



## Spatial Patterns and Intensity of the Surface Storm Tracks in CMIP5 Models

JAMES F. BOOTH

*City College of the City University of New York, New York, New York*

YOUNG-OH KWON

*Woods Hole Oceanographic Institution, Woods Hole, Massachusetts*

STANLEY KO

*Rutgers, The State University of New Jersey, New Brunswick, New Jersey*

R. JUSTIN SMALL

*National Center for Atmospheric Research, Boulder, Colorado*

RYM MSADEK

*CNRS/CERFACS, CECI UMR 5318, Toulouse, France*

(Manuscript received 17 March 2016, in final form 17 March 2017)

### ABSTRACT

To improve the understanding of storm tracks and western boundary current (WBC) interactions, surface storm tracks in 12 CMIP5 models are examined against ERA-Interim. All models capture an equatorward displacement toward the WBCs in the locations of the surface storm tracks' maxima relative to those at 850 hPa. An estimated storm-track metric is developed to analyze the location of the surface storm track. It shows that the equatorward shift is influenced by both the lower-tropospheric instability and the baroclinicity. Basin-scale spatial correlations between models and ERA-Interim for the storm tracks, near-surface stability, SST gradient, and baroclinicity are calculated to test the ability of the GCMs' match reanalysis. An intermodel comparison of the spatial correlations suggests that differences (relative to ERA-Interim) in the position of the storm track aloft have the strongest influence on differences in the surface storm-track position. However, in the North Atlantic, biases in the surface storm track north of the Gulf Stream are related to biases in the SST. An analysis of the strength of the storm tracks shows that most models generate a weaker storm track at the surface than 850 hPa, consistent with observations, although some outliers are found. A linear relationship exists among the models between storm-track amplitudes at 500 and 850 hPa, but not between 850 hPa and the surface. In total, the work reveals a dual role in forcing the surface storm track from aloft and from the ocean surface in CMIP5 models, with the atmosphere having the larger relative influence.

### 1. Introduction

Atmospheric storm tracks are very important for climate dynamics. They indicate regions of maximum transient poleward energy transport and zonal momentum transport (Chang et al. 2002) and play an important role in setting the dynamical response of the mid-latitudes to global warming through their radiative forcing (Voigt and Shaw 2015). Storm tracks are

generally calculated as the standard deviation of atmospheric data that has been filtered in the time domain to isolate synoptic variability (Blackmon 1976). Typical variables used to calculate storm tracks are meridional wind, eddy kinetic energy, or geopotential height, at a fixed vertical level. This metric represents the climatology of baroclinic wave activity (i.e., high and low pressure systems), but for historical reasons has been termed "storm track" [see Wallace et al. (1988) for more discussion]. Following Chang et al. (2002), we consider each ocean basin as having its own storm track. Storm

*Corresponding author:* James F. Booth, jbooth@ccny.cuny.edu

DOI: 10.1175/JCLI-D-16-0228.1

© 2017 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the [AMS Copyright Policy](http://www.ametsoc.org/PUBSReuseLicenses) ([www.ametsoc.org/PUBSReuseLicenses](http://www.ametsoc.org/PUBSReuseLicenses)).

tracks offer a reasonable proxy for climatological activity of extratropical cyclones (Hoskins and Hodges 2002), and their maxima occur over the oceans, in the vicinity of ocean western boundary currents (WBCs) and their extensions (e.g., Fig. 1b).

WBCs are unique regions of air–sea coupling: ocean currents in these regions generate strong ocean heat flux convergence, which can dictate spatial and temporal variability in air–sea fluxes [see reviews by Kwon et al. (2010) and Kelly et al. (2010)]. The North Atlantic and North Pacific WBCs, the Gulf Stream, and Kuroshio–Oyashio Extension (KOE) influence the atmosphere through the entire troposphere during spring and summer (Minobe et al. 2008, 2010; Xu et al. 2011; Sasaki et al. 2012) and modify low-level atmospheric baroclinicity, shifting the free-tropospheric storm track and altering the poleward heat and moisture transport (e.g., Tokinaga et al. 2009; Frankignoul et al. 2011; Ogawa et al. 2012; Taguchi et al. 2012; Kwon and Joyce 2013; O’Reilly and Czaja 2015). In the Southern Ocean, south of the Indian Ocean, the Agulhas Return Current (ARC) helps to anchor the climatological location of the free-tropospheric storm track (Nakamura et al. 2004). This causes the region to have a consistent storm track throughout the year, which, for the Southern Ocean storm track, is a trait that is unique to the ARC region.

These examples of the oceans influencing the storm tracks primarily focus on the free-tropospheric storm tracks (e.g., the filtered geopotential at 500 hPa or the filtered meridional winds at 850 hPa). However, one can also analyze the surface storm tracks based on meridional winds at 10 m. Booth et al. (2010) show that the spatial patterns of storm tracks at 10 m differ from the free-tropospheric storm tracks due to the influence of ocean WBCs. Booth et al. (2010) used physical arguments proposed by Sweet et al. (1981) to suggest that the warm water in WBC creates regions with stronger atmospheric instability during cold air outbreaks associated with extratropical cyclones. The greater instability on the warm side of the WBC increases vertical mixing of momentum in these unstable regions creating stronger surface winds (a so-called momentum-mixing mechanism; see also Wallace et al. 1989). This preferential vertical mixing of momentum causes surface storm tracks to have a maximum in a region that differs from the maximum aloft. In addition, Joyce et al. (2009) showed that the surface storm tracks covary with the WBC at the interannual-to-decadal time scale.

The momentum-mixing mechanism is one element of forcing at the WBC. It is also known that an atmospheric pressure gradient force created by strong ocean fronts can accelerate winds blowing from the cold to the warm side of the sea surface temperature (SST) front (Lindzen and

Nigam 1987; Chelton et al. 2004). Thus, in the regions of surface storm tracks, it is possible that the spatial gradient in momentum mixing and the pressure gradient force, both associated with WBCs, could influence the surface winds. In the Gulf Stream region, both mechanisms have been shown to play some role in at least one general circulation model (GCM) (Brachet et al. 2012). However, other work suggests that the pressure gradient mechanism, which was created for the tropics, may not be very strong in high-wind regimes of the storm tracks (Spall 2007; Small et al. 2008; Schneider and Qiu 2015). Additionally, recent analysis by Liu et al. (2013) shows that the momentum-mixing mechanism tends to dominate on shorter time scales, such as those captured by the storm tracks.

In addition to the momentum mixing and pressure gradient physics, the storm tracks near the WBC need to be considered because the WBCs are extratropical cyclone genesis regions in the Northern Hemisphere (Hoskins and Hodges 2002). Because the storms typically grow due to the merging of a surface and upper-level disturbance, the near-surface behavior at the WBC region may be indicative of storm genesis. Nakamura and Shimpo (2004) examined the Southern Ocean storm track and showed that the SST gradient at the ARC is important for maintaining low-level baroclinicity. Hoskins and Hodges (2005) show that the genesis region for the Indian Ocean storm track is in the Andes mountain region; however, the large amount of secondary cyclogenesis in the Southern Ocean suggests that baroclinic anchoring by the ARC would still be important for storm genesis. Booth et al. (2010) showed that for JJA in the Indian Ocean, there are two active regions in the surface storm track: one near the ARC and another near the sea ice edge. Related to this, Nakamura and Shimpo (2004) emphasize that the ARC helps maintain a strong storm track during SH summer (DJF).

Given the climatological importance of storm tracks and the role of WBCs in forcing surface storm tracks, it stands to reason that surface storm tracks in GCMs are a good variable to analyze to check model biases and better understand coupled model physics. In particular, the biases in the surface storm tracks, as compared to the biases in the free-tropospheric storm tracks, may inform on model issues regarding the WBCs and the modeled momentum mixing in the midlatitudes. It is already known that GCMs often have issues in representing the separation of the WBCs from the coastlines in the Northern Hemisphere, in particular, for the non-eddy-resolving ocean models (e.g., Gent et al. 2011; Schoonover et al. 2016). Coupled models with eddy-resolving oceans better represent the strength, width, and path of the WBCs, but can still exhibit overshooting of the path (e.g., Small et al. 2014; Griffies et al. 2015).

Therefore, an analysis of the surface storm tracks in coupled GCMs tests the physics of the ocean and atmosphere as well as their coupling. With this as motivation, the present study examines free-tropospheric and surface storm tracks along with SST in the WBC regions using 12 CMIP5 models.

Previous work has analyzed the free-tropospheric storm tracks in the CMIP5 models with a focus on future projections (Chang et al. 2012). Here, we instead focus on the historical runs, to determine the model's ability to represent the surface storm tracks. We ask the following questions: 1) Can models capture differences in the locations and amplitudes of the free-tropospheric and surface storm tracks? 2) What factors determine the strength of the surface storm tracks in the models? 3) What are the relative influences of the free troposphere and the ocean surface in determining the modeled location of the surface storm tracks? To address these questions, we examine the storm tracks and SST at the global and ocean-basin scale. The physics that we are interested in, such as momentum mixing affecting the location of the surface storm track, have already been discussed in previous papers. Here, we are attempting to use the CMIP5 models to determine if these same physical processes cause biases in the SST to be manifest as biases in the surface storm track.

## 2. Data and methods

### a. Models and data

The variables we analyze are meridional winds at 10 m (V10), 850 hPa (V850), and 500 hPa (V500), as well as surface temperature (TS) and a rough estimate of atmospheric stability in the lower troposphere, hereafter,  $T_{\text{DIFF}}$ , defined as TS minus 850-hPa air temperature. Note that TS is exactly equal to SST over the oceans, except in regions of sea ice. The reanalysis data utilized here are from ERA-Interim (hereinafter ERA-I; Dee et al. 2011) and have been shown to perform as well as any other recent reanalyses at capturing midlatitude storm activity (Hodges et al. 2011). We use the SST provided with the reanalysis (which is based on merged SST observations, discussed in the next paragraph) so that 1) we use the SST that reanalysis variables were driven by, and 2) all of the reanalysis variables are on the same grid. The epoch used for this study is 1979–2005, which is the overlap of ERA-I and the time period of the historical integrations according to the CMIP protocol.

We note that the horizontal resolution of the SST used to drive ERA-I has been changed three times, which can have some impact on surface winds (e.g., Chelton 2005; Masunaga et al. 2015). However, the spatial and

TABLE 1. The reanalysis and models used in this study. (Expansions of acronyms are available online at <http://www.ametsoc.org/PubsAcronymList>.)

Reanalysis and models	Abbreviation	Resolution (lat × lon)
ERA-Interim	ERA-I	$0.67^\circ \times 0.67^\circ$
GISS-E2 (AMIP)	GISS AMIP	$2^\circ \times 2.5^\circ$
GFDL CM3 (AMIP)	GFDL AM3	$2^\circ \times 2.5^\circ$
BCC_CSM1	BCC	$2.8125^\circ \times 2.8125^\circ$
CanESM2	CANESM	$2.8125^\circ \times 2.8125^\circ$
CESM1-LE	CESM1-LE	$0.9424^\circ \times 1.25^\circ$
GFDL CM3	GFDL CM3	$2^\circ \times 2.5^\circ$
CNRM-CM5	CNRM	$1.4062^\circ \times 1.40625^\circ$
CSIRO Mk3.6.0	CSIRO	$1.875^\circ \times 1.875^\circ$
GISS-E2-R	GISS ER	$2^\circ \times 2.5^\circ$
INM-CM4.0	INM	$1.5^\circ \times 2^\circ$
IPSL-CM5A-MR	IPSL	$1.25^\circ \times 2.5^\circ$
MIROC5	MIROC	$1.4062^\circ \times 1.40625^\circ$
MPI-ESM-LR	MPI	$1.875^\circ \times 2.5^\circ$
NorESM1-M	NORESM	$1.875^\circ \times 2.5^\circ$

temporal scales analyzed in those studies differ from those of interest in the present work. Additionally, we find that surface storm tracks in ERA-I are very similar in spatial pattern and intensity to that in the NCEP CFSR (Saha et al. 2010) and NASA MERRA (Rienecker et al. 2011) (not shown).

