Mitigating climate biases in the mid-latitude North Atlantic by increasing model resolution: SST gradients and their relation to blocking and the jet

Article

Accepted Version

Mitigating climate biases in the mid-latitude North Atlantic by increasing model resolution: SST gradients and their relation to blocking and the jet.

Panos J. Athanasiadis*

CMCC, Bologna, Italy

Fumiaki Ogawa

University of Bergen and BCCR, Bergen, Norway, and Hokkaido University, Sapporo, Japan

Nour-Eddine Omrani

University of Bergen and BCCR, Bergen, Norway

Noel Keenlyside

University of Bergen, NERSC and BCCR, Bergen, Norway

Reinhard Schiemann, Alexander J. Baker, Pier Luigi Vidale

NCAS and University of Reading, Reading, UK

Alessio Bellucci

CMCC, Bologna, Italy, and CNR-ISAC, Bologna, Italy

Early Online Release: This preliminary version has been accepted for publication in *Journal of Climate*, may be fully cited, and has been assigned DOI 10.1175/JCLI-D-21-0515.1. The final typeset copyedited article will replace the EOR at the above DOI when it is published.

Paolo Ruggieri

Department of Physics and Astronomy, Alma Mater Studiorum, University of Bologna, Bologna, Italy, and CMCC, Bologna, Italy

Rein Haarsma

KNMI, Utrecht, Netherlands

Malcolm Roberts

Met Office, Exeter, UK

Chris Roberts

ECMWF, Reading, UK

Lenka Novak

Caltech, Pasadena, USA

Silvio Gualdi

CMCC, Bologna, Italy

[*] *Corresponding author*: Panos Athanasiadis, panos.athanasiadis@cmcc.it

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-21-0515.1.

ABSTRACT

Starting to resolve the oceanic mesoscale in climate models is a step change in model fidelity. This study examines how certain obstinate biases in the midlatitude North Atlantic respond to increasing resolution (from 1° to 0.25° in the ocean) and how such biases in sea surface temperature (SST) affect the atmosphere. Using a multi-model ensemble of historical climate simulations run at different horizontal resolutions, it is shown that a severe cold SST bias in the central North Atlantic, common to many ocean models, is significantly reduced with increasing resolution. The associated bias in the time-mean meridional SST gradient is shown to relate to a positive bias in low-level baroclinicity, while the cold SST bias causes biases also in static stability and diabatic heating in the interior of the atmosphere. The changes in baroclinicity and diabatic heating brought by increasing resolution lead to improvements in European blocking and eddy-driven jet variability. Across the multi-model ensemble a clear relationship is found between the climatological meridional SST gradients in the broader Gulf Stream Extension area and two aspects of the atmospheric circulation: the frequency of high-latitude blocking and the southern-jet regime. This relationship is thought to reflect the two-way interaction (with a positive feedback) between the respective oceanic and atmospheric anomalies. These North Atlantic SST anomalies are shown to be important in forcing significant responses in the midlatitude atmospheric circulation, including jet variability and the stormtrack. Further increases in oceanic and atmospheric resolution are expected to lead to additional improvements in the representation of Euro-Atlantic climate.

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-21-0515.1.

1. Introduction

Making and assessing improvements in the realism of climate models with increasing spatial resolution has been a constant effort in climate modeling. In different phases of the Climate Model Intercomparison Project (CMIP) the representation of a wide variety of physical processes has been improved, partly thanks to increases in resolution, leading to a reduction of various model biases. Certain biases, however, have persisted over different model generations, including a cold sea surface temperature (SST) bias in the broader subpolar gyre area and an underestimation of atmospheric blocking in the Euro-Atlantic domain relating directly to a misrepresentation of the North Atlantic eddy-driven jet (mean position and variability). Is there a link between these obstinate biases occurring in different model realms? And what is the impact of further increasing model resolution on these biases?

More specifically, on the one hand, generations of stand-alone ocean simulations [e.g. Tsujino et al. (2020)] show an intense cold bias near the so-called Northwest Corner at the beginning of the North Atlantic Current (NAC). Similar biases can be seen also in CMIP3 and CMIP5 coupled historical simulations. On the other hand, a number of studies [Anstey et al. (2013); Masato et al. (2013); Davini and D'Andrea (2016)] have reported significant negative biases in European blocking frequency in CMIP5 models. Such biases are dynamically linked to biases in the eddy-driven jet, which, in most models lacks the observed trimodality in the jet latitude distribution [Woollings et al. (2010a); Anstey et al. (2013); Iqbal et al. (2018)] and exhibits a reduced poleward tilt in respect to the observed climatological jet [Zappa et al. (2013)]. Such biases in blocking and the jet are present, yet weaker, also in atmosphere-only simulations forced with observed SSTs and, therefore, may be partly attributed to limitations in the respective atmospheric models [e.g., Berckmans et al. (2013); Schiemann et al. (2017); Jiang et al. (2019)].

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-21-0515.1.

Nevertheless, previous studies have found evidence for a possible causal link between such SST biases and European blocking biases. First, Scaife et al. (2011) documented a significant reduction in both of these biases by increasing model resolution (from 1° in the ocean to 0.25°). Keeley et al. (2012) took the wintertime SST biases (similar to those of the previous study) exhibited in the extratropical North Atlantic by a coupled model and ran atmosphere-only simulations forced first with observed SSTs and then with the same SSTs modified by these biases (up to $-8\,\mathrm{K}$ at the beginning of the NAC). Their results showed that "the SST bias in the North Atlantic is likely the major cause of the coupled model atmospheric circulation bias in the North Atlantic and European region." Finally, Drews et al. (2015) dealt with a similar strong SST bias in the Kiel Climate Model, largely eliminating it by flow-field correction and flux-correction techniques along a historical simulation, yet without reducing the biases in the surface heat fluxes seen by the atmospheric model. As a result, they observe only a minor reduction in the negative pressure bias to the west of the British Isles (indicative of a deficit in European blocking).

Apart from the above-discussed cold SST bias at the beginning of the NAC, ocean models (OGCMs) and coupled models (CGCMs) tend to suffer from a misrepresentation of the Gulf Stream and the associated SST front, both at the separation point and farther downstream in the so-called Gulf Stream Extension (GSE) area. More specifically, most CMIP5 and (as shown in this study) many CMIP6 models exhibit a biased SST front, both in terms of the maximum meridional SST gradient (magnitude) and its location (meridional shift). The impacts of such SST anomalies (an oceanic forcing to the atmosphere) have been discussed by numerous and diverse studies over the last two decades, some focusing on the meridional position of the SST front and its variability [e.g., Joyce et al. (2009, 2019)] and others on its sharpness, namely the magnitude of the associated meridional SST gradient [e.g., Woollings et al. (2010b); Small et al. (2014); O'Reilly et al. (2016, 2017)]. All of these studies have documented significant impacts on the North Atlantic large-scale

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-21-0515.1.

circulation, such as changes in the configuration and the variability of the eddy-driven jet and in blocking frequency and duration.

Another recent study [Lee et al. (2018)] also analyzed the impact of SST biases at the Gulf Stream separation area on the atmospheric circulation and found a significant poleward jet shift in the perturbed atmosphere-only simulation comparing to the control simulations forced with observed SSTs. In another perturbed experiment including additional cold and warm biases straddling the GSE area they found the simulated jet distribution closer to the observed.

Even though the above-mentioned studies clearly indicate the significance of SST anomalies / biases along the Gulf Stream and the NAC in forcing notable responses in the North Atlantic eddy-driven jet and blocking, detailed analyses have been restricted to two single-model studies [Keeley et al. (2012); Lee et al. (2018)]. Moreover, no consensus has been reached yet regarding the causal chain connecting specific SST biases and their impacts on the midlatitude atmospheric circulation over the North Atlantic. One of the aims of the present analysis is to provide a multi-model view on such impacts, along with a detailed diagnostic analysis.

Another aim of this work is to assess the impact of increasing model resolution on such SST biases and consequently on common atmospheric circulation biases, such as the under-representation of European blocking. After better resolving the oceanic mesoscale in CGCMs (namely running at eddy-permitting resolutions comparable to 0.25°) became —only recently— possible for centennial simulations, a number of studies have started reporting diverse benefits from this significant step change. Haarsma et al. (2019) showed that the midlatitude interannual atmospheric variability and mean climate in an ensemble of seasonal hindcasts are improved, respectively, by increasing the oceanic resolution (from 1° to 0.25°) and the atmospheric resolution (from 80 km to 40 km). Roberts et al. (2020), using a hierarchy of model simulations each performed at two different resolutions, reported that various mean biases are reduced by increasing the oceanic resolution

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-21-0515.1.

(from 1° to 0.25°, including North Atlantic SST and blocking biases), particularly beyond the seasonal timescale. Czaja et al. (2019) provide a comprehensive review of focused studies [among which: Willison et al. (2013); Parfitt et al. (2016); Siqueira and Kirtman (2016); Foussard et al. (2019)] on the diverse benefits of better representing air–sea interactions and processes internal to each realm via increasing model resolution.

The present study assesses changes in complex phenomena, such as blocking and the occurrence of atmospheric circulation regimes with a focus on the Euro-Atlantic domain, using a multi-model set of coordinated experiments. The study assesses the role of changes (improvements with increasing model resolution) in meridional SST gradients associated with large-scale biases rather than differences in the representation of sharp, mesoscale ocean fronts. The following section (2) introduces the model simulations and all the data used, including observations. Then, section 3 examines such changes in European blocking and the northern-jet regime resulting from improvements in SST bias at the central North Atlantic. Section 4 assesses the relationship between SST biases farther upstream, at the broader GSE area, high-latitude blocking and the southern-jet regime. The concluding section (5) includes a synopsis of the findings and some discussion on their significance.

2. Model simulations, data and methods

a. PRIMAVERA models

This multi-model study is based on the historical coupled simulations (1950–2014) performed for the European Horizon 2020 PRIMAVERA (PRocess-based climate sIMulation: AdVances in high resolution modeling and European climate Risk Assessment) project following the HighResMIP protocol (Haarsma et al. 2016). In this protocol, these simulations are referred to as "hist-1950".

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-21-0515.1.

The respective models are documented as follows: CMCC-CM2 (Cherchi et al. 2019), CNRM-CM6 (Voldoire et al. 2019), EC-Earth3P (Haarsma et al. 2020), ECMWF-IFS (Roberts et al. 2018), HadGEM3-GC31 (Roberts et al. 2019), MPI-ESM1-2 (Gutjahr et al. 2019), AWI-CM-1.0 (Sein et al. 2017). Each of these models was run in its coupled configuration, provided with observed greenhouse-gas and aerosol concentrations, in at least two different resolutions, with most of the models increasing simultaneously both the atmospheric and the oceanic resolution, while two models only increased the atmospheric resolution (CMCC-CM2 and MPI-ESM-2), and two other models (ECMWF-IFS and HadGEM3-GC31) also run configurations at intermediate resolutions aiming to assess separately the impact of increasing the oceanic and the atmospheric resolution. Table 1 shows the different configurations and the respective nominal resolutions in midlatitudes.

As documented in Schiemann et al. (2020), the majority of coupled PRIMAVERA simulations were run in a single realization. Even though this might present a limitation to certain studies dealing with natural variability, the present study is based on long-term climatological averages (1950–2014) and thus is less sensitive to sampling internal variability. Therefore, for consistency across the multi-model ensemble, the present analysis is conducted using a single realization from each model. As discussed later, the main findings and the conclusions of this study are found to be insensitive to this choice.

Following the HighResMIP protocol, the coupled historical simulations ("hist-1950") were initialized from multi-decadal spinup simulations referred to as "spinup-1950". One may expect, however, that model drift continued in the "hist-1950" simulations. For this reason the stationarity of the examined midlatitude North Atlantic SST biases was assessed by comparing the latter between different epochs (1950–1974 and 1990–2014). It is reported (results not shown) that for all models the differences were found to be insignificant (less than 1°C) compared to the biases themselves. Roberts et al. (2019) have thoroughly assessed model drift for the HadGEM3-

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-21-0515.1.

GC3.1 model configurations used in PRIMAVERA (Table 1) using the respective "control-1950" centennial simulations with constant external forcings and found that severe drifts were limited to the spinup period. Referring to the same simulations, Moreno-Chamarro et al. (2021) comment that despite the relatively short spinup, the control simulations show a relatively stable surface climate with any left-over drifts being rather insignificant compared to the magnitude of the climate biases.

Two different paired sub-ensembles have been used in this study: LR / HR and LOW-RES / HIGH-RES. The former (LR / HR) serves to categorize all available model configurations into two classes: "LR" includes the coarsest configuration of each model, while "HR" includes all the remaining configurations with increased resolution either in atmosphere, or both in ocean and atmosphere. Thus, the characterization of each model configuration as high-resolution or low-resolution is relative. The second paired sub-ensemble (LOW-RES / HIGH-RES) was formed to cope with some missing data (no daily data other than geopotential height were available for the AWI-CM configurations) and to allow a multi-model comparison between LR and HR configurations in which —to avoid double-counting of similar biases— every oceanic model configuration participates only once. Thus, the increased-resolution configurations ECMWF-IFS-HR and HadGEM3-GC31-HM are excluded from "HIGH-RES" as their oceanic model configurations are represented by ECMWF-IFS-MR and HadGEM3-GC31-MM. In contrast, HadGEM3-GC31-HH has been included in "HIGH-RES" because its ultra-high (8 km) oceanic resolution gives a mean climate that is considered sufficiently different to both HadGEM3-GC31-MM and HadGEM3-GC31-HM, which share the same oceanic resolution (25 km). In the spirit of the HighResMIP protocol, all models (with the exception of the ECMWF models) have been tuned in their LR version, while HR versions are obtained by increasing the resolution with no additional tuning, except shorter timesteps for numerical stability.

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-21-0515.1.

b. Data and methods

This study focuses on the boreal winter season (DJF) when air–sea contrasts and the oceanic forcing on the atmosphere are stronger. However, most results for the spring season (not shown) were found to be quite similar to winter. Daily data were used for geopotential height at 500 hPa to assess blocking. Also, daily data for air-temperature at 850 hPa and the horizontal wind components at 850 hPa and 250 hPa were used to compute the meridional eddy heat fluxes and the horizontal E-vector components, as well as the jet latitude (u-wind at 850 hPa). The maximum Eady growth rate (EGR) was computed between the 850 hPa and 700 hPa levels using monthly mean data of air-temperature, geopotential height and horizontal wind components. Further details on methods are provided at the end of this section. Monthly mean fields were used also to assess biases in sea surface temperature (SST), surface sensible and latent turbulent heat fluxes (FLUX), total precipitation (PREC) and the zonal wind at 850 hPa (U850). The respective daily mean data from the ERA5 [Hersbach et al. (2020)] reanalysis were used as an observational reference for the above-mentioned diagnostics. Also, the corresponding monthly mean data from HadISST2 [Kennedy et al. (2017)], OAFlux [Yu et al. (2008)] and ERA5 were used to compute model biases and observational climatologies.

For the blocking analysis, the two-dimensional detection method developed by Scherrer et al. (2006) was adopted, as in Anstey et al. (2013); Athanasiadis et al. (2014); Schiemann et al. (2017) and Schiemann et al. (2020). More specifically, in this analysis the 5-day persistence criterion was dropped (thus assessing instantaneous blocking) so as to better relate results to the daily position of the jet. As discussed later, this choice does not affect the interpretation of our results and the validity of our conclusions. The results shown here are based on a joint ERA-40 [Uppala and co-authors (2005)] and ERA-Interim [Dee and co-authors (2011)] reanalysis described in Schiemann

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-21-0515.1.

et al. (2020). It should be noted that blocking climatologies from different reanalyses agree very well in the post-1950 period [Rohrer et al. (2019)].

The position of the North Atlantic eddy-driven jet was assessed via the jet latitude index as defined in Woollings et al. (2010a). Here daily mean zonal winds at 850 hPa are used for this purpose. To account for the fact that this pressure level intersects major orographic features, a mask was applied prior to the zonal averaging, excluding areas with surface elevation higher than 1300 m (essentially Greenland and the Atlas mountains in the area of interest). This refinement does not affect the trimodal character of the observed jet-latitude distribution.

The average meridional SST gradient within indicated geographical areas was computed by zonally averaging the equatorward SST gradient along the associated SST front, the latter being identified at each longitude through the meridional location of the maximum SST gradient. For comparability, prior to the computation of gradients the SST fields from all (LR and HR) models and the observations have been re-gridded onto a common $1^\circ \times 1^\circ$ grid. Instead, when comparing biases in SST gradient between high-resolution models, as in Fig. 9 (panels c, d), for a more detailed view, a common $0.25^\circ \times 0.25^\circ$ grid has been used matching that of the high-resolution HadISST2 data.

For the computation of the horizontal E-vector components ($\overline{v'^2 - u'^2}$, $\overline{-u'v'}$) and the meridional eddy heat flux ($\overline{v'T'}$), high-frequency transients (denoted by a prime) were defined via a high-pass Lanczos filter retaining periods lower than 10 days, as in Novak et al. (2015). Following Hoskins et al. (1983), these two variables are used to assess the impact of baroclinic eddies on the mean westerly flow. The atmospheric fields (u, v, T) from all models and the observations have been re-gridded onto a common $1^\circ \times 1^\circ$ grid matching the one used for SST fields.

The maximum Eady growth rate is computed as $\sigma = 0.31 \frac{f}{N} \frac{\partial U}{\partial Z}$, where f is the Coriolis parameter, N is the static stability parameter [as in James (1994)], Z is the geopotential height

and $U=\sqrt{u^2+v^2}$ is the magnitude of the horizontal wind. The vertical derivative is computed between the 850 hPa and 700 hPa isobaric levels. Regarding EGR climatologies, using monthly mean instead of daily mean data was found to make no significant difference.

Finally, horizontal gradients of SST and air-temperature at 850 hPa were computed using central differences scheme on the sphere.

3. Reduction of SST biases in the central North Atlantic leading to better blocking and jet

a. Improvements with resolution

A thorough examination of biases in the North Atlantic sector for all PRIMAVERA historical coