Published in final edited form as:
Geophys Res Lett. 2021 October 16; 48(19): . doi:10.1029/2021gl094364.

On Assessing ERA5 and MERRA2 Representations of Cold-Air Outbreaks Across the Gulf Stream

C. Seethala[1], **Paquita Zuidema**[1], **James Edson**[2], **Michael Brunke**[3], **Gao Chen**[4], **Xiang-Yu Li**[5], **David Painemal**[4,6], **Claire Robinson**[4], **Taylor Shingler**[6], **Michael Shook**[6], **Armin Sorooshian**[3,7], **Lee Thornhill**[6], **Florian Tornow**[8,9], **Hailong Wang**[5], **Xubin Zeng**[3], **Luke Ziemba**[6]

[1]Rosenstiel School of Marine and Atmospheric Science, University of Miami, Miami, FL, USA

[2]Woods Hole Oceanographic Institution, Woods Hole, MA, USA

[3]Department of Hydrology and Atmospheric Sciences, University of Arizona, Tucson, AZ, USA

[4]Science Systems and Applications, Inc., Hampton, VA, USA

[5]Pacific Northwest National Laboratory, Richland, WA, USA

[6]NASA Langley Research Center, Hampton, VA, USA

[7]Department of Chemical and Environmental Engineering, University of Arizona, Tucson, AZ, USA

[8]Earth Institute, Columbia University, New York City, NY, USA

[9]NASA Goddard Institute for Space Sciences, New York City, NY, USA

Abstract

The warm Gulf Stream sea surface temperatures strongly impact the evolution of winter clouds behind atmospheric cold fronts. Such cloud evolution remains challenging to model. The Gulf Stream is too wide within the ERA5 and MERRA2 reanalyses, affecting the turbulent surface fluxes. Known problems within the ERA5 boundary layer (too-dry and too-cool with too strong westerlies), ascertained primarily from ACTIVATE 2020 campaign aircraft dropsondes and secondarily from older buoy measurements, reinforce surface flux biases. In contrast, MERRA2 winter surface winds and air-sea temperature/humidity differences are slightly too weak, producing surface fluxes that are too low. Reanalyses boundary layer heights in the strongly forced winter cold-air-outbreak regime are realistic, whereas late-summer quiescent stable boundary layers are too shallow. Nevertheless, the reanalysis biases are small, and reanalyses adequately support their use for initializing higher-resolution cloud process modeling studies of cold-air outbreaks.

Plain Language Summary

Correspondence to: C. Seethala and P. Zuidema, seethala.chellappan@rsmas.miami.edu; pzuidema@miami.edu.

Supporting Information:
Supporting Information may be found in the online version of this article.

The Gulf Stream is a narrow band of warm water to the east of continental North America. As air moves eastward off of the continent, the warm ocean temperatures transfer moisture and heat that help develop and modify the marine low clouds. This transfer, particularly during cold-air outbreaks present significant modeling challenges that contribute uncertainty to temperature projections for a world with more carbon dioxide. Simulations seeking to represent the details of such shallow clouds must rely on initializations and forcings that originate from coarser-resolution reanalyses. Here we explore how well two major reanalyses and a commonly used flux product represent these fluxes and the boundary layer, using ocean buoy measurements and new *in-situ* observations from an aircraft campaign. We find that the reanalyses are adequate for the purpose of initializing higher-resolution modeling of the cold-air outbreak clouds. In particular, the winter boundary layer heights are realistic. These heights are important for capturing winter cloud-environmental interactions correctly. Late summer boundary layers are too shallow. Known biases do remain present, and impact the surface flux errors differently in the two reanalyzes examined.

1. Introduction

A prominent feature of the northwest Atlantic is the Gulf Stream, a western boundary current transporting water to the north that is up to 10°C warmer than its surroundings (Figure 1). During the off-summer months, eastward-moving mid-latitude synoptic disturbances exchange warm, low-latitude air with continentally cooled air. The evolution of the low marine clouds in the post-frontal regions where strong subsidence and large air-sea fluxes prevail is described in Grossman and Betts (1990); Kolstad et al. (2009); Liu et al. (2014); Fletcher et al. (2016a, 2016b); McCoy et al. (2017), and Painemal et al. (2021). Leading questions remain realistic representations of the roll cloud circulations (Honnert et al., 2020; Skyllingstad & Edson, 2009), and the correct partitioning between the liquid and ice phases (Abel et al., 2017; Field & Heymsfield, 2015; Field et al., 2014; Mülmenstadt et al., 2015). The latter contributes to the cloud feedback uncertainty in climate models (Sherwood et al., 2020; Zelinka et al., 2020).

Turbulent fluxes can exceed 1000 W m^{-2} during wintertime cold-air outbreak (CAO) events (Bane & Osgood, 1989; Bigorre et al., 2013; Marshall et al., 2009). The adjustment of the boundary-layer air temperature and humidity to the underlying surface can establish a thermally direct circulation, in which horizontal pressure and boundary-layer height gradients drive a surface wind convergence on the warmer flank of the Gulf Stream (Liu et al., 2014; Minobe et al., 2008; Plagge et al., 2016). The increase in surface winds is aided by a downward transfer of momentum and the shear mediates an adjustment to the altered boundary layer stratification as well (Small et al., 2008).

The northwest Atlantic is a strategic environment for improving the understanding and modeling of CAOs through observationally aided process studies, aided by proximity to the eastern North American seaboard. This is a key objective of the NASA Earth Venture Suborbital-3 Aerosol-Cloud-meTeorology Interactions oVer the western Atlantic Experiment (ACTIVATE; Sorooshian et al., 2019). The process modeling activities rely on reanalysis data for initialization (Li et al., 2021; Tornow et al., 2021), and include subsequent nudging

to above-inversion values (Tornow et al., 2021). The ACTIVATE winter 2020 campaign included two dropsonde circles to explicitly derive vertical velocities following Bony and Stevens (2019), out of concern that reanalysis-derived vertical motion might not be adequate for model forcings. Reanalysis fluxes provide a tempting alternative to derived *in-situ* fluxes requiring long, low-altitude level legs that compete with the gathering of new cloud microphysical information. Reanalyses, in combination with satellite datasets, also provide useful longer-term context (Painemal et al., 2021).

This study addresses the following two questions: (a) How accurate are the surface fluxes from the latest major reanalyses (the fifth-generation ECMWF (ERA5) and Modern-Era Retrospective Analysis for Research and Applications-2 (MERRA2)) and the Objectively Analyzed air-sea Heat Fluxes (OAFLUX) in the presence of wintertime CAOs over the Gulf Stream? This extends prior assessments based on coarser-resolution products (Jin & Yu, 2013; Moore & Renfrew, 2002). (b) Can ERA5 and MERRA2 provide a realistic depiction of the Gulf Stream-affected boundary layer? We rely on buoy and ACTIVATE dropsonde data for reference. The questions are relevant beyond the scope of the ACTIVATE campaign, and recognize the challenges inherent to representing strong air-sea coupling events. Reanalysis products can be an alternative to observations in weather and climate studies and are also applied to climate model assessments - sometimes without fully understanding their performance.