Our analysis focuses on CMIP5-type models, which were run using observed atmospheric forcing (i.e., the “historical” run in the CMIP5 protocol). The coupled models used, along with their acronyms, are detailed in Table 1. The 12 models used in this analysis were chosen based on the availability of the variables used in the analysis, with the daily (or finer temporal resolution) surface winds often being the limiting factor. Some of our analysis also examines atmosphere-only versions of the GFDL and GISS models, referred to here as GFDL AM3 and GISS AMIP, respectively. These models are also driven by historical atmospheric radiative forcing, but they have prescribed SSTs (which are based on observations); that is, they are AMIP-type models. The horizontal resolution of each of the models is given in Table 1. For each GCM listed in the table, we analyze a single ensemble member of the model.

For TS and 850-hPa air temperature, we used monthly data to calculate the climatology. Daily data were used for V10, V850, and V500, and for models for which data were not available, 3- or 6-hourly data (if available) were averaged to create a proxy for the daily value before calculating the storm tracks.

V10 was not available for CESM1 Large Ensemble (CESM1-LE) and NorESM1-M. Therefore, we use the meridional wind at the lowest model level ( $V_{\text{BOT}}$ ), which is located at 55–70 m for CESM1-LE, and at a similar height for NorESM1-M. It may be questioned whether 55–70 m is really a representative height for a “surface”

storm track. Over the land, this would be a big issue, but over the oceans, especially in the unstable regions of the WBCs, the difference between 10- and 60-m winds is most likely small. In separate non-CMIP5 simulations with CESM1 [described in [Small et al. \(2014\)](#)], with additional output data, it was found that typical ratios of lowest model level wind to 10-m wind were from 0.9 to 0.95 in the NH winter; that is, 10-m winds were slightly stronger than bottom level, in very unstable conditions (surface ocean temperature minus 2-m air temperature was greater than 4°C). Conversely, in more stable conditions of the Southern Ocean in austral summer, the lowest model level wind was typically 1.05 times the 10-m wind. Also, we found that climatological values of  $V_{10}$  and  $V_{\text{BOT}}$  differ by less than  $0.4 \text{ m s}^{-1}$  and the differences can be either positive or negative (not shown). Finally, spatial patterns of the 10-m storm track are very similar to model bottom level storm track, and we refer to both as “surface storm track.” The conclusions of this paper are thus not sensitive to whether we actually used bottom-level wind or 10-m wind.

Another issue regarding  $V_{10}$  is the question of whether the modeling centers report the “real”  $V_{10}$  or the neutral equivalent  $V_{10}$ , and this was not always clear from the provided documentation. The CESM simulations of [Small et al. \(2014\)](#) mentioned above show that the storm track based on  $V_{10}$ -neutral and that based on  $V_{10}$ -real never differ by more than 2% in the WBC regions (not shown). This is because substantial differences (e.g., of 10% or more) between the actual wind and the neutral wind only occur in quite low wind speed regimes (e.g., weaker than  $5 \text{ m s}^{-1}$ ) under strongly unstable or stable conditions ([Liu and Tang 1996](#)), but the storm-track regions have strong winds ( $>10 \text{ m s}^{-1}$ ). Therefore we do not need to distinguish between neutral wind and actual wind in the analysis below. To put the differences between  $V_{\text{BOT}}$ ,  $V_{10}$ , and neutral equivalent  $V_{10}$  in context, results shown below (see [Fig. 9](#)) reveal most models have surface storm tracks that are 20%–30% weaker than at 850 hPa. This is a much larger difference than between the different surface-wind variables used for calculating the surface storm track.

Each of the models and the reanalysis were generated and saved on their own grids ([Table 1](#)). However, for our analysis we use two-dimensional interpolation via a cubic spline method to project all of the data to the same grid. We choose to use the most often occurring grid from our set of models, which is  $2^\circ$  latitude by  $2.5^\circ$  longitude. All results are shown on this grid.

Throughout the analysis, we will refer to the climatological locations of the WBCs. The locations of the Gulf Stream and Kuroshio Extension have been estimated through an analysis of the observed sea surface height using the satellite altimeters [provided by K. Kelly and

S. Dickinson; see [Kelly et al. \(2010\)](#) for details]. The ARC has been defined as the equatorward edge of the location of observed maximum SST gradient for the climatology of SST for 1981–2005, which we calculate using a  $0.25^\circ$  blended SST product based on satellite measurements (OISSTv2; [Reynolds et al. 2007](#)).

### b. Analysis methods

Here we focus on DJF for both the Northern and Southern Hemisphere. We also carried out an analysis of JJA in the Southern Hemisphere and a discussion of those results is included. As highlighted in the introduction, previous work by [Nakamura and Shimpou \(2004\)](#) suggests that the influence of the ocean surface on the storm track in the Indian Ocean sector of the Southern Ocean is more apparent in DJF than JJA. This, in combination with our findings herein, has led to our decision to present results for DJF only for this region (hereafter the Indian Ocean).

As mentioned in the introduction, it is common practice to time-filter data to isolate synoptic-scale variability (e.g., [Blackmon 1976](#)). In the literature, there are two methods commonly used to filter the data: 1) applying a 2–8-day (or in some cases 2–6 day) bandpass filter to 6-hourly or daily data and 2) calculating 24-h differences of daily mean data. In the latter case, using daily averages removes the diurnal cycle and the 24-h differencing removes variability beyond 5 days. [Wallace et al. \(1988\)](#) discuss the comparison of the two methods and show that dividing daily differenced data by two provides a close match to the amplitude reduction generated by a bandpass filter. Therefore, we define our time-filtered transient eddies as

$$\tilde{v}(t) = \frac{v_D(t+1) - v_D(t)}{2}, \quad (1)$$

where  $v_D$  represents the daily averaged meridional winds, either  $V_{10}$ ,  $V_{850}$ , or  $V_{500}$ . We define the storm-track value at each latitude–longitude grid as the climatology of each season’s standard deviation of  $\tilde{v}$ :

$$v^{\text{ST}} = \frac{\sum_{n=1979}^{2005} (\sigma_n^{\text{SEASON}})}{N_{\text{YEARS}}}, \quad (2)$$

where

$$\sigma^{\text{SEASON}} = \sqrt{\frac{1}{J} \sum_{j=1}^J (\tilde{v}_j)^2 / J}. \quad (3)$$

The  $j$  index on  $\sigma^{\text{SEASON}}$  represents the days in the season of interest. Thus, the surface storm track is  $v_{10}^{\text{ST}}$ , and the free-tropospheric storm track is  $v_{850}^{\text{ST}}$ .

There are two advantages to using the daily difference filtering method: 1) daily outputs are easier to save, and 2) the analysis can be coded into GCMs so that in future simulations monthly storm track statistics can be saved. All that is required is to keep daily averages from the day before and the accumulated storm-track value as the month proceeds. This yields a finescale temporal resolution metric that does not require copious model output. This filtering method can also be used on observations that are available only at a daily resolution (e.g., Guo et al. 2009).

For one component of the analysis, we utilize the technique of Booth et al. (2010) to calculate an estimated surface storm track defined as the region of overlap of the upper quantiles of  $v_{850}^{\text{ST}}$  and  $T_{\text{DIFF}}$ . Note that Booth et al. (2010) used the difference between TS and 2-m air temperature to define  $T_{\text{DIFF}}$ . We use the TS minus  $T$  at 850 hPa. This is because 2-m air temperature was not available for all the models, and a comparison of results using the two definitions of  $T_{\text{DIFF}}$  produced negligible differences (not shown). This is not meant to imply that the 2-m temperature fields are identical to the 850-hPa temperature fields, but instead that the model-to-model variability of the two temperature fields is similar.

Here we refine the Booth et al. (2010) method to make it more suitable to apply it to different datasets. We start by identifying the region of the strongest  $v_{10}^{\text{ST}}$  [hereafter  $v_{10}^{\text{ST}}(\text{TOP})$ ], and we calculate a similar term at 850 hPa, defined as the area with values above the top  $M^*$  percentile. We define  $A_{\text{TOP}}$  as the area contained in this region. [Here we define area based on number of grid points (after regridding); because all of these components occur in a similar latitude range, the convergence of meridians does not create a notable impact.] Then, we consider the region of overlap of the top  $M$  percent for  $v_{850}^{\text{ST}}$  and  $T_{\text{DIFF}}$  (where  $M$  is initially equal to  $M^*$ ) and define this as the estimated surface storm track  $v_{10}^{\text{ESTIMATE}}$ . If the size of  $v_{10}^{\text{ESTIMATE}}$  is smaller than  $A_{\text{TOP}}$ , then we increase the value of  $M$  (i.e., increase the size of the  $v_{850}^{\text{ST}}$  and  $T_{\text{DIFF}}$  regions used for defining  $v_{10}^{\text{ESTIMATE}}$ ), iteratively, until the areal size of  $v_{10}^{\text{ESTIMATE}}$  is equal to  $A_{\text{TOP}}$ .

In addition to creating an estimated storm track using the overlap of  $v_{850}^{\text{ST}}$  and  $T_{\text{DIFF}}$ , we create estimated storm tracks using the overlap of  $v_{850}^{\text{ST}}$  and SST gradient  $|\nabla\text{SST}|$ , as well as  $v_{850}^{\text{ST}}$  and the baroclinicity at 850 hPa. We define the baroclinicity in a manner similar to Nakamura and Yamane (2009):

$$\sigma_{\text{BI}} = \frac{-g}{N\theta} \frac{\partial\theta}{\partial y}. \quad (4)$$

The notation in (4) is standard, with  $N$  denoting the Brunt–Väisälä frequency and  $g$  the gravitational constant. The units shown in the figures for  $\sigma_{\text{BI}}$  are  $\text{day}^{-1}$ .

As Nakamura and Shimpo (2004) discuss, the definition in (4) is very similar to the Eady growth rate. Unlike the Eady growth rate, we do not include the scaling coefficient of 0.31, and therefore our maximum values of  $\sigma_{\text{BI}}$  are close to  $3 \text{ day}^{-1}$  and not the near  $1 \text{ day}^{-1}$  values seen for the Eady growth rate (e.g., Hoskins and Valdes 1990). Nakamura and Shimpo (2004) defined a lower-tropospheric growth rate using the layer between 700 and 850 hPa. Here we center the growth rate at 850 hPa, and use potential temperature  $\theta$  on the 925- and 700-hPa layers to calculate  $N$ . For simplicity, we plot the negative of  $\sigma_{\text{BI}}$  for the Southern Hemisphere.

The three separate definitions of the estimated storm track are motivated by the question of which temperature-related factors might affect the offset in location of the surface storm track compared to that at 850 hPa: 1) momentum mixing, and hence  $T_{\text{DIFF}}$ , 2) wind acceleration related to the SST gradient, or 3) the genesis of storms at the WBC regions in the presence of upper-tropospheric perturbations (e.g., Cione et al. 1993), and hence the baroclinicity. We acknowledge that the SST and  $|\nabla\text{SST}|$  anomalies may not be perfectly collocated with wind anomalies, due to horizontal advection, but the coarse grid resolution used in this analysis means that the lack of collocation will likely be no more than one grid cell. We also note that there is potential for indirect impacts from  $T_{\text{DIFF}}$  and  $|\nabla\text{SST}|$ , since they can also contribute to stronger storms through diabatic heating in the storms and/or frontogenesis [see Booth et al. (2012) for more discussion].

After generating the estimated storm track, we compare the spatial locations of the top  $M^*$  percent of  $v_{10}^{\text{ST}}$  and  $v_{10}^{\text{ESTIMATE}}$ . To quantify this, we calculate the amount of overlap in the locations of the different fields of maximum. Note also, for this analysis we only consider the grid points over the ocean. This way if  $v_{10}^{\text{ST}}(\text{TOP})$  occupies 50 grid points, we know that the same is true for  $v_{850}^{\text{ST}}(\text{TOP})$ , and from its definition above, we know that  $v_{10}^{\text{ESTIMATE}}$  will be the same size as well. Then, if 30 of the grid points occupied by  $v_{10}^{\text{ST}}(\text{TOP})$  are also occupied by  $v_{850}^{\text{ST}}(\text{TOP})$ , their overlap would be 60%, and similar percentages are calculated for the other variables.

