data description

(1992–2017). *Earth System Science Data*, 10, 1491-1501, https://doi.org/10.5194/essd-10-1491-2018, 2018.

Durack, P. J.and S. E. Wijffels, 2010: Fifty-Year Trends in Global Ocean Salinities and Their Relationship to Broad-Scale Warming. *J. Climate*, **23**, 4342–4362.

Edson, J. B., Fairall, C. W., Larsen, S. E., and Mestayer, P. G. (1991). A study of the inertial-dissipation technique for computing air–sea fluxes. *Journal of Geophysical Res*earch, 96, 10689–10711.

Edson, J. B., Hinton, A. A., Prada, K. E., Hare, J. E., and Fairall, C. W. (1998). Direct covariance flux estimates from mobile platforms at sea. *Journal of Atmospheric and Oceanic Technology,* 15, 547-562.

Edson, J., Crawford, T., Crescenti, J., Farrar, T., Frew, N., Gerbi, G., Helmis, C., Hristov, T., Khelif, D., Jessup, A., Jonsson, H., Li, M., Mahrt, L., McGillis, W., Plueddemann, A., Shen, L., Skyllingstad, E., Stanton, T., Sullivan, P., Sun, J., Trowbridge, J., Vickers, D., Wang, S., Wang, Q., Weller, R., Wilkin, J., William III, A. J., Yue, D. K. P., Zappa, C. (2007). The coupled boundary layers and air-sea transfer experiment in low winds. *Bulletin of the American Meteorology Society*, 88, 341-356.

Edson, J. B., Jampana, V., Weller, R. A., Bigorre, S., Plueddemann, A. J., Fairall, C. W., Miller, S. D., Mahrt, L., Vickers, D., and Hersbach, H. (2013). On the exchange of momentum over the open ocean. *Journal of Physical Oceanography*, 43, 1589–1610.

Fairall, C. W., and Larsen, S. E. (1986). Inertial dissipation methods and turbulent fluxes at the air ocean interface. *Boundary-Layer Meteorology*, 34, 287–301.

Fairall, C. W., Bradley, E. F., Rogers, D.P., Edson, J. B., and Young, G. S. (1996). Bulk parameterization of air-sea fluxes for TOGA COARE. *Journal of Geophysical Research*, 101, 3747-3764.

Fairall, C. W., White, A. B., Edson, J. B., and Hare, J. E. (1997). Integrated shipboard measurements of the marine boundary layer. *Journal of Atmospheric Oceanic Technology*, 14, 338-359.

Fairall, C. W., Hare, J. E., Edson, J. B., and McGillis, W. R. (2000). Parameterization and micrometeorological measurement of air-sea gas transfer. *Boundary-Layer Meteorology*, 96, 63-105.

Fairall, C. W., Bradley, E. F., Hare, J. E., Grachev, A. A., and Edson, J. B. (2003). Bulk parameterization of air–sea fluxes: Updates and verification for the COARE algorithm. *Journal of Climate*, 16, 571–591.

Fairall, C. W., Taneil Uttal, Duane Hazen, Jeffrey Hare, Meghan F. Cronin, Nicholas Bond, and Dana Veron. Observations of cloud, radiation, and surface forcing in the equatorial eastern pacific. J. of Climate, 21, 655-673, 2008.

Fairall, C. W., Yang, M., Bariteau, L., Edson, J. B., Helmig, D., McGillis, W, Pezoa, S, Hare, J. E., Huebert, B., and Blomquist, B. (2011). Implementation of the COARE algorithm with O3, CO2 and DMS. *Journal of Geophysical Research*, 116, C00F09, doi:10.1029/2010JC006884.

Farrar, J. T., Rainville, L., Plueddemann, A. J., Kessler, W. S., Lee, C., Hodges, B. A., Schmitt, R.W., Edson, J. B., Riser, S. C., Eriksen, C. C., and Fratantoni, D. M. (2015). Salinity and temperature balances at the SPURS central mooring during fall and winter. *Oceanography*, 28, 56-65.

Finnigan, J. J., Clement, R., Malhi, Y., Leuning, R. and Cleugh, H.A. (2003). Re-evaluation of long-term flux measurement techniques. Part I: Averaging and coordinate rotation. *Boundary-Layer Meteorology*, 107(1), 1-48.