2. Datasets and Method

Buoy measurements including direct covariance buoyancy fluxes from the CLIMODE (Climate Variability and Predictability Mode Water Dynamics Experiment; Marshall et al., 2009) campaign provide absolute reference values. The CLIMODE buoy data were collected at 38°N, 65°W (Figure 1) during November 2005 through February 2007; only the data for 2006 are incorporated into this study. The buoy was to the north of the Gulf Stream in February-March, 2006, and within and south of the Gulf Stream during August–September, 2006 (Figure 1). Twenty-minute averages of temperature (T), relative humidity (RH) and wind speed (WS) were measured at approximately 3 m above the waterline and calibrated, with drifts and biases corrected (Bigorre et al., 2013; Weller et al., 2012). The wind speed was adjusted to the 10 m level using the COARE v3.0 algorithm (Bigorre et al., 2013), but not T and RH. Both a surface-skin SST (*SST_skin*), to which the surface fluxes are responsive, and a sub-surface foundation SST (*SST_foundation*) were measured.

The ACTIVATE campaign sampled on both sides of the northern Gulf Stream SST gradient (Figure 1). During February-March, the Gulf Stream meandered to the north, more noticeable west of 73°W, altering the local SST by more than 4K in places (Figure 1). The more quiescent synoptic conditions in August-September over a more uniform area support an assessment of boundary-layer depictions less influenced by strong air-sea interactions. Dropsondes were launched from the UC-12 King Air flying at approximately 9 km. Of the 13 King Air flight days in February–March, 2020, eight coincided, by choice, with CAOs identified both visually and using a measure of boundary layer instability (Papritz et al., 2015), encompassing 43 of the 59 winter dropsondes (Table S1 lists the individual flight days and their designation as CAO/non-CAO days). Dropsonde circles provide intensive

sampling of two CAOs and are the subject of detailed simulations (Li et al., 2021). The 18 August–September flight days include three (weaker) CAOs and deployed 107 dropsondes total. Neither the buoy nor the dropsonde data are assimilated in either reanalysis or flux product.

The satellite-derived Group for High-Resolution SST (GHRSST), at 9 km spatial resolution, provides spatial context and input to bulk flux calculations. The provided *SST_foundation* at 1m depth is derived from the measured *SST_skin* using a diurnal model. Satellite data do not indicate a diurnal cycle in *SST_skin* within the wintertime northwest Atlantic (Clayson & Edson, 2019), suggesting the GRHSST *SST_foundation* is likely set equal to the *SST_skin* (Donlon et al., 2002).

The ERA5 Reanalysis is a publicly available reanalysis possessing a horizontal grid spacing of 31 km, 137 vertical levels, and an hourly temporal resolution. A systematic bias in the partitioning of ERA5's global wind kinetic energy, with an excessive mean zonal flow coupled with weak meridional flow, is attributed to difficulty in representing high-frequency transient atmospheric events (Rivas & Stoffelen, 2019). The mid-latitude boundary layer maintains a cold and dry bias (Hersbach et al., 2020). The assimilated SST product prior to 2007 was based on the Hadley Centre Sea Ice and Sea Surface Temperature data set (Had-ISST2; Titchner & Rayner, 2014), a 0.25° × 0.25° pentad product too coarse to resolve Gulf Stream SST gradients well (Chelton & Risien, 2016). After 2007 it was based on the Operational Sea Surface Temperature and Sea Ice Analysis (OSTIA, Donlon et al., 2012). OSTIA is produced daily on a higher-resolution 0.05° × 0.05° grid, and includes satellite microwave measurements, which are less affected by non-precipitating clouds than are satellite infrared measurements. Systematic differences between the ERA5 SST still remain from other SST climatologies, attributed to spatial resolution (Hersbach et al., 2020). ERA5 provides both *SST_skin* and *SST_foundation*. The 2020 ERA5 *SST_skin* is typically cooler than the *SST_foundation* by 0.2–0.4K (Figure S2) consistent with infrared cooling (Fairall et al., 1996; Minnett et al., 2019).

MERRA2 spatial resolution is 0.625° × 0.5° longitude by latitude, with 72 vertical levels, and a three-hour temporal resolution, although the surface latent and sensible heat fluxes and lowest-model-level meteorological variables are available at a higher one-hour interval. Previous comparisons to soundings over the Beaufort Sea indicate a MERRA2 warm bias near the surface, and no humidity bias (Rozenhaimer et al., 2018). MERRA2 winds appear too weak, by up to 2 m s^{-1}, in the ACTIVATE region (Carvalho, 2019). MERRA2 only provides *SST_skin*. OAFLUX, at a 1° × 1° spatial resolution, is included because it provides further insight into the impact of spatial resolution.

The COARE v3.0 bulk parameterization values, applied to the buoy *T*, *q*, and WS values, compare well against the CLIMODE buoy direct covariance buoyancy fluxes (Bigorre et al., 2013; Edson et al., 2013) for wind speeds between 4 and 12 m s^{-1}. Modifications for wind speeds > 12 m s^{-1} based on the CLIMODE measurements led to the COARE v3.5 bulk flux algorithm (Edson et al., 2013). Overall this indicates that if the near-surface parameters are known, then the buoyancy fluxes can be estimated with little bias except for too low winds, even within CAOs. The dropsonde temperature and specific humidity

measurements are referenced to the 2 m (T_{2m} and q_{2m}) values and the wind speeds to 10 m (WS_{10m}) using the COARE v3.5 algorithm, from which the dropsonde surface fluxes are also calculated. An *SST_skin* is estimated as GHRSST *SST_foundation* + (ERA5 *SST_skin*–ERA5 *SST_foundation*).

The reanalysis comparisons to the dropsondes are nearest in time and space. Instantaneous values captured by the dropsondes from up- and downdrafts will increase the variability of the dropsonde-calculated fluxes beyond those of the coarser-resolution reanalysis fluxes. The reanalysis comparisons of T_{2m}, RH_{2m}, WS_{10m} to the CLIMODE buoy values of T and RH at 3 m and WS at 10 m are of the daily and monthly mean values, ignoring diurnal variations in wind speed (Dai & Deser, 1997) and near-surface humidity (Clayson & Edson, 2019). Further details can be found in the Supplement.

3. Surface Flux Representations

3.1. 2006 CLIMODE

The buoy was located within an SST gradient of approximately 8K over a mere 100 km in January–April of 2006 (Figure S1). This constitutes a challenging regime for any reanalysis. The monthly mean SSTs exceed buoy values by up to 5K during February–April, indicating reanalyses depictions of the Gulf Stream that are broader than in nature (Figure 2), more noticeable as spatial-resolution degrades. Consistent with this, the reanalyses T_{2m} are too warm, and the saturated specific humidity (q_s) too high. The winds are too strong, which physically can be related to the too-warm ocean surface (Small et al., 2008). Monthly mean reanalyses and OAFLUX buoyancy fluxes differ significantly from the CLIMODE direct covariance values (Figure S3), most notably in February–March, when buoyancy fluxes exceed the buoy values by 40–60 W m^{-2} (see also Table S2), an overestimate of >80%. Since the bulk flux calculations are validated at the wind speeds dominating the observations (Edson et al., 2013), most of the reanalyses flux biases must reflect their SST representation. This conclusion is in line with Jin and Yu (2013), extended here to newer reanalyses possessing higher spatial resolutions.