### 3. Results

#### a. Global scale

For both the Northern and Southern Hemispheres, the storm tracks' maxima occur over the ocean during DJF and are evident in Fig. 1 for  $v_{10}^{\text{ST}}$  and  $v_{850}^{\text{ST}}$ . At the global scale, all 12 of the climate models succeed in their representation of the geographical placement of  $v_{850}^{\text{ST}}$  (Fig. 1; here we show two representative models, GFDL

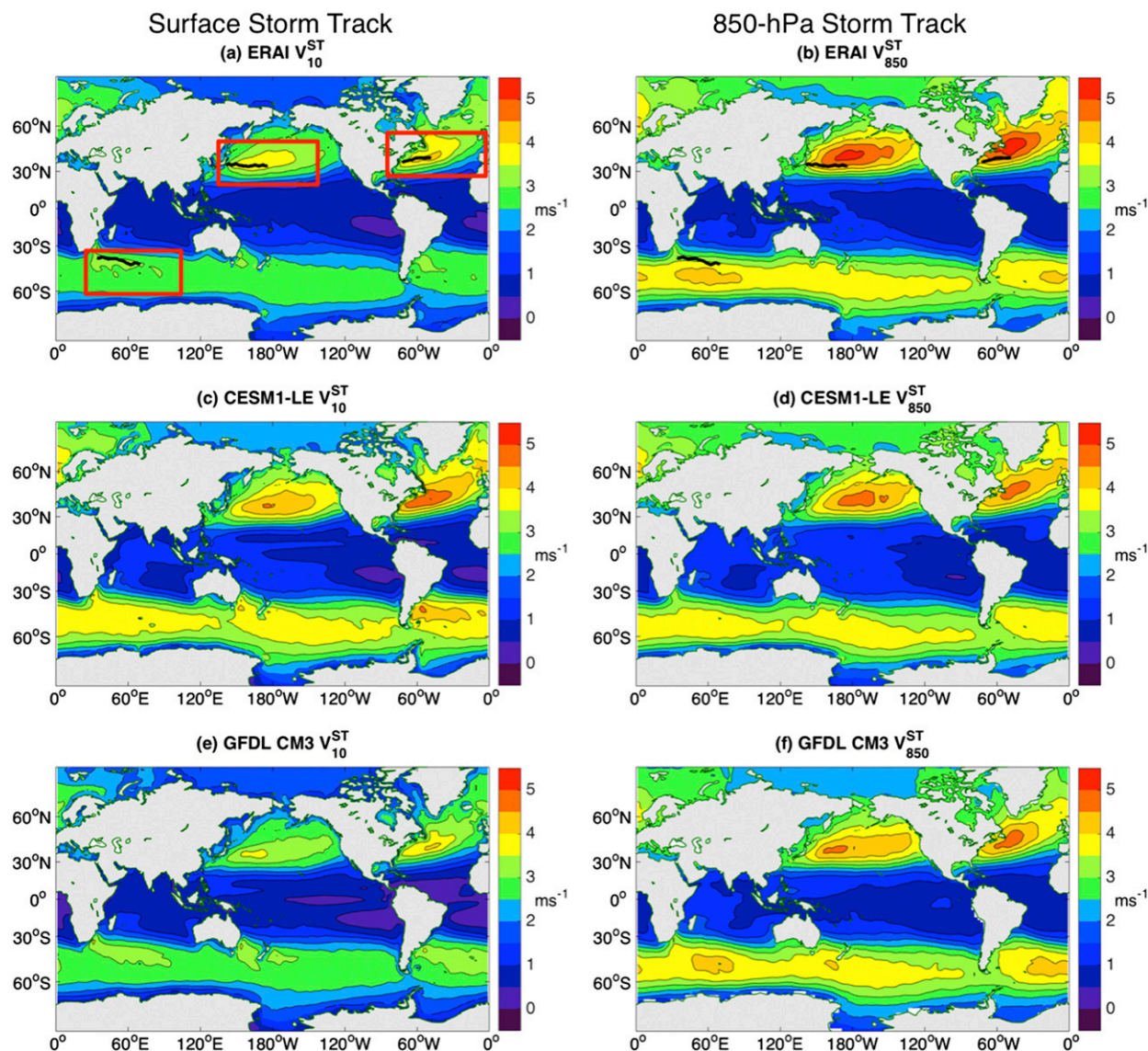


FIG. 1. (a),(c),(e) Surface and (b),(d),(f) free-tropospheric wintertime storm tracks in DJF. Color scales are the same for the surface and free troposphere. The Gulf Stream, Kuroshio Extension, and ARC locations based on observations of SSH and SST (black lines) are shown in (a),(b). The red boxes in (a) indicate the regions of focus for this remainder of the study.

CM3 and CCSM1-LE). Figure 1 also shows that the Southern Ocean DJF  $v_{850}^{ST}$  maximum, with respect to longitude, occurs south of the Indian Ocean, which is consistent with Nakamura and Shimpo (2004), and the same is true for JJA (not shown). The spatial pattern of the 500-hPa storm tracks looks very similar to those at 850 hPa and hence is not shown.

For each ocean basin, the location of the surface storm track maximum differs from that of  $v_{850}^{ST}$ , with  $v_{10}^{ST}$  displaced equatorward toward the WBC. For the reanalysis data (Figs. 1a,b), the spatial structure of  $v_{10}^{ST}$  differs slightly from that in Booth et al. (2010) due to interannual variability, because the analysis here uses

1979–2005 whereas Booth et al. used 1999–2006. For the models shown, the displacement of  $v_{10}^{ST}$  is greater in the North Atlantic and Southern Ocean than it is in the North Pacific. From a global viewpoint, GFDL CM3 captures the observed strength of  $v_{10}^{ST}$  reasonably well (Fig. 1e), while CESM1-LE is too strong (Fig. 1c). The surface wind bias in CESM1-LE also exists in the unfiltered zonal and meridional winds (not shown) and surface wind stress (Small et al. 2014), and is most likely due to either excessive vertical mixing in the boundary layer scheme or weak frictional damping in the surface layer scheme. Because we are using  $V_{BOT}$  and not V10 for CESM1-LE, we also examined monthly output from

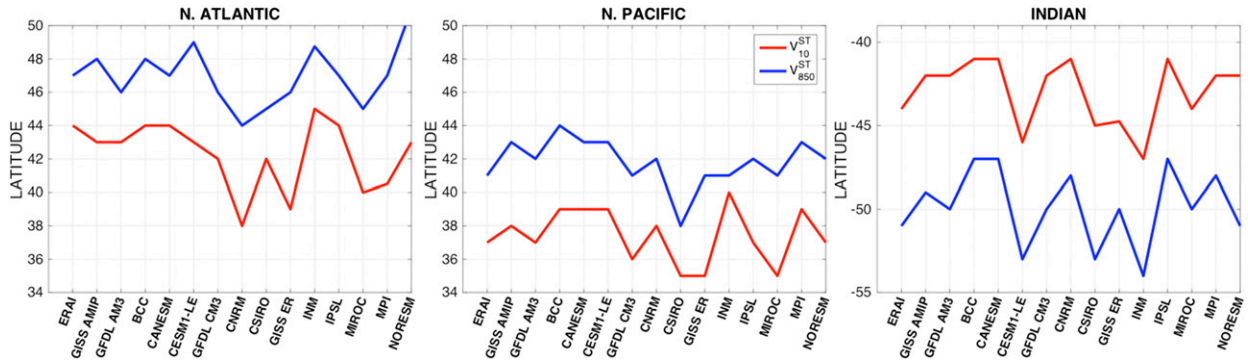


FIG. 2. Mean latitude for the top 10th percentile of the storm tracks at surface (red) and 850 hPa (blue). All panels show DJF. The legend shown in the North Pacific panel holds for all basins.

CESM1-LE (for which 10-m wind speed is available) to compare 10-m wind speed for the model and ERA-I. We find that the model’s climatological 10-m winds are also systematically stronger than those of ERA-I (not shown), consistent with Small et al. (2014) and our surface storm track results.

Figure 2 shows that all models analyzed capture the equatorward shift in the location of the  $v_{10}^{ST}$  maxima as compared to those at 850 hPa by showing the average latitude for the region occupied by the top 10% of  $v_{10}^{ST}$  and  $v_{850}^{ST}$ . Note that similar results occur when either larger or smaller percentiles are used for the analysis. Figure 2 also shows that the mean locations of the storm track maxima at the two levels covary from model to model. This is a result that will be shown in subsequent analysis as well.

*b. Spatial location of the surface storm tracks*

Next we examine the spatial correlations between the models and reanalysis. The variables examined for spatial correlations are  $v_{10}^{ST}$ ,  $v_{850}^{ST}$ ,  $T_{DIFF}$ ,  $|\nabla SST|$ , and  $\sigma_{BI}$ . For this, we only consider grid points that are over the oceans. Also, because the physics that we are interested

in occur near the WBCs, we limit the region used in the correlation (as shown in Figs. 4a, 6a, and 7a). We also examined the correlations for the entire region. The main results of the analysis do not change. However, the spatial correlations of the storm tracks increase when the larger region is considered due to the fact that the storm tracks universally weaken at the north and south edges of the regions (e.g., Fig. 1).

Figure 3 shows the results of the spatial correlation analysis per basin. For all basins except the Indian Ocean in DJF, the storm tracks have the strongest correlations. The higher skill for the storm tracks is partially attributable to the significant SST biases in the climate models (as discussed in the introduction), which impacts  $T_{DIFF}$ ,  $|\nabla SST|$ , and  $\sigma_{BI}$ . However, it is also the case that all of the correlations above 0.6 are statistically significant at the 95 percentile, based on the Student’s *t* test after a Fisher transformation of a Pearson correlation coefficient. The temperature related fields are significant despite their lower values because the degrees of freedom for those fields are greater than those for the storm tracks. This is due to the fact that the storm tracks are a spatially smooth

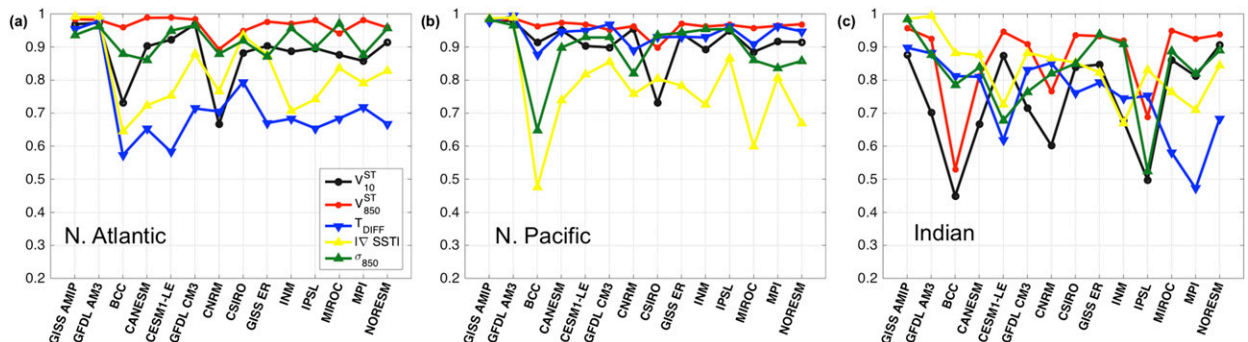


FIG. 3. Spatial correlations between ERA-I and each model for  $v_{10}^{ST}$ ,  $v_{850}^{ST}$ ,  $T_{DIFF}$ ,  $|\nabla SST|$ , and  $\sigma_{BI}$  (see text for details). All panels show DJF. The legend shown in the North Atlantic panel holds for all basins. (The regions used for the correlations are shown in Figs. 4a, 6a, and 7a.)

TABLE 2. Model-to-model correlations of the spatial correlations shown in Fig. 3 between the pair variables listed. Asterisks indicate the correlations that are significant at the 95% based on the Student's  $t$  test after a Fisher transformation of the Pearson correlation coefficient.

