5-5-2016

Measuring and Modeling Oceanic Air-Sea Fluxes

Aaron M. Rosenberg
aaron.m.rosenberg@uconn.edu

Measuring and Modeling Oceanic Air-Sea Fluxes

Aaron Rosenberg

B.A., Boston University, 2011

M.S., University of Connecticut, 2014

A Thesis

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Master of Science

At the

University of Connecticut

2016

APPROVAL PAGE

Masters of Science Thesis

Measuring and Modeling Oceanic Air-Sea Fluxes

Presented by Aaron Rosenberg, M.S.

Major Advisor_____

James B. Edson

Associate Advisor_____

Kelly Lombardo

Associate Advisor_____

Michael Whitney

University of Connecticut

2016

Table of Contents

Abstract

There have been numerous studies evaluating model representation of the latent heat flux (LHF) over terrestrial surfaces due to LHF's role in weather prediction, heat balance, and the hydrological cycle. However, LHF model representation over the ocean, where 86% of global evaporation occurs, has been largely untested due to the scarcity of in-situ measurements and difficulties associated with open ocean observations. This study evaluates the Weather Research and Forecasting Model (WRF) latent heat, sensible heat and momentum surface fluxes over the Sub-tropical North Atlantic Ocean from September 16 to October 30, 2012 under various surface layer and planetary boundary layer (PBL) parameterization schemes. WRF output is validated against bulk and direct covariance flux observations collected during the NASA Salinity Processes in the Upper-Ocean Regional Study (SPURS) from a highly instrumented surface mooring and surveying research vessel. WRF is also compared to the OAFlux hindcast product at an interpolated 1-day, 1-degree resolution.

WRF produced a persistent positive bias in LHF when evaluated against buoy and ship measurements at respective locations in all native parameterization schemes. Modifications to surface layer schemes were employed to mimic the functionality of the COARE3.5 bulk flux algorithm. These modifications reduced root mean square error (RMSE) and model bias in LHF for the default surface layer scheme known as MM5. The inclusion of COARE3.5 functionality had a slightly negative impact on a preferred surface physics scheme known as MYNN, which had already been modified to use the COARE3.0 algorithm. The MYNN scheme with COARE3.0 provided the minimum RMSE between model and observations. Results for momentum flux agreed with these findings. OAFlux closely agreed with SPURS validation data and produced lower surface LHF than WRF across the domain throughout the test case.

This study has found that the WRF overestimation of surface fluxes is the result of inaccurate model physics; e.g., inappropriate convective velocity scale approximation and the lack of an oceanic cool-skin correction. Other factors contributing to the bias and RMSE are a positive bias in the SST used to initialize and force the model and large variability in wind speed, respectively.

1. Introduction

The exchange of solar energy stored in the world's oceans is accomplished mainly through the flux of sensible and latent heat. Latent heat is stored in water molecules during phase change from liquid to vapor at the Earth's surface, which is transferred above the molecular sublayer though turbulent mixing until released during condensation. This process provides energy to the atmosphere from the oceans and land surfaces (mainly through evapotranspiration over vegetation) as latent heat flux (LHF). LHF is the largest non-radiative transfer of energy at the earth's surface, cooling, on average, by 80 Wm^{-2} (Trenberth et al. 2009). LHF also quantifies the transfer of mass, as water vapor (i.e., evaporation), through the air-land/sea interface and is fundamental to the global hydrological cycle. LHF and can be related to the moisture or evaporative flux (E) through

$$E = \frac{1}{L_v} LHF \tag{1}$$

in which L_v is the latent heat of vaporization.

Studies have revealed evidence of a strengthening global water cycle apparent in ocean salinity over the past 50 years, i.e., increased salinities in regions dominated by evaporation and freshening in regions dominated by precipitation (Yu 2007, Wijffels 2010). This strengthening is expected as the troposphere warms and increases its ability to store and transport moisture – a 7% increase in atmospheric moisture storage capacity per 1°C warming due to the exponential nature of the Clausius-Clayperon Relationship (Held and Soden 2006, Schmitt 2008). Since 1989 alone, there has been an increase of tropospheric water vapor concentration at a rate of 1.2% per decade (Trenberth 2009), accompanied by predictions of increased stress on water supplies from global climate and population models (Lagerloef 2010, Raneesh 2014).

A comprehensive understanding of current water pathways is necessary to begin investigating future water availability scenarios, to close the global water budget, and to ultimately produce reliable estimates of precipitation and evaporation in all regions. Much work has been done in observing and validating terrestrial evaporation estimates (Mu et al. 2007, Jung et al. 2010, Wand and Dickinson 2012); however, there have been significantly fewer investigations of moisture flux over the ocean, where 97% of free water is stored and 86% of global evaporation occurs (Schmitt 1995). As a result, the ocean, particularly the sub-tropical ocean, is the main source of atmospheric water vapor (Schmitt 1995) (Figure 1). The importance of this moisture source has been reasserted in recent studies connecting anomalies in sub-tropical Atlantic evaporation, via atmospheric rivers, to flooding in the UK (Lavers et al. 2011, 2013). Consequently, characterizing the fluxes in these regions is imperative to understanding the ocean's role in global climate and Sea Surface Salinity (SSS) regulation, and moisture input for convective weather systems in the tropics as well as mid-latitude cyclones. Further, ocean circulation, an integral part of climate regulation, is affected by surface evaporation through changes in near surface salinity, temperature and, resultantly, density (Taylor 2010).

There are several notable efforts aimed at creating gridded flux products over the ocean for these investigations. A good example is the OAFlux project that integrates observation, remote sensing, and model reanalysis to derive well-validated estimates of evaporation over the ocean; however, OAFlux is unavoidably coarse in both space (1° x 1°) and time (daily averages) and does not extend to future prediction (Yu 2007, Yu et al. 2008). While these products are invaluable for process studies, the coarse resolution limits their use in the present study type, e.g. rapidly developing systems and numerical modeling of future climate. Developing a more detailed picture of the spatial and temporal structure of ocean evaporation is essential for improving the predictive

capabilities of global climate and hydrological models, and, further, for short term weather forecasts in which the planetary boundary layer (PBL) is strongly affected by surface energy balance and the input of moisture.

a. Background

Building an observational database of turbulent fluxes with great spatial extent over the ocean has proven difficult. This is particularly true for long-term direct measurement of the LHF where technical issues such as motion contamination, flow distortion and mechanical wear experienced by turbulent flux instrumentation are compounded by additional fast response humidity sensor requirements including high power and clean optics (Fairall et al. 1990; Edson 2001). Additionally, even where high-resolution time series are available from research vessels (Fujitani 1985; Fairall et al. 1997; Edson et al 2004) and, more recently, surface moorings (e.g., Weller et al. 2012; Bigorre et al. 2013), they provide spatial information on fluxes that are limited to a point or a relatively small survey area.

This study attempts to overcome these limitations by combining high-resolution time series with simulations from the Weather Research and Forecasting (WRF) model as a tool to obtain surface flux fields over the ocean. Like most numerical weather prediction models, WRF cannot directly resolve turbulent moisture exchange at the air-surface interface and through the marine boundary layer due to limitations in resolution and model physics. To account for processes that operate on small spatial and temporal scales, e.g., turbulence, WRF consists of modular code that utilizes separate physics components with user selectable parameterization schemes. This allows for sensitivity tests with multiple WRF native surface physics schemes, and easy modification and

addition of new flux modules. Model LHF is calculated at the air-sea interface using bulk aerodynamic formulae with general form

$$LHF = \rho L_v C_E (q_s - q_a) U \tag{2}$$

in which C_E is the moisture exchange coefficient (Dalton Number) a function of surface roughness and atmospheric stability, $(q_s - q_a)$ is the difference between the saturated surface and atmospheric specific humidity, and U is the surface-relative wind speed. Gustiness is incorporated into the near surface wind speed using a parameterization that is a function of parameters that account for convective instability. This parameterization is described in greater detail in section 2.

This analysis addresses errors in both the mean input variables, e.g., $(q_s - q_a)$ and U, and flux model (C_E), in order to understand the limit of parameterization improvements, and to properly partition future efforts. Ultimately, this study contributes to a better understanding of WRF physics options for the prediction of the LHF.

b. SPURS Field Program

Model output is validated with a dataset collected as part of the NASA Salinity Processes in the Upper-Ocean Regional Study (SPURS) in the Sub-tropical North Atlantic Ocean centered about a buoy located at 38.5ºW longitude, 24ºN latitude (Figure 2). SPURS was conducted to gain an improved understanding of the processes modulating sea surface salinity (SSS). The program provided oceanic salinity, temperature and velocity structure measurements from a surface mooring, research vessels, and autonomous vehicles (e.g., subsurface gliders and wave riders) from September 2012 to October 2013 (Farrar et al. 2015). The surface mooring was equipped with a direct covariance flux system (DCFS, Edson et al. 1998) to provide turbulent fluxes on an

hourly basis. The surface mooring also provided continuous time series of atmospheric and sea surface variables at one-minute resolution to compute means of wind speed and direction, air and sea temperature, humidity, precipitation, solar and IR radiation, and fluxes of heat, moisture and momentum using bulk formulae.

The evaporative flux (E) was combined with precipitation measurements (P) to estimate the net freshwater flux through the ocean surface, i.e., P-E. The SPURS region, site of the climatological salinity maximum, was chosen for its dominance of E relative to P. It is also a region of relatively homogeneous horizontal salinity gradients, which reduces the impact of oceanic freshwater convergence and simplifies freshwater budgeting (Lindstrom et al., 2015); however, the surface fluxes measured by the buoy near the center of the ~1500x1500 km SPURS domain represent the only in situ measurements of the evaporative fluxes during SPURS. As a result, horizontal homogeneity has been assumed to estimate the net freshwater flux over the entire domain for many of the salinity budget studies during SPURS (e.g., Farrar et al. 2015). The impact of spatial variability on this assumption has been the impetus for evaluating WRF fluxes over the open ocean.

c. OAFlux Product

To contrast point observations made as part of the SPURS field campaign, The Objectively Analyzed Air-Sea Heat Fluxes (OAFlux) will provide daily, spatially varying validation data spanning the North Atlantic Basin at 1-degree horizontal resolution. The OAFlux product is similar to model and buoy measurements in that is it calculated via a bulk algorithm - in it's most current version: COARE 3.0. However, it distinguishes itself through the synthesis of meteorological input data from ships, satellite retrievals and atmospheric reanalysis; a technique

intended to reduce error via multiple data sources. The phrase objective analysis refers to the process of combining non-uniform data and accounting for the varying associated errors to create an optimized product, i.e., minimized error variance (Yu 2007). Both field measurements from SPURS and WRF output do not use multiple inputs or optimization, making OAFlux a unique indicator of the value provided by input data assimilation.

2. WRF Investigations

There have been multiple studies examining the effects of PBL physics on the prediction of surface variables (Zhang et al. 2004, Hu et al. 2010, Wang et al. 2014); however, more work is necessary to characterize the sensitivity of surface fluxes to PBL and surface layer schemes over the ocean – a region where studies are scarce. To determine which common WRF physics modules produce realistic surface variables, two popular surface flux schemes with their accompanying PBL schemes are evaluated: 1) the fifth generation mesoscale model surface physics (MM5, Paulson 1970, Dyer and Hicks 1970, Webb 1970, Beljaars 1994, Zhang and Anthes 1982) with the Yonsei University scheme (YSU, Hong, Noh and Dudhia 2006) and Asymmetrical Convective Model 2 (ACM2, Pleim 2007) non-local closure PBL, and 2) the Mellor-Yamada Nakanishi and Niino surface physics (MYNN) and both the MYNN Level 2.5 and 3.0 TKE local closure PBL (Nakanishi and Niino 2006, 2009) . In both cases, WRF is first set up to run over the SPURS domain with the original surface flux modules. The surface flux modules are then modified to incorporate the newest version of the COARE algorithm, a commonly used model for the calculation of fluxes over the open ocean (Brunke et al. 2003, Yu 2007, Yu et al. 2008). In so doing, this study addresses three primary goals: (1) to evaluate WRF model fluxes using the unique, open-ocean SPURS dataset to determine the best physics options

for the prediction of oceanic surface LHF; (2) to modify existing surface flux parameterizations with COARE updates and assess the effect on flux prediction; (3) to begin analysis of model error to determine potential sources of inaccuracy in model simulation (i.e., parameterization versus input variables), in order to guide future work. This effort is the necessary precursor for spatial and temporal investigations of surface fluxes for the SPURS process studies field campaign utilizing WRF output fields.

a. The WRF test case

The evaluation of model surface fluxes over the open ocean was achieved through the use of a suite of numerical model hindcasts using the WRF Advanced Research Core v3.5. The model test case domain covered most of the North Atlantic Basin with 10 km horizontal resolution (Figure 2). The SPURS region of interest was located centrally within the model field and operationally defined as all cells within 15˚ to 30˚ N latitude and 30˚ to 45˚ W longitude.

The model resolves 30 vertical levels with linearly increasing pressure coordinates (i.e., spatially concentrated near the air-land/sea interface) with a model top of 50 mb. The physical parameterizations used in all test runs were WRF Single-Moment 3-class scheme microphysics (Hong, Dudhia and Chen 2004), Rapid Radiative Transfer Model longwave physics (Mlawer et al. 1997), Dudhia Scheme shortwave physics (Dudhia 1989), the Noah Land Surface Model (Niu et al. 2011, Yang et al, 2011), and the Betts-Miller-Janjic Cumulus scheme (Janjic 1994). A more detailed description of WRF run settings is found in Table 1 and 2.

Relevant WRF test case options		
Setting	**Description**	**Namelist Value**
run_hours	run time in hours	36

interval_seconds	time interval between incoming real data at lateral boundary conditions	21600
history_interval	history output file interval (minutes)	60
time_step	time step for integration in integer seconds	60
max_dom	number of domains	1
e_we	index in E-W direction	480
e_sn	index in N-S direction	480
e_vert	number of vertical levels	30
p_top_requested	pressure top (Pa) to use in the model	5000
num_metgrid_levels	levels on incoming forcing data	38
dx	grid length in x direction (meters)	10000
dy	grid length in y direction (meters)	10000

Table 1: WRF Namelist test case options for all test cases.

Parameterization Options			
Setting	**Description**	**Namelist Value**	**Option**
mp_physics	microphysics option	1	WRF Single-Moment 3-class scheme
ra_lw_physics	longwave phsyics option	1	Rapid Radiative Transfer Model longwave physics

ra_sw_physics	shortwave physics option		Dudhia Scheme shortwave physics
radt	minutes between radiation physics calls	30	
cu_physics	cumulus option	1	Betts-Miller-Janjic Cumulus scheme
cudt	minutes between cumulus physics calls	5	
num_soil_layers	number of soil layers in land surface model	4	Noah Land Surface Model
isfflx	heat and moisture fluxes from surface (binary)	1	On
sst_update	option to use time-varying SST at boundary forcing time scale (binary)	1	On

Table 2: WRF physics parameterization modules used in all test case schemes.

The initial state of the atmosphere and boundary forcing were derived from the National Centers for Environmental Prediction (NCEP) CFSv2 Reanalysis Data (Saha 2010, http://rda.ucar.edu/datasets/ds094.0/). In all cases, the Sea Surface Temperature (SST) update was enabled to allow SST as a lower boundary condition to refresh every 6 hours from the NCEP reanalysis. The test case consisted of a series of 36 hour runs initiated daily at 1200 UTC from September 15th through October 29th, 2012 for each set of surface layer schemes and their respective PBL schemes (Table 3). These dates were chosen to correspond to the initial deployment of the SPURS surface mooring when measurements are most reliable before sensor degradation and calibration drift (Edson 2003). The first 12 hours of each run, 1200 to 0000 UTC,

were considered spin up time and disregarded. The next 24 hours, 0000 to 0000 UTC, were used for analysis. The closest model grid point to the surface mooring, along with all bordering model grid points, were averaged to approximate the hourly footprint of the mean wind. This was concluded via an argument from Taylor's Hypothesis, i.e., the frozen turbulence hypothesis (Taylor 1938):

$$\Delta x = u\Delta t \tag{3}$$

in which Δx is the advection length scale, calculated as the product of the mean buoy wind speed (u) and averaging time for observed fluxes (Δt). This hypothesis is based on the assertion that due to the high power of energy in larger turbulent scales, small-scale turbulence can be viewed as frozen in a statistical sense and advected across the appropriate length scale. The advective length scale was found to be between 20 and 30km based on the observed average wind speed of 5.8 ± 2.7 m/s. Therefore, the 30x30km (3x3 model grid points at 10km spacing) box was averaged to create model time series data. The 45 days of 3x3 averaged WRF data were concatenated and used as the model time series for observational comparison.

b) Test case environmental conditions

The test case, taking place during the initial phases of the SPURS observational field campaign from September 16 to October 30, 2012, occurred during a period of typical conditions for a summertime sub-tropical gyre. 95% of hourly mean air temperatures fell between 24.4 and 27.4 C, with sporadic small dips in temperature spread throughout the 45-day test case (Figure 3). Throughout most of the test case the ocean exceeded air temperatures, resulting in an unstable atmosphere, driving consistent convective gusts (Figure 6). The 10-meter wind speeds fell into

two regimes; light convective winds, and increases that coincide with the presence of low-pressure systems (e.g., Tropical Storm Oscar (October 3-5, Figure 4) and Hurricane Nadine (September 17, Figure 5)). The presence of Tropical Storm Oscar corresponds with wind speed outliers in the buoy time series (values > 3*σ) seen in Figure 3.

Precipitation was, as expected, nearly non-existent during the test period besides one episode on 9/26 (Figure 7). The large moisture transfer, combined with small rainfall results in a negative integrated fresh water flux, which is the principal driver of the large local SSS signal.

c) Validation dataset

Data for the evaluation of the WRF surface physics options was collected during the NASA SPURS field campaign from multiple observational platforms and from the gridded OAFlux hindcasts. Meteorological measurements were made from a heavily instrumented air-sea mooring in the center of the SPURS domain, deployed September 14th, 2014 at 24.35˚ N latitude x 38.00˚ W longitude as shown in Figure 8. This region is characteristic of an open ocean gyre, with high salinities, steady winds, and predominantly unstable atmospheric conditions. All systems were mounted between 2 and 3 meters above the buoy deck. The buoy measured a full suite of mean meteorological variables including the wind speed and direction, temperature, specific humidity, and the downwelling solar and infrared radiation (Figure 9). This allowed for a continuous time series of hourly bulk formula fluxes, spanning the entirety of the WRF 45-day test case. Additionally, the R/V Knorr, in close proximity to the central mooring from September 16th to October 6th, 2012 during the initial deployment cruise, was equipped with Gill sonic anemometer/thermometers and fast response humidity sensors to directly measure fluxes using the eddy correlation (EC) method after correction for platform motion (Edson et al. 1998) The physical

setup of these sensors in shown in Figure 10. EC calculates fluxes by taking the average of the product of the vertical fluctuations in wind velocity and transported quantity to calculate, e.g., the latent and sensible heat fluxes as

$$LHF_{EC} = r L_v \overline{w'q'} \tag{4}$$

$$SHF_{EC} = \rho c_p \overline{w'T'} \tag{5}$$

where c_p is the specific heat at constant pressure, w', T' and q' are the instantaneous fluctuations of vertical wind velocity, temperature and specific humidity, respectively; and the overbar represents time averaging over periods of 10 minutes to an hour (Wand and Dickinson 2012). The fluxes were calculated with hourly intervals to match the output time chosen for WRF model variables, and evaluated against a model time series derived from an algorithm that matched ship GPS to the nearest model node at each time step (Figure 11).