Daily mean differences between the reanalyses/product and CLIMODE buoy values (Figure 2) clarify the atmospheric consequences of the SST misrepresentations. The overestimated SST skews reanalysis wind speeds to positive values (Figure 2f). The ERA5 *SST_foundation* and T_{2m} deviate the least from the buoy values (Figures 2a and 2b). Both reanalyses match the buoy wind speeds well (Figure 2f). The surface q_s is elevated for all reanalyses/product, as expected. A dry bias in ERA5's q_{2m} contrasts with a moist bias for MERRA2's q_{2m}, both by about 1 g kg^{-1}. In combination, the air-sea thermodynamic differences are smaller for MERRA2 on most days, compared to ERA5 (Figures 2g and 2h), compensating for MERRA2's poorer T_{2m} (Figure 2a). This allows the MERRA2 buoyancy fluxes to ultimately compare better to the CLIMODE values, than the ERA5 values (Figure 2e). Overall, this comparison suggests ERA5 provides a more accurate depiction of the Gulf Stream near-surface meteorology, likely in part because of an improved resolution, but compensations within MERRA2 model physics may be improving the fluxes. OAFLUX, with the coarsest resolution, has flux, SST and wind speed values that diverge the most of the three products from CLIMODE buoy values (Figures 2b, 2e and 2f). Further support for

these conclusions is available in the supplement: Table S2 provides numerical daily mean bias estimates, Figure S3 indicates the annual cycle in monthly mean buoy and reanalysis values, and Figure S4 shows the large range of the absolute daily mean values.

During the summer months, the SST differences are smaller and more evenly distributed about zero (Figure 2j). Buoyancy fluxes remain consistently overestimated (Figure 2m), most noticeable by ERA5 because of its too-cool T_{2m} and too-dry (by 1–2 g kg^{-1}) q_{2m}. The too-dry ERA5 q_{2m} bias for both seasons indicates a common bias source. In contrast, a too-warm T_{2m} in winter and too-cool T_{2m} in summer suggests differing underlying causes.

3.2. 2020 ACTIVATE

ERA5 *SST_foundation* exceed GHRSST values by up to 2K during February–March of 2020 at the Gulf Stream boundaries, while slightly underestimating the cooler southward-flowing coastal Labrador Current temperatures (Figure 1d). The ERA5 SST bias exists even though its assimilated SST product has a slightly finer spatial resolution than GHRSST (0.05° vs. 9 km), reflecting the coarsening needed to match the ERA5 resolution of 31 km (Hersbach et al., 2020).

The hourly mean ERA5 reanalysis sensible and latent heat fluxes overestimate during the more severe CAOs, by up to 100 W m^{-2} (Figure 3a) or more for the latent heat fluxes (Figure 3b). The cause is linked to both SST and wind speed overestimates. In contrast, the MERRA2 fluxes always underestimate, because of underestimates in the air-sea temperature (Figure 3c) and humidity (Figure 3d) differences, and weaker MERRA2 wind speeds (Figure 3e). Differences in temporal/spatial resolution (instantaneous vs. hourly mean values over a larger spatial domain) seem unlikely to explain the underestimates because these are systematic biases, and instead point to a near-surface boundary layer that is too close in thermodynamic equilibrium with the ocean.

In August–September 2020, both reanalyzes slightly underestimate the fluxes relative to those calculated from the dropsondes, with ERA5 performing better than MERRA2 (Table S3). Air-sea humidity differences are more realistically captured by ERA5 than by MERRA2. MERRA2 consistently underestimates all inputs into the flux calculations, although the biases are small (Table S3).

4. Thermodynamic Vertical Structure

The reanalyses capture the main features of the lower tropospheric structure for both February–March and August–September of 2020 (Figure 4, with Figure S5 indicating the differences as histograms at different altitudes). Consistent with the near-surface analysis, the wintertime ERA5 boundary layer is slightly too cold and too dry, with RH and q averages indicating underestimates of 5% (ranging up to 30%) and 0.5 g kg^{-1} (ranging up to 2 g kg^{-1}), respectively. Locations with ERA5-SST–GHRSST > 1K reveal mean ERA5 boundary layer potential temperature (θ) profiles that are 0.2K warmer, ranging up to 1–2K. Where ERA5-SST–GHRSST < -1K, the ERA5 θ profiles are almost 1K cooler than the dropsonde values (inset plot in Figure 4a).

The mean ERA5 wind biases are small, with a slight overestimation (1 m s^{-1}, or 10%) that primarily comes from the zonal component. The lower free troposphere in ERA5 does not resolve the elevated moisture layer between 800 and750 hPa, but the main inversion top at approximately 850 hPa, identified using an RH threshold, is adequately captured. The winter MERRA2 thermodynamic structure compares more closely to the *in-situ* values. Interestingly, the sign of the MERRA2 wind bias contrasts with that from ERA5. An underestimate of the near-surface zonal winds is consistent with an underestimate at higher altitudes.

The ERA5 wind biases are smaller in August-September than February-March, while ERA5 θ remains depressed by 0.2–0.3 K (ranging up to 3K) and q is biased low by up to 0.5 g kg^{-1}. These compensate to generate realistic RH values near the surface. Specific humidity underestimates above the surface-based mixed layer are slightly larger for MERRA2 than ERA5, permeating into the relative humidity. Both ERA5 and MERRA2 struggle with capturing the cloud layer between 900 and 800 hPa (Figure 4h inset). This is not linked to a pronounced bias in the winds for ERA5, while MERRA2 winds are clearly too weak above 1 km.

In contrast to ERA5, the RH and q-mean MERRA2 profiles agree well with the dropsondes during February-March, while the zonal winds are consistently weaker, by 1–2 m s^{-1} near the surface, increasing (mostly) with altitude. During the late summer, MERRA2 is more likely to be drier within 0.4–2.0 km than the *in-situ* measurements, consistent with a known underestimate of low cloud cover (Miao et al., 2019).

During February–March, both ERA5 and MERRA2 capture the inversion height of approximately 1.7 km reasonably well (estimated from the RH profiles). During August–September, the inversion is naturally lower, at approximately 1.1 km (similarly estimated). Both reanalyses often fail to capture the cloud layer in late summer. The mean MERRA2 boundary layer height is lower than that from ERA5, with a drier cloud layer. The lower boundary layer height is even more pronounced after the few September CAO cases are excluded (not shown), indicating the issue may be a similar difficulty in representing stable boundary layers as for ERA5. The MERRA2 winds, both zonal and meridional, are also weaker.

5. Discussions and Conclusions

The too-dry ERA5 boundary layer, evident here in both seasons, coincides with a warming of the lower troposphere in Hersbach et al. (2020). The assimilation of microwave radiances are shown to warm and dry ERA5 at 850 hPa over the ocean (Geer et al., 2017), through unclear mechanisms (Hersbach et al., 2020); the additional information can nevertheless not fully constrain the thermodynamic profile (e.g., Pincus et al., 2017; Zhang et al., 2018), introducing trade-offs.

The reasonable depiction of the ERA5 wintertime boundary layer depth, whereas the late summer cloudy boundary layer is too shallow, may reflect a choice to artificially enhance the turbulent diffusion towards improving the depiction of synoptic cyclones (Sandu et al.,

2013). The ACTIVATE campaign selectively sampled CAO conditions during its winter campaigns, for which the ERA5 turbulent diffusion choices may be more optimal. In addition, the too-strong wintertime westerlies in ERA5 will generate an anomalous wind convergence near the surface (Rivas & Stoffelen, 2019), which should act to raise the boundary layer height, all else being equal. In contrast, both reanalyses have more difficulty in maintaining a cloudy stable layer during the more quiescent late summer time period. The artificial enhancement in the ERA5 diffusion parameters was also intended to improve a near-surface cold temperature bias (by encouraging the entrainment of warmer air aloft); we find a small (~0.2K) ERA5 cold temperature bias still remains during both winter and late summer.