Fisher, J. B., Tu, K. P., and Baldocchi, D. D. (2008). Global estimates of the land–atmosphere water flux based on monthly AVHRR and ISLSCP-II data, validated at 16 FLUXNET sites. *Remote Sensing of Environment*, 112, 901–919.

Fisher, J. B., Middleton, E., Melton, F., Anderson, M., Hook, S., Hain, C., Allen, R., McCabe, M., Lagouarde, J. P., Tu, K., Baldocchi, D., Townsend, P., Kilic, A., Perret, J., Miralles, D, G., Waliser, D., French, D., Famiglietti, J., Schimel, D., and Wood, E. F. (2017). The future of evapotranspiration: Global requirements for ecosystem functioning, carbon and climate feedbacks, agricultural management, and water resources, *Water Resources Research*, doi: 10.1002/2016WR020175.

Flügge, M., Paskyabi, M. B., Reuder, J., Edson, J. B., and Plueddemann, A. J. (2016). Comparison of direct covariance flux measurements from an offshore tower and a buoy. *Journal of Atmospheric and Oceanic Technology,* 33, 873-890.

Foltz, G.R., Evan, A.T., Freitag, H.P., Brown, S., & McPhaden, M.J. (2013). Dust Accumulation Biases in PIRATA Shortwave Radiation Records. Journal of Atmospheric and Oceanic Technology, 30, 1414-1432.

Ghilain, N., Arboleda, A., and Gellens-Meulenberghs, F. (2011). Evapotranspiration modelling at large scale using near-real time MSG SEVIRI derived data. *Hydrology and Earth System Sciences*, 15(3), 771–786.

Gimeno, L., Drumond, A., Nieto, R., Trigo, R.M., and Stohl, A. (2010). On the origin of continental precipitation. *Geophysical Research Letters*, 33, L13804.

Gloor, M., Sarmiento, J. L., and N. Gruber (2010). What can be learned about carbon cycle climate feedbacks from the CO2 airborne fraction? *Atmos. Chem. Phys.,* 10, 7739 - 7751.

Grare, L., Lenain, L., and Melville, W. K. (2013a). Wave-coherent airflow and critical layers over ocean waves. *Journal of Physical Oceanography*, 43, 2156–2172.

Grare, L., Lenain, L., and Melville, W. K. (2018). Vertical Profiles of the Wave-Induced Airflow above Ocean Surface Waves. *Journal of Physical Oceanography*, 48, 2901–2922.

Gulev, S. K., and Coauthors, 2010: Surface energy, CO_2 fluxes and sea ice. Proceedings of OceanObs'09: Sustained Ocean Observations and Information for Society, Vol. 1, J. Hall, D. E. Harrison, and D. Stammer, Eds., ESA Publ. WPP-306, 19 pp.

Hara, T. and Sullivan, P.P. (2015) Wave boundary layer Turbulence over surface waves in a strongly forced condition. *Journal of Physical Oceanography*, 45, 868–883.

Heil, P., Hutchings, J. K., Worby, A. P., Johansson, M., Launiainen, J., Haas, C., and Hibler III, W.D. (2008). Tidal forcing on sea-ice drift and deformation in the western Weddell Sea in early austral summer, 2004, *Deep-Sea Research II*, 55, 943-962, doi:10.1016/j.dsr2.2007.12.026.

Henjes, K., Taylor, P. K., and Yelland, M. J. (1999). Effect of pulse averaging on sonic anemometer spectra. *Journal of Atmospheric Oceanic Technology*, 16, 181-184.

Hersbach, H., de Rosnay, P., Bell, B., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Alonso-Balmaseda, M., Balsamo, G., Bechtold, P., Berrisford, P., Bidlot, J. R., de Boisséson, E., Bonavita, M., Browne, P., Buizza, R., Dahlgren, P., Dee, D., Dragani, R., Diamantakis, M., Flemming, J., Forbes, R., Geer, A. J., Haiden, T., Hólm, E., Haimberger, L., Hogan, R., Horányi, A., Janiskova, M., Laloyaux, P., Lopez, P., Munoz-Sabater, J., Peubey, C., Radu, R., Richardson, D., Thépaut, J. N., Vitart, F., Yang, X., Zsótér, E., Zuo, H. (2018). Operational global reanalysis: progress, future directions and synergies with NWP, ECMWF report, ERA report series, https://www.ecmwf.int/en/elibrary/18765-operational-global-reanalysis-progress-future-directions-and-synergies-nwp.

Houtekamer, P. L. and Zhang, F., (2016). Review of the ensemble Kalman filter for atmospheric data assimilation. *Monthly Weather Review*, 144(12), 4489-4532.

Jackson, D. L., Wick, G. A., and Robinson, F. R., (2009). Improved multisensor approach to satellite-retrieved near-surface specific humidity observations. *Journal of Geophysical Research*, 114, D16303. doi:10.1029/2008JD011341.

Jackson, D. L. and Wick, G. A. (2010). Near surface air temperature retrieval derived from AMSU-A and sea surface temperature observations. *Journal Technologies*, doi:10.1175/2010JTECHA1414.1.

Janssen, P. A. E. M. (1999). On the Effect of Ocean Waves on the Kinetic Energy Balance and Consequences for the Inertial Dissipation Technique. *Journal of Physical Oceanography*, 29, 530-534.

Jiménez, C., Prigent, C., and Aires, F. (2009). Toward an estimation of global land surface heat fluxes from multisatellite observations. *Journal of Geophysical Research*, 114(D6), D06305.

Jimenez, C., Prigent, C., Mueller, B., Seneviratne, S. I., McCabe, M. F., Wood, E. F., Rossow, W. B., Balsamo, G., Betts, A. K., Dirmeyer, P. A. and Fisher, J. B. (2011). Global intercomparison of 12 land surface heat flux estimates. *Journal of Geophysical Research*: *Atmospheres*, 116(D2).

Jin, X., Yu, L., Jackson, D., and Wick, G. A. (2015) An improved near-surface specific humidity and air temperature climatology for the SSM/I satellite period. *Journal of Atmospheric and Oceanic Technology,* 32(3), doi: 10.1175/JTEC-D-14-00080.1.

Jung, M., Reichstein, M., and Bondeau., A. (2009). Towards global empirical upscaling of FLUXNET eddy covariance observations: validation of a model tree ensemble approach using a biosphere model. *Biogeosciences Discussion*, 6, 5271–5304.

Jung, M., Reichstein, M., Ciais, P., Seneviratne, S. I., Sheffield, J., Goulden, M. L., Bonan, G., Cescatti, A., Chen, J., de Jeu, R., Dolman, J., Eigster, W., Gerten, D., Gianelle, D., Gobron, N., Heinke, J., Kimball, J., Law, B. E., Montagnani, L., Mu, Q., Mueller, B., Oleson, K., Papalle, D., Richardson, A. D., Roupsard, O., Running, S., Tomelleri, E., Viovy, N., Weber, U., Williams, C., Wood, E., Zaehle, S., and Zhang, K. (2010). Recent decline in the global land evapotranspiration trend due to limited moisture supply. *Nature*, 467, 951-954.

Jung, M., Reichstein, M., Margolis, H. A., Cescatti A., Richardson, A. D., Arain, M. A., Arneth, A., Bernhofer, C., Bonal, D., Chen, J.,Gianelle, D., Gobron, N., Kiely, G., Kutsch, W., Lasslop, Law, B. E., Lindroth, A., Merbold, L., Montagni, L., Moors, E. J., Papale, D., Sottocornola, M., Vaccari, F., and Williams, C. (2011). Global patterns of land-atmosphere fluxes of carbon dioxide, latent heat, and sensible heat derived from eddy covariance, satellite, and meteorological observations. *Journal of Geophysical Research - Biogeosciences*, 116, G00J07, doi:10.1029/2010JG001566.

Jung, M., Koirala, S., Weber, U., Ichii, K., Gans, F., Camps-Valls, G., Papale, D., Schwalm, C., Tramontana, G., and Reichstein, M. (2019). The FLUXCOM ensemble of global land-atmosphere energy fluxes, *Scientific Data,* in press.

Kalnay, E. (2003). Atmospheric modeling, data assimilation and predictability. Cambridge University Press.

Keenan, T.F., I.C. Prentice, J.G. Canadell, C.A. Williams, H. Wang, M. Raupach, and G.J. Collatz. 2016. Recent pause in the growth rate of atmospheric CO_2 due to enhanced terrestrial carbon uptake. *Nature Communications* 7.

L'Ecuyer, T. S. et al. (2015), The observed state of the energy budget in the early twenty-first century, *J. Clim.*, *28*(21), 8319–8346, doi:10.1175/JCLI-D-14-00556.1.

Lee, C. M., Thomson, J., and the Marginal Ice Zone and Arctic Sea State Teams, (2017). An autonomous approach to observing the seasonal ice zone in the Western Arctic. *Polar Oceanography*, 30 (2), 56-68.

Lees, K. J., Quaife, T., Artz, R. R. E., Khomik, M., and J. M. Clark, 2018: Potential for using remote sensing to estimate carbon fluxes across northern peatlands - A review. *Science of the Total Environment,* 615, 857-874.

Le Quéré, C., Rodenbeck, C., Buitenhuis, E. T., Conway, T. J., Langenfelds, R., Gomez, A., Labuschagne, C., Ramonet, M. Nakazawa, T., Metzl, N., Gillett, N., and M. Heimann (2007). Saturation of the Southern Ocean CO2 sink due to recent climate change. *Science*, 316(5832), 1735-1738.

Le Quéré, C., Takahashi, T., Buitenhuis, E. T., Rodenbeck, C., and S. C. Sutherland (2010). Impact of climate change and variability on the global oceanic sink of CO2. *Global Biogeochemical Cycles*, 24, doi:10.1029/2009GB003599.

Le Quéré, C., et al. (2018). Global carbon budget 2018. *Earth Sys. Sci. Data*, 10, 2141-2194.

Leuning, R., van Gorsel, E., Massman, W.J. and Isaac, P.R. (2012a). Reflections on the surface energy imbalance problem. *Agricultural and Forest Meteorology*, 156: 65-74.

Liman, J., Schroder, M., Fennig, K., Andersson, A., and R. Hollmann, 2018: Uncertainty characterization of HOAPS 3.3 latent heat-flux-related parameters. Atmos. Meas. Tech., 11, 1793-1815.

Liu, W. T., Katsaros, K. B., and Businger, J. A. (1979). Bulk parameterization of air-sea exchanges of heat and water vapor including the molecular constraints at the interface. *Journal of Atmospheric Sciences*, 36, 1722-1735.

Liu, W. T. (1986). Statistical relation between monthly mean precipitable water and surface-level humidity over global oceans. *Monthly Weather Review*, 114, 1591–1602.

Liu, C., Allan, R.P., Berrisford, P., Mayer, M., Hyder, P., Loeb, N., Smith, D., Vidale, P.L. and Edwards, J.M. (2015). Combining satellite observations and reanalysis energy transports to estimate global net surface energy fluxes 1985–2012. *Journal of Geophysical Research: Atmospheres,* 120(18), 9374-9389.

McArthur, L. J. B.: Baseline Surface Radiation Network (BSRN), WMO/TD-No. 879, WCRP/WMO, hdl:10013/epic.52032.d001, 2005.

Mahrt, L., Vickers, D., Edson, J., Sun, J., Højstrup, J., Hare, J., and Wilczak, J.M. (1998). Heat flux in the coastal zone. *Boundary-Layer Meteorology*, 86, 421-446.

Mahrt, L., Vickers, D., Edson, J. B., Wilczak, J. M., and Hare, J. (2001). Vertical structure of turbulence in offshore flow during RASEX. *Boundary-Layer Meteorology*, 100, 47-61.

Mahrt, L., Vickers, D., Sun, J., Crawford, T. L., Crescenti, G., and Frederickson, P. (2001). Surface stress in offshore flow and quasi-frictional decoupling. *Journal of Geophysical Research*, 106, 20629-20639.

Mammarella, I., Pumpanen, J., Rannik, U., and Vesala, T. (2009). Relative Humidity Effect on the High-Frequency Attenuation of Water Vapor Flux Measured by a Closed-Path Eddy Covariance System. *Journal of Atmospheric and Oceanic Technology*, 26(9): 1856-1866.

Martens, B., Miralles, D.G., Lievens, H., van der Schalie, R., De Jeu, R.A.M., Fernandez-Prieto, D., Beck, H.E., Dorigo, W.A., and Verhoest, N.E.C. (2017). GLEAM v3: satellite-based land evaporation and root- zone soil moisture. *Geoscientific Model Development Discuss*ion, doi: 10.5194/gmd-2016-162.