Lastly, The OAFlux product was also used as a validation dataset on a daily-averaged interval. OAFlux is coarse in space; therefore, the flux products were linearly interpolated to the SPURS mooring coordinates for assessment against buoy time series and WRF output. For these comparisons, mooring and WRF data were both averaged daily starting from 0000 UTC.

d. COARE Bulk Algorithm

The COARE bulk flux algorithm is used to modify the surface flux parameterizations within WRF code. COARE's genesis was the result of the TOGA COARE field program in the Tropical Pacific Ocean (Fairall et al. 1996a). It has since been verified and improved for a larger range of wind speeds during numerous field campaigns led by NOAA's Physical Science Division (Fairall et al. 1997, 2003) and using data from the Office of Naval Research's MBL/FLIP and RASEX programs

and the National Science Foundation's CLIMODE program as described by Edson et al. (2013). The COARE algorithm calculates the bulk fluxes using model estimates of the required variables and computes the exchange of momentum, heat, and mass by employing Monin-Obukhov Similarity scaling (MOS, e.g., Panofsky and Dutton 1984; Edson and Fairall 1998, Edson et al. 2004). MOS works through the assumption that the behavior of many turbulent statistics in a horizontally homogenous, stationary, and constant surface flux layer [approximately 10% of the fully developed daytime atmospheric boundary layer (Wang and Dickinson 2012)] are universal functions of z/L

$$-\frac{z}{L} = \frac{zkg}{\overline{T}_v}\frac{\overline{w'T_v'}}{u_*^3} \tag{6}$$

where z is the height about the free surface, L is the Monin-Obukov (MO) length, κ is the von Karman's constant, g is gravitational acceleration, T_v is the virtual temperature, $\overline{w'T'_v}$ represents the buoyancy flux, and u_* is the friction velocity. The MO length is the height at which generation of mechanical and buoyancy generated turbulence are equal. The friction velocity is defined as

$$u_* \equiv \sqrt{-\overline{u'w'}} \tag{7}$$

where u' is the along-wind (longitudinal) velocity fluctuations. The buoyancy flux is approximated by

$$BHF_{EC} = \rho c_p \overline{w'T'_v} \cong \rho c_p \left[\overline{w'T'} + 0.61\bar{T}\overline{w'q'}\right] \tag{8}$$

and includes the effects of both temperature and moisture on the density of a fluid parcel (Stull 1988). The virtual temperature fluctuations required for its calculation are closely approximated by the temperature output from a sonic anemometer (Larsen et al. 1993, Edson and Fairall 1998).

e. Parameterization Modification

Both surface layer parameterization schemes (MM5 and MYNN) were modified with COARE elements before being tested again against observations. The extent of these modifications is described below.

First, the WRF MM5 surface flux module (sf_module_sfclay.F) was edited to resemble the functioning of COARE 3.5. Although the original surface flux routine was not iterative, unlike COARE, the multiple WRF time steps in the 12 hour spin-up allowed for the module to stabilize. The most significant change to the code was the form of the gustiness parameterization, which accounts for convective motion and "gustiness" in driving turbulent exchange of momentum, heat and mass. The gusts can contribute significantly to turbulent transfer and are accounted for in the mean wind speed magnitude as follows:

$$U^2 = u_x^2 + u_y^2 + w_g^2 \tag{9}$$

where U is the total horizontal wind speed, u_x and u_y are the vector-averaged wind components (i.e., $u^2 = u_x^2 + u_y^2$), and w_g is the gustiness (Fairall et. al 1996a). These effects are important in light wind conditions when the average winds are small, but the variability (i.e., gustiness) can remain significant. The WRF domain over land uses a gustiness parameterization developed by Godfrey and Beljaars (1991) that is based on the convective velocity scale proposed by Dierdorff (1970)

$$w_g = \beta w_* \tag{10}$$

where β is the gustiness parameter and w_* is the convective velocity scale defined as

$$w_* = \left(\frac{g}{T_v} \overline{w'T_v'} z_i \right)^{1/3} \tag{11}$$

where z_i is the boundary layer height in convective conditions. Over water, the original gustiness parameterization contains a switch to an air-sea temperature difference formulation derived from the older MM5 model and vaguely cited in the module to work by Wyngaard:

$$w_g = 2(\Delta T_{sea-air})^{1/2} \tag{12}$$

This form of the parameterization over the ocean can generate unrealistically large values of the gustiness, particularly over the western boundary currents such as the Gulf Stream, or any region with a large air-sea temperature difference. Therefore, this switch was removed to allow water masked model nodes to employ the convective velocity-based parameterization. This parameterization has been validated over the ocean through multiple observational and modeling studies (Beljaars 1995, Sykes 1993), and was used in the COARE algorithm (Fairall et al., 1996). The removal of this artificial barrier assumes similarity in the convective structure and its dependence on w_* whether over land or over the ocean.

The calculation of surface roughness was modified to compare the default formulation used over water with the latest parameterization found in COARE 3.5 (Edson et al. 2013). These formulations are fairly similar, but differ significantly from the formulation used over terrestrial surfaces. Over land, surface roughness is often constant or varying very slowly on timescales longer than weather simulations; however, in the ocean, the surface roughness is largely a function of the wave field, which is closely coupled to wind forcing. In all the WRF surface flux schemes

tested and COARE, the surface roughness over water is calculated as a combination of 'smooth' and 'rough' elements, accounting, respectively, for viscous drag and wave-induced form drag:

$$z_0 = z_0^{smooth} + z_0^{rough} = z_0^{smooth} + \alpha \frac{u_*^2}{g} \tag{13}$$

where α is known as the Charnock coefficient. In WRF, the Charnock coefficient is set to a constant value of $\alpha = 0.018$ and the roughness length for smooth flow is also set to a constant value of $z_0^{smooth} = 1.59 \times 10^{-5}$ m. In COARE, the smooth portion is parameterized as

$$z_0^{smooth} = g\frac{n}{u_*} \tag{14}$$

where v is the kinematic viscosity and $\gamma = 1/9$ is the value of the roughness Reynolds number for smooth flow from laboratory studies.

The surface roughness for rough flow also differs between WRF and COARE. While the default parameterization defines the Charnock coefficient as a constant, studies have concluded that the surface roughness is better represented when the Charnock coefficient is a variable dependent on wind speed or wave field characteristics (Edson et al. 2013). For WRF, we have embedded a 10-meter wind speed dependent parameterization in the updated module.

$$z_0^{rough} = a(U_{10})\frac{u_*^2}{g} \tag{15}$$

These roughness lengths are used to define the neutral stability value of drag coefficient (i.e., the transfer coefficient for momentum) as

$$C_{DN} = \left(\frac{\kappa}{\ln(z / z_0)} \right)^2 \tag{16}$$

Likewise, the "thermal roughness" lengths z_{0q} and z_{0T} are used to define the transfer coefficients for heat and mass transfer, respectively. For example, the neutral value of the Dalton number for mass (moisture) transfer is defined as

$$C_{EN} = C_{DN}^{1/2} C_{qN} = \left(\frac{\kappa}{\ln(z/z_0)} \right) \left(\frac{\kappa}{\ln(z/z_{0q})} \right) \tag{17}$$

where C_{qN} represents the scalar component of the transfer coefficient. The default parameterization of the thermal roughness lengths are described in Jimenez et al. (2012), while the COARE parameterization is detailed in Fairall et al. (2003). The actual value of the drag coefficient, C_D, and Dalton number, C_E, are a function of atmospheric stability. The stability effects are incorporated using the universal functions of z/L mentioned above and described in detail by Fairall et al. (1996), Grachev et al. (2000) and Edson et al. (2004) .

Additionally, the MM5 surface flux routine was to modify the stability functions for both stable and unstable atmospheric regimes to match COARE3.5. This was accomplished through changes to the formulations of the code in relevant physics modules, followed by recompilation of the WRF model. COARE profiles have been shown to work well over both land and water (Businger, 1988; Edson et al. 2004; Edson et al. 2013). Lastly, the kinematic form of the sensible heat flux, $\overline{w'T'}$, is replaced by the buoyancy flux, $\overline{w'T_v'}$, in WRF calculations of the Monin-Obukhov length as required by (6). As demonstrated by (8), this includes the impact of moisture on the buoyancy flux, which can have a significant impact over the ocean.

The MYNN surface layer parameterization (module_sf_mynn.F) contained many of the updates previously described for the MM5 surface module (Equations 9-11, 13-15). Changes were made to align the module with the recently updated COARE 3.5. Moisture was added to the MO Length by changing the sensible heat flux to the buoyancy flux, to include the additional buoyancy

of moist parcels of air due to surface evaporation. The surface roughness was modified to use the COARE3.5 speed dependent Charnock parameterization. Updated stability functions and thermal roughness length formulations were inserted into the module. The MYNN surface physics already contained the heat flux parameterizations of COARE 3.0, which are nearly identical to the parameterizations in COARE 3.5. Therefore, differences in the heat fluxes are expected to be minimal expect at winds above 15 m/s where the formulations begin to exhibit significant differences.

f) Model Evaluation

In analyzing model output there are four objectives: (1) obtain model performance metrics for multiple parameterization schemes using observational validation data sets collected during the SPURS process studies field program; (2) explore root causes of model inaccuracy; (3) provide estimates for open ocean evaporation; (4) Determine optimized WRF parameterization configurations for budgeting studies and weather/climate prediction.

The RMSE is the principal statistic for quantifying the error in a model's prediction of observed values. It is defined as the square root of the mean square error:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(X_i^m - X_i^o)^2 \tag{13}$$

$$RMSE = \sqrt{MSE} \tag{14}$$

for any variable X, between model (*m*) and observation (*o*) where n is the number of samples. Model bias was also employed to identify persistent bias in modeled LHF. Model bias is calculated as the difference of the mean values of model and observations:

$$m_b = \overline{X}^m - \overline{X}^o \tag{15}$$

The effects of COARE updates to surface flux modules on all variables are evaluated and given a skill score (SS) following the techniques of Oke et al. 2002, in which the model SS is defined as:

$$SS = 1 - \frac{MSE}{MSE_R} \tag{16}$$

where MSE_R is the mean square error of a reference model output, taken in this study as the 45 day test case with the unmodified modules. *MSE* refers to the test case once COARE modifications are added to surface layer parameterizations.

3. Results

a. Observations

Before WRF evaluation, COARE3.5 was validated in the SPURS region. SPURS transfer coefficients for moisture (Dalton Number) and coefficients derived from previously collected datasets show good agreement when compared as a stability corrected function of wind speed. SPURS Dalton numbers do not demonstrate a significant variation from past observations (Figure 12). EC measured LHF from the R/V Knorr DCFS was used to validate bulk fluxes derived from observed mean meteorological variables while the vessel operated in the SPURS domain (Figure 13). Average values from motion corrected DCFS and bulk-derived fluxes, when both available (i.e., accounting for wind directions for fluxes and rainy periods), were nearly identical: DCFS and bulk LHF averages of 112.26 and112.28 W/m^2 respectively. There

was small discrepancy between the two datasets, resulting in good fit between the datasets (r^2 = 0.87) and a RMSE of 22.0 W/m^2. This close agreement substantiates the use of bulk algorithms as a proxy for in-situ moisture flux where eddy covariance fluxes are unavailable. Ship-based bulk measurements were compared to the buoy bulk measurements when the R/V Knorr was within a 30 km radius of the buoy (this occurred approximately 123 hours, comprising 123 hour-averaged flux measurements) [Figure 14]. During times of close proximity, the two observational platforms produced similar shapes (r^2 = 0.88), a RMSE of 22.2 W/m^2, and a difference in average LHF of 14.2 W/m^2 ($\overline{LHF}_{Buoy} = 97.7$ W/m^2, $\overline{LHF}_{Knorr} = 111.9$ W/m^2).

Meteorological variables measured at different heights (16m and 2.84m on the R/V Knorr and moored buoy respectively) were adjusted using a semi-logarithmic profile to a common 10-meter height for comparison. As a result of the common algorithm, discrepancy between bulk fluxes from ship and buoy can be attributed to differences in measurement of scalar values (Figure 16).

b. WRF model output

Model output from all surface parameter schemes produced persistent and easily explained large-scale patterns of LHF over the North Hemisphere as shown in Figure 16; e.g., large fluxes over the Gulf Stream (large air-sea specific humidity difference) and small fluxes over Saharan Africa (low moisture availability). Across the domain, hourly mean WRF LHF ranged from 8.7 W/m^2 to 551.6 W/m^2

The WRF buoy location time series spanned 45 days of hourly average LHF values. For test cases employing original (not COARE-modified) surface parameterizations, LHF values ranged from 60.0 W/m^2 to 438.0 W/m^2 with a standard deviation of 73.5 W/m^2. The range of LHF

increased with COARE modified surface layer schemes: 8.7 W/m^2 - 551.5 W/m^2 (Figure 17). Correlation between mean WRF LHF and mean WRF winds was high (r^2 = 0.722); while correlation between mean WRF LHF and mean WRF Δq (r^2 = 0.097) (Figure 18). This suggests a relationship between variability in wind speed input and variability in WRF LHF.

c. WRF vs. buoy bulk

The eight WRF surface layer and PBL parameterization configurations were evaluated against buoy bulk time series (Figure 19). In all parameterization schemes tested, WRF produced larger average LHF than the buoy bulk fluxes. Non-local PBL schemes (YSU and ACM2) generally produced positive anomalies that were dependent on wind speed (Figure 19; b-d), i.e., the positive anomaly became larger with increased wind speed. The magnitude of error is less sensitive to wind speed in local PBL schemes; MYNN2.5 and MYNN3.0 (Figure 19; e-h).

The parameterization scheme with the smallest LHF model bias was the MM5 COARE updated surface layer and the YSU PBL (MY5-C). This result may be misleading because we are getting the right statistical value for the wrong reason. The cause for the low total model bias was not a better mechanistic or parametric representation of air-sea transfer processes, but rather, counteracting biases: a low bias at low wind speeds, and a high bias at high wind speeds which approximately balance each other (Figure 19; b). The variance in error is also very large, and it has poor fit to the buoy time series with one of the lowest r^2 values of all parameterization configurations. The lowest model biases indicative of sound model flux representation were produced by the MM2.5 and MM3.0 parameterization schemes. These surface flux modules already contained significant COARE functionality, resulting in negligible changes with COARE updates ('-C' is indicative of a COARE modified module) (Table 3).

Test Case	Surface Layer	PBL	RMSE (W/m^2)	Mbias (W/m^2)	R^2
M5Y	*MM5*	*YSU*	75.94	63.10	0.6704
M5Y-C	*MM5-C*	*YSU*	61.30	24.12	0.6132
M5A	*MM5*	*ACM2*	76.47	38.60	0.6396
M5A-C	*MM5-C*	*ACM2*	69.96	32.57	0.6376
MM2.5	*MYNN*	*MYNN2.5*	51.62	33.82	0.7076
MM2.5-C	*MYNN-C*	*MYNN2.5*	59.52	42.80	0.6958
MM3.0	*MYNN*	*MYNN3.0*	53.22	34.60	0.6926
MM3.0-C	*MYNN-C*	*MYNN3.0*	59.80	42.53	0.6856

Table3: Comparison of Buoy Bulk and WRF model output LHF fluxes at the buoy location. September 16, 2012 – October 30, 2012. Average Buoy Bulk LHF = 130.25 W/m^2

For the prediction of LHF, RMSE was the main metric to assess the quality of the WRF model output against buoy fluxes. MM2.5 produced LHF output with the smallest error: 51.6 W/m^2. In all cases, the MYNN surface layer (MM2.5, MM3.0) produced smaller RMSE (51.62 and 53.22 W/m^2 respectively) than the MM5 surface layer and its PBL schemes, both before and after COARE adjustments. Like model bias, RMSE slightly increased using COARE updates. MM5 surface flux runs before COARE updates (M5Y, M5A) produced the largest RMSE (75.94 and 76.47 W/m^2 respectively). These RMSE values were larger than 50% of the mean value of the validation dataset (130.25 W/m^2).

d. WRF vs. R/V Knorr Bulk

As a result of the inherent limitations of research vessels, the R/V Knorr time series was only 43% the length of the buoy and WRF times series (468 hourly fluxes vs. 1080 hourly fluxes from the moored buoy and WRF). MM2.5 produced the smallest RMSE, further bolstering the case for

MM2.5 as the optimal scheme of those tested by corroborating the buoy validation while adding spatial variability within the larger SPURS domain. M5Y again had the largest RMSE. In all WRF parameter configurations, the R/V Knorr dataset agreed better and produced lower RMSE than buoy derived data (Table 3 and Table 4). This is expected and an artifact of the time in which the Knorr was in the SPURS domain; a period when the buoy values also produced lower RMSE before more biased fluxes in the later half of October when the ship was not collecting data in the region. The Knorr time series differed from the buoy time series in that the positive bias was not as persistent across parameterization schemes and was negative for one local PBL scheme run (M5Y-C) (Table 4).

Test Case	**Surface Layer**	**PBL**	**RMSE (W/m^2)**	**Mbias (W/m^2)**	**R^2**
M5Y	*MM5*	*YSU*	71.46	52.63	0.5091
M5Y-C	*MM5-C*	*YSU*	70.33	-0.55	0.4035
M5A	*MM5*	*ACM2*	79.94	15.08	0.4466
M5A-C	*MM5-C*	*ACM2*	73.89	11.65	0.4591
MM2.5	*MYNN*	*MYNN2.5*	51.40	19.72	0.5429
MM2.5-C	*MYNN-C*	*MYNN2.5*	57.55	27.55	0.5304
MM3.0	*MYNN*	*MYNN3.0*	53.77	20.97	0.5224
MM3.0-C	*MYNN-C*	*MYNN3.0*	58.39	27.30	0.5130

Table 4: Comparison of R/V Knorr and WRF model output fluxes at R/V Knorr location. Average Buoy Bulk LHF = 125.08 W/m^2

The change in sign can be identified in the integrated time series of moisture flux at buoy location. The ship, in the SPURS region September 16th through October 4th, 2012, measured fluxes in a period when the positive bias was minimized across measurement platforms. It is assumed that with a longer time series, the positive bias would dominate the short-term variability. The Knorr

dataset also produced lower average r^2 values. The decrease could be a result of the spatial interpolation method used in calculating the Knorr-following WRF time series or differences in the space or time frame within the SPURS domain and test case.

e. WRF vs. OAFlux

Also evaluated was the OAFlux hindcast product. OAFlux was compared to buoy bulk fluxes and used as a validation dataset for two representative WRF test cases (MY5 and MM2.5, Table 6).