	$v_{10}^{\text{ST}}$	$v_{850}^{\text{ST}}$	$T_{\text{DIFF}}$	$ \nabla\text{SST} $
North Atlantic				
$v_{850}^{\text{ST}}$	0.80*	—	—	—
$T_{\text{DIFF}}$	0.53*	0.25	—	—
$ \nabla\text{SST} $	0.64*	0.24	0.88*	—
$\sigma_{\text{BI}}$	0.61*	0.24	0.49	0.52*
North Pacific				
$v_{850}^{\text{ST}}$	0.95*	—	—	—
$T_{\text{DIFF}}$	0.39	0.43	—	—
$ \nabla\text{SST} $	0.37	0.31	0.87*	—
$\sigma_{\text{BI}}$	0.14	0.16	0.76*	0.83*
Indian Ocean				
$v_{850}^{\text{ST}}$	0.90*	—	—	—
$T_{\text{DIFF}}$	-0.09	-0.09	—	—
$ \nabla\text{SST} $	0.06	-0.03	0.82*	—
$\sigma_{\text{BI}}$	0.59*	0.54*	0.33	0.32

field compared to the temperature fields, and thus have larger serial correlations.

Next we use the spatial correlations to consider physical forcing of  $v_{10}^{\text{ST}}$  from aloft and from the surface. If we assume that any physical link between the spatial patterns between  $v_{10}^{\text{ST}}$ ,  $v_{850}^{\text{ST}}$ ,  $T_{\text{DIFF}}$ ,  $|\nabla\text{SST}|$ , and  $\sigma_{\text{BI}}$  exists on the spatial scale of the WBC regions, then we can use the model-to-model variability in the spatial correlations to test for relationships between these variables. Table 2 shows the intermodel correlation of the spatial correlations, that is,  $R(\text{var1}, \text{var2}) = \text{corr}[r^{\text{SP}}(\text{var1}), r^{\text{SP}}(\text{var2})]$ , where var1 and var2 are variables listed above and  $r^{\text{SP}}$  indicates a spatial correlation for that variable. The metric  $R(\text{var1}, \text{var2})$  will be referred to as the model-to-model covariability. The interpretation of this metric is shown with the following example: if the models that generate  $v_{850}^{\text{ST}}$  well do the same for  $v_{10}^{\text{ST}}$ , and those that do poorly on  $v_{850}^{\text{ST}}$  also do poorly on  $v_{10}^{\text{ST}}$ , then there will be strong model-to-model covariability of the spatial correlations, and we argue that this suggests a physical link (in the models) between the two variables. On the other hand, if, for example, the models capture  $v_{10}^{\text{ST}}$  well despite doing a poor job of capturing  $|\nabla\text{SST}|$  then  $R(v_{10}^{\text{ST}}, |\nabla\text{SST}|)$  will be small, and that suggests that  $|\nabla\text{SST}|$  does not have a strong influence on the storm tracks in the models.

The results in Table 2 indicate that across the ocean basins the strongest model-to-model covariability for  $v_{10}^{\text{ST}}$  occurs with  $v_{850}^{\text{ST}}$ . In each basin, the model-to-model correlations between the surface and 850-hPa spatial correlation with ERA-I are statistically significant at the 99th percentile. To derive the statistical significance of the correlations, the Student's  $t$  test is applied to the

Fisher-transformed correlation coefficients, which are Pearson correlations.

The strong covariability of the differences with ERA-I for the surface and 850 hPa implies that any surface forcing would be a secondary influence on the surface storm tracks. This secondary forcing can be seen in the CMIP5 models in the North Atlantic, where the model-to-model correlations of the surface storm track and  $T_{\text{DIFF}}$ , and  $|\nabla\text{SST}|$ , and  $\sigma_{\text{BI}}$  are all statistically significant at the 95%. This result appears to be related to SST biases in the Gulf Stream extension and North Atlantic Current region and is discussed in detail in section 4.

Table 2 also shows that there is high model-to-model covariability between  $T_{\text{DIFF}}$  and  $|\nabla\text{SST}|$  biases in all of the ocean basins. This result suggests that the SST is strongly reflected in the spatial patterns of the lower-tropospheric stability: biases in the spatial pattern of SST translate to biases in the spatial pattern of  $T_{\text{DIFF}}$ . One might ask if the forcing is the other direction:  $T_{\text{DIFF}}$  bias generating surface flux biases that change the SST. However, our analysis of the surface fluxes found that climatological biases in the fluxes were acting to dampen the SST biases (not shown).

The baroclinicity, despite being calculated at 850 hPa, also has a strong model-to-model covariability with  $T_{\text{DIFF}}$  and  $|\nabla\text{SST}|$  in the North Pacific. As will be shown below, in the North Pacific the maxima for these three variables are located close to one another. Thus, if a model has a bias that affects one, it will most likely impact all three.

In the Indian Ocean for DJF, there a strong model-to-model covariability in the spatial correlations of  $\sigma_{\text{BI}}$  and the storm track, which agrees with Nakamura and Shimpo (2004). This result is much weaker in JJA (not shown), due to the additional influence of the baroclinicity and low-level stability associated with sea ice near Antarctica [discussed in Booth et al. (2010)].

The next analysis focuses on the region of the strongest storm track per basin. Figure 4 shows  $T_{\text{DIFF}}$ ,  $|\nabla\text{SST}|$ , and  $\sigma_{\text{BI}}$  for the North Atlantic (Figs. 4a–c) for ERA-I. Figures 4d–f shows the estimated surface storm track (defined in section 2) using each of the variables from Figs. 4a–c. The  $v_{10}^{\text{ESTIMATE}}$  using  $T_{\text{DIFF}}$  is able to capture more of the hooklike shape of the top 10th percentile of the  $v_{10}^{\text{ST}}$  location, as compared to the other two variables. However, it also predicts that the surface storm track should extend farther north than it does. Both  $|\nabla\text{SST}|$  and  $\sigma_{\text{BI}}$  generate an estimated storm track that is south of the 850-hPa storm track. However, they predict a storm track in the shelf water region north of the Gulf Stream, which is incorrect.

We quantify the skill of the estimated storm tracks by calculating their overlap with  $v_{10}^{\text{ST}}$  (see section 2 for



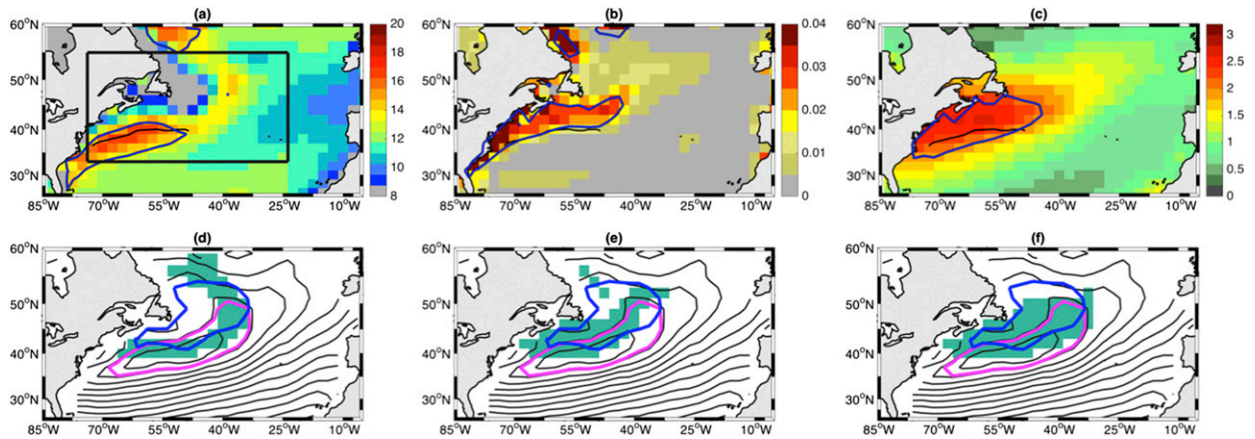


FIG. 4. Shown are (a)  $T_{\text{DIFF}}$  (K), (b)  $|\nabla\text{SST}|$  ( $\text{K km}^{-1}$ ), and (c)  $\sigma_{\text{BI}}$  ( $\text{day}^{-1}$ ) for the North Atlantic in ERA-I for DJF. (d)–(f) Actual (black contours) and estimated storm tracks (blue-green shading) using each variable from (a)–(c), respectively. Blue contour shows the top 10th percentile for  $v_{850}^{\text{ST}}$ . Magenta contour shows the top 10th percentile for  $v_{10}^{\text{ST}}$ . The cut-off percentile used so that the estimated storm track is equal in size to the top 10th percentile is 24% for  $T_{\text{DIFF}}$ , 24% for  $|\nabla\text{SST}|$ , and 21% for  $\sigma_{\text{BI}}$ . In (a)–(c), the dark blue contour shows the top 10th percentile, and the black line shows the Gulf Stream path. The box in (a) shows the region used for spatial correlations. Because the actual storm tracks are shown only for comparison with the estimated storm track, we exclude values for the contours.

details). For comparison, we also calculate the overlap of  $v_{10}^{\text{ST}}$  and  $v_{850}^{\text{ST}}$ . To aid in comparison with the reanalysis, in Fig. 5 we show the difference between the overlap of  $v_{10}^{\text{ST}}$  and the estimated storm track with that of  $v_{10}^{\text{ST}}$  and  $v_{850}^{\text{ST}}$  for ERA-I. In the models, nearly all of the estimated storm tracks overlap with the location of the actual storm track  $v_{10}^{\text{ST}}$  more than  $v_{850}^{\text{ST}}$  does (as the yellow bars have the smallest values), and the estimates using  $T_{\text{DIFF}}$  or  $\sigma_{\text{BI}}$  perform best in most cases (Fig. 5).

For the North Pacific, Fig. 6 shows  $T_{\text{DIFF}}$ ,  $|\nabla\text{SST}|$ , and  $\sigma_{\text{BI}}$  from ERA-I for reference, as well as the predicted storm track for each variable. Unlike the Gulf Stream, there is no northward moving current at the terminus of the KOE, and as such the SST and the atmospheric stability above the KOE are very zonal. Figure 5b shows that the estimated storm track using  $T_{\text{DIFF}}$  and  $\sigma_{\text{BI}}$  has a stronger overlap with the location of the surface storm track in comparison to  $v_{850}^{\text{ST}}$ . This result is consistent with the physical forcing associated with momentum mixing occurring preferentially in the less stable regions (as in Booth et al. 2010). Figure 5b also shows negative values for overlap of  $v_{10}^{\text{ST}}$  versus  $v_{850}^{\text{ST}}$  for the models as compared to reanalysis in the North Pacific; however, the overlap in the reanalysis is stronger than in any other basin (47% vs 30% in the North Atlantic and 7% in the Indian Ocean). Thus, for our purposes here, the important result is that the overlap is more realistic for the estimated storm track than the actual surface storm track in each model. This implies that the spatial pattern of the surface storm track in the models resembles a combination of the 850-hPa storm track and the SST related variable more than it resemble the 850-hPa storm track alone.

We also note that in the North Pacific the maximum in the ocean current is not collocated with the strongest SST gradient (e.g., Yasuda 2003). This is apparent in Fig. 6b, which shows no overlap in the location of the KOE based on altimetry data (which captures the current) and the maximum in  $|\nabla\text{SST}|$ . Because our analysis examines the SST, we cannot comment directly on the ocean currents, but previous research has shown that coarse-resolution models like those used in this study produce a single, merged front that has both the strong ocean current and the SST front rather than having separate Kuroshio and Oyashio Extension fronts (e.g., Thompson and Kwon 2010). In the analysis presented here, the strong collocation of the  $T_{\text{DIFF}}$ ,  $|\nabla\text{SST}|$ , and  $\sigma_{\text{BI}}$  may be a result of the merged locations of the SST gradient and the ocean currents in the CMIP5 models, and for ERA-I it relates to the reduced spatial resolution we use for the analysis.