The length scale of ocean mesoscale eddies is 20–30 km at the latitude of the Gulf Stream, set to first-order by the Rossby radius. The ERA5 horizontal grid spacing is best able to represent the majority of the ocean mesoscale activity of the three products examined, though still missing the smallest eddies. A wintertime western boundary current that is too wide in the reanalyses could imply that the boundary layer adjustment for air coming from the west might be affected earlier within the reanalysis than in nature. In addition CAO air flows first over cooler coastal waters north of 35°N generated by the Labrador Current, whose ERA5 reanalysis temperatures are too cool, potentially further energizing the adjustment process of the boundary layer to the warmer Gulf Stream waters. Furthermore, all of the ERA5 biases contribute to exaggerating its surface heat fluxes during CAOs, which will also contribute to elevating the inversion. Perhaps because of these characteristics, the wintertime ERA5 boundary layer depth, thermodynamic and dynamic structure is broadly representative of the observations. ERA5 vertical motion fields have also been shown to compare well to those derived from the two dropsonde circles (Li et al., 2021), lending further confidence in the ability of ERA5 to depict the strongly forced cold-air outbreak regime. Interestingly, the biases in the MERRA2 reanalysis tend to too-weak surface fluxes, in contrast ERA5. Nevertheless the boundary layer depth depiction is similar: approximately realistic during the winter, and too shallow during the summer. A correct boundary layer depth is a critical parameter for shallow clouds, as the depth affects the coupling to the ocean surface. The momentum transport to the surface by ERA5 can increase the surface wind speed, deepening the boundary layer more quickly and encouraging a faster cloud transition, than in nature (Saggiorato et al., 2020).

The robustness of the bias in ERA5 q and T suggests an observationally determined correction factor could be applied and improved as more dropsonde data become available. Additional dropsonde data will also support analysis of how much reanalysis profiles deviate from observations as a function of distance from shore. Future work will incorporate space-based lidar and radar data and *in-situ* measurements to evaluate the CAO cloud structure evolution over the Gulf Stream.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by NASA grant 80NSSC19K0390 to ACTIVATE, a NASA Earth Venture Suborbital-3 (EVS-3) investigation funded by NASA's Earth Science Division and managed through the Earth System Science Pathfinder Program Office. The Pacific Northwest National Laboratory (PNNL) is operated for the US Department of Energy (DOE) by Battelle Memorial Institute under Contract DE-AC06-76RLO 1830.

Data Availability Statement

ACTIVATE Dropsonde Data are available at https://doi.org/10.5067/ASDC/ACTIVATE_MetNav_Aircraft-InSitu_KingAir_Data_1. Microwave OI SST data are produced by Remote Sensing Systems, available at https://remss.com/measurements/sea-surface-temperature/. ERA5 data are available to download for the registered users at Copernicus Data Store: hourly data on pressure levels are available at https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-pressure-levels?tab=overview and hourly data on single level from 1979 to present are available at https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels?tab=overview. MERRA-2 data are available via Global Modeling and Assimilation Office (GMAO) (2015), version 5.12.4, Greenbelt, MD, USA: Goddard Space Flight Center Distributed Active Archive Center (GSFC DAAC); MERRA2 3D Instantaneous 3-hourly (p-coord, 0.625 × 0.5L42) Meteorology (inst3_3d_asm_Np) data can be accessed at doi:https://doi.org/10.5067/QBZ6MG944HW0, 2D Instantaneous 1-hourly single-level (0.625 × 0.5) Meteorology (inst1_2d_asm_Nx) data can be accessed at doi:https://doi.org/10.5067/3Z173KIE2TPD, and 1-hourly Time-averaged (0.625 × 0.5) Flux (tavg1_2d_flx_Nx) data can be accessed at doi:https://doi.org/10.5067/7MCPBJ41Y0K6. CLIMODE buoy measurements are available to download at http://www.opal.sr.unh.edu/data/airsea_flux.shtml. The OAFlux data are available through https://oaflux.whoi.edu/data-access/. We thank Ludovic Bariteau for his Python implementation of the COARE 3.5 flux algorithm, made publicly available through http://doi.org/10.5281/zenodo.5110991.

References

Abel SJ, Boutle IA, Waite K, Fox S, Brown PRA, Cotton R, et al. (2017). The Role of precipitation in controlling the transition from stratocumulus to cumulus clouds in a Northern Hemisphere Cold-Air Outbreak. Journal of the Atmospheric Sciences, 74(7), 2293–2314. 10.1175/jas-d-16-0362.1

Bane JM Jr, & Osgood KE (1989). Wintertime air-sea interaction processes across the Gulf Stream. Journal of Geophysical Research, 94(C8), 10755–10772. 10.1029/JC094iC08p10755

Bigorre S, Weller RA, Lord J, Edson JB, & Ware JD (2013). A surface mooring for air–sea interaction research in the Gulf Stream. Part 2: Analysis of the observations and their accuracies. Journal of Atmosheric and Oceanic Technology, 30, 450–469. 10.1175/JTECH-D-12-00078.1

Bony S, & Stevens B (2019). Measuring area-averaged vertical motions with dropsondes. Journal of the Atmospheric Sciences, 76(3), 767–783. 10.1175/jas-d-18-0141.1

Carvalho D (2019). An assessment of NASA's GMAO MERRA-2 reanalysis surface winds. Journal of Climate, 32, 8261–8281. 10.1175/JCLI-D-19-0199.1

Chelton DB, & Risien CM (2016). *Zonal and meridional discontinuities and other issues with the HadISST1.1 dataset:* Oregon State University. Technical Report. 10.13140/RG.2.1.4503.0168

Clayson CA, & Edson JB (2019). Diurnal surface flux variability over western boundary currents. Geophysical Research Letters, 46, 9174–9182. 10.1029/2019GL082826

Dai A, & Deser C (1997). Diurnal and semidiurnal variations in global surface wind and divergence fields. Journal of Geophysical Research, 104, 31109–31125. 10.1029/1999JD900927

Donlon CJ, Martin M, Stark J, Roberts-Jones J, Fiedler E, & Wimmer W (2012). The Operational Sea Surface Temperature and Sea Ice Analysis (OSTIA) system. Remote Sensing of Environment, 116, 140–158. 10.1016/j.rse.2010.10.017

Donlon CJ, Minnett PJ, Gentemann C, Nightingale TJ, Barton IJ, Ward B, & Murray J (2002). Toward improved validation of satellite sea surface skin temperature measurements for climate research. Journal of Climate, 15, 353–369. 10.1175/1520-0442(2002)015<0353:tivoss>2.0.co;2

Edson JB, Jampana V, Weller RA, Bigorre SP, Plueddemann AJ, Fairall CW, et al. (2013). On the exchange of momentum over the open ocean. Journal of Physical Oceanography, 43, 1589–1610. 10.1175/JPO-D-12-0173.110.1175/jpo-d-12-0173.1

Fairall CW, Bradley EF, Godfrey JS, Edson JB, Young GS, & Wick GA (1996). Cool skin and warm layer effects on the sea surface temperature. Journal of Geophysical Research, 101, 1295–1308. 10.1029/95jc03190

Field PR, Cotton RJ, McBeath K, Lock AP, Webster S, & Allan RP (2014). Improving a convection-permitting model simulation of a cold air outbreak. Quarterly Journal of the Royal Meteorological Society, 140, 124–138. 10.1002/qj.2116

Field PR, & Heymsfield AJ (2015). Importance of snow to global precipitation. Geophysical Research Letters, 42, 9512–9520. 10.1002/2015GL065497

Fletcher JK, Mason SL, & Jakob C (2016). The climatology, meteorology, and boundary layer structure of marine cold air outbreaks in both hemispheres. Journal of Climate, 29, 1999–2014. 10.1175/JCLI-D-15-0268.1

Fletcher JK, Mason SL, & Jakob C (2016). A climatology of clouds marine cold air outbreaks in both hemispheres. Journal of Climate, 29, 6677–6692. 10.1175/JCLI-D-15-0783.1

Geer AJ, Baordo F, Bormann N, English SJ, Kazumori M, Lawrence H, et al. (2017). The growing impact of satellite observations sensitive to humidity, cloud and precipitation. Quarterly Journal of the Royal Meteorological Society, 143, 3189–3206. 10.1002/qj.3172

Grossman RL, & Betts AK (1990). Air–sea interaction during an extreme cold air outbreak from the eastern coast of the United States. Monthly Weather Review, 118, 324–342. 10.1175/1520-0493(1990)118<0324:aidaec>2.0.co;2

Hersbach H, Bell B, Berrisford P, Hirahara S, Horányi A, Muñoz-Sabater J, et al. (2020). The ERA5 global analysis. Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049. 10.1002/qj.3803

Honnert R, Efstathiou G, Beare R, Ito J, Lock A, Neggers R, et al. (2020). The atmospheric boundary layer and the “gray zone” of turbulence: A critical review. Journal of Geophysical Research: Atmospheres, 125, e2019JD030317. 10.1029/2019JD030317

Jin X, & Yu L (2013). Assessing high-resolution analysis of surface heat fluxes in the Gulf Stream region. Journal of Geophysical Research Oceans, 118, 5353–5375. 10.1002/jgrc.20386

Kolstad EW, Bracegirdle TJ, & Seierstad IA (2009). Marine cold-air outbreaks in the North Atlantic: Temporal distribution and associations with large-scale atmospheric circulation. Climate Dynamics, 33, 187–197. 10.1007/s00382-008-0431-5

Li X-Y, Wang H, Chen J, Endo S, George G, Cairns B, et al. (2021). Large-eddy simulations of marine boundary-layer clouds associated with cold air out-breaks during the ACTIVATE campaign-part 1: Case setup and sensitivities to large-scale forcings. arXiv e-prints, arXiv:2107.06193, 2107.06193.

Liu J, Xie S, Norris JR, & Zhang S (2014). Low-level cloud response to the Gulf Stream front in winter using CALIPSO. Journal of Climate, 27(12), 4421–4432. 10.1175/jcli-d-13-00469.1

Marshall J, Andersson A, Bates N, Dewar W, Doney S, Edson J, et al. (2009). The CLIMODE field campaign: Observing the cycle of convection and restratification over the Gulf Stream. Bulletin of the American Meteorological Society, 90, 1337–1350. 10.1175/2009BAMS2706.1

McCoy I, Wood R, & Fletcher JK (2017). Identifying meteorological controls on open and closed mesoscale cellular associated with marine cold air outbreaks. Journal of Geophysical Research, 122(21), 11678–11702. 10.1002/2017JD027031

Miao H, Wang X, Liu Y, & Wu G (2019). An evaluation of cloud vertical structure in three reanalyses against CloudSat/cloud-aerosol lidar and infrared pathfinder satellite observations. Atmospheric Science Letters, 20, e906. 10.1002/asl.906

Minnett PJ, Alvera-Azcárate A, Chin TM, Corlett GK, Gentemann CL, Karagali I, et al. (2019). Half a century of satellite remote sensing of sea-surface temperature. Remote Sensing of Environment, 233, 111366. 10.1016/j.rse.2019.111366

Minobe S, Kuwano-Yoshida A, Komori N, Xie SP, & Small RJ (2008). Influence of the Gulf Stream on the troposphere. Nature, 452, 206–209. 10.1038/nature06690 [PubMed: 18337820]

Moore GWK, & Renfrew IA (2002). An assessment of the surface turbulent heat fluxes from the NCEP–NCAR reanalysis over the western boundary currents. Journal of Climate, 15, 2020–2037. 10.1175/1520-0442(2002)015<2020:aaotst>2.0.co;2

Mülmenstadt J, Sourdeval O, Delano J, & Quaas J (2015). Frequency of occurrence of rain from liquid-, mixed-, and ice-phase clouds derived from A-Train satellite retrievals. Geophysical Research Letters, 42, 6502–6509. 10.1002/2015GL064604

Painemal D, Corral AF, Sorooshian A, Brunke MA, Seethala C, Gorooh VA, et al. (2021). An overview of atmospheric features over the Western North Atlantic Ocean and North American East Coast–Part 2: Circulation, boundary layer, and clouds. Journal of Geophysical Research, 126, e2020JD033423. 10.1029/2020JD033423

Papritz L, Pfahl S, Sodemann H, & Wernli H (2015). A climatology of cold air outbreaks and their impact on air–sea heat fluxes in the high-latitude South Pacific. Journal of Climate, 28, 342–364. 10.1175/JCLI-D-14-00482.1

Pincus R, Beljaars A, Buehler SA, Kirchengast G, Ladstaedter F, & Whitaker JS (2017). The representation of tropospheric water vapor over low-latitude oceans in (Re-)analysis: Errors, Impacts, and the ability to exploit current and prospective observations. Surveys in Geophysics, 38, 1399–1423. 10.1007/s10712-017-9437-z

Plagge A, Edson JB, & Vandemark D (2016). In situ and satellite evaluation of air–sea flux variation near ocean temperature gradients. Journal of Climate, 29(4), 1583–1602. 10.1175/jcli-d-15-0489.1

Rivas MB, & Stoffelen A (2019). Characterizing ERA-Interim and ERA5 surface wind biases using ASCAT. Ocean Science, 15, 831–852. 10.5194/os-15-831-2019

Rozenhaimer M, Barton N, Redemann J, Schmidt S, LeBlanc S, Anderson B, et al. (2018). Bias and sensitivity of boundary layer clouds and surface radiative fluxes in MERRA-2 and airborne observations over the Beaufort Sea during the ARISE campaign. Journal of Geophysical Research, 123, 6565–6580. 10.1029/2018JD028349

Saggiorato B, Nuijens L, Siebesma AP, de Roode S, Sandu I, & Papritz L (2020). The influence of convective momentum transport and vertical wind shear on the evolution of a cold air outbreak. Journal of Advances in Modeling Earth Systems, 12. 10.1029/2019MS001991

Sandu I, Beljaars A, Bechtold P, Mauritsen T, & Balsamo G (2013). Why is it so difficult to represent stably stratified conditions in numerical weather prediction (NWP) models? Journal of Advances in Modeling Earth Systems, 5(2), 117–133. 10.1002/jame.20013

Sherwood SC, Webb MJ, Annan JD, Armour KC, Forster PM, Hargreaves JC, et al. (2020). An assessment of Earth's climate sensitivity using multiple lines of evidence. Reviews of Geophysics, 58, e2019RG000678. 10.1029/2019RG000678

Skyllingstad ED, & Edson JB (2009). Large-eddy simulation of moist convection during a cold-air outbreak over the Gulf Stream. Journal of the Atmospheric Sciences, 66(5), 1274–1293. 10.1175/2008JAS275510.1175/2008jas2755.1

Small RJ, deSzoeke SP, Xie SP, O'Neill L, Seo H, Song Q, et al. (2008). Air-sea interaction over ocean fronts and eddies. Dynamics of Atmospheres and Oceans, 45, 274–319. 10.1016/j.dynatmoce.2008.01.001

Sorooshian A, Anderson B, Bauer SE, Braun RA, Cairns B, Crosbie E, et al. (2019). Aerosol-cloud-meteorology interaction airborne field investigations: Using lessons learned from the U.S. West Coast in the design of ACTIVATE off the U.S. East Coast. Bulletin of the American Meteorological Society, 100(8), 1511–1528. 10.1175/BAMS-D-18-0100.1 [PubMed: 33204036]

Titchner HA, & Rayner NA (2014). The Met Office Hadley Centre sea ice and sea surface temperature data set, version 2: 1. Sea ice concentrations. Journal of Geophysical Research Atmospheres, 119, 2864–2889. 10.1002/2013JD020316

Tornow F, Ackerman AS, & Fridlind AM (2021). Preconditioning of overcast-to-broken cloud transitions by riming in marine cold air outbreaks. Atmospheric Chemistry and Physics Discussions, 21, 12049–12067. 10.5194/acp-2021-82

Weller RA, Bigorre SP, Lord J, Ware JD, & Edson JB (2012). A Surface Mooring for Air–Sea Interaction Research in the Gulf Stream. Part I: Mooring Design and Instrumentation. Journal of Atmosheric and Oceanic Technology, 29(9), 1363–1376. 10.1175/JTECH-D-12-00060.1

Zelinka M, Myers TA, McCoy DT, Po-Chedley S, Caldwell PM, Ceppi P, et al. (2020). Causes of higher climate sensitivity in CMIP6 models. Geophysical Research Letters, 47, e2019GL085782. 10.1029/2019GL085782

Zhang J, Zuidema P, Turner D, & Cadeddu M (2018). Surface-based microwave humidity retrievals over the equatorial Indian Ocean: Applications and challenges. Journal of Applied Meteorology and Climatology, 57, 1765–1782. 10.1175/JAMC-D-17-0301.1

References From the Supporting Information

Bosilovich MG, & Coauthors (2015). MERRA-2: Initial Evaluation of the Climate. Technical Report Series on Global Modeling and Data Assimilation, NASA Technical Report NASA/TM–2015–104606 (Vol. 43, pp. 139). Available online at https://gmao.gsfc.nasa.gov/pubs/docs/Bosilovich803.pdf

Fairall CW, Bradley E, Hare J, Grachev A, & Edson J (2003). Bulk parameterization of air-sea fluxes: Updates and verification for the COARE algorithm. Journal of Climatology, 16, 571–591.

Gelaro R, McCarty W, Suárez MJ, Todling R, Molod A, Takacs L, et al. (2017). The Modern-Era Retrospective analysis for Research and Applications, version 2 (MERRA-2). Journal of Climate, 30(14), 5419–5454. 10.1175/JCLI-D-16-0758.1 [PubMed: 32020988]

Remote Sensing Systems (REMSS) (2008). GHRSST Level 4 MW_IR_OI Global Foundation Sea Surface Temperature analysis version 5.0 from REMSS (GDS versions 1 and 2). NOAA National Centers for Environmental Information. 10.5067/GHMWI-4FR05

Reynolds RW, Smith TM, Liu C, Chelton DB, Casey KS, & Schlax MG (2007). Daily high-resolution-blended analyses for sea surface temperature. Journal of Climate, 20(22), 5473–5496. 10.1175/2007jcli1824.1

Wick GA, Hock TF, Neiman PJ, Vömel H, Black ML, & Spackman JR (2018). The NCAR–NOAA Global Hawk Dropsonde System. Journal of Atmospheric and Oceanic Technology, 35, 1585–1604. 10.1175/JTECH-D-17-0225.1

Wu W-S, Purser RJ, & Parrish DF (2002). Three-dimensional variational analysis with spatially inhomogeneous covariances. Monthly Weather Review, 130, 2905–2916. 10.1175/1520-0493(2002)130,2905:TDVAWS.2.0.CO;2

Yu L, Jin X, & Weller RA (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:* Woods Hole Oceanographic Institution. OAFlux Project Technical Report.OA-2008-01, 64pp. Woods Hole.

Key Points:

- Reanalysis surface fluxes and boundary layers are representative of observations to first-order, sufficient for higher-resolution model initialization
- Reanalyses represent the Gulf Stream more broadly than is seen in nature, contributing to turbulent flux and boundary layer biases
- Previously noted thermodynamic and dynamic biases reinforce (ERA5) or compensate (MERRA2) surface fluxes but support realistic winter boundary layer heights

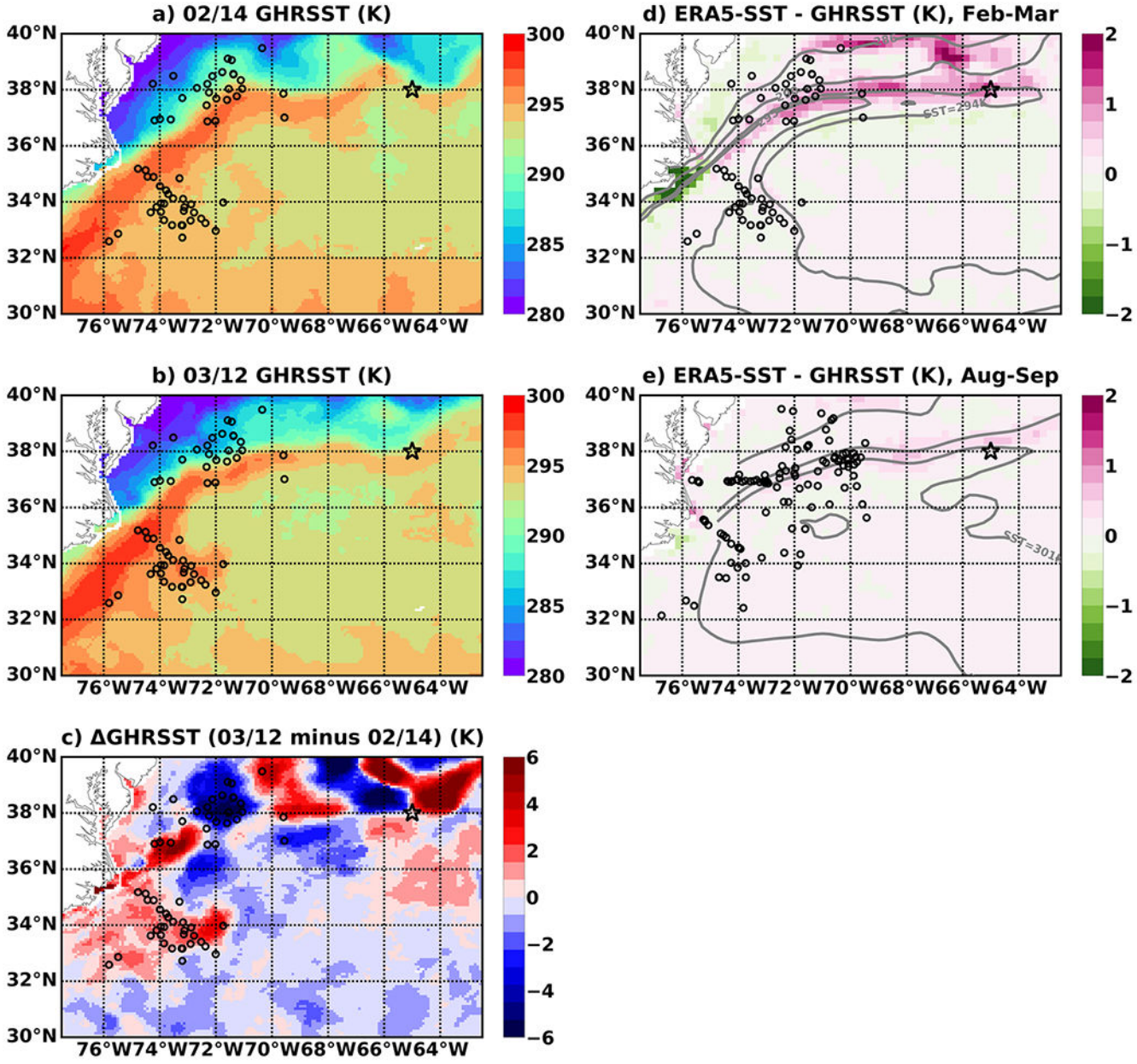

Figure 1.

Sea surface temperatures (GHRSST) on (a) February 14, 2020, (b) March 12, 2020, and (c) March 12–February 14, 2020. Open circles indicate ACTIVATE dropsonde locations, and the star denotes the CLIMODE buoy location. (ERA5-SST–GHRSST) differences for (d) February–March 2020 and (e) August–September 2020. Gray contours in (d) and (e) correspond to ERA5 SSTs of 286, 290, 294, and 295K in panel (d) and 301K in panel (e).

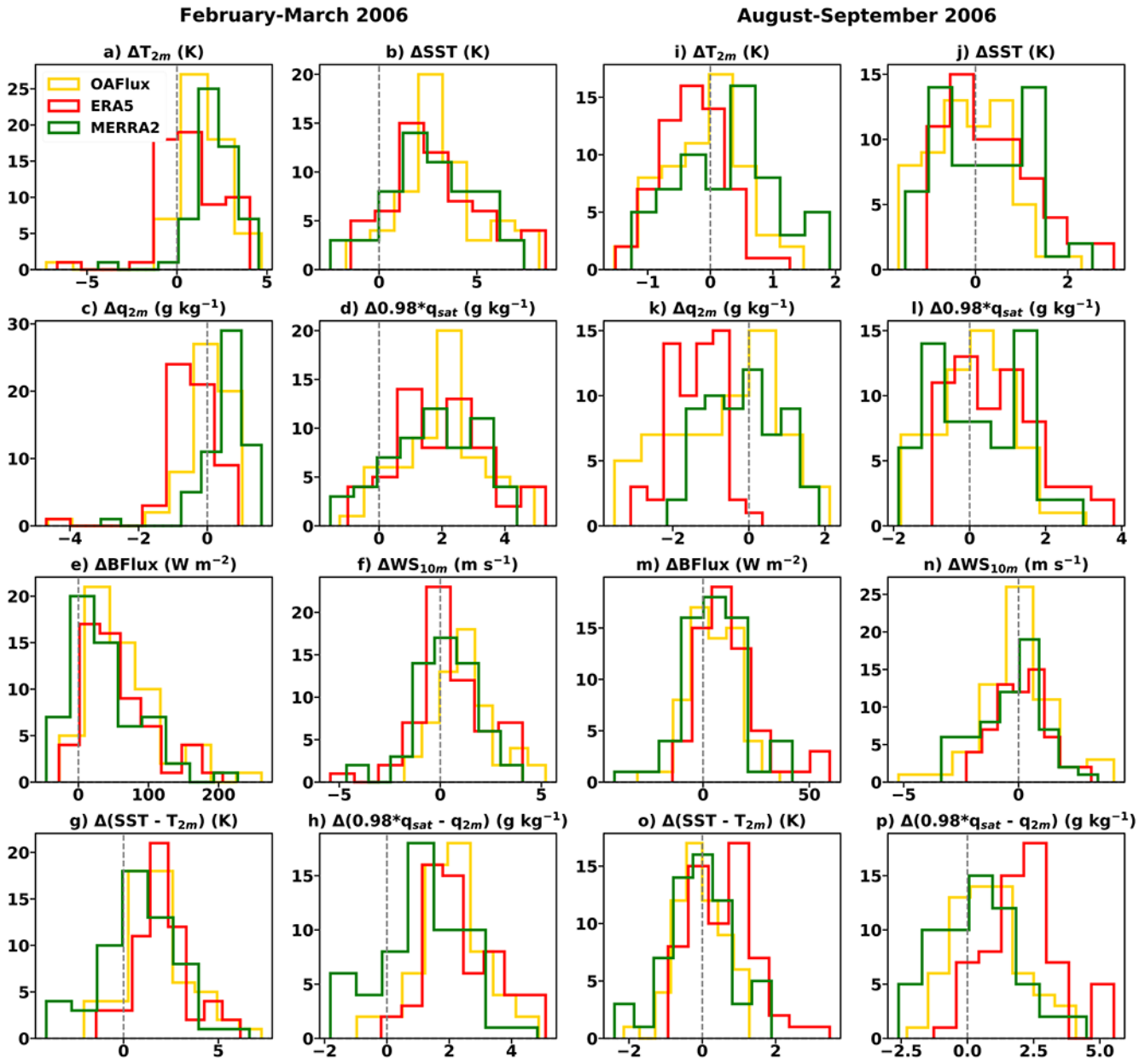

Figure 2.

Comparison of ERA5 (red), MERRA2 (green), and OAFLUX (yellow) daily mean differences from CLIMODE buoy measurements (reanalysis-buoy) for February–March 2006 for (a) near-surface air temperature, (b) sea surface temperature (SST), (c) near-surface specific humidity, (d) saturated specific humidity at the SST, (e) buoyancy flux, (f) wind speed at 10m, and the air-sea (g) temperature and (h) humidity differences. (i–p): same as (a–h) but for August-September 2006. Note changes in x-scale range. MERRA2 SST is a skin value, while the buoy, ERA5, and OAFLUX SSTs are foundation SST. OAFLUX WS_{10m} is the neutral wind speed.