Data Set for Comparison	**RMSE (W/m^2)**	**Mean Diff (W/m^2)**	**R^2**
MY5	78.60	67.73	0.6898
MM2.5	54.25	38.46	0.7079
Buoy Bulk	21.81	-4.63	0.8346

Table 5: Statistical Analysis of OAFlux output against WRF output and buoy measurements.

The OAFlux product, a daily value linearly interpolated to SPURS mooring location, showed good agreement with the daily averaged buoy bulk LHF observational data set (Table 5). OAFlux RMSE was smaller than that of any WRF model runs when compared to buoy observations and with minimal bias (< |5 W/m^2|). Both WRF runs estimate larger fluxes than the OAFlux hindcast (mean difference of +67.7 and +38.5 W/m^2 for MY5 and MM2.5 respectively) (Figure 15).

Spatially varying surface fields of daily mean LHF produced from OAFlux and WRF were compared as 1 degree linearly interpolated fields (native resolution to OAFlux, and a reduction of resolution from 10 km for WRF) to determine overall WRF surface flux behavior over the oceans. WRF produced an average daily positive anamoly of 50.7 W/m^2 and 30.9 W/m^2 for MY5 and MM2.5 respectively (Figure 22 and 23). This anomaly was greatest in tropics near the inter-

tropical convergence zone. This again supports the hypothesis of overestimation of surface fluxes from the WRF ARW model in these common parameterization schemes.

Test Case	Surface Layer	PBL	RMSE	Mbias	R^2
M5Y	*MM5*	*YSU*	0.0949	0.0691	0.6488
M5Y-C	*MM5-C*	*YSU*	0.0743	0.0195	0.6812
M5A	*MM5*	*ACM2*	0.0763	0.0201	0.6701
M5A-C	*MM5-C*	*ACM2*	0.0744	**0.0142**	0.6538
MM2.5	*MYNN*	*MYNN2.5*	**0.0666**	0.0183	**0.6827**
MM2.5-C	*MYNN-C*	*MYNN2.5*	0.0705	0.0243	0.6661
MM3.0	*MYNN*	*MYNN3.0*	0.0689	0.0182	0.6602
MM3.0-C	*MYNN-C*	*MYNN3.0*	0.0702	0.0243	0.6688

Table 6: Comparison of Buoy Bulk and WRF model output friction velocity (u) at buoy location. September 16, 2012 – October 30, 2012. Average Buoy Bulk LHF = 130.25 W/m^2*

4. Analysis

a. COARE updates

COARE modifications to the WRF surface layer physics modules resulted in both improved and diminished agreement with measured values of LHF and diagnostic variables of interest (Table 7).

M5Y, the default WRF parameterization scheme and setup with the largest LHF RMSE and model bias evaluated against buoy bulk measurements, i.e., the setup with the worst agreement with observations, had the largest LHF skill score with COARE modification. The large improvement in representation of LHF is from the combination of changes to flux algorithm (i.e., gustiness, buoyancy in MO length), and from the positive feedback of improved meteorological input variables (u*, U10, Q2), which were output as diagnostic variables and evaluated against height-adjusted buoy, measured values (Figure 24 and 25). This feedback highlights the

importance of high-quality parameterizations and data sources in the prediction of non-linear model surface variables. The MM5 surface layer and ACM2 PBL (M5A) was improved with COARE modifications, but to a lesser extent than M5Y, resulting in a smaller skill score. The MYNN surface layer parameterizations (MM2.5 and MM3.0) already contained significant COARE3.5 functionality, and changes to these modules may have disturbed tuning parameters, resulting in negative skill scores with COARE updates (Table 7). The positive skill scores in the MM5 surface layer and not the MYNN surface layer runs infers the added COARE functionality improves the parametric representation of oceanic surface fluxes.

Surface Layer	**PBL**	**LHF**	**u***	**U10**	**q_{2m}**
MM5	*YSU*	0.348	0.387	0.081	0.446
MM5	*ACM2*	0.163	0.048	0.011	-0.085
MYNN	*MYNN 2.5*	-0.386	-0.113	-0.079	-0.160
MYNN	*MYNN 3.0*	-0.263	-0.040	-0.042	-0.093

Table 7: Skill Scores (Oke et al. 2002) for COARE Modifications to WRF surface layer physics parameterization modules for relevant surface variables.

b. Error analysis

An additional advantage to modifying WRF physics modules is the ability to separate the factors influencing model inaccuracy. In mimicking the functionality of the COARE3.5 turbulent flux parameterization in the WRF surface physics modules, which was validated in the test region, the transfer coefficients become virtually identical, and the differences between model and bulk fluxes can be attributed to discrepancy in mean input variables and gradients, e.g., U10, q and T (Figure 19).

These input variable differences and the compounded feedbacks over the 36-hour run are the cause of the model bias and RMSE. While there is variability in the biases of U10, air temperature ($Tair_{2m}$) and specific humidity (Q_{2m}), the most persistent bias is seen in SST. WRF model SST at the skin, provided via the NCEP reanalysis, is on average 0.52 °C warmer than the skin temperature measured in-situ. One reason for this could be that NCEP data has a positive bias. This is unlikely as NCEP ocean temperature is satellite derived and includes a wide range of assimilated ocean observations, e.g., ship measurements and buoys (Saha 2010). It is possible that the bias is a result of a discrepancy in definition of ocean bulk temperature compared to sea surface skin temperature. This seems feasible as WRF uses the same value for SST and sea surface skin temperature. The distinction is important for surface flux parameterizations, which use skin temperature to derive the saturation humidity value at the lower boundary for the calculation of the specific humidity gradient.

The effect of the ocean surface temperature bias can be eliminated by forcing SST input to both the model and bulk algorithm to agree. This is accomplished through the addition of the mean difference between measured SST and model diagnostic SST (i.e., the mean value of the SST bias) to the bulk algorithm SST input. This new SST time series can be re-ingested into the COARE3.5 bulk algorithm to calculate an artificial time series, which will eliminate the effect of SST on LHF comparisons. Adding the 0.52 °C SST bias increased the average value of LHF by 17.13 W/m^2, eliminating 51% of the model bias observed in the MM2.5 test case and reducing RMSE in all cases.

Test Case	**Surface Layer**	**PBL**	**dRMSE (W/m^2)**	**dMbias (W/m^2)**	**dR^2**
M5Y	*MM5*	*YSU*	-13.67	-17.13	0.0034
M5Y-C	*MM5-C*	*YSU*	-6.37	-17.13	0.0184

M5A	*MM5*	*ACM2*	-9.19	-17.13	0.0140
M5A-C	*MM5-C*	*ACM2*	-8.13	-17.13	0.0135
MM2.5	*MYNN*	*MYNN2.5*	-9.58	-17.13	0.0050
MM2.5-C	MYNN-C	MYNN2.5	-11.31	-17.13	0.0047
MM3.0	MYNN	MYNN3.0	-9.47	-17.13	0.0039
MM3.0-C	MYNN-C	MYNN3.0	-11.07	-17.13	0.0040

Table 8: Changes to buoy-model comparison statistics as a result of forced equality of SST time series

Further tests with internally modified WRF SST would be necessary to directly quantify error to this artifact; however, the COARE test indicates a significant contribution to error from disagreement in sea surface skin temperature.

c. Evaporative Flux

The inconsistency between modeled and measured LHF; and therefore, moisture flux, could have a significant effect on freshwater budgeting efforts in the open ocean where moisture flux through the air-sea interface is the largest ocean freshwater sink (Schmitt 1995). Figure 20 presents integrated evaporation at buoy location for all WRF parameterization configurations, observations, and OAFlux hindcasts. As could be garnered from the LHF model biases, all of the WRF parameterization schemes produced more evaporation than buoy bulk fluxes over the test period. The buoy bulk methodology predicted the evaporation of 224.07 mm H_2O over the 45-day test case (4.98 mm/day). The default WRF model simulations predicted total evaporation ranging from 282.27 mm H_2O to 332.63 mm H_2O (6.27 - 7.39 mm/day) (Figure 26). If the COARE modified parameterizations are taken into account the range increases to 5.90 - 7.39 mm/day (Table 9).

Test Case	E (mm/day)	σ (mm/day)
M5Y	7.40	2.81
M5Y-C	5.91	3.40
M5A	6.46	3.92
M5A-C	6.23	3.73
MM2.5	6.28	2.76
MM2.5-C	6.62	2.86
MM3.0	6.31	2.79
MM3.0-C	6.61	2.86

Table 9: WRF model daily average evaporation. Buoy observed daily evaporation = 4.98 mm/day

The large range of moisture flux rates highlights the importance of parameterization choice when using model derived surface forcing fields in budgeting experiments.

Although there does appear to be heterogeneity in moisture flux over the area of interest (Figure 28), spatial variability of mean daily evaporation remains within the standard deviation of test case evaporation rates inside the SPURS domain for all WRF parameterization schemes. This suggests the assumption that a central measurement could represent evaporative flux over the region of interest does hold in this 45-day test period if fluctuations on shorter timescales are not of interest. Although model output was shown to contain a persistent positive bias, test case MM2.5, the best performing model parameterization scheme, displayed homoscedastic model bias as a function of time and wind speed. The constant overestimation implies that although the magnitude is biased, the variation in space is likely an accurate estimation of horizontal heterogeneity in the region surrounding the SPURS buoy, i.e., the SPURS domain (Figure 29).

d. Parameter choice recommendation

This research is the first step in identifying the optimal WRF parameterization setup for the accurate prediction of air-sea latent heat and moisture transfer. For prediction of meteorological dynamics and daily weather, the MYNN surface layer and MYNN 2.5 PBL produced the smallest RMSE, the largest r^2 and small positive bias when compared to all observational datasets. For studies interested in accurate freshwater budgeting, the MM5 surface layer with COARE updates and YSU PBL produced the minimum model bias. If resolution is not an issue and the time of interest is in the past, the OAFlux hindcast product produced a very small bias compared to buoy measurements with daily, global 1°x1° coverage.

5. Discussion

The LHF bias of 63.1 W/m^2 associated with the default WRF surface layer and PBL revealed in this study has significant implications for surface energy balance and hydrological studies.

If we extrapolate the average bias over the global ocean, it would result in an increase in a 9.2 Sv increase in global evaporation. This is an 59% increase from the base estimate of 15.7 Sv from Schimitt (1995). This simulation of the atmosphere would be deeply flawed as a result of the poor bottom boundary flux conditions. As a result, WRF regional climate simulations over areas that contain significant portions of ocean, using default surface physics and PBL, would poorly represent the energy balance and the hydrological cycle of the study area. Further, WRF simulations of cyclogensis over oceans, e.g., nor'easters, are prone to overestimate buoyancy flux into storms and will poorly represent atmospheric dynamics. In the example of nor'easter prediction, the advantages of hindcasts such as OAFlux are not applicable; therefore, WRF surface physics must be improved if it is to be a useful storm forecasting tool.

Before working to improve the model, it is important to determine which errors are resolvable through changes to model physics (contained in C_E), and which are due to variability in

input variables (e.g., U_r and Δq). The lack of similarity in gustiness calculations between land and water model nodes in the MM5 surface layer is a clear example of a systemic choice that leads to error in WRF transfer coefficients. Assuming similarity between land and water reduces a substantial portion of the error. Modules that have been adjusted to closely mimic the functionality of the COARE surface flux module still produce errors, which implies that the remaining, significant errors may be a result of inaccurate surface physics input variables (e.g., SST, U10) or non-linear model behavior. An example which highlights the importance of input variables is the OAFlux product, also based on COARE3.0, which performs measurably better than WRF when compared to SPURS observations. One reason for the success of OAFlux is the use of objective analysis and the optimized combination of multiple meteorological input data sources (Yu et al. 2007).

This raises the question: how do we overcome the remaining errors we see in WRF fluxes? A first step for forecasting will be to use the physics modules that are shown to perform best in validation. For this study over the open ocean, that is the MYNN surface layer and the MYNN 2.5 PBL. If there are reasons to use the MM5 surface layer, such as the preference for the associated non-local PBL schemes, the user should modify the MM5 surface flux module to employ the added COARE functionality used in this study, especially the updated convective velocity scale parameterization. For hindcasts, the answer likely lies in data assimilation. This analysis demonstrated that wind speed is the variable most correlated to WRF flux magnitude; therefore, wind speed accuracy is greatly important to producing the variability seen in surface fluxes. OAFlux assimilates three remotely sensed wind products: two from passive radiometers, and the active, QuickSCAT scatterometer. In historical runs, modelers interested in WRF surface

fluxes will greatly benefit from assimilation and nudging the model towards these measurements. A possible methodology could assimilate two sources and reserve one source for model validation.

6. Summary

In this work, WRF model simulations, combined with in-situ buoy and ship based flux observations, were used to evaluate a number of native and modified surface flux parameterizations. The results of a 45-day WRF test case from September 2012 through October 2012 revealed a persistent positive model bias in modeled oceanic LHF. COARE modifications to surface flux modules increased the skill of parameterizations lacking much existing COARE functionality, while the modifications slightly decreased the skill of modules already containing significant COARE functionality. The parameterization scheme that produced the minimum RMSE, the main metric for assessment of model output, was the MYNN surface layer with the MYNN 2.5 PBL. The model output showed heterogeneity in LHF values across the SPURS domain; however, in all test cases, the variability remained within a standard deviation of the buoy location value for mean daily evaporation rates, thus validating the SPURS assumption in measuring latent heat flux at a central location for budgeting efforts if integrated over a significant enough time. Due to the congruence of transfer coefficient formulations between bulk algorithm and WRF model fluxes in modified test cases, substantial model error was shown to be the result of disagreement between measured and model input variables, most notably SST, rather than lacking model formulations.

Future work will determine the necessary time period in which assumptions of spatial homogeneity are validated and will assess the sensitivity of WRF to input variables (e.g., SST and wind speed) to advance understanding of model error and bias. Further, WRF surface flux fields

will be used to identify flux variability and high-resolution spatial scales of heterogeneity in the open ocean. Understanding this variability will help in simulating ocean surface dynamics through improved surface forcing, as well as aid in spatial planning for observational arrays in future field campaigns to better collect spatially representative measurements.

Citations

Alice K. Duvivier and John J. Cassano. Evaluation of WRF Model Resolution on Simulated Mesoscale Winds and Surface Fluxes near Greenland. *Monthly Weather Review*. 2012

Beljaars, A., 1995. The parametrization of surface fluxes in large- scale models under free convection. *Quarterly Journal of the Royal Meteorological Society*, 121(522), pp.255-270.

Beljaars, A.C.M., 1994. The parameterization of surface fluxes in large-scale models under free convection. *Quart. J. Roy. Meteor. Soc.*, 121, 255–270.

Bigorre, S.P., Weller, R.A., Edson, J.B. and Ware, J.D., 2013. A surface mooring for air–sea interaction research in the Gulf Stream. Part II: Analysis of the observations and their accuracies. *Journal of Atmospheric and Oceanic Technology*, 30(3), pp.450-469.

Brunke, M.A., Fairall, C.W., Zeng, X., Eymard, L. and Curry, J.A., 2003. Which bulk aerodynamic algorithms are least problematic in computing ocean surface turbulent fluxes?. *Journal of Climate*, 16(4), pp.619-635.

Businger, J.A., 1988. A note on the Businger-Dyer profiles. *In Topics in Micrometeorology*. A Festschrift for Arch Dyer (pp. 145-151). Springer Netherlands.

Cassano, J., M. Higgins, and M. Seefeldt, 2011. Performance of the Weather Research and Forecasting Model for month-long pan-Arctic simulations. *Mon. Wea. Rev.*, 139, 3469–3488.

Charnock, H., 1955. Wind stress on a water surface. *Quart. J. Roy. Meteor. Soc.*, 81, 639–640

Deardorff, J.W., 1970. Convective velocity and temperature scales for the unstable planetary boundary layer and for Rayleigh convection. *Journal of the atmospheric sciences*, 27(8), pp.1211-1213.

Dudhia, J., 1989. Numerical study of convection observed during the Winter Monsoon Experiment using a mesoscale two–dimensional model. *J. Atmos. Sci.*, 46, 3077–3107.

Durack, P.J. and Wijffels, S.E., 2010. Fifty-year trends in global ocean salinities and their relationship to broad-scale warming. *Journal of Climate*,23(16), pp.4342-4362.

Dyer, A. J., and B. B. Hicks, 1970. Flux–gradient relationships in the constant flux layer. *Quart. J. Roy. Meteor. Soc*., 96, 715–721.

Edson, J. B., 2001. Sensors for micrometeorological flux measurements. *Elements of Physical* Oceanography.

Edson, J.B. and Fairall, C.W., 1998. Similarity relationships in the marine atmospheric surface layer for terms in the tke and scalar variance budgets*.*Journal of the atmospheric sciences*, 55(13), pp.2311-2328.

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(2), pp.547-562.

Edson, J.B., Jampana, V., Weller, R.A., Bigorre, S.P., 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(8), pp.1589-1610.

Edson, J.B., Zappa, C.J., Ware, J.A., McGillis, W.R. and Hare, J.E., 2004. Scalar flux profile relationships over the open ocean. *Journal of Geophysical Research*: *Oceans*, 109(C8).

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

Fairall, C.W., Bradley, E.F., Rogers, D.P., Edson, J.B. and Young, G.S., 1996. Bulk parameterization of air-sea fluxes for tropical ocean- global atmosphere coupled-ocean atmosphere response experiment. *Journal of Geophysical Research: Oceans*, 101(C2), pp.3747-3764.

Fairall, C.W., Edson, J.B., Larsen, S.E. and Mestayer, P.G., 1990. Inertial-dissipation air-sea flux measurements: a prototype system using realtime spectral computations. *Journal of Atmospheric and Oceanic Technology*,7(3), pp.425-453.

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 and Oceanic Technology*, 14(3), pp.338-359.

Farrar, J.T., L. Rainville, A.J. Plueddemann, W.S. Kessler, C. Lee, B.A. Hodges, R.W. Schmitt, J.B. Edson, S.C. Riser, C.C. Eriksen, and D.M. Fratantoni. 2015. Salinity and temperature balances at the SPURS central mooring during fall and winter. *Oceanography* 28(1):56–65,

Farrar, J.T., L. Rainville, A.J. Plueddemann, W.S. Kessler, C. Lee, B.A. Hodges, R.W. Schmitt, J.B. Edson, S.C. Riser, C.C. Eriksen, and D.M. Fratantoni. 2015. Salinity and temperature balances at the SPURS central mooring during fall and winter. *Oceanography* 28(1):56–65, http://dx.doi.org/10.5670/oceanog.2015.06.

Fujitani, T., 1985. Method of turbulent flux measurement on a ship by using a stable platform system. *Papers in Meteorology and Geophysics*, 36(3), pp.157-170.

Godfrey, J.S. and Beljaars, A.C.M., 1991. On the turbulent fluxes of buoyancy, heat and moisture at the air-sea interface at low wind speeds. *Journal of Geophysical Research: Oceans*, 96(C12), pp.22043-22048.

Grachev, A.A., Fairall, C.W. and Bradley, E.F., 2000. Convective profile constants revisited. *Boundary-layer meteorology*, 94(3), pp.495-515.

Held, I.M. and Soden, B.J., 2006. Robust responses of the hydrological cycle to global warming. *Journal of Climate*, 19(21), pp.5686-5699.

Hong, Song–You, Jimy Dudhia, and Shu–Hua Chen, 2004. A revised approach to ice microphysical processes for the bulk parameterization of clouds and precipitation. *Mon. Wea. Rev*., 132, 103–120.

Hong, Song–You, Yign Noh, Jimy Dudhia, 2006. A new vertical diffusion package with an explicit treatment of entrainment processes. *Mon. Wea. Rev*., 134, 2318–2341.

Hu, X.M., Klein, P.M. and Xue, M., 2013. Evaluation of the updated YSU planetary boundary layer scheme within WRF for wind resource and air quality assessments. *Journal of Geophysical Research: Atmospheres*,118(18).

Janjic, Zavisa I., 1994. The Step–Mountain Eta Coordinate Model: Further developments of the convection, viscous sublayer, and turbulence closure schemes. *Mon. Wea. Rev*.,122, 927–945.

Jiménez, P.A., Dudhia, J., González-Rouco, J.F., Navarro, J., Montávez, J.P. and García-Bustamante, E., 2012. A revised scheme for the WRF surface layer formulation. *Monthly Weather Review*, 140(3), pp.898-918.

Jung, M., Reichstein, M., Ciais, P., Seneviratne, S.I., Sheffield, J., Goulden, M.L., Bonan, G., Cescatti, A., Chen, J., De Jeu, R. and Dolman, A.J., 2010. Recent decline in the global land evapotranspiration trend due to limited moisture supply. *Nature*, 467(7318), pp.951-954.

Lagerloef, G., Schmitt, R., Schanze, J. and Kao, H.Y., 2010. The ocean and the global water cycle. *Oceanography*.

Larsen, S.E., Edson, J.B., Fairall, C.W. and Mestayer, P.G., 1993. Measurement of temperature spectra by a sonic anemometer. *Journal of Atmospheric and Oceanic Technology*, 10(3), pp.345-354.

Lavers, D.A., Allan, R.P., Villarini, G., Lloyd-Hughes, B., Brayshaw, D.J. and Wade, A.J., 2013. Future changes in atmospheric rivers and their implications for winter flooding in Britain. *Environmental Research Letters*,8(3), p.034010.

Lavers, D.A., Allan, R.P., Wood, E.F., Villarini, G., Brayshaw, D.J. and Wade, A.J., 2011. Winter floods in Britain are connected to atmospheric rivers. *Geophysical Research Letters*, 38(23).

Lindstrom, E., F. Bryan, and R. Schmitt. 2015. SPURS: Salinity Processes in the Upperocean Regional Study—The North Atlantic Experiment. *Oceanography* 28(1):14–19, http://dx.doi.org/10.5670/oceanog.2015.01.

Lindstrom, E., F. Bryan, and R. Schmitt. 2015. SPURS: Salinity Processes in the Upperocean Regional Study—The North Atlantic Experiment. *Oceanography* 28(1):14–19.

Mlawer, Eli. J., Steven. J. Taubman, Patrick. D. Brown, M. J. Iacono, and S. A. Clough (1997), Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated–k model for the longwave. *J. Geophys.* Res.,102, 16663–16682.

Mu, Q., Heinsch, F.A., Zhao, M. and Running, S.W., 2007. Development of a global evapotranspiration algorithm based on MODIS and global meteorology data. *Remote Sensing of Environment*, 111(4), pp.519-536.

Nakanishi, M., and H. Niino, 2006. An improved Mellor–Yamada level 3 model: its numerical stability and application to a regional prediction of advecting fog. *Bound. Layer Meteor*. 119, 397–407.

Nakanishi, M., and H. Niino, 2009. Development of an improved turbulence closure model for the atmospheric boundary layer. *J. Meteor. Soc. Japan*, 87, 895–912.

Niu, Guo–Yue, Zong–Liang Yang, Kenneth E. Mitchell, Fei Chen, Michael B. Ek, Michael Barlage, Anil Kumar, Kevin Manning, Dev Niyogi, Enrique Rosero, Mukul Tewari, Youlong Xia, 2011. The community Noah land surface model with multiparameterization options (Noah–MP): 1. Model description and evaluation with local–scale measurements. *J. Geophys. Res*., 116, D12109.

Oke, P.R., Allen, J.S., Miller, R.N., Egbert, G.D., Austin, J.A., Barth, J.A., Boyd, T.J., Kosro, P.M. and Levine, M.D., 2002. A modeling study of the three-dimensional continental shelf circulation off Oregon. Part I: Model-data comparisons. *Journal of Physical Oceanography*, 32(5), pp.1360-1382.

Panofsky, H.A., Dutton, J.A., 1984. Atmospheric Turbulence: Models and Methods for Engineering Applications. New York: *Wiley*, 1984. Print.

Paulson, C. A., 1970. The mathematical representation of wind speed and temperature profiles in the unstable atmospheric surface layer. *J. Appl. Meteor*., 9, 857–861.

Pleim, Jonathan E., 2007. A Combined Local and Nonlocal Closure Model for the Atmospheric Boundary Layer. Part I: Model Description and Testing. *J. Appl. Meteor. Climatol*.,46, 1383–1395.

Raneesh, K.Y., 2014. Impact of Climate Change on Water Resources. *Journal of Earth Science & Climatic Change*, 5(3), p.1.

Saha, S., Moorthi, S., Pan, H.L., Wu, X., Wang, J., Nadiga, S., Tripp, P., Kistler, R., Woollen, J., Behringer, D. and Liu, H., 2010. The NCEP climate forecast system reanalysis. *Bulletin of the American Meteorological Society*,91(8), pp.1015-1057.

Schmitt, R.W., 1995. The Ocean Component of the Global Water Cycle. *Reviews of Geophysics*, 33(S2) pp 1395-1409

Schmitt, R.W., 2008. Salinity and the Global Water Cycle. *Oceanography*,21(1), p.12.

Skamarock, W. C., 2004. Evaluating mesoscale NWP models using kinetic energy spectra. *Mon. Wea. Rev*., 132, 3019–3032.

Sproson, D. A. J., I. A. Renfrew, and K. J. Heywood, 2008. Atmospheric conditions associated with oceanic convection in the south-east Labrador Sea. *Geophys. Res. Lett.*, 35, L06601, doi:10.1029/2007GL032971.

Stull, R.B., 2012. An introduction to boundary layer meteorology (Vol. 13). *Springer* Science & Business Media.

Sykes, R.I., Henn, D.S. and Lewellen, W.S., 1993. Surface- layer description under free-convection conditions. *Quarterly Journal of the Royal Meteorological Society*, 119(511), pp.409-421.

Taylor, G.I., 1938, February. The spectrum of turbulence. *In Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* (Vol. 164, No. 919, pp. 476-490). The Royal Society.

Trenberth, K.E., Fasullo, J.T. and Kiehl, J., 2009. Earth's global energy budget. *Bulletin of the American Meteorological Society*, 90(3), pp.311-323.

Wang, K. and Dickinson, R.E., 2012. A review of global terrestrial evapotranspiration: Observation, modeling, climatology, and climatic variability. *Reviews of Geophysics*, 50(2).

Wang, Z., Duan, A. and Wu, G., 2014. Impacts of boundary layer parameterization schemes and air-sea coupling on WRF simulation of the East Asian summer monsoon. *Science China Earth Sciences*, 57(7), pp.1480-1493.

Webb, E. K., 1970. Profile relationships: The log-linear range, and extension to strong stability. *Quart. J. Roy. Meteor. Soc*., 96, 67–90.

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 I: Mooring design and instrumentation. *Journal of Atmospheric and Oceanic Technology*, 29(9), pp.1363-1376.

Yang, Z.–L., G.–Y. Niu, K. E. Mitchell, F. Chen, M. B. Ek, M. Barlage, L. Longuevergne, K. Manning, D. Niyogi, M. Tewari, and Y. Xia, 2011. The community Noah land surface model with multiparameterization options (Noah–MP): 2. Evaluation over global river basins. *J. Geophys. Res*., 116, D12110.

Yu, L, 2007. Global variations in oceanic evaporation (1958–2005): the role of the changing wind speed. *J. Climate,* 20, 5376–5390.

Yu, L., X. Jin, and R. A. Weller, 2008. Multidecade Global Flux Datasets from the Objectively Analyzed Air-sea Fluxes (OAFlux) Project: Latent and sensible heat fluxes, ocean evaporation, and related surface meteorological variables. OA-2008-1, *Woods Hole Oceanographic Institution*, 64 pp.

Zhang, D.L. and Zheng, W.Z., 2004. Diurnal cycles of surface winds and temperatures as simulated by five boundary layer parameterizations. *Journal of Applied Meteorology*, 43(1), pp.157-169.

Zhang, D.–L., and R.A. Anthes, 1982. A high–resolution model of the planetary boundary layer–sensitivity tests and comparisons with SESAME–79 data. *J. Appl. Meteor*., 21, 1594–1609.

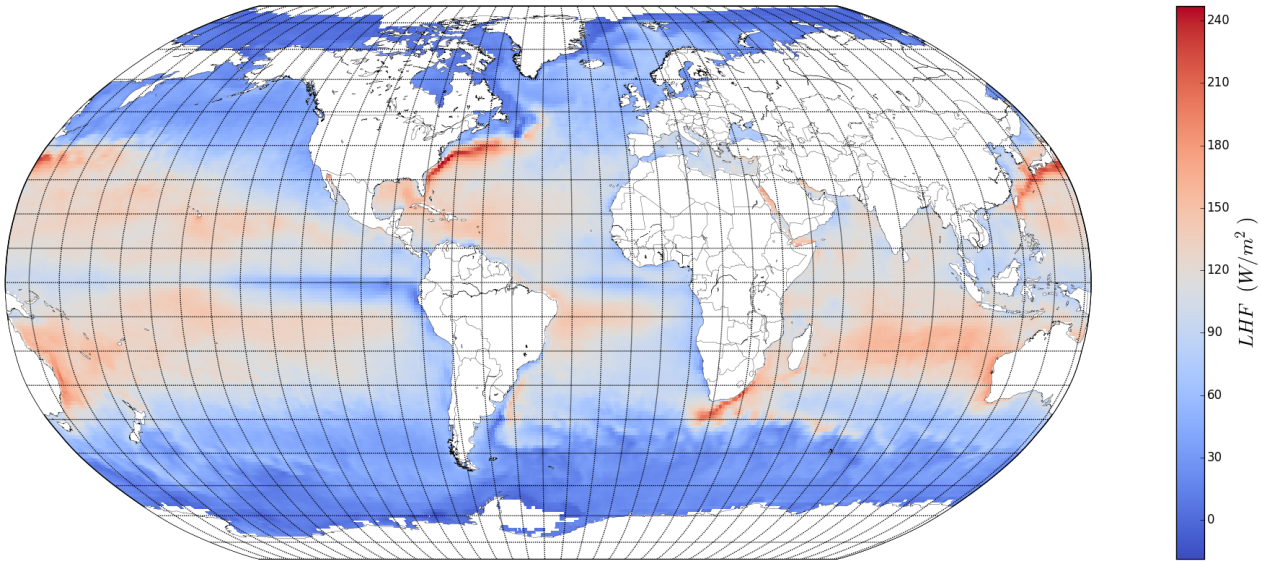

Figure 1: Average Oceanic LHF (W m^{-2}): Jan 1, 2012 – Dec 31, 2012. Data source: OAFlux 2012 lhf

Figure 2: Model Grid and Field Campaign Spatial Extent

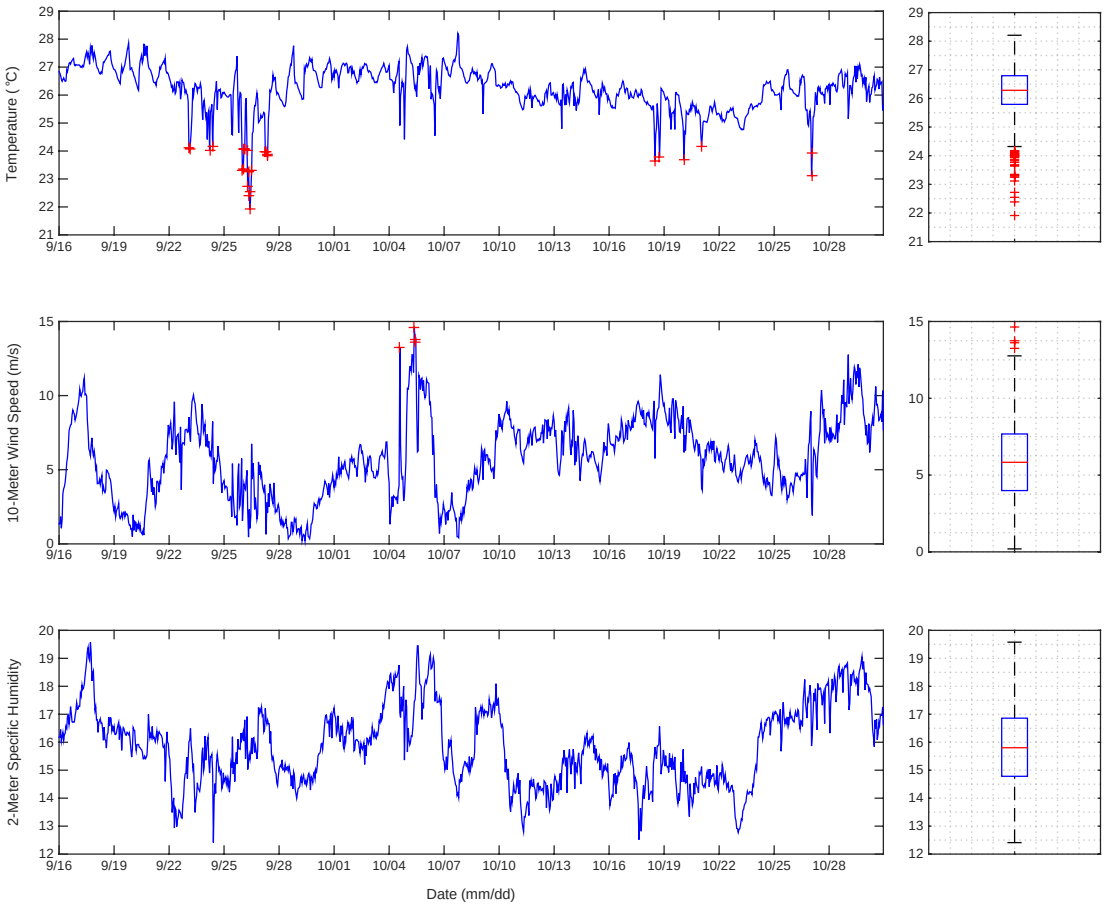

Figure 3: Left: Time series of meteorological variables at buoy location. Right: box and whisker plots of meteorological data. Red dots denote outliers from the population of samples.