In the Indian Ocean, the maximum in  $T_{\text{DIFF}}$  extends from the coast of South Africa toward the ARC (Fig. 7a). The maximum in  $|\nabla\text{SST}|$  is situated along the ARC (Fig. 7b), while the  $\sigma_{\text{BI}}$  maximum is located farther south (Fig. 7c), due in part to the large weakened atmospheric stability over the cold water south of the ARC. The surface storm track maximum is almost completely dislocated from the maximum for  $v_{850}^{\text{ST}}$  (Fig. 7d). Both the SST gradient and the baroclinicity estimates are able to capture the surface storm track, while  $T_{\text{DIFF}}$  instead creates a pattern that includes a strong storm track to the south of the  $v_{850}^{\text{ST}}$  maximum. This pattern resembles the surface storm track in JJA (not shown). However, the percent

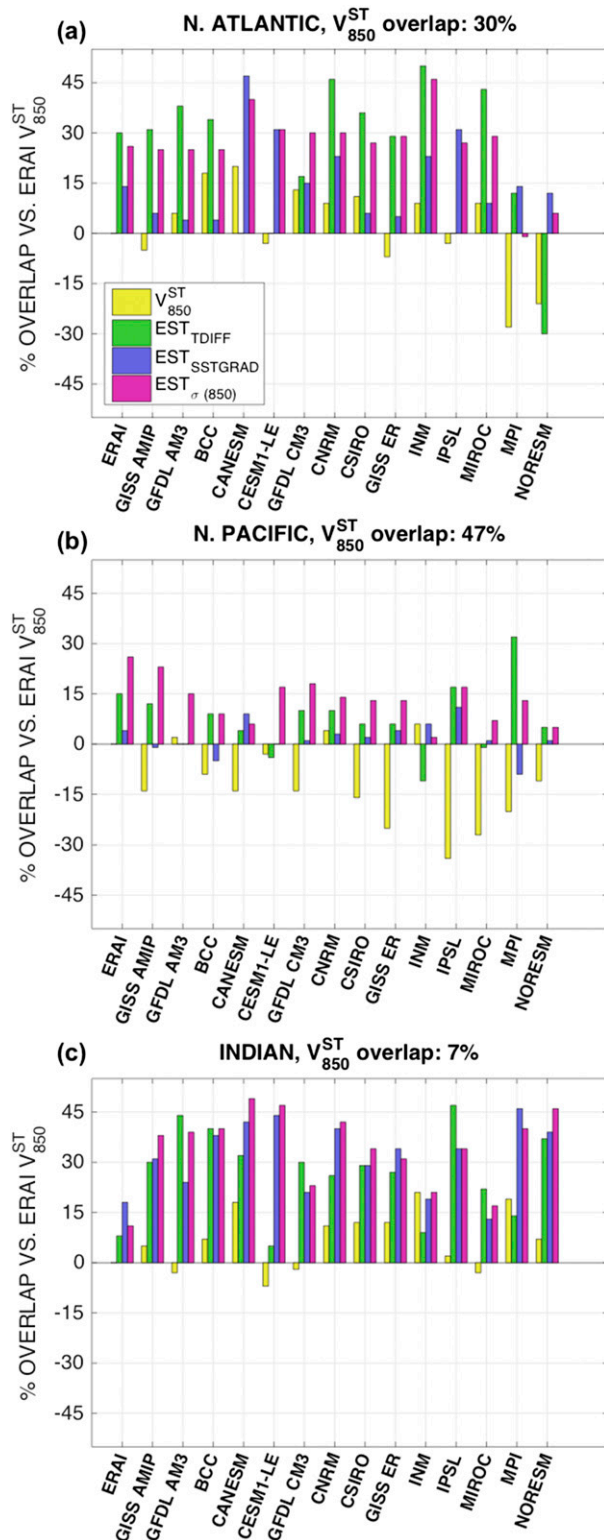


FIG. 5. Percent overlap of each estimated storm track and the top 10th percentile of  $v_{10}^{\text{ST}}$ . Values shown are anomalies with respect to the overlap between the top 10th percentile of  $v_{850}^{\text{ST}}$  and  $v_{10}^{\text{ST}}$  for ERA-I per basin. The value of this overlap is given in the title of each panel: 30% for North Atlantic, 47% for North Pacific, and 7% for Indian Ocean. For the estimated storm tracks, a region of equivalent size to the top 10th percentile is used (see section 2 for details). All panels show DJF.

overlap scores show that there is skill added by using any of the estimated storm tracks (Fig. 5d). In JJA, the observed and modeled surface storm track has two maxima, one near the ARC and another near the sea ice edge. The estimated storm tracks are able to capture this pattern.

### c. Amplitude of $v_{10}^{\text{ST}}$

The global maps of the storm tracks (Fig. 1) show that GFDL CM3 and CCSM1-LE differ significantly in their representation of the strength of  $v_{10}^{\text{ST}}$ , with GFDL CM3 more closely matching ERA-I. This leads to one of the motivations for this research: what sets the strength of the surface storm track? To help answer this, we calculate the average value of the top 10% for the storm track at the surface and in the free troposphere (for both 850 and 500 hPa) per ocean basin and use it as a measure of storm track strength. We have repeated the analysis using the top 5% and top 25% and the results presented below remain the same.

Figure 8 shows the strength of the storm track for  $v_{500}^{\text{ST}}$  versus  $v_{850}^{\text{ST}}$ , and  $v_{850}^{\text{ST}}$  versus  $v_{10}^{\text{ST}}$ , per model. Focusing first on the comparison of the storm track aloft, one can see a strong linear relationship. In each basin the correlation is statistically significant at the 95% level using the Student's  $t$  test and indicates that the strength of the storm track at 500 hPa is a good estimator for that at 850 hPa and vice versa. Moving now to the comparison of  $v_{850}^{\text{ST}}$  versus  $v_{10}^{\text{ST}}$ , Fig. 8 shows that four models create surface storm tracks that are stronger than the storm tracks at 850 hPa. There are no outliers in the  $v_{500}^{\text{ST}}$  versus  $v_{850}^{\text{ST}}$  relationship, which implies that the surface storm track bias in the four models that are outliers is a boundary layer problem.

A linear analysis of  $v_{850}^{\text{ST}}$  versus  $v_{10}^{\text{ST}}$  strength excluding the outlier models (i.e., models 5, 6, 9, and 15 in Fig. 8) shows that there is no relationship between the strength of the storm tracks aloft and that at the surface in the North Atlantic (Fig. 8a). In the North Pacific and Indian Ocean, a weak linear relationship exists. The correlation in the North Pacific is not statistically significant beyond the 95%. In the Indian Ocean, the statistically significant correlation coefficient is a result of two models only (3 and 11 in Fig. 8), with the other models clustered together in no relationship. Thus, the general result here is that the amplitude of surface storm track in the vicinity of its maximum cannot be predicted by the strength of the storm track at 850 hPa, which implies significant influence from the SST as was the case for the spatial pattern in the previous subsection.

For  $v_{10}^{\text{ST}}$ , outlier models in all three ocean basins are the same: CANESM, CESM1-LE, CSIRO, and NORESM.

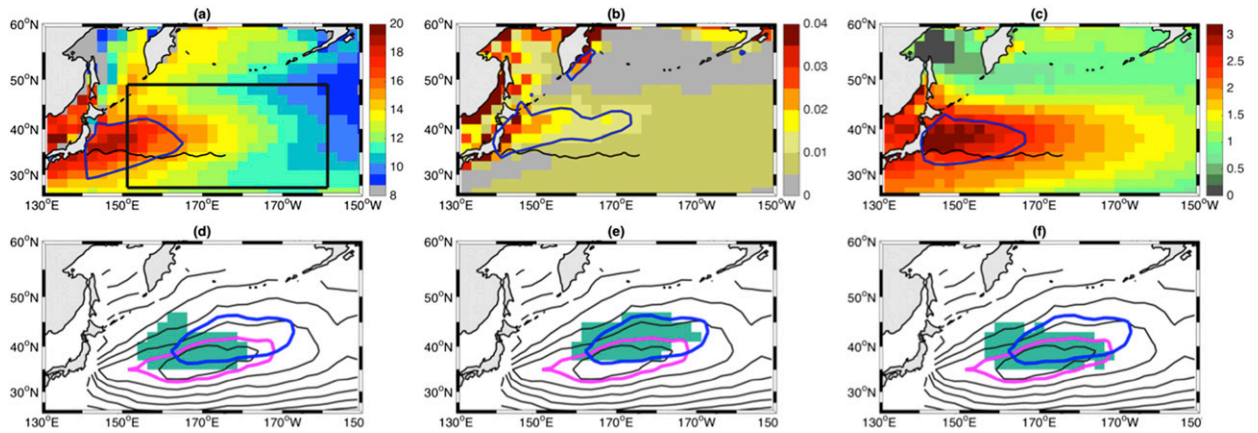


FIG. 6. As in Fig. 4, but for the North Pacific. The cut-off percentile used so that the estimated storm track is equal in size to the top 10th percentile is 35% for  $T_{DIFF}$ , 20% for  $|\nabla SST|$ , and 29% for  $\sigma_{BI}$ . In (a)–(c), the dark blue contour shows the top 10th percentile, and the black line shows the KOE path.

In all of these models, the mean of the top 10% of  $v_{10}^{ST}$  is near equal to or stronger than that of  $v_{850}^{ST}$ . The first three models listed also generate monthly-mean 10-m zonal winds that are too strong as compared to reanalysis (not shown). For two of these models, CESM1-LE and NORESM, the data provided to the CMIP5 archive are the wind at the lowest model level, rather than the 10-m winds. As discussed above, this is unlikely to be the only or dominant cause of the wind strength bias, because we know from separate studies that CESM1-LE is consistently too strong in its surface winds. We note that NORESM (Bentsen et al. 2013) is based on the Community Climate System Model, version 4, which is the predecessor to CESM1-LE.

Figure 8 also shows that the majority of the models are weaker than the reanalysis in their storm track maximum at 850 and 500 hPa (e.g., examine the values along the  $x$  axis vs model name in Fig. 8). This result has been shown previously by Chang et al. (2012); however, we mention it here as a contrast to the storm-track strength at the surface.

We also examined the strength of the top 10th percentile for the North Atlantic compared to the North Pacific, per model (Fig. 9). The strong linear relationship at 850 hPa suggests that the model-to-model variability in storm track strength is mostly independent of the basin (i.e., the model differences span the hemisphere). On the other hand, there is only a weak linear

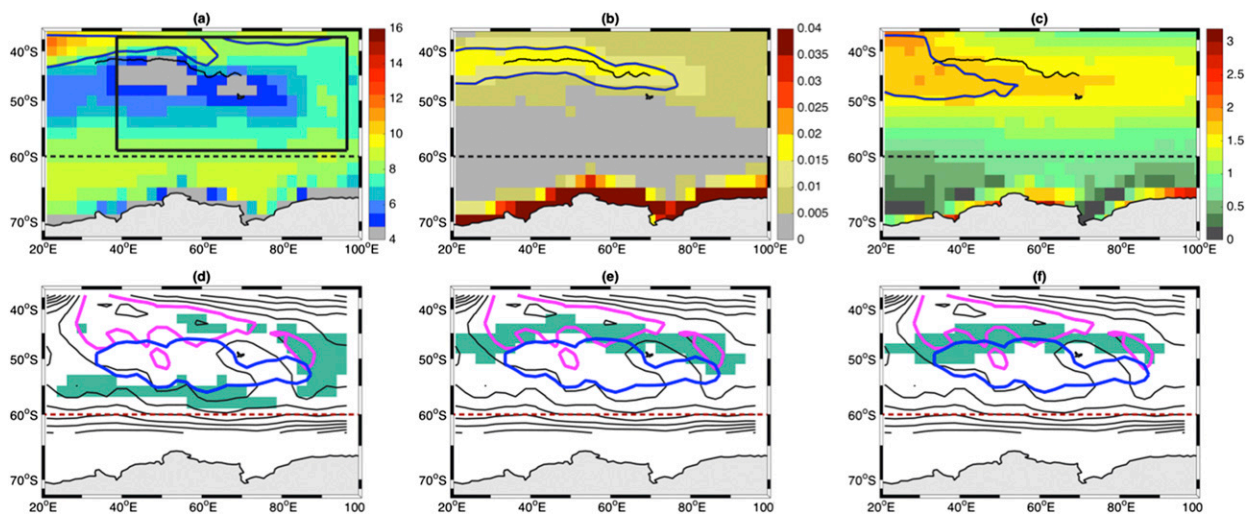


FIG. 7. As in Fig. 4, but for the Indian Ocean. The cut-off percentile used so that the estimated storm track is equal in size to the top 10th percentile is 57% for  $T_{DIFF}$ , 48% for  $|\nabla SST|$ , and 44% for  $\sigma_{BI}$ . In (a)–(c), the dark blue contour shows the top 10th percentile, and the black line shows the ARC.