Martens, B., de Jeu, R., Verhoest, N., Schuurmans, H., Kleijer, J. and Miralles, D. (2018). Estimating land evaporation at field scales using GLEAM. *Remote Sensing*, 10(11), 1720–25, doi:10.3390/rs10111720, 2018

Mao, J., Fu, W., Shi, X., Ricciuto, D. M., Fisher, J. B., Dickinson, R. E., et al. (2015). Disentangling climatic and anthropogenic controls on global terrestrial evapotranspiration trends. *Environmental Research Letters*, 10(9), 1–13.

McCabe, M. F., Ershadi, A., Jiménez, C., Miralles, D. G., Michel, D., and Wood, E. F. (2016). The GEWEX LandFlux project: evaluation of model evaporation using tower-based and globally-gridded forcing data. *Geoscientific Model Development*, 9, 283–305.

McCabe, M. F., Rodell, M., Alsdorf, D. E., Miralles, D. G., Uijlenhoet, R., Wagner, W., Lucieer, A., Houborg, R., Verhoest, N. E. C., Franz, T. E., Shi, J., Gao, H. and Wood, E. F. (2017). The future of Earth observation in hydrology. *Hydrology and Earth System Sciences*, 21(7), 3879–3914, doi:10.5194/hess-21-3879-2017.

McGillis, W. R., Edson, J. B., Hare, J. E., and Fairall, C. W. (2001) Direct covariance air-sea CO2 fluxes. *Journal of Geophysical Research*, 106, 16729-16745.

McKinley GA, Fay AR, Takahashi T, Metzl N. 2011. Convergence of atmospheric and North Atlantic carbon dioxide trends on multidecadal timescales. *Nat. Geosci.* 4:606–10

Meissner, T. and F. J. Wentz, (2012). The emissivity of the ocean surface between 6 and 90 GHz over a large range of wind speeds and earth incidence angles. *IEEE Trans, Geos. Rem. Sens.* , 50(8), 3004-3026.

Michalsky, J., E. G. Dutton, D. Nelson, J. Wendell, S. Wilcox, A. Andreas, P. Gotseff, D. Myers, I. Reda, T. Stoffel, K. Behrens, T. Carlund, W. Finsterle, and D. Halliwell (2011), An Extensive Comparison of Commercial Pyrheliometers under a Wide Range of Routine Observing Conditions, J. Atmos. Ocean. Tech., 28 (6), 752-766

Michel, D., Jiménez, C., Miralles, D. G., Jung, M., Hirschi, M., Ershadi, A., Martens, B., McCabe, M. F., Fisher, J. B., Mu, Q., Seneviratne, S. I., Wood, E. F., and Fernández-Prieto, D. (2016). The WACMOS-ET project – Part 1: Tower-scale evaluation of four remote sensing-based evapotranspiration algorithms. *Hydrology and Earth System Sciences*, 20, 803–822.

Mikaloff Fletcher SE, Gruber N, Jacobson AJ, Doney SC, Dutkiewicz S, et al. 2006. Inverse estimates of anthropogenic CO2 uptake, transport, and storage by the ocean. *Glob. Biogeochem. Cycles* 20:GB2002.

Miller, S., C. Friehe, T. Hristov, and J. Edson, (2008). Platform motion effects on measurements of turbulence and air-sea exchange over the open ocean, *J. Atmos. Oceanic Tech.,* 25, 1683-1694.

Miralles, D. G., Holmes, T. R. H., De Jeu, R. A. M., Gash, J. H., Meesters, A. G. C. A., and Dolman, A. J. (2011). Global land-surface evaporation estimated from satellite-based observations. *Hydrology and Earth System Sciences*, 15, 453-469.

Miralles, D. G., van den Berg, M. J., Gash, J. H., Parinussa, R. M., De Jeu, R. A. M., Beck, H. E., Holmes, T., Jiménez, C., Verhoest, N. E. C., Dorigo, W. A., Teuling, A. J., and Dolman, A. J. (2014). El Niño–La Niña cycle and recent trends in continental evaporation. *Nature Climate Change*, 4, 122–126

Miralles, D.G., Teuling, A.J., van Heerwaarden, C.C., and Vilá-Guerau de Arellano, J. (2014). Mega-heatwave temperatures due to combined soil desiccation and atmospheric heat accumulation. *Nature Geoscience*, 7, 345–349.

Miralles, D. G., Jiménez, C., Jung, M., Michel, D., Ershadi, A., McCabe, M. F., Hirschi, M., Martens, B., Dolman, A. J., Fisher, J. B., Mu, Q., Seneviratne, S. I., Wood, E. F., and Fernández-Prieto, D. (2016). The WACMOS-ET project - Part 2: Evaluation of global terrestrial evaporation data sets. *Hydrology and Earth System Sciences*, 20, 823–842.

Meissner, T. and F. J. Wentz, (2012). The emissivity of the ocean surface between 6 and 90 GHz over a large range of wind speeds and earth incidence angles. *IEEE Trans, Geos. Rem. Sens.* , **50(8)**, 3004-3026.

Monteith, J. L. (1965). Evaporation and environment. *Symposia of the Social Experimental Biology*, 19, 205–234.

Moum, J. N., de Szoeke, S. P., Smyth, W. D., Edson, J. B., DeWitt, H. L., Moulin,A. J., Thompson, E. J., Zappa, C. J., Rutledge, S. A., Johnson, R. H., and Fairall, C. W. (2014). Air-sea interactions from westerly wind bursts during the November 2011 MJO in the Indian Ocean. *Bulletin of the American Meteorological Society*, 95, 1185-1199.

Mu, Q., Zhao, M., and Running, S. W. (2011). Improvements to a MODIS global terrestrial evapotranspiration algorithm. *Remote Sensing of Environment*, 115, 1781–1800.

Mueller, B., Hirschi, M., Jimenez, C., Ciais, P., Dirmeyer, P. A., Dolman, A. J., Fisher, J. B., Jung, M., Ludwig, F. Maignan, F., Miralles, D. G., McCabe, M. F., Reichstein, M., Sheffield, J., Wang, K., Wood, E. F., Zhang, Y., and Seneviratne, S. I. (2013). Benchmark products for land evapotranspiration: LandFlux-EVAL multi- dataset synthesis, *Hydrology and Earth System Sci*ences, 10, 769–805.

Mueller, J.A. and F. Veron, (2014). Impact of Sea Spray on Air–Sea Fluxes. Part II: Feedback Effects. *J. Phys. Oceanogr.,* 44, 2835–2853, https://doi.org/10.1175/JPO-D-13-0246.1

Nakai, T., Van Der Molen, M. K., Gash, J. H. C. and Kodama, Y. (2006). Correction of sonic anemometer angle of attack errors. *Agricultural and Forest Meteorology*, 136(1-2), 19-30.

National Academies of Sciences, Engineering, and Medicine. 2018. Thriving on Our Changing Planet: A Decadal Strategy for Earth Observation from Space. Washington, DC: The National Academies Press. https://doi.org/10.17226/24938.

Nilsson, E., Bergström, H., Rutgersson, A., Podgrajsek, E., Wallin, M.B., Bergström, G., Dellwik, E., Landwehr, S., and Ward, B. (2018). Evaluating Humidity and Sea Salt Disturbances on CO2 Flux Measurements. *Journal of Atmospheric and Oceanic Technology*, 35, 859–875, https://doi.org/10.1175/JTECH-D-17-0072.1.

Norman, J. M., Kustas, W. P., Prueger, J. H., and G. R. Diak, 2000: Surface flux estimation using radiometric temperature: A dual-temperature-difference method to minimize measurement errors. *Water Resour. Res.*, 36, 2263-2274, doi:10.1029/2000WR900033.

O'Sullivan, N., Landwehr, S., and Ward, B. (2015) Air-flow distortion and wave interactions: An experimental and numerical comparison. *Methods of Oceanography*, 12, 1–17.

Ohmura, A., Dutton, E. G., Forgan, B., Fröhlich, C., Gilgen, H., Hegner, H., Heimo, A., König-Langlo, G., McArthur, B., Müller, G., Philipona, R., Pinker, R., Whitlock, C. H., Dehne, K., and Wild, M. (1998). Baseline surface radiation network (BSRN/WCRP): New precision radiometry for climate research. *Bulletin of American Meteorological Society*, 79, 2115–2136, https://doi.org/10.1175/1520-0477079

Pedreros, R., Dardier, G., Dupuis, H., Graber, H.C., Drennan, W. M., Weill, A., Guerin, C., and Nacass, P. (2003). Momentum and heat fluxes via eddy correlation method on the r/v L'Atalante and an asis buoy. *Journal of Geophysical Research*, 108, 3339.

Price, J. C. (1982). Estimation of regional scale evapotranspiration through analysis of satellite thermal- infrared data. *IEEE Transactions on Geoscience and Remote Sensing*, 3, 286–292.

Priestley, C., and Taylor, R. (1972). On the assessment of surface heat flux and evaporation using large- scale parameters. *Monthly Weather Review*, 100, 81–92.

Prytherch, J., Kent, E. C., Fangohr, S., and Berry, D. I. (2014). A comparison SSM/I-derived global marine surface-specific humidity datasets, *International Journal of Climatol*ogy, 35(9), 2359–2381, doi:10.1002/joc.4150.

Renfrew, I. A., Moore, G. W. K., Guest, P. S., and Bumke, K. (2002). A comparison of surface layer and surface turbulent flux observations over the Labrador Sea with ECMWF analyses and NCEP reanalyses. *Journal of Physical Oceanography*, 32(2), 383-400.

Roberts, J. B., C. A. Clayson, F. R. Robertson, and D. L. Jackson (2010), Predicting near-surface atmospheric variables from Special Sensor Microwave/Imager using neural networks with a first-guess approach, *J. Geophys. Res. Atmos.*, *115*(19), doi:10.1029/2009JD013099.

Roberts, J. B., Clayson, C. A., and Robertson, F. (2019). Improving near-surface retrievals of surface humidity over the global oceans from passive microwave observations. *Earth and Space Science*, revised.

Rodell, M. et al. (2015), The observed state of the water cycle in the early twenty-first century, *J. Clim.*, *28*(21), 8289–8318, doi:10.1175/JCLI-D-14-00555.1.

Roesch, A., M. Wild, A. Ohmura, E. Dutton, C.N. Long, and T. Zhang (2011), Assessment of the completeness of BSRN radiation records and the consequence for the computation of monthly means, Atmos. Meas. Tech., 4, 339–354 .

Rutgersson, A., Norman, M., Schneider, B., Pettersson, H., and Sahlée, E. (2008). The annual cycle of carbon-dioxide and parameters influencing the air-sea carbon exchange in the Baltic Proper. *Journal of Marine Systems*, 74, 381-394.

Sabine C..L, Feely R., Gruber N., Key R., Lee K., et al. 2004. The oceanic sink for anthropogenic CO2. *Science,* 305:367–71

Sahlée, E., Drennan, W. M., Potter, H., and Rebozo, M. A. (2012). Waves and air-sea fluxes from a drifting ASIS buoy during the Southern Ocean Gas Exchange experiment. *Journal of Geophysical Research*, 117, C08003, doi: 10.1029/2012JC008032.

Taylor, P. C., Hegyi, B. M., Boeke, R. C., and Boisvert, L. N. (2018). On the increasing importance of air-sea exchanges in a Thawing Arctic: A review. *Atmosphere*, 9(2), 41; doi:10.3390/atmos9020041.

Teuling, A. J., Van Loon, A. F., Seneviratne, S. I., Shner, I., Aubinet, M., Heinesch, B., et al. (2013). Evapotranspiration amplifies european summer drought. *Geophysical Research Letters*, 40(10), 2071-2075.

Timmermans, W. J., Kustas, W. P., Anderson, M. C., & French, A. N. (2007). An intercomparison of the Surface Energy Balance Algorithm for Land (SEBAL) and the Two-Source Energy Balance (TSEB) modeling schemes. *Remote Sensing of the Environment*, 108, 369-384. doi:10.1016/j.rse.2006.11.028.

Tomita, H., T. Hihara, and M. Kubota (2018), Improved Satellite Estimation of Near-Surface Humidity Using Vertical Water Vapor Profile Information, *Geophys. Res. Lett.*,*45*(2), 899–906, doi:10.1002/2017GL076384.

Tomita, H., Hihara, T., Kako, S., Kubota, M., and Kutsuwada, K., (2019). An introduction to J-OFURO3, a third-generation Japanese ocean flux data set using remote-sensing observations. *Journal of Oceanography*, 75, 171-194.

Tramontana, G., et al. (2016). Predicting carbon dioxide and energy fluxes across global FLUXNET sites with regression algorithms. *Biogeosciences*, 13(14), 4291-4313.

Trenberth, K. E., Fasullo, J. T. and Kiehl, J. (2009). Earth's global energy budget. *Bulletin of the American Meteorological Society*, 90(3), 311-324.

Trenberth, K. E., and Fasullo, J. T. (2017), Atlantic meridional heat transports computed from balancing Earth's energy locally, *Geophys. Res. Lett.*, 44, 1919– 1927, doi:10.1002/2016GL072475.

Uttal, T., and Coauthors, 2002: Surface Heat Budget of the Arctic Ocean. *Bull. Amer. Meteor. Soc.*, **83,** 255–275.

Vickers, D., M. Gockede, and B. E. Law (2010). Uncertainty estimates for 1-h averaged turbulence fluxes of carbon dioxide, latent heat and sensible heat. Tellus, 62(2), 87-99.

Vihma, T., Johansson, M. M., and Launiainen, J. (2009). Radiative and turbulent surface heat fluxes over sea ice in the western Weddell Sea in early summer. *Journal of Geophysical Research*, 114(C4), doi:10.1029/2008JC004995.

von Schuckmann, K., Palmer, M. D., Trenberth, K. E., Cazenave, A., Chambers, D., Champollion, N., Hansen, J., Josey, S. A., Loeb, N., Mathieu, P. P., Meyssignac, B., and Wild, M. (2016). An imperative to monitor Earth's energy imbalance. *Nature Climate Change*, 6(138-144). doi:10.1038/nclimate2876

Wanninkhof, R., and Trinanes, J. (2017). Impact of changing wind speeds on gas transfer and its effect on global air-sea CO2 fluxes. *Journal of Geophysical Research*, 31(6), 961-974, doi:10.1002/2016GB005592.

WCRP Report No. 19, "GEWEX Radiative Flux Assessment (RFA) Volume 1: Assessment", Ed. Ehrhard Rashke, Stefan Kinne, Paul Stackhouse, 2012. (https://www.wcrp-climate.org/documents/GEWEX%20RFA-Volume%201-report.pdf)

Weller R. A., Bigorre, S. P., Lord, J., Ware, J. D., and Edson, J.B. (2012). A surface mooring for air-sea interaction research in the Gulf Stream. Part 1: Mooring design and instrumentation. *Journal of Atmospheric and Oceanic Technology*, 29, 1363-1376. Rep., 5, 19124.

Wentz, F. J. et al., (2017). Evaluating and Extending the Ocean Wind Climate Data Record. *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens*, **10 (5)**, 2165-2185, doi:10.1109/JSTARS.2016.2643641.

Wild, M et al. (2017): The Global Energy Balance Archive (GEBA) version 2017: a database for worldwide measured surface energy fluxes. Earth System Science Data, 9(2), 601-613, https://doi.org/10.5194/essd-9-601-2017.

Wielicki B.A., Barkstrom, B. R., Harrison, E. F., Lee, R. B., Smith, G. L., and Cooper, J.E. (1996). Clouds and the earth's radiant energy system (CERES): an earth observing system experiment. *Bulletin of the American Meteorology Society*, 77(5), 853–868.

Wilson, K. B., et al. (2002). Energy partitioning between latent and sensible heat flux during the warm season at FLUXNET sites. *Water Resources Research*, 38(12), doi:10.1029/2001WR000989.

Yang, P. Z. Jing, and L. Wu (2018). An assessment of representation of oceanic mesoscale eddy-atmosphere interaction in the current generation of general circulation models and reanalyses. *Geophs. Res. Lett.*, doi.org:/10.1029/2018GL080678.

Yu, L., and Weller, R. A., (2007). Objectively Analyzed Air–Sea Heat Fluxes for the Global Ice-Free Oceans (1981–2005). *Bulletin of the American Meteorological Society*, 88, 527–540.

Zhang, Y., Peña-Arancibia, J. L., McVicar, T. R., Chiew, F. H. S., Vaze, J., Liu, C., Lu, X., Zheng, H., Wang., Y., Liu, Y. Y., Miralles, D. G., and Pan M. (2016). Multi-decadal trends in global terrestrial evapotranspiration and its components. *Scientific Reports*, 6, 19124.

Zscheischler, J. et al. (2017). Reviews and syntheses: An empirical spatiotemporal description of the global surface-atmosphere carbon fluxes: opportunities and data limitations. *Biogeosciences*, 14, 3685-3703.