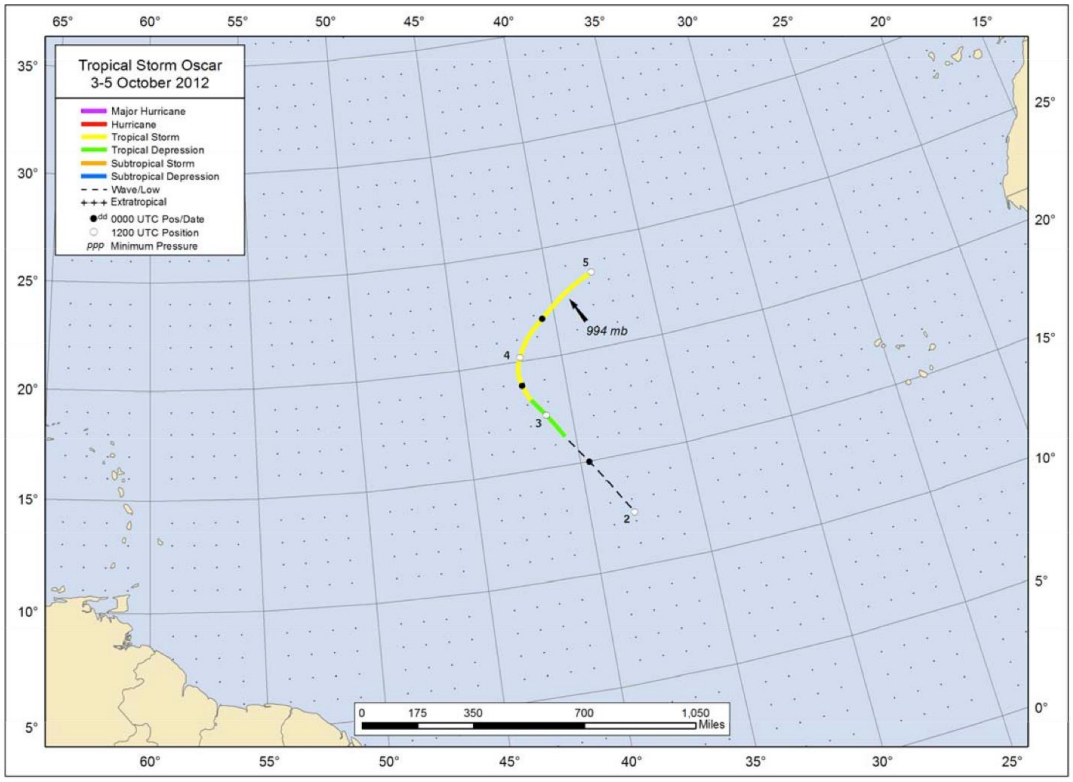

Figure 4: Tropical Storm Oscar track from NOAA National Hurricane Center

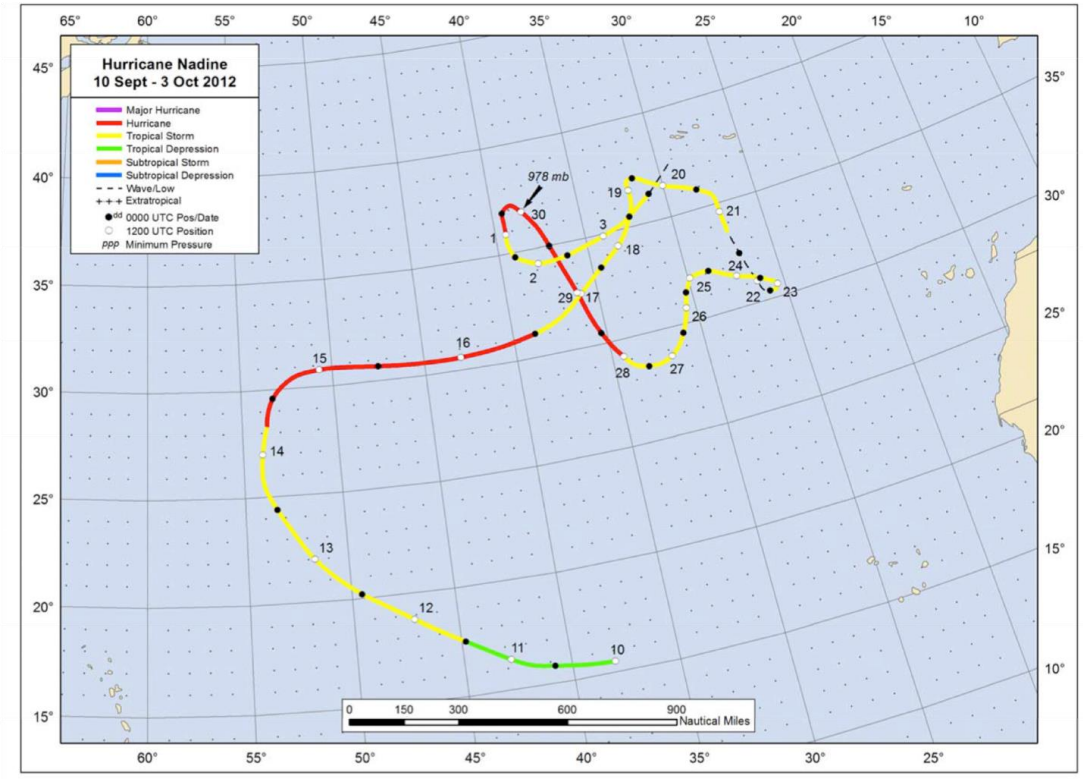

Figure 5: Hurricane Nadine track from NOAA National Hurricane Center

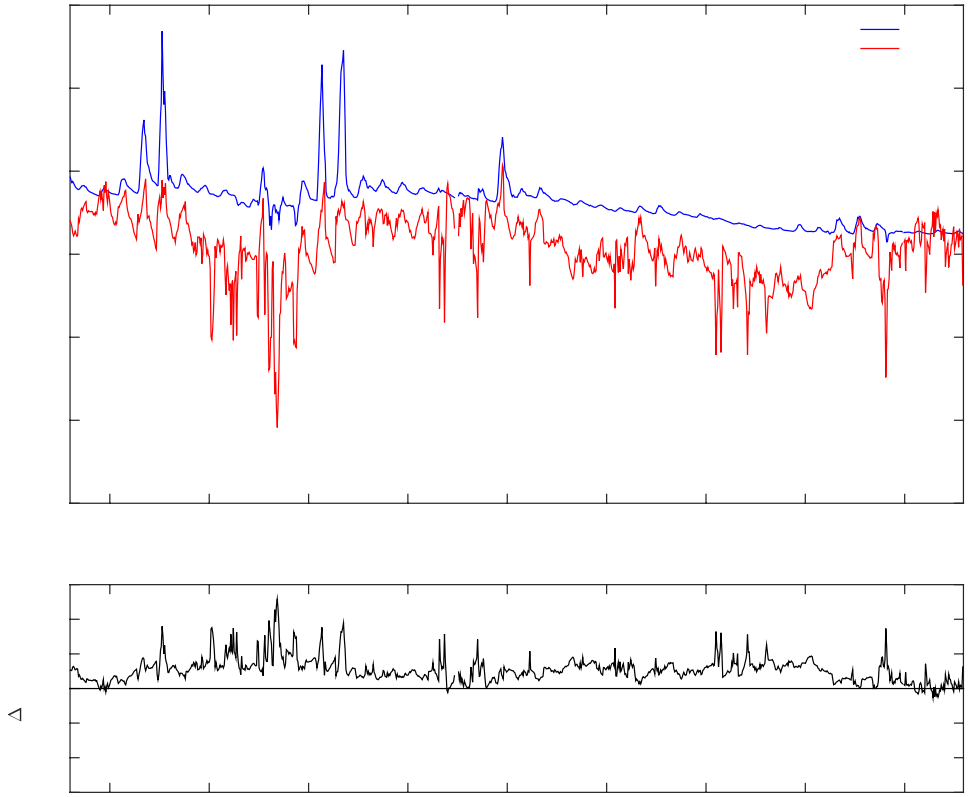

Figure 6: Top: Buoy location ocean and atmosphere temperature time series. Bottom: Air-sea temperature difference.

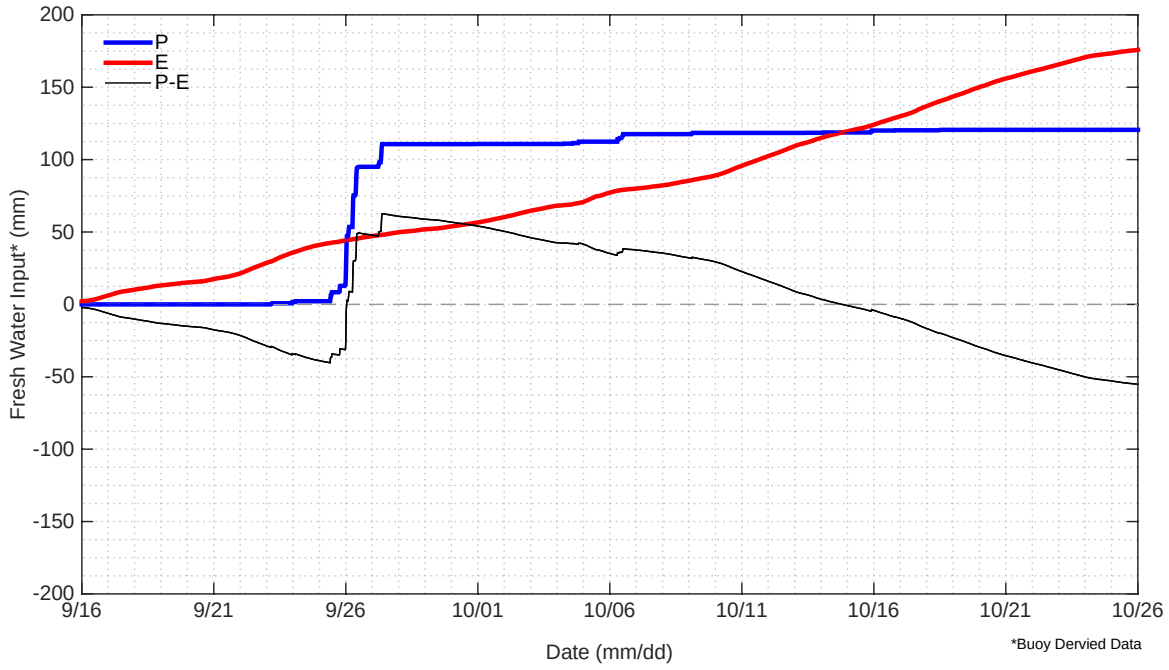

Figure 7: Buoy location precipitation (P) and evaporation (E) throughout the test case.

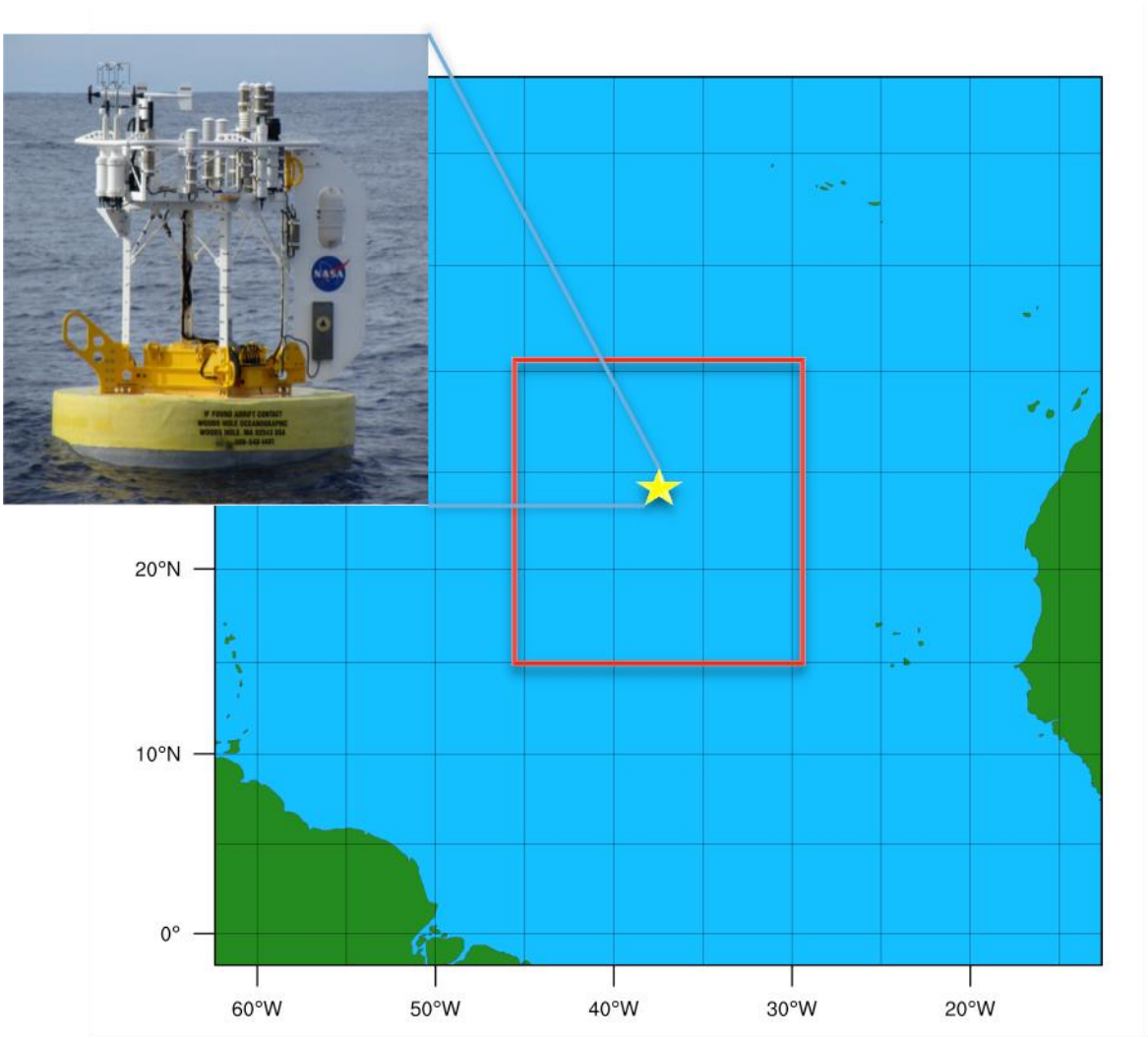

Figure 8: Deployment location of SPURS meteorological mooring. Red square denotes region of interest for the SPURS field campaign.

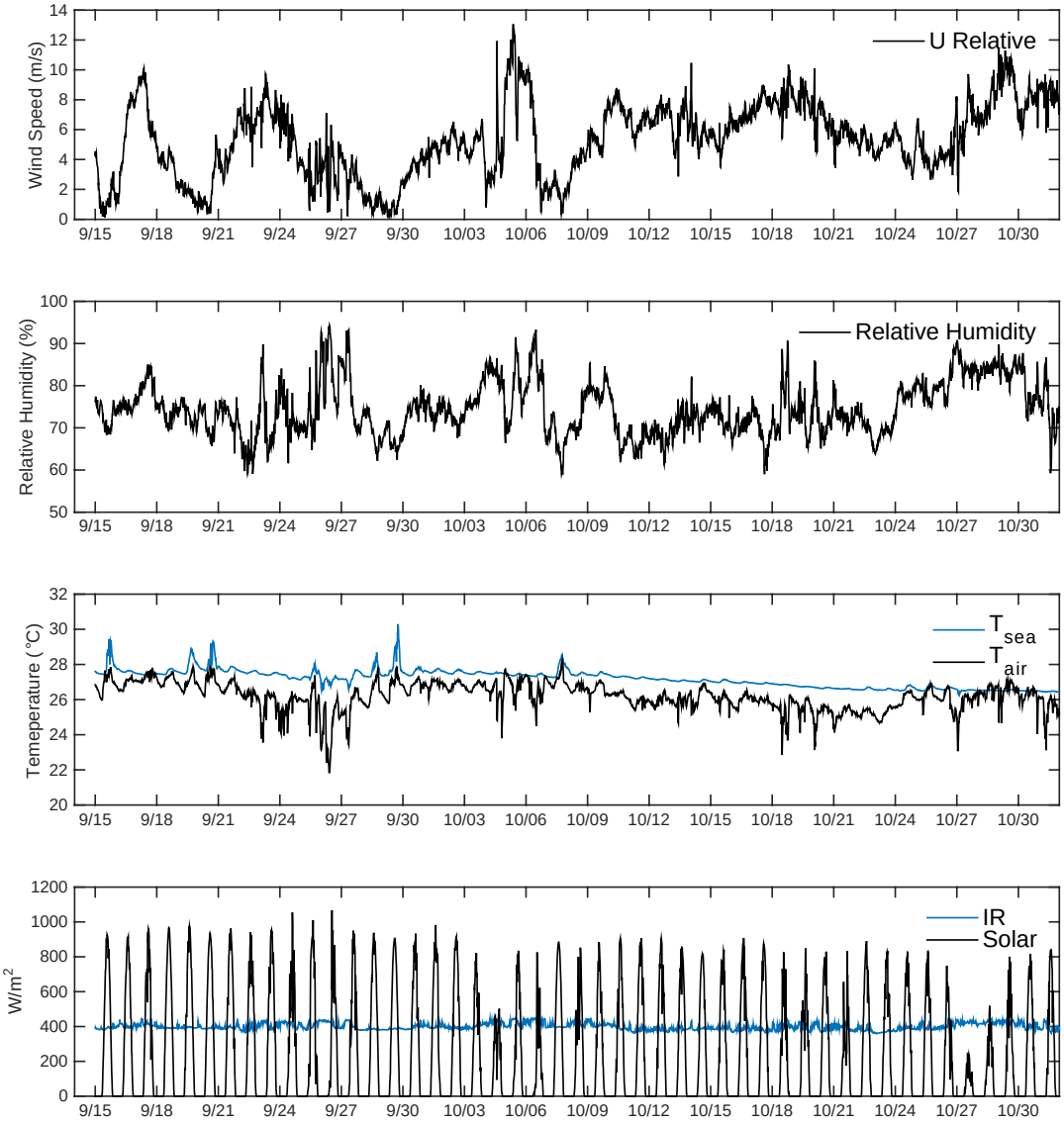

Figure 9: Buoy Time series of meteorological variables during SPURS deployment.

Figure 10: Photograph: Bow of RV Knorr with meteorological mast raised. Inset: DCFS set-up on its side before being raised on bow of R/V Knorr in preparation for the September 2012 deployment cruise.

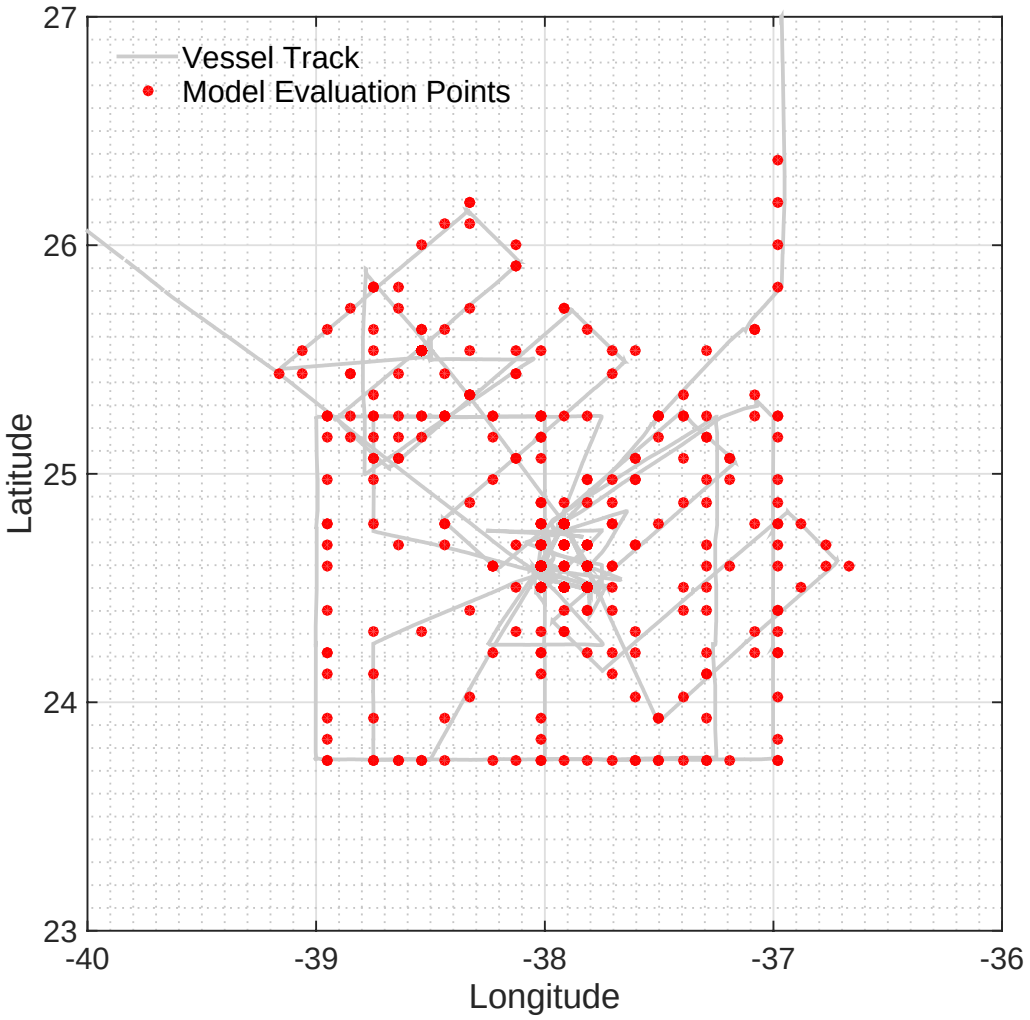

Figure 11: RV Knorr cruise track (grey) and coordinates of model nodes (red) used for the construction of a Knorr following WRF time series for comparison with shipboard DCFS and bulk fluxes.

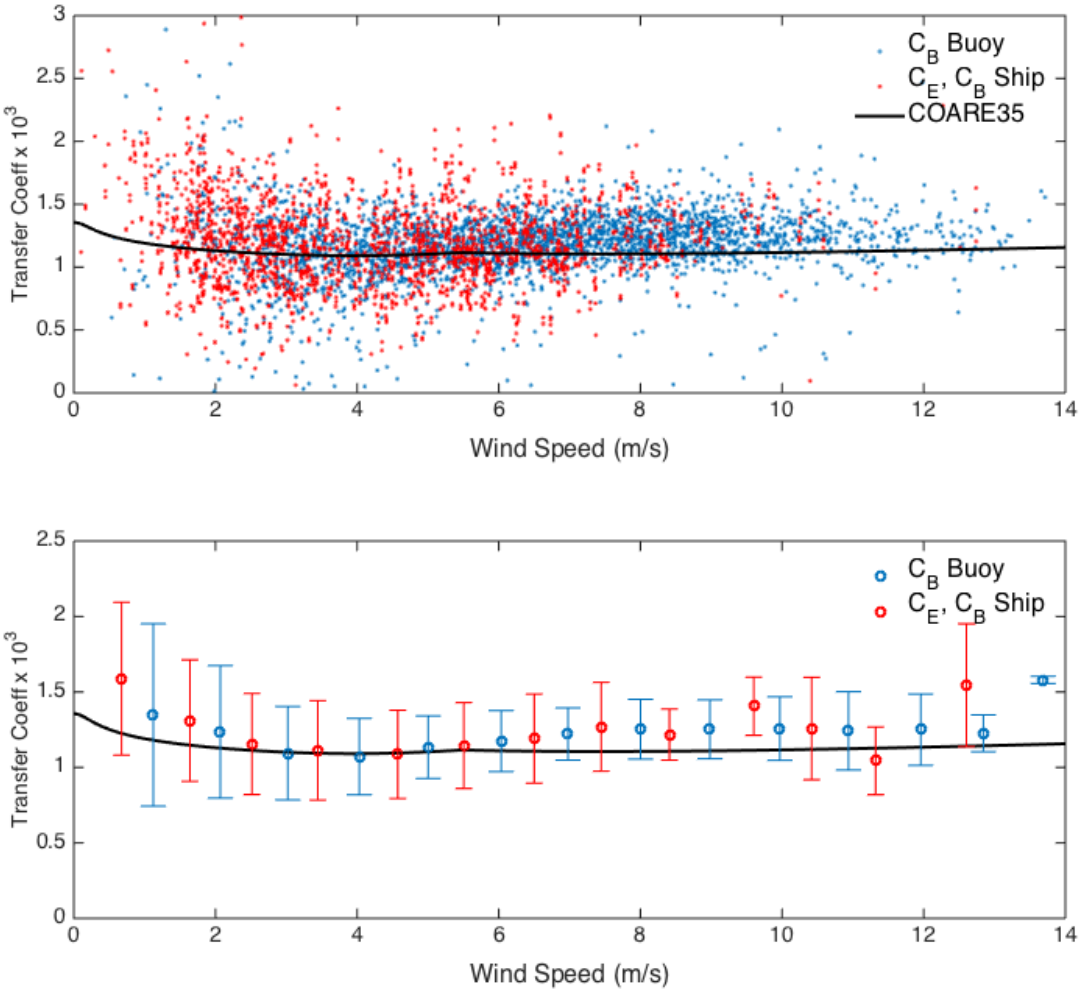

Figure 12: Moisture exchange coefficient as a function of wind speed. Upper frame: raw flux data; lower frame: 0.5 m/s bin averaged data with standard deviation bars.

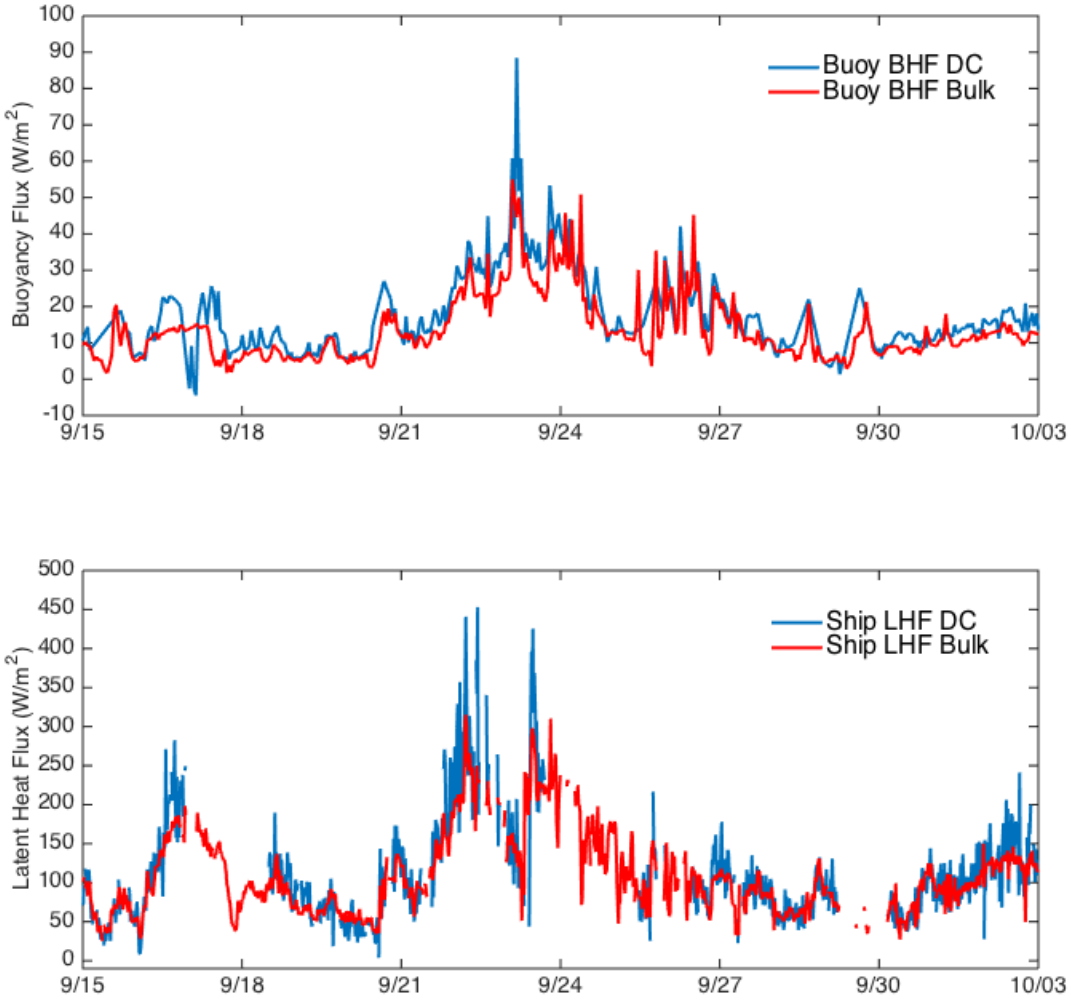

Figure 13: Comparison of R/V Knorr eddy covariance (red) and COARE bulk fluxes (blue).

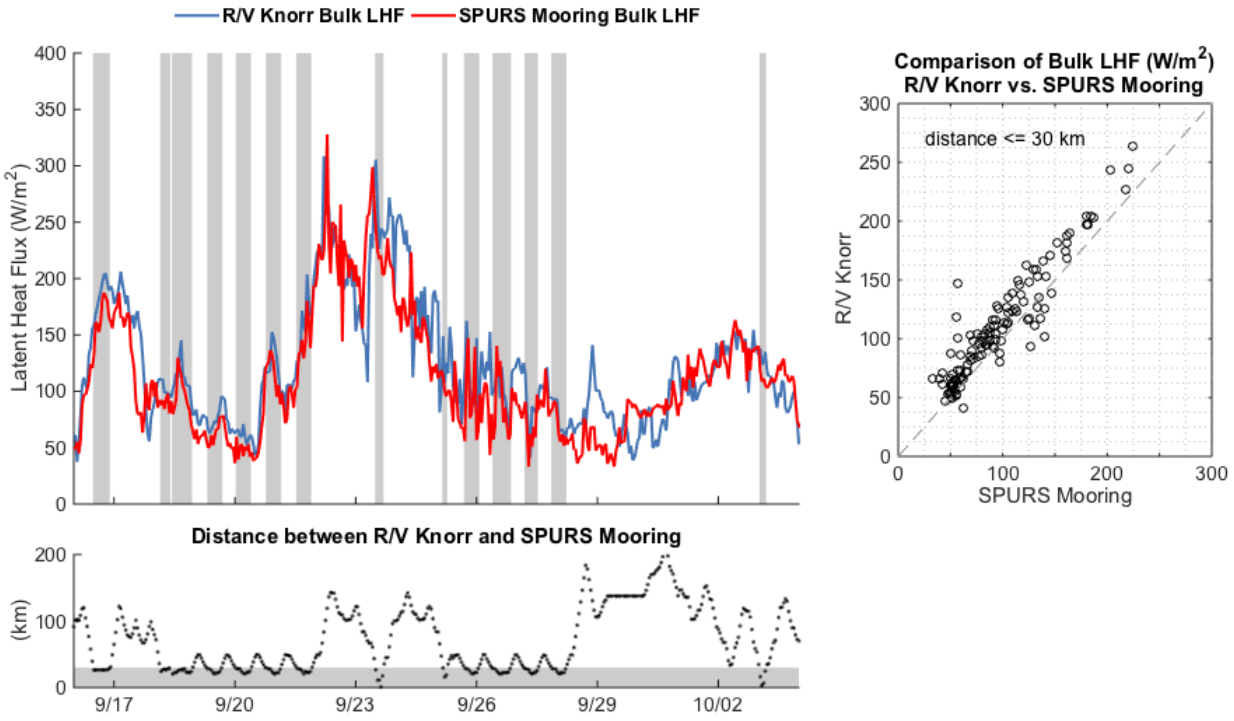

Figure 14: Time series and scatter plot comparison of bulk-formula derived fluxes using mean observations from ship (R/V Knorr) and Buoy (SPURS Mooring). Gray regions indicate data within the 30 km proximity threshold set for comparison between the datasets. The bottom panel show's the full time series of Ship-Mooring distance in kilometers.

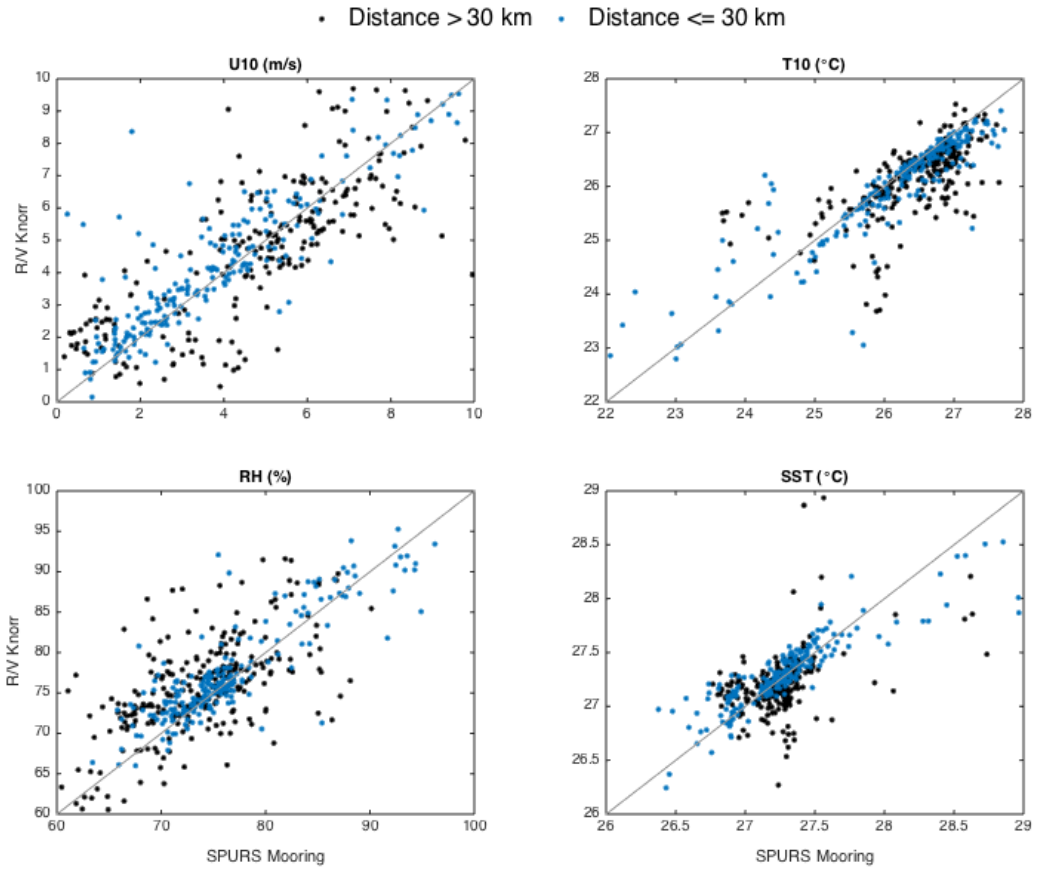

Figure 15: Diagnostic variable comparison between SPURS mooring (x-axis) and the R/V Knorr (y-axis).

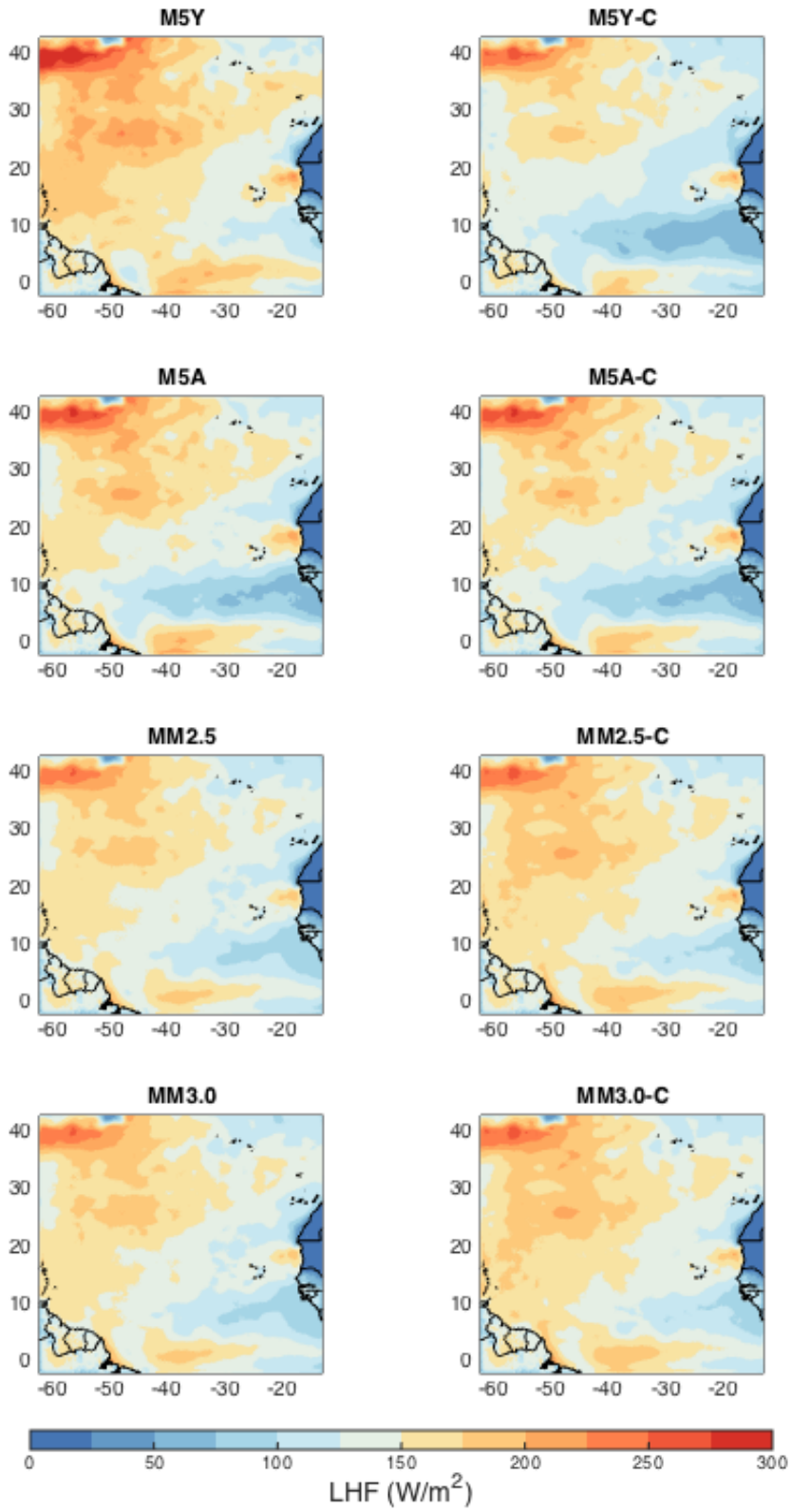

Figure 16: 45-day average of hourly latent heat flux values for all parameterization schemes

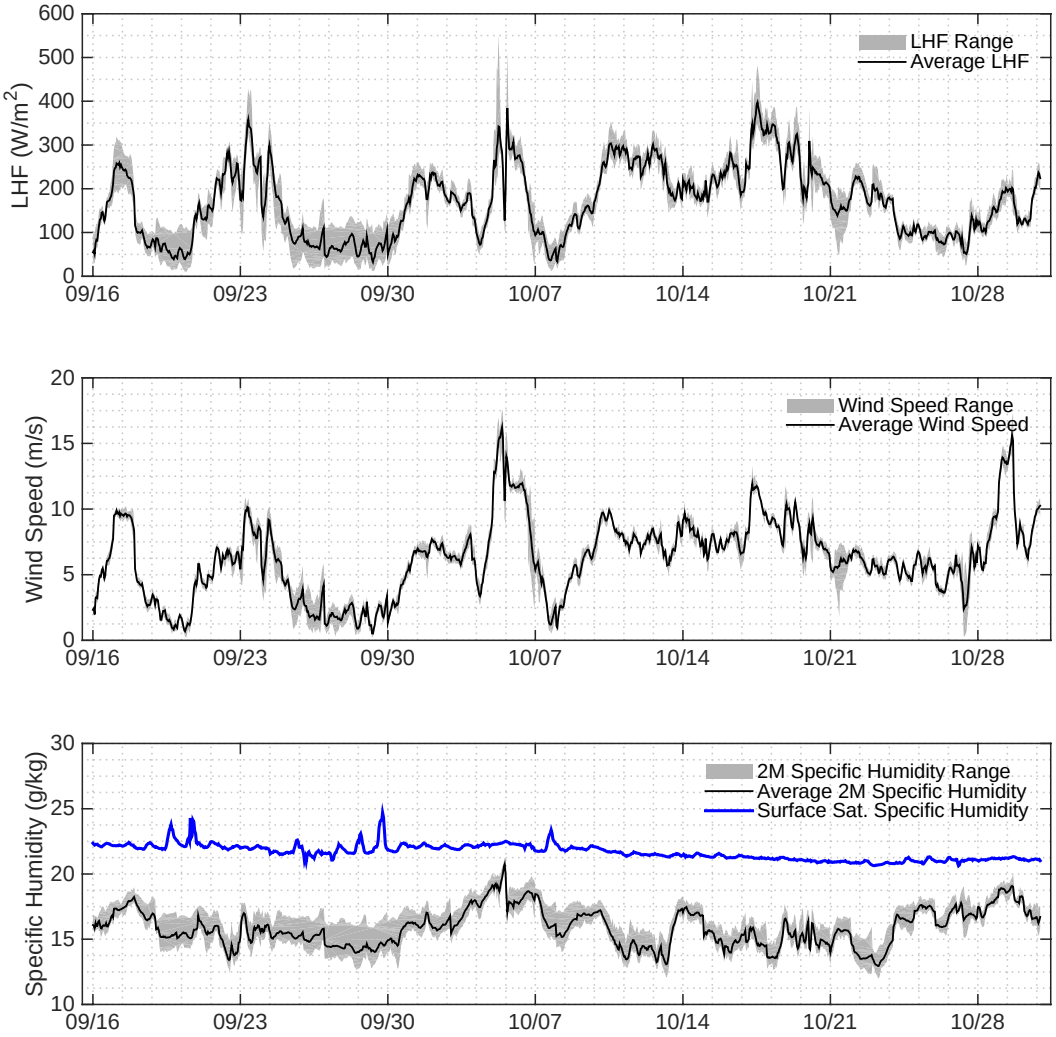

Figure 17: Top: Range of WRF LHF at buoy location in all parameter configuration output (grey) and mean LHF (black). Middle: Range of WRF wind speed at buoy location in all parameter configuration output (grey) and mean wind speed (black). Bottom: Range of 2-Meter specific humidity at buoy location in all parameter configuration output (grey), mean 2-Meter specific humidity (black), and surface saturation specific humidity (blue).