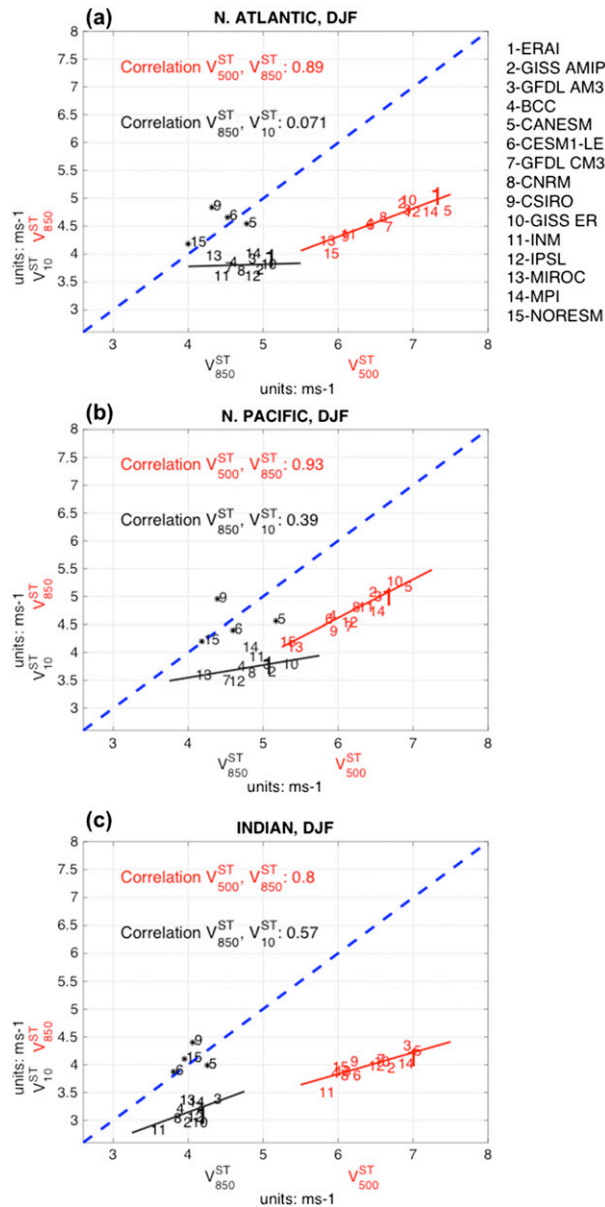


FIG. 8. Scatterplots for the storm track intensities ( $\text{m s}^{-1}$ ) at surface vs 850 hPa (black letters) and 850 vs 500 hPa (red letters) in the (a) North Pacific, (b) North Atlantic, and (c) Indian Ocean. The blue dashed lines indicate 1:1 relationship. The black lines are the least squares fit to the 850-hPa vs 10-m data, excluding models numbered 5, 6, 9, and 15 in the legend. The red lines are the least squares fit to the 500- vs 850-hPa data. Models with surface storm tracks that are too large have a black asterisk next to their number. For all of the other models, the bottom-right corner of the number corresponds to the location of the data point. Note that the correlation between  $v_{10}^{ST}$  and  $v_{850}^{ST}$  in (c) is zero if models numbered 3 and 11 are excluded.

relationship at the surface when excluding the four outliers, which implies some influence from the ocean surface, which is likely distinct between the two basins, on the strength of  $v_{10}^{ST}$ . We did not find a significant

correlation between storm track strength in the Northern and Southern Hemispheres at 850 hPa (not shown).

#### 4. Discussion

The model-to-model correlation analysis (i.e., Table 2) suggests that the modeled North Atlantic surface storm track was influenced by biases in the modeled SST (albeit secondarily compared to the influence of the 850-hPa storm track). Figure 10 explores this issue by analyzing the multimodel means for SST,  $T_{\text{DIFF}}$ , and the storm tracks as compared to reanalysis. Because of the findings shown in Fig. 8 regarding the large bias in the strength of the surface storm tracks in four of the GCMs, the multimodel mean in Fig. 10 excludes those four models and the AMIP models.

Figure 10a shows that the models are too warm in the shelf water region, indicative of the Gulf Stream separation problem in the models. The models are also too cold in the North Atlantic Current (NAC) region, most likely because they do not have the proper northward warm advection generated by the NAC due to an overly zonal and southerly NAC path. Figure 10b shows that  $T_{\text{DIFF}}$  has many of the same biases as SST. In the multimodel mean, the differences in the surface turbulent heat fluxes (compared to reanalysis) showed the surface fluxes in the models act to dampen the SST biases (not shown). Thus, the SST and  $T_{\text{DIFF}}$  biases are related to ocean circulation issues, as highlighted in previous work (e.g., Kelly et al. 2010; Kwon et al. 2010).

Meanwhile, the difference plots for the storm tracks (Figs. 10c,d) show that the models are too weak on the poleward flank of the storm tracks and too strong in the Azores region. This difference between the storm tracks and reanalysis partially relates to a long-standing issue of the GCM storm tracks being too zonal (e.g., Ulbrich et al. 2008). However, the biases for  $v_{10}^{ST}$  and  $v_{850}^{ST}$  differ in the region of the shelf water. The 850-hPa storm track is too weak and it is statistically significant, whereas the surface storm track is too strong, although it is not strong enough to be ruled different from random error.

These differences in the shelf water can be examined through a different perspective, by considering the error in the storm tracks normalized by the error in  $v_{850}^{ST}$ . Given our findings that show that the 850-hPa storm track has a strong influence on the surface storm track, one might consider whether biases in  $v_{850}^{ST}$  show up as biases in  $v_{10}^{ST}$ . If this is the case, analyzing  $v_{10}^{ST}$  normalized by  $v_{850}^{ST}$  per model might give a better representation of surface relationship. Therefore, we introduce a new metric: the ratio of  $v_{10}^{ST}$  to  $v_{850}^{ST}$  (Fig. 10e). If the ratio is larger than 1, the surface storm track is stronger, and vice versa.

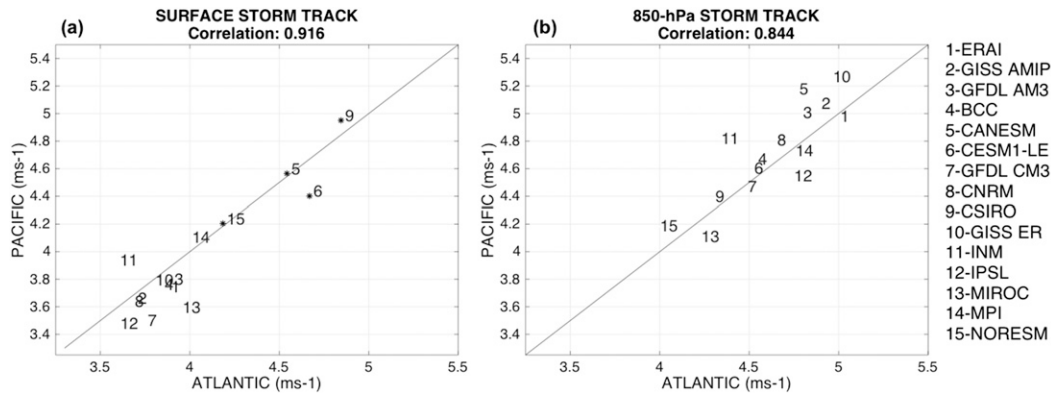


FIG. 9. Scatterplots for the North Pacific vs North Atlantic in terms of the storm track intensities ( $\text{m s}^{-1}$ ) at (a) surface and (b) 850 hPa. Models with surface storm tracks that are too large have a black asterisk before their number in (a).

If momentum mixing is the dominant physical mechanism creating differences in the spatial location of the surface and 850-hPa storm track (as suggested in our results above), then we might expect that in regions where the models mix too much the surface storm track is biased too strong (relative to that model's  $v_{850}^{\text{ST}}$ ). Figure 10f shows the difference between the storm-track ratio from the multimodel mean and ERA-I. Locations where Fig. 10f are positive indicate that the models have a stronger surface storm track (when normalized by their 850-hPa storm track) than the reanalysis. Given that momentum mixing affects the strength of the surface storm track (Booth et al. 2010; Liu et al. 2013), we interpret the similarities between Figs. 10f and 10b, especially off the U.S. East Coast, the Labrador Sea, and northwestern Europe, as a strong indication that model biases in  $T_{\text{DIFF}}$  create biases in momentum mixing and these impact the surface storm tracks. Similar results are not apparent in multimodel mean biases for the North Pacific (not shown). However, biases in the multimodel mean  $T_{\text{DIFF}}$  are negligible there. This might also explain the lack of significance in the model-to-model correlations for the temperature variables and the surface storm track in the North Pacific (Table 2). In the Indian Ocean during DJF, we find a result that is similar to the North Atlantic: biases in  $T_{\text{DIFF}}$  are collocated with biases in the storm track ratio (not shown). However, the model-to-model correlations shown in Table 2 suggest that forcing from the SST bias is primarily detectable in the North Atlantic. Taken together, Table 2 plus the analysis presented in Fig. 10 give a strong indication that SST biases in the North Atlantic are apparent in the biases in the spatial patterns of the models surface storm track. The physical causes of the SST bias are not explored herein; however, it is highly likely to be caused by biases

in the ocean currents, specifically the Gulf Stream and North Atlantic Current path, associated with the biases in the modeled wind stress and coarse-resolution bathymetry.

Given the fact that climatologically WBC regions have the largest turbulent surface heat fluxes out of the ocean, we also explore the relationship between the surface storm track,  $T_{\text{DIFF}}$ , 10-m zonal wind ( $U_{10}$ ), and the fluxes. Turbulent heat flux includes latent and sensible heating, however, the sensible heat flux (SHF) is directly proportional to  $T_{\text{DIFF}}$  and more likely to reflect the local surface heating that would drive boundary layer momentum mixing. With this in mind, we focus here on SHF. (However, the results do not change drastically if we use total turbulent flux.) The North Atlantic and Indian Ocean have a weak linear relationship for the top 10% in  $v_{10}^{\text{ST}}$  versus the top 10% for SHF. However, any meaningful statistical significance is lost if we remove the four models in which the surface storm track amplitudes are biased too strong. In the North Pacific, there is no linear relationship between the SHF and the surface storm track strength. Given the lack of statistically significant results, we do not show figures from this analysis. A similar result holds if we examine  $T_{\text{DIFF}}$  versus the fluxes or  $U_{10}$  versus the fluxes. Thus, the GCMs that have surface winds that are too strong are not similarly impacted in their turbulent flux fields. Part of the reason for this may be differences in the surface drag coefficients for the momentum and for the heat fluxes. If these are parameterized differently in the models, then biases in the surface winds would not necessarily correlate with biases in the fluxes.