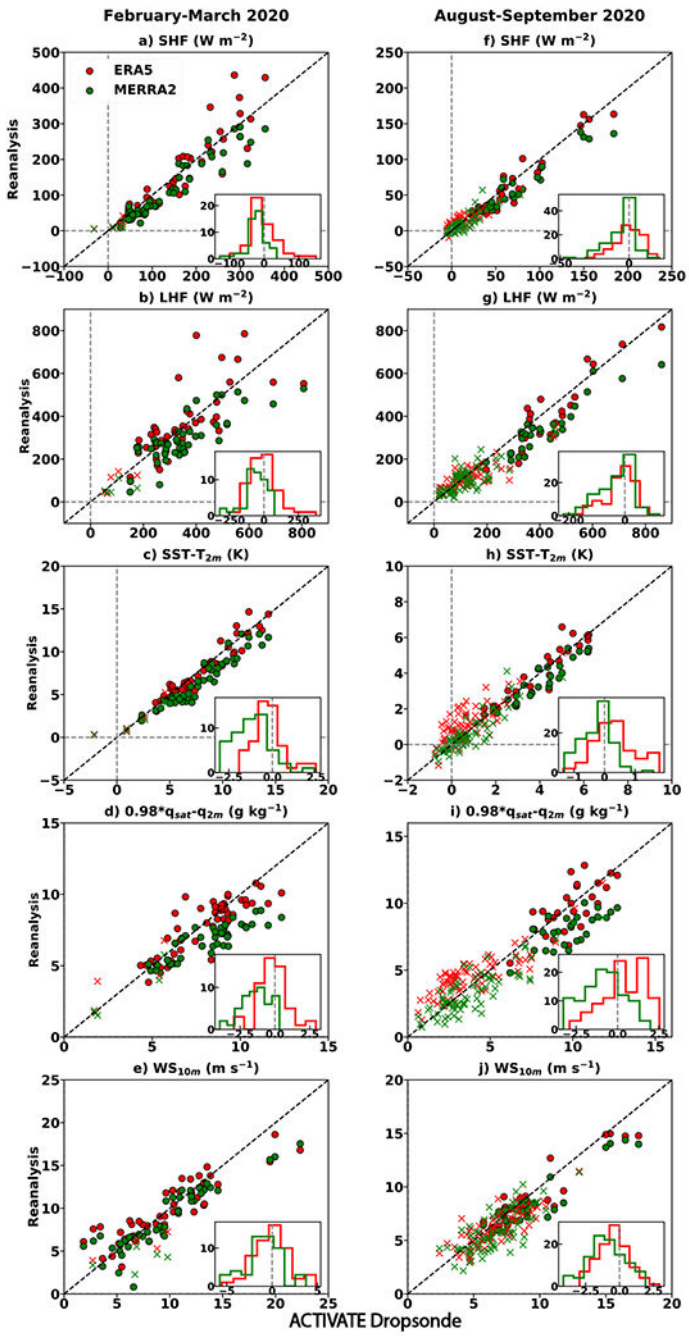

Figure 3.

Comparison of nearest-in-space-and-time hourly mean ERA5 (red) and three-hourly-mean MERRA2 (green) reanalysis to instantaneous ACTIVATE dropsonde-calculated values for February–March 2020 of (a) sensible heat flux, (b) latent heat flux, calculated using the COARE v3.5 algorithm, (c) $SST\text{-}T_{2m}$, (d) $0.98{*}q_s{-}q_{2m}$, and (e) WS_{10m}. (f–j): same as (a–e) but for August–September 2020; note change in range on both axes. Insets within each panel are histograms of the (reanalysis-dropsonde) differences. Filled circles represent cold-air outbreak conditions, and "x" markers signify non-CAO conditions, using $\theta_{\mathrm{SST\_skin}}{-}\theta_{900hPa}$

> 0 to define whether an individual dropsonde represented a cold-air outbreak (CAO) (see supplement). GHRSST values are corrected to represent a “skin” sea surface temperatures.

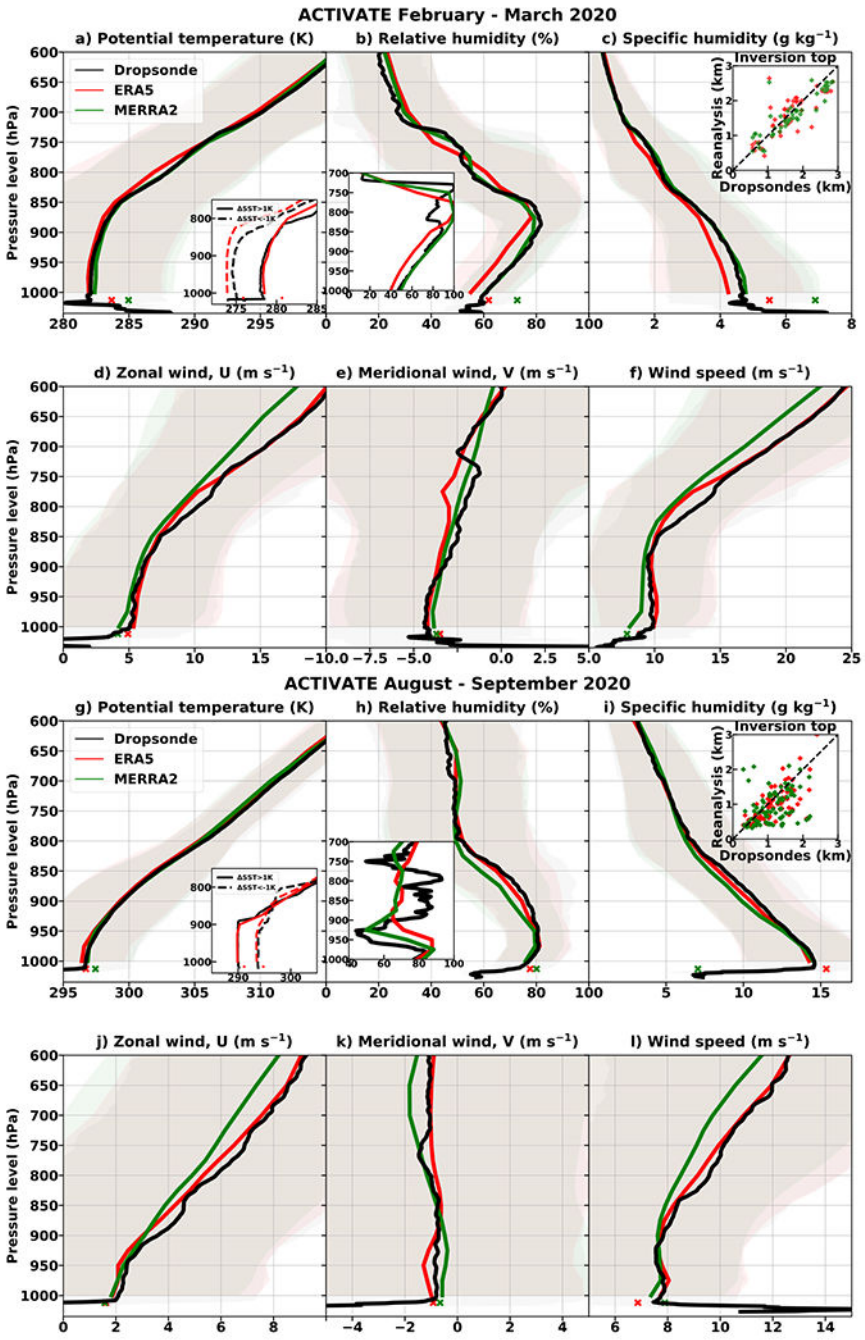

Figure 4.

Mean vertical profiles from dropsondes (black), ERA5 (red), and MERRA2 (green) of (a) potential temperature θ, (b) relative humidity, (c) specific humidity, (d) zonal wind (U), (e) meridional wind (V), and (f) wind speed for February–March 2020. Colored shading indicates the standard deviation. Small "x" markers indicate θ_{2m}, RH_{2m} and q_{2m}, and U, V and wind speed at 10m. (f–l): same as (a–f) but for August–September 2020. Insets in (a) and (g) represent the mean θ profiles for ERA5-SST–GHRSST > 1K and ERA5-SST–GHRSST < −1K. The inset profile in panel (b) is from 28 February (profile#24) and in panel

(h) from 20 August (profile#12). Insets in (c) and (i) indicate inversion top heights estimated from the relative humidity profiles.