Figure 18: Standard Deviation of all WRF test case model runs as a function of instantaneous wind speed.

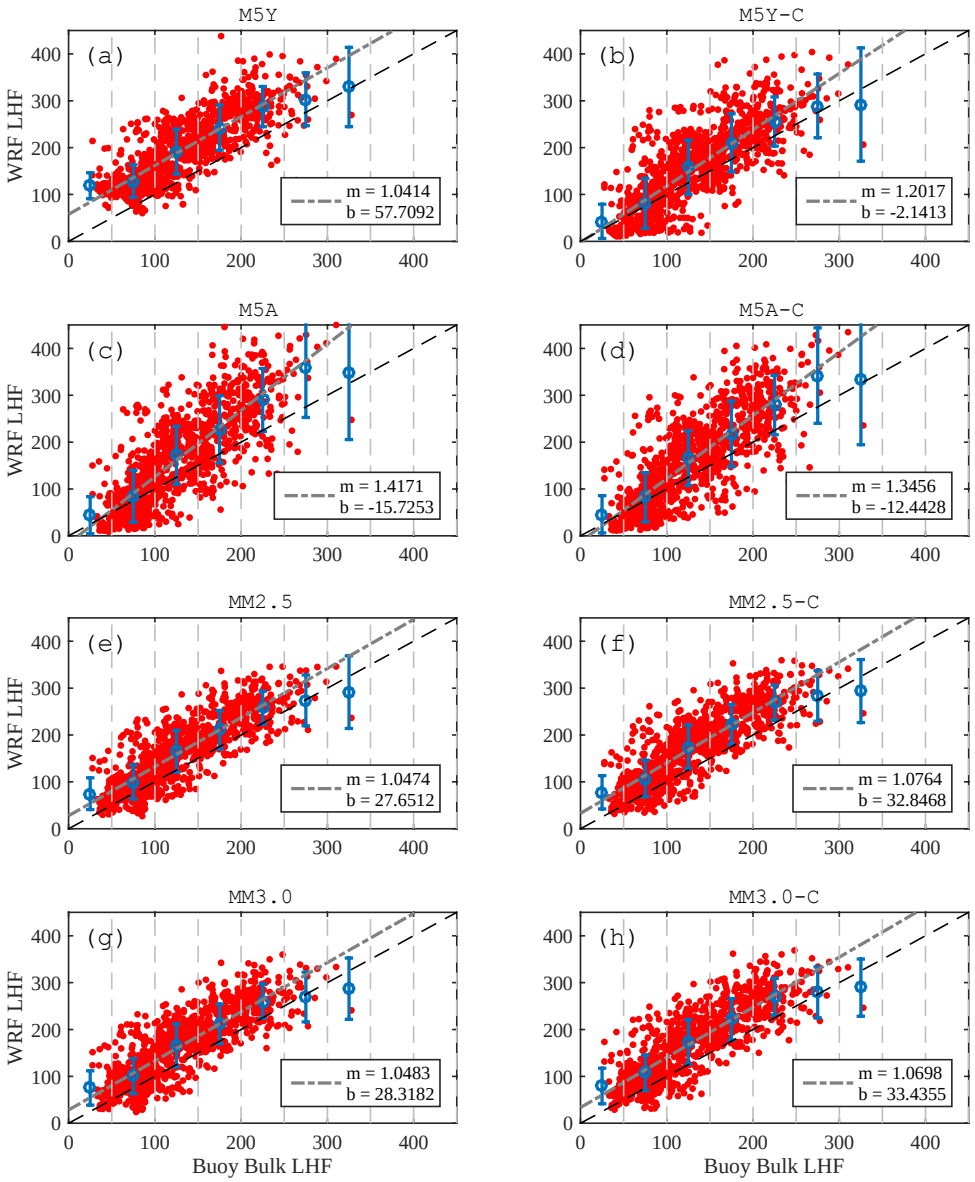

Figure 19: Comparison of WRF and Buoy Bulk LHF. Blue dots indicate 50 Wm^{-2} binned averages, and blue bars represent standard deviation within bins. Ordinary Least Squares regression results are plotted as green dashed line where m = slope and b = y-intercept.

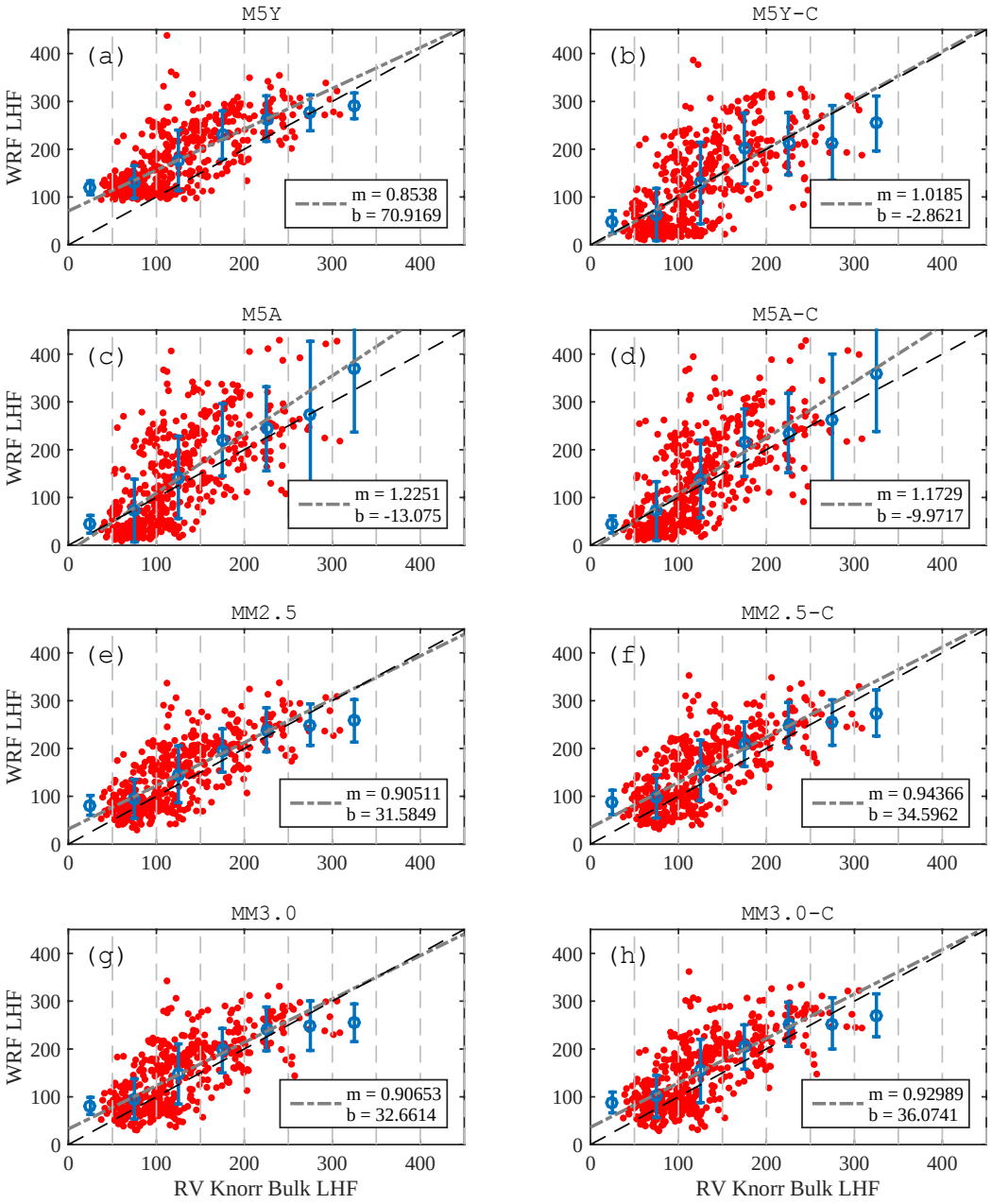

Figure 20: Comparison of Model and R/V Knorr Measured DCFS LHF. Blue dots indicate 50 W/m^2 binned averages, and blue bars represent standard deviation within bins. Ordinary Least Squares regression results are plotted as green dashed line where m = slope and b = y-intercept.

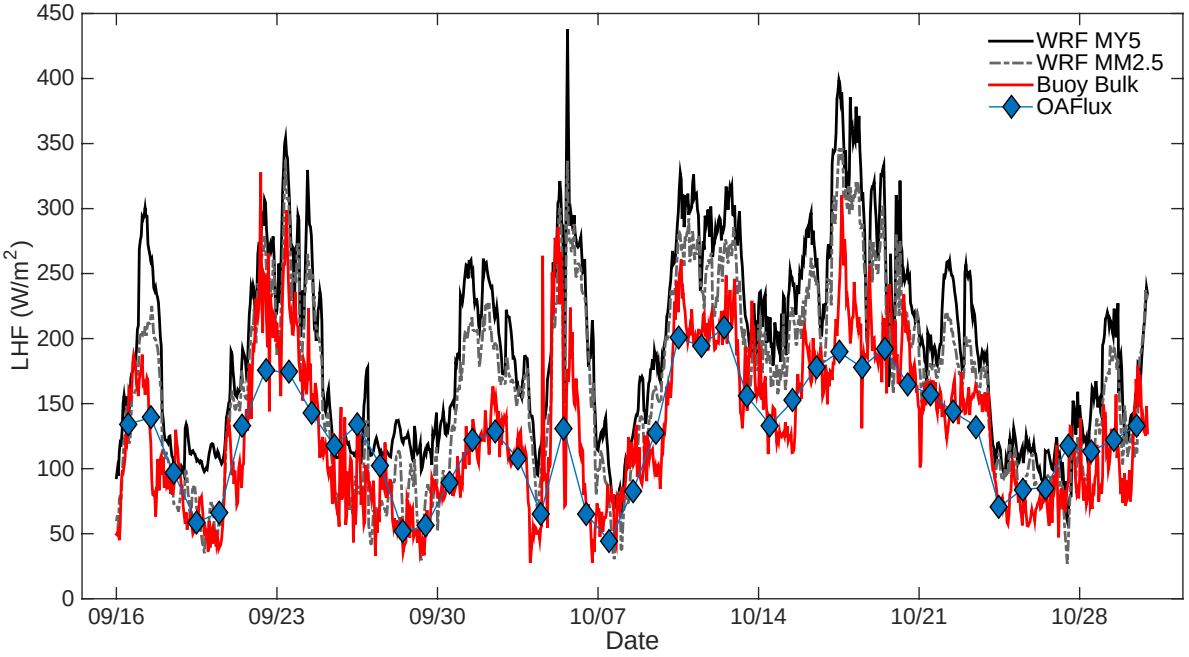

Figure 21: Time series between two closest OAFlux nodes and WRF MM2.5

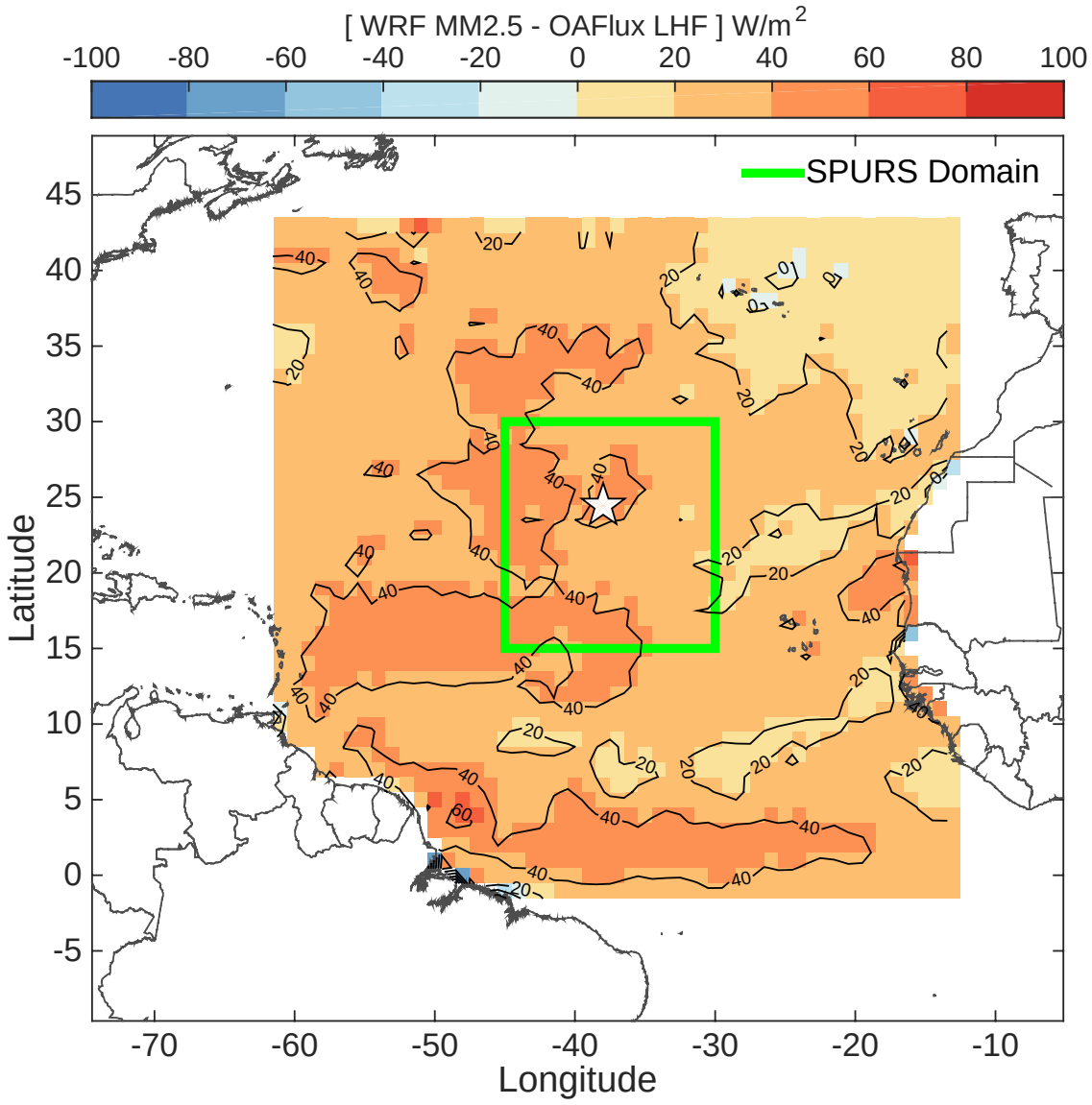

Figure 22: Difference between mean surface fields (MM2.5 – OAFlux) throughout 45-day test case.

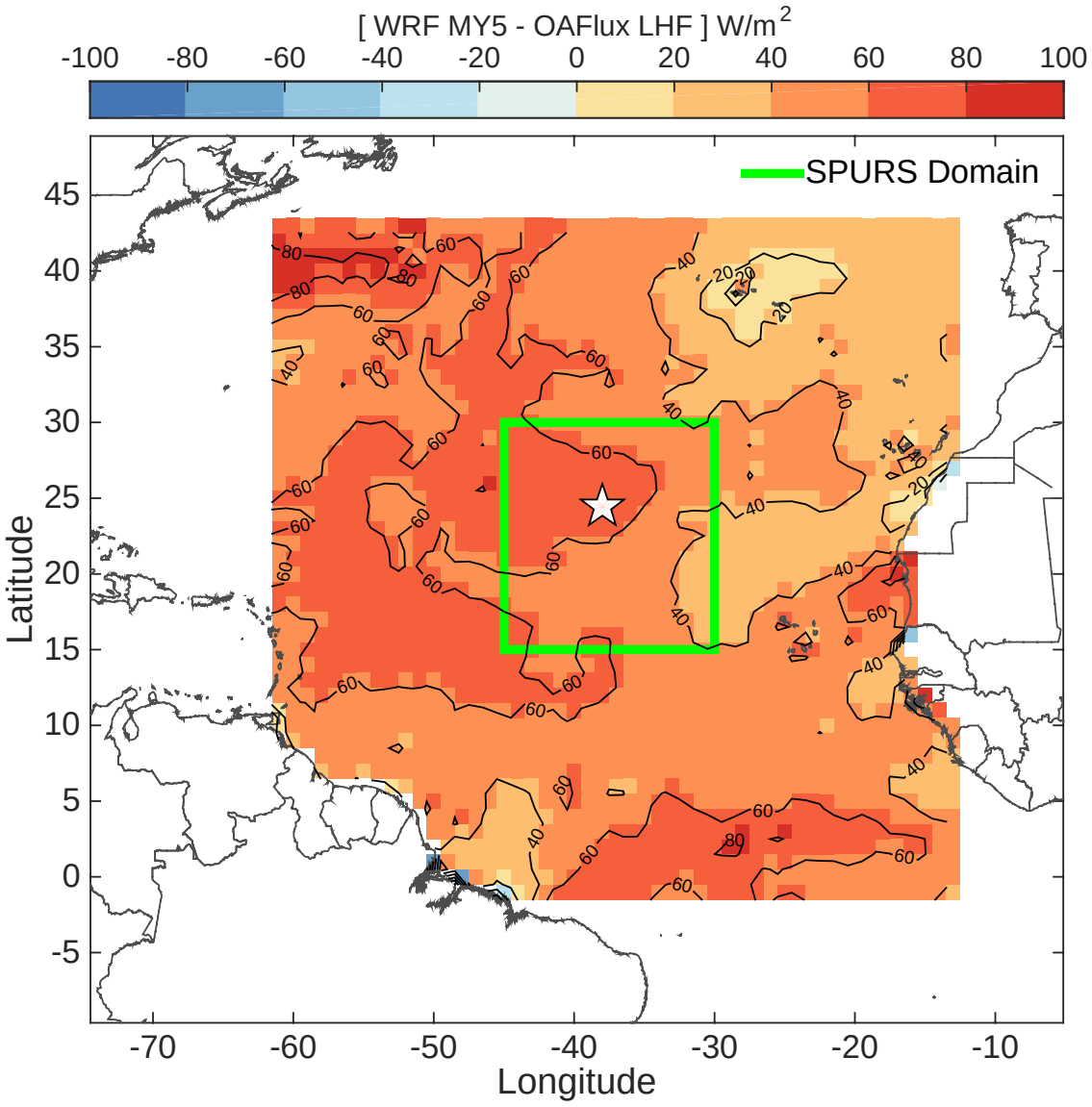

Figure 23: Difference between mean surface fields (MY5 – OAFlux) throughout 45-day test case.

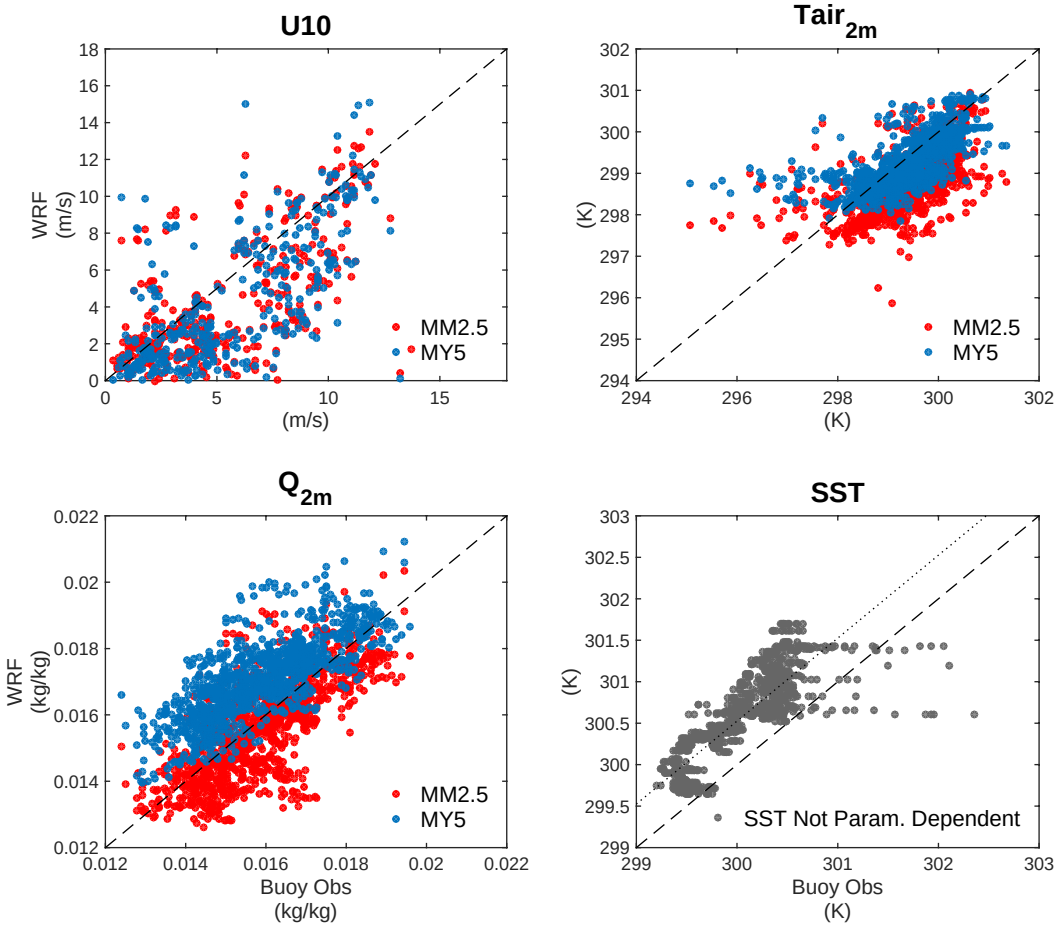

Figure 24: Differences between WRF diagnostic variables and buoy observations for default parameterizations.

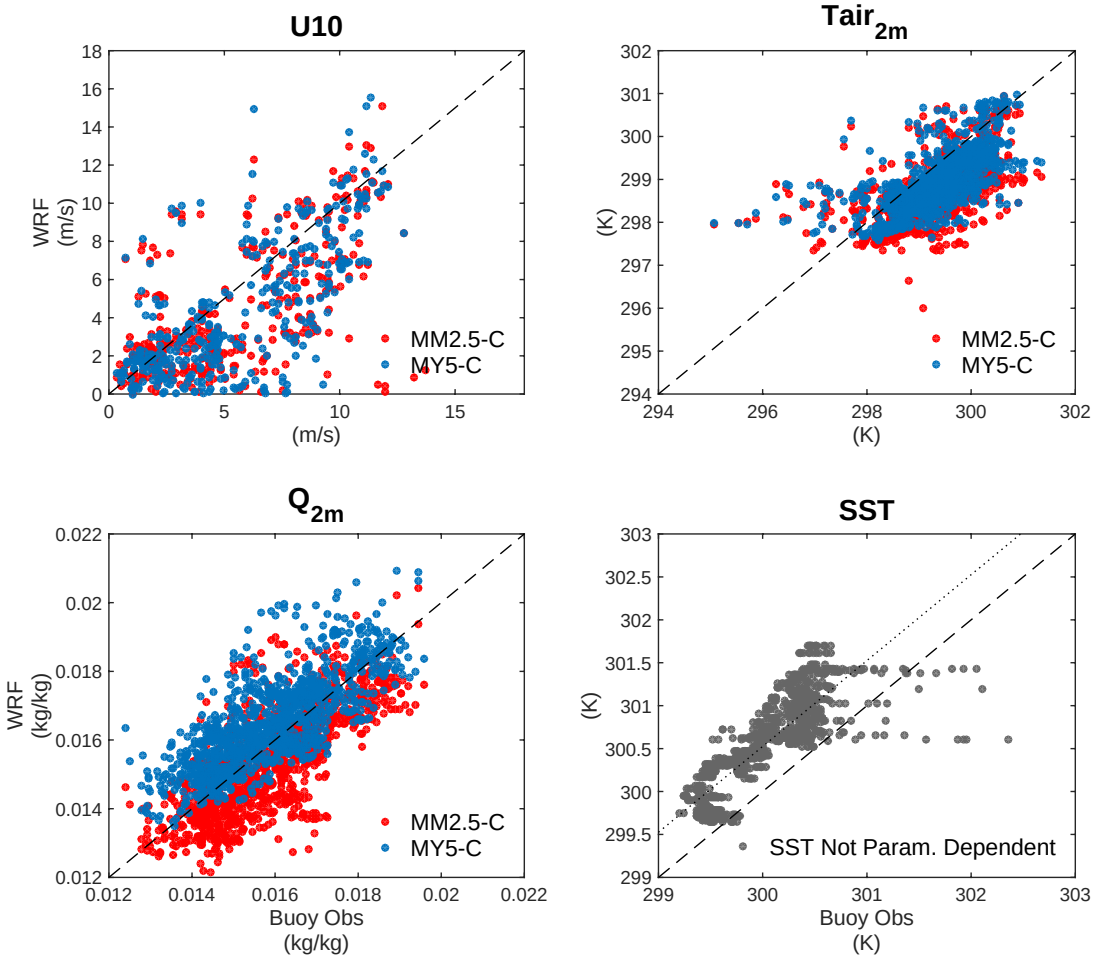

Figure 25: Differences between WRF diagnostic variables and buoy observations for COARE modified parameterizations.

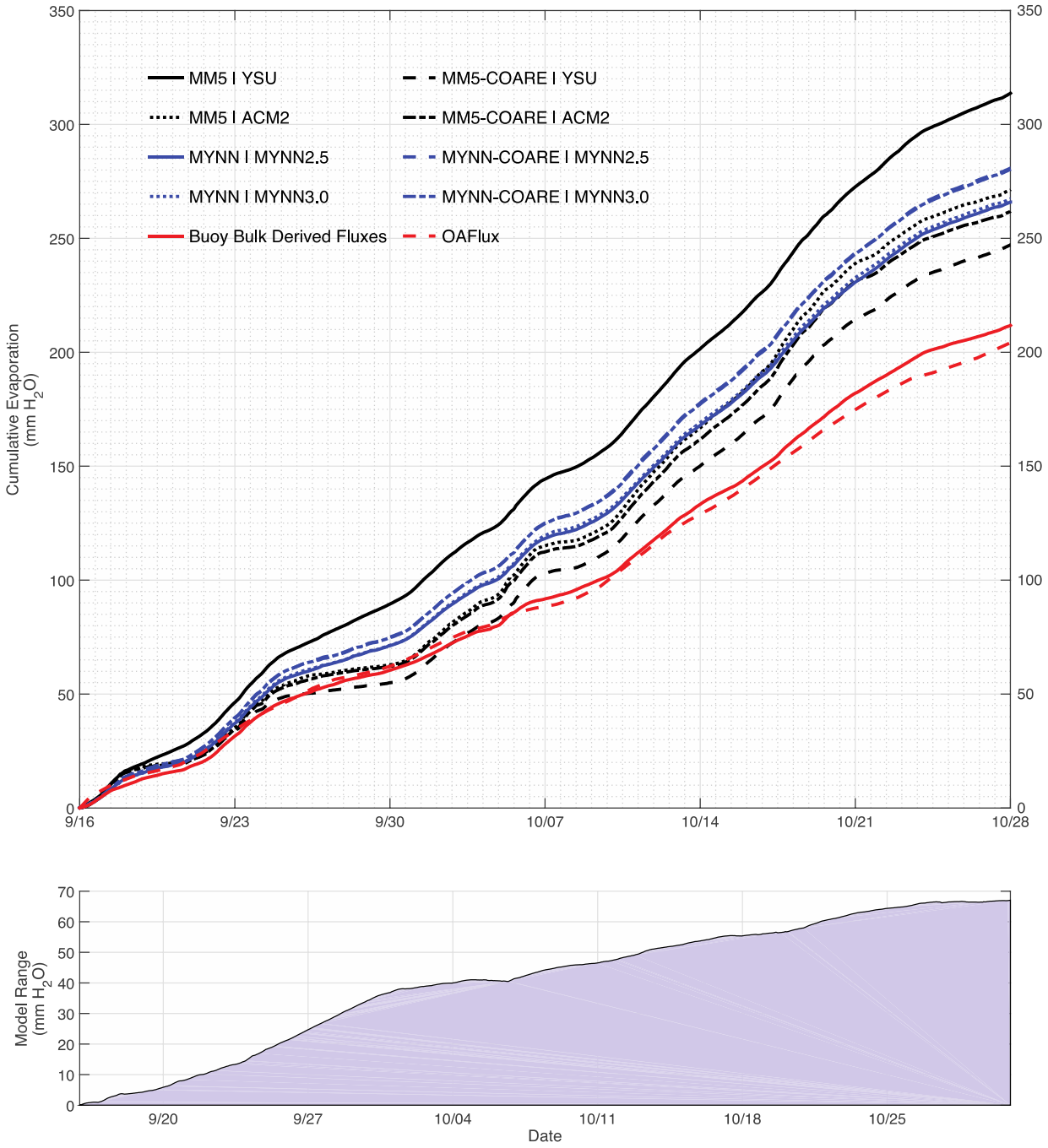

Figure 26: Cumulative Evaporation at the SPURS buoy location. Top: time series accumulation of evaporation in mm H_2O. Bottom: range of accumulated evaporation between WRF test cases.

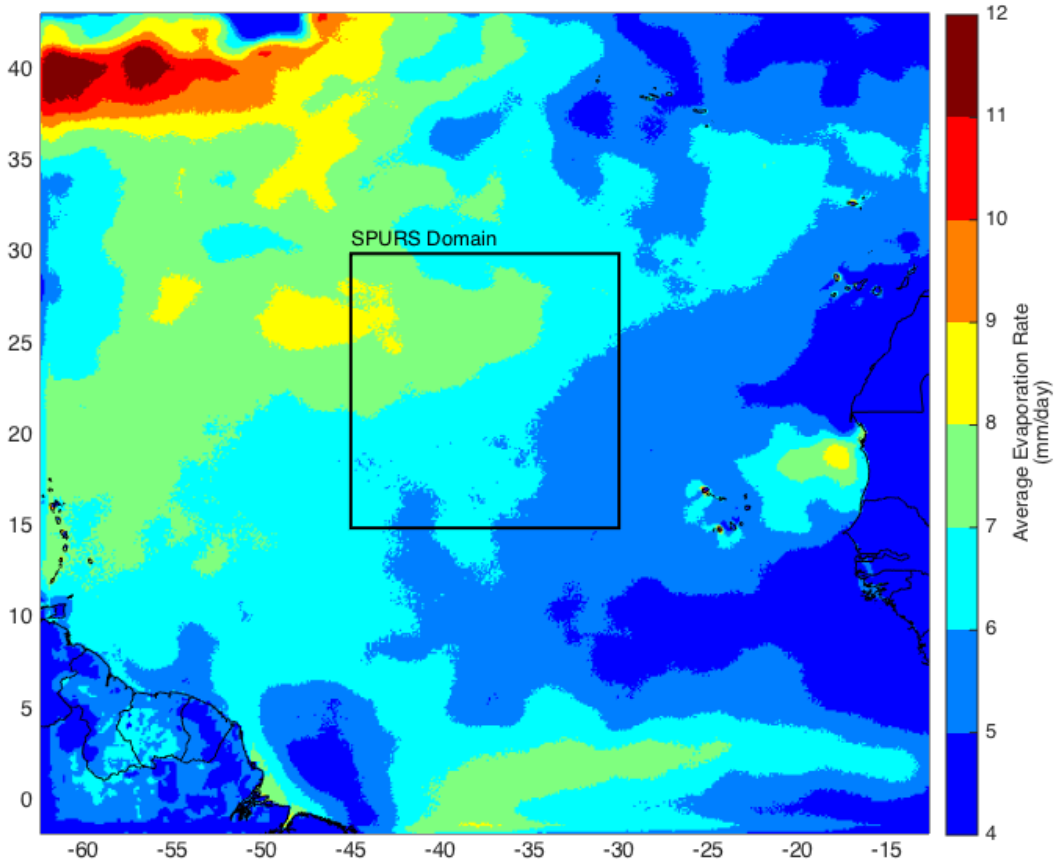

Figure 28: Daily average evaporation rate map: model suite average.

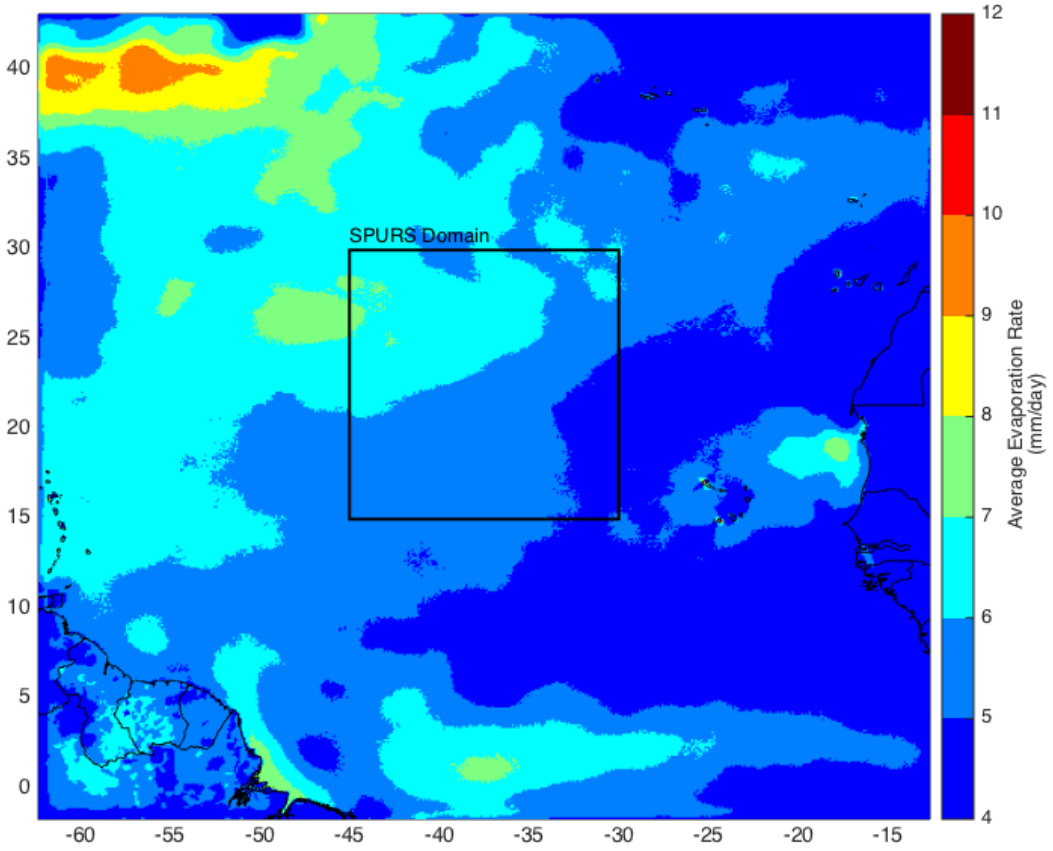

Figure 29: Daily average evaporation rate map: test case MM2.5