Separate from the flux issue, we note that the AMIP models included in the analysis were able to capture the storm track with more fidelity than the CMIP models for

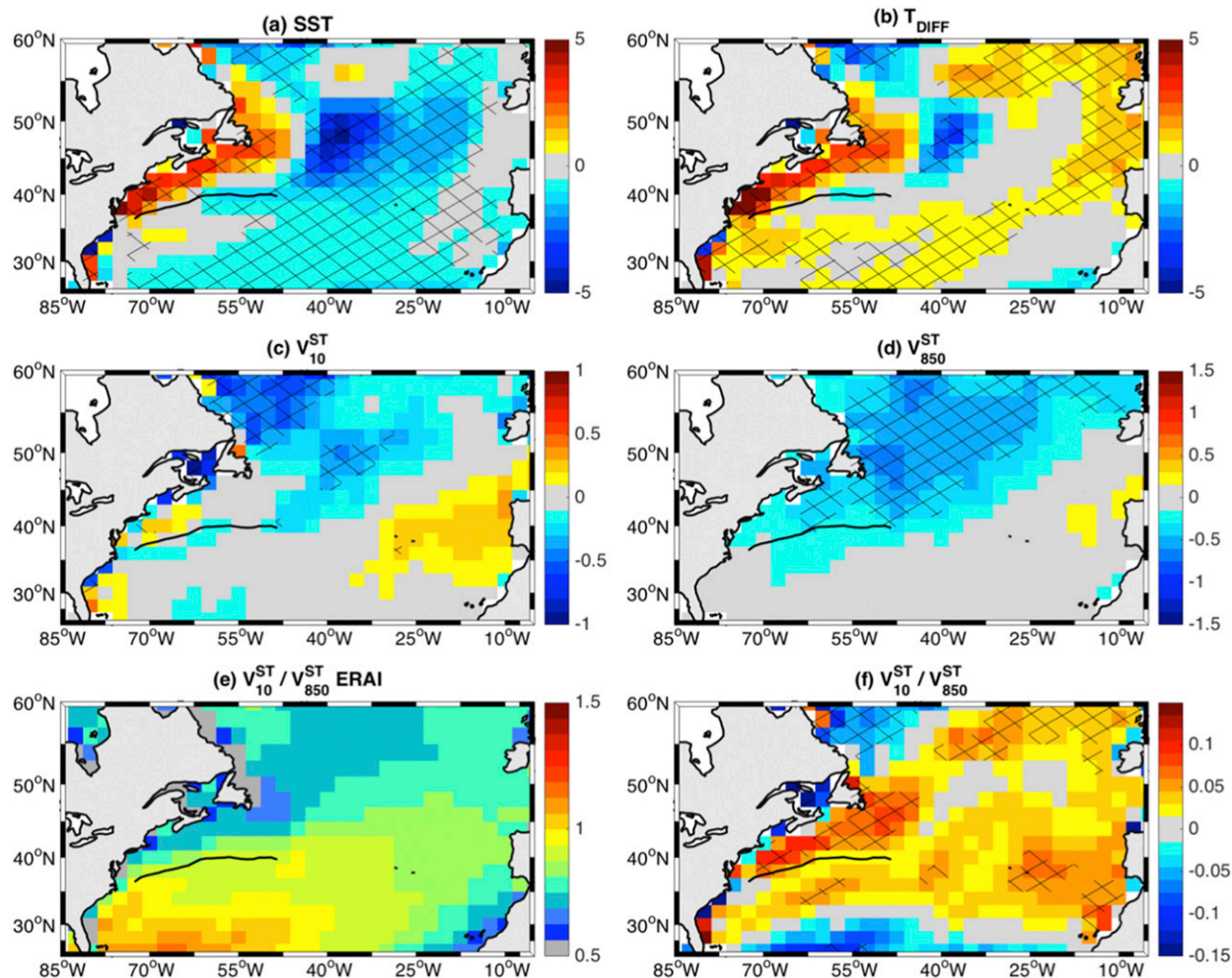


FIG. 10. Multimodel mean minus reanalysis for (a) SST (K), (b)  $T_{\text{DIFF}}$  (K), (c)  $v_{10}^{\text{ST}}$  ( $\text{m s}^{-1}$ ), (d)  $v_{850}^{\text{ST}}$  ( $\text{m s}^{-1}$ ), and (f) the ratio  $v_{10}^{\text{ST}}/v_{850}^{\text{ST}}$  and (e) the ratio  $v_{10}^{\text{ST}}/v_{850}^{\text{ST}}$  for ERA-I in the North Atlantic. In (a)–(d) and (f), cross hatching indicates statistically significant differences using a threshold at 95%. The black line indicates the location of the Gulf Stream. The multimodel mean is based on CMIP models only and excludes the four models with large biases in the strength of the surface storm track (see Fig. 8).

the North Atlantic and North Pacific (Fig. 3). For the Indian Ocean, this was true in JJA; however, in DJF the GFDL AM3 model performed worse than some of the CMIP models. The AMIP models also capture the 850-hPa baroclinicity more realistically than the coupled models (Fig. 3), which suggests that the SST has an appreciable influence on the spatial distribution of the 850-hPa baroclinicity.

Finally, our analysis found no relationship between the strength or spatial representation of the storm track and atmospheric model horizontal resolutions (e.g., Table 1). Studies focused on a single model found that the strength of the storm track increases with finer resolution (e.g., Champion et al. 2011). The horizontal resolution of CMIP5 GCMs may not be fine enough to properly capture physics within the storms

(Willison et al. 2013). However, the lack of a relationship between the grid spacing and the storm tracks might also be impacted by other factors that influence storm track location, such as the stationary wave pattern (Brayshaw et al. 2009).

## 5. Summary

Analysis of the surface storm tracks in the CMIP5 models shows that the models capture the equatorward shift in the location of storm track maximum relative to the storm track maximum at 850 hPa. The result holds for all ocean basins in DJF, however in the Indian Ocean in JJA, the pattern is obscured by the influence of the sea ice margin on the storm track. To analyze what might generate the equatorward shift in region of the

maximum values we refine the definition of the estimated storm track metric and define a skill score to quantify its relationship to the actual storm track. If the estimated storm track is generated based on the overlapping region between the 850-hPa storm track and the 850-hPa baroclinicity, then it captures the equatorward shift in location. For many of the models, an estimated storm track in which the 850-hPa baroclinicity is replaced by the temperature difference between the surface and 850-hPa is equally successful, suggesting an influence of atmospheric stability in driving this equatorial shift. Thus, the models' surface storm tracks spatial pattern more closely resembles a combination of the 850-hPa storm track and the baroclinicity or  $T_{\text{DIFF}}$  rather than the 850-hPa storm track alone. Replacing the baroclinicity with the SST gradient for the calculation of the estimated storm track degrades the skill of the estimated storm track. This suggests that the air–sea stability and stratification influence, rather than the influence of the SST gradient, generate the physical mechanism that shifts the surface storm tracks equatorward in the models.

Analysis of the amplitude of the storm tracks shows that the modeled 850-hPa storm track is stronger than that at the surface. However, there are four outlier models that generate a surface storm track whose strength exceeds that of the 850-hPa storm track. This bias in strength also occurs in unfiltered surface winds in the models, but it does not translate to large biases in the surface turbulent heat fluxes. No statistically significant relationship is found between the strength of the surface and 850-hPa storm tracks, even if the outlier models are excluded. However, there is a strong linear relationship across models between the strength of the storm tracks at 500 and 850 hPa, and there are no outlier models. These analyses suggest that the strength of the surface storm track maxima is controlled by more than just the strength of the free-tropospheric storm track in the majority of the CMIP5 models, and that the issues in the boundary layer or surface physics in the models most likely cause the surface storm-track biases in the outlier models.

We analyzed the spatial correlations between the models and ERA-I for the storm tracks as well as on a set of temperature-related fields that are influenced by the SST in the WBC regions. Our analysis indicates that the models capture the spatial patterns of the storm tracks with more fidelity than they do for  $T_{\text{DIFF}}$ ,  $|\nabla\text{SST}|$ , and  $\sigma_{\text{BI}}$ ; however, in most of the cases, the spatial correlations were statistically significant. A subsequent study of the model-to-model covariability of the spatial correlations shows that 1) models with strong or weak biases in the spatial pattern of the 850-hPa storm tracks tend to have

similar biases in the surface storm tracks, and 2) for the North Atlantic, the across-model covariability of the spatial regressions of  $T_{\text{DIFF}}$  and/or  $\sigma_{\text{BI}}$  and the surface storm tracks is also strong. Thus, in the North Atlantic we find indicators suggesting that the biases in the SST create dominant biases in the surface storm track.

An analysis of the multimodel mean using only the CMIP model without the large bias in the surface storm track also shows forcing from the SST biases impact the surface storm track in the North Atlantic. Along the shelf water region, the models' SST is too warm. This creates a weakened surface stability, which creates more momentum mixing. However, the impact that this has on the surface storm tracks is only apparent when we consider a new metric: the ratio of the surface storm track to the storm track at 850 hPa. This is because the primary forcing of the surface storm tracks is the storm track aloft. Thus, significant momentum mixing in a region with a warm SST bias will strengthen the surface storm track in a model. However if the model has a biased weak storm track at 850 hPa, the surface storm track may still appear to be bias weak.

The work here provides metrics for testing the climatology of the surface storm tracks. However, more work is needed using a perturbed physics analysis of a single model, and our group is pursuing such work. Additionally, the analysis here does not isolate individual storms, nor does it focus on the different dynamic and thermodynamic conditions within the warm and cold sectors of storms. Such studies could help in interpreting the relative influence of baroclinicity and the momentum-mixing mechanism.

*Acknowledgments.* We thank ECMWF for providing ERA-Interim via a public server. We acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate modeling groups (see [Table 1](#)) for producing and making available their model output. For CMIP the U.S. Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals. CESM1-LE data were downloaded from the Earth System Grid at NCAR. JFB was partially supported by the NOAA Climate Program Office's Modeling, Analysis, Predictions, and Projections program (Grant NA15OAR4310094). Y-OK was supported by NSF Division of Atmospheric and Geospace Science Climate and Large-scale Dynamics Program (AGS-1355339), NASA Physical Oceanography Program (NNX13AM59G), and DOE Office of Biological and Environmental Research Regional and

Global Climate Modeling Program (DE-SC0014433). RJS was supported by DOE Office of Biological and Environmental Research (DE-SC0006743) and NSF Directorate for Geosciences Division of Ocean Sciences (1419584). We thank the reviewers for useful suggestions and comments that have significantly improved this manuscript.

## REFERENCES

- Bentsen, M., and Coauthors, 2013: The Norwegian Earth System Model, NorESM1-M—Part 1: Description and basic evaluation. *Geosci. Model Dev.*, **6**, 687–720, doi:10.5194/gmdd-5-2843-2012.
- Blackmon, M. L., 1976: A climatological spectral study of the 500 mb geopotential height of the Northern Hemisphere. *J. Atmos. Sci.*, **33**, 1607–1623, doi:10.1175/1520-0469(1976)033<1607:ACSSOT>2.0.CO;2.
- Booth, J. F., L. Thompson, J. Patoux, K. A. Kelly, and S. Dickinson, 2010: The signature of the midlatitude tropospheric storm tracks in the surface winds. *J. Climate*, **23**, 1160–1174, doi:10.1175/2009JCLI3064.1.
- , —, —, and —, 2012: Sensitivity of midlatitude storm intensification to perturbations in the sea surface temperature near the Gulf Stream. *Mon. Wea. Rev.*, **140**, 1241–1256, doi:10.1175/MWR-D-11-00195.1.
- Brachet, S., F. Codron, Y. Feliks, M. Ghil, H. Le Treut, and E. Simonnet, 2012: Atmospheric circulations induced by a midlatitude SST front: A GCM study. *J. Climate*, **25**, 1847–1853, doi:10.1175/JCLI-D-11-00329.1.
- Brayshaw, D. J., B. Hoskins, and M. Blackburn, 2009: The basic ingredients of the North Atlantic storm track. Part I: Land–sea contrast and orography. *J. Atmos. Sci.*, **66**, 2539–2558, doi:10.1175/2009JAS3078.1.
- Champion, A. J., K. I. Hodges, L. O. Bengtsson, N. S. Keenlyside, and M. Esch, 2011: Impact of increasing resolution and a warmer climate on extreme weather from Northern Hemisphere extratropical cyclones. *Tellus*, **63A**, 893–906, doi:10.1111/j.1600-0870.2011.00538.x.
- Chang, E. K. M., S. Lee, and K. L. Swanson, 2002: Storm track dynamics. *J. Climate*, **15**, 2163–2183, doi:10.1175/1520-0442(2002)015<02163:STD>2.0.CO;2.
- , Y. Guo, and X. Xia, 2012: CMIP5 multimodel ensemble projection of storm track change under global warming. *J. Geophys. Res.*, **117**, D23118, doi:10.1029/2012JD018578.
- Chelton, D. B., 2005: The impact of SST specification on ECMWF surface wind stress fields in the eastern tropical Pacific. *J. Climate*, **18**, 530–550, doi:10.1175/JCLI-3275.1.
- , M. G. Schlax, M. H. Freilich, and R. F. Milliff, 2004: Satellite measurements reveal persistent small-scale features in ocean winds. *Science*, **303**, 978–983, doi:10.1126/science.1091901.
- Cione, J. J., S. Raman, and L. J. Pietrafesa, 1993: The effect of Gulf Stream-induced baroclinicity on U.S. East Coast winter cyclones. *Mon. Wea. Rev.*, **121**, 421–430, doi:10.1175/1520-0493(1993)121<0421:TEOGSI>2.0.CO;2.
- Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation systems. *Quart. J. Roy. Meteor. Soc.*, **137**, 553–597, doi:10.1002/qj.828.
- Frankignoul, C., N. Sennéchal, Y.-O. Kwon, and M. A. Alexander, 2011: Influence of the meridional shifts of the Kuroshio and the Oyashio Extensions on the atmospheric circulation. *J. Climate*, **24**, 762–777, doi:10.1175/2010JCLI3731.1.
- Gent, P., and Coauthors, 2011: The Community Climate System Model version 4. *J. Climate*, **24**, 4973–4991, doi:10.1175/2011JCLI4083.1.
- Griffies, S. M., and Coauthors, 2015: Impacts on ocean heat from transient mesoscale eddies in a hierarchy of climate models. *J. Climate*, **28**, 952–977, doi:10.1175/JCLI-D-14-00353.1.
- Guo, Y., E. K. M. Chang, and S. S. Leroy, 2009: How strong are the Southern Hemisphere storm tracks? *Geophys. Res. Lett.*, **36**, L22806, doi:10.1029/2009GL040733.
- Hodges, K. I., R. W. Lee, and L. Bengtsson, 2011: A comparison of extratropical cyclones in recent reanalyses ERA-Interim, NASA MERRA, NCEP CFSR, and JRA-25. *J. Climate*, **24**, 4888–4906, doi:10.1175/2011JCLI4097.1.
- Hoskins, B. J., and P. J. Valdes, 1990: On the existence of storm tracks. *J. Atmos. Sci.*, **47**, 1854–1864, doi:10.1175/1520-0469(1990)047<1854:OTEOST>2.0.CO;2.
- , and K. I. Hodges, 2002: New perspectives on the Northern Hemisphere winter storm tracks. *J. Atmos. Sci.*, **59**, 1041–1061, doi:10.1175/1520-0469(2002)059<1041:NPOTNH>2.0.CO;2.
- , and —, 2005: A new perspective on Southern Hemisphere storm tracks. *J. Climate*, **18**, 4108–4129, doi:10.1175/JCLI3570.1.
- Joyce, T. M., Y.-O. Kwon, and L. Yu, 2009: On the relationship between synoptic wintertime atmospheric variability and path shifts in the Gulf Stream and Kuroshio Extension. *J. Climate*, **22**, 3177–3192, doi:10.1175/2008JCLI2690.1.
- Kelly, K. A., R. J. Small, R. M. Samelson, B. Qiu, T. M. Joyce, Y.-O. Kwon, and M. Cronin, 2010: Western boundary currents and frontal air–sea interaction: Gulf Stream and Kuroshio Extension. *J. Climate*, **23**, 5644–5667, doi:10.1175/2010JCLI3346.1.
- Kwon, Y.-O., and T. M. Joyce, 2013: Northern Hemisphere winter atmospheric transient eddy heat fluxes and the Gulf Stream and Kuroshio–Oyashio Extension variability. *J. Climate*, **26**, 9839–9859, doi:10.1175/JCLI-D-12-00647.1.
- , M. A. Alexander, N. A. Bond, C. Frankignoul, H. Nakamura, B. Qiu, and L. Thompson, 2010: Role of Gulf Stream and Kuroshio–Oyashio systems in large-scale atmosphere–ocean interaction: A review. *J. Climate*, **23**, 3249–3281, doi:10.1175/2010JCLI3343.1.
- Lindzen, R. S., and R. S. Nigam, 1987: On the role of sea surface temperature gradients in forcing low-level winds and convergence in the tropics. *J. Atmos. Sci.*, **44**, 2418–2436, doi:10.1175/1520-0469(1987)044<2418:OTROSS>2.0.CO;2.
- Liu, J.-W., S.-P. Zhang, and S.-P. Xie, 2013: Two types of surface wind response to the East China Sea Kuroshio Front. *J. Climate*, **26**, 8616–8627, doi:10.1175/JCLI-D-12-00092.1.
- Liu, T., and W. Tang, 1996: Equivalent neutral wind. Jet Propulsion Laboratory Publ. 96-17, 22 pp.
- Masunaga, R., H. Nakamura, T. Miyasaka, K. Nishii, and Y. Tanimoto, 2015: Separation of climatological imprints of the Kuroshio Extension and Oyashio fronts on the wintertime atmospheric boundary layer: Their sensitivity to SST resolution prescribed for atmospheric reanalysis. *J. Climate*, **28**, 1764–1787, doi:10.1175/JCLI-D-14-00314.1.
- Minobe, S., A. Kuwano-Yoshida, N. Komori, S. P. Xie, and R. J. Small, 2008: Influence of the Gulf Stream on the troposphere. *Nature*, **452**, 206–209, doi:10.1038/nature06690.
- , M. Miyashita, A. Kuwano-Yoshida, H. Tokinaga, and S.-P. Xie, 2010: Atmospheric response to the Gulf Stream: Seasonal variations. *J. Climate*, **23**, 3699–3719, doi:10.1175/2010JCLI3359.1.
- Nakamura, H., and A. Shimpo, 2004: Seasonal variations in the Southern Hemisphere storm tracks and jet streams as revealed



- in a reanalysis dataset. *J. Climate*, **17**, 1828–1844, doi:10.1175/1520-0442(2004)017<1828:SVITSH>2.0.CO;2.
- , T. Sampe, Y. Tanimoto, and A. Shimpo, 2004: Observed associations among storm tracks, jet streams, and midlatitude oceanic fronts. *Earth's Climate: The Ocean–Atmosphere Interaction*, *Geophys. Monogr.*, Vol. 147, Amer. Geophys. Union, 329–345.
- Nakamura, M., and S. Yamane, 2009: Dominant anomaly patterns in the near-surface baroclinicity and accompanying anomalies in the atmosphere and oceans. Part I: North Atlantic basin. *J. Climate*, **22**, 880–904, doi:10.1175/2008JCLI2297.1.
- Ogawa, F., H. Nakamura, K. Nishii, T. Miyasaka, and A. Kuwano-Yoshida, 2012: Dependence of the climatological axial latitudes of the tropospheric westerlies and storm tracks on the latitude of an extratropical oceanic front. *Geophys. Res. Lett.*, **39**, L05804, doi:10.1029/2011GL049922.
- O'Reilly, C. H., and A. Czaja, 2015: The response of the Pacific storm track and atmospheric circulation to Kuroshio Extension variability. *Quart. J. Roy. Meteor. Soc.*, **141**, 52–66, doi:10.1002/qj.2334.
- Reynolds, R. W., T. M. Smith, C. Liu, D. B. Chelton, K. S. Casey, and M. G. Schlax, 2007: Daily high-resolution-blended analyses for sea surface temperature. *J. Climate*, **20**, 5473–5496, doi:10.1175/2007JCLI1824.1.
- Rienecker, M. M., and Coauthors, 2011: MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications. *J. Climate*, **24**, 3624–3648, doi:10.1175/JCLI-D-11-00015.1.
- Saha, S., and Coauthors, 2010: The NCEP Climate Forecast System Reanalysis. *Bull. Amer. Meteor. Soc.*, **91**, 1015–1057, doi:10.1175/2010BAMS3001.1.
- Sasaki, Y. N., S. Minobe, T. Asai, and M. Inatsu, 2012: Influence of the Kuroshio in the East China Sea on the early summer (baiu) rain. *J. Climate*, **25**, 6627–6645, doi:10.1175/JCLI-D-11-00727.1.
- Schneider, N., and B. Qiu, 2015: The atmospheric response to weak sea surface temperature fronts. *J. Atmos. Sci.*, **72**, 3356–3377, doi:10.1175/JAS-D-14-0212.1.
- Schoonover, J., and Coauthors, 2016: North Atlantic barotropic vorticity balances in numerical models. *J. Phys. Oceanogr.*, **46**, 289–303, doi:10.1175/JPO-D-15-0133.1.
- Small, R. J., and Coauthors, 2008: Air–sea interaction over ocean fronts and eddies. *Dyn. Atmos. Oceans*, **45**, 274–319, doi:10.1016/j.dynatmoce.2008.01.001.
- , and Coauthors, 2014: A new synoptic scale resolving global climate simulation using the Community Earth System Model. *J. Adv. Model. Earth Syst.*, **6**, 1065–1094, doi:10.1002/2014MS000363.
- Spall, M. A., 2007: Midlatitude wind stress–sea surface temperature coupling in the vicinity of oceanic fronts. *J. Climate*, **20**, 3785–3801, doi:10.1175/JCLI4234.1.
- Sweet, W., R. Fett, J. Kerling, and P. LaViolette, 1981: Air–sea interaction effects in the lower troposphere across the north wall of the Gulf Stream. *Mon. Wea. Rev.*, **109**, 1042–1052, doi:10.1175/1520-0493(1981)109<1042:ASIEIT>2.0.CO;2.
- Taguchi, B., H. Nakamura, M. Nonaka, N. Komori, A. Kuwano-Yoshida, K. Takaya, and A. Goto, 2012: Seasonal evolutions of atmospheric response to decadal SST anomalies in the North Pacific subarctic frontal zone: Observations and a coupled model simulation. *J. Climate*, **25**, 111–139, doi:10.1175/JCLI-D-11-00046.1.
- Thompson, L., and Y.-O. Kwon, 2010: An enhancement of low-frequency variability in the Kuroshio–Oyashio Extension in CCSM3 owing to ocean model biases. *J. Climate*, **23**, 6221–6233, doi:10.1175/2010JCLI3402.1.
- Tokinaga, H., Y. Tanimoto, S.-P. Xie, T. Sampe, H. Tomita, and H. Ichikawa, 2009: Ocean frontal effects on the vertical development of clouds over the western North Pacific: In situ and satellite observations. *J. Climate*, **22**, 4241–4260, doi:10.1175/2009JCLI2763.1.
- Ulbrich, U., J. G. Pinto, H. Kupfer, G. C. Leckebusch, T. Spanghel, and M. Reyers, 2008: Changing Northern Hemisphere storm tracks in an ensemble of IPCC climate change simulations. *J. Climate*, **21**, 1669–1679, doi:10.1175/2007JCLI1992.1.
- Voigt, A., and T. A. Shaw, 2015: Circulation response to warming shaped by radiative changes of clouds and water vapour. *Nat. Geosci.*, **8**, 102–106, doi:10.1038/ngeo2345.
- Wallace, J. M., G.-H. Lim, and M. L. Blackmon, 1988: Relationship between cyclone tracks, anticyclone tracks and baroclinic waveguides. *J. Atmos. Sci.*, **45**, 439–462, doi:10.1175/1520-0469(1988)045<0439:RBCTAT>2.0.CO;2.
- , T. P. Mitchell, and C. Deser, 1989: The influence of sea surface temperature on sea surface wind in the eastern equatorial Pacific: Seasonal and interannual variability. *J. Climate*, **2**, 1492–1499, doi:10.1175/1520-0442(1989)002<1492:TIOSST>2.0.CO;2.
- Willison, J., W. A. Robinson, and G. M. Lackmann, 2013: The importance of resolving mesoscale latent heating in the North Atlantic storm track. *J. Atmos. Sci.*, **70**, 2234–2250, doi:10.1175/JAS-D-12-0226.1.
- Xu, H., M. Xu, S.-P. Xie, and Y. Wang, 2011: Deep atmospheric response to the spring Kuroshio over the East China Sea. *J. Climate*, **24**, 4959–4972, doi:10.1175/JCLI-D-10-05034.1.
- Yasuda, I., 2003: Hydrographic structure and variability in the Kuroshio–Oyashio transition area. *J. Oceanogr.*, **59**, 389–402, doi:10.1023/A:1025580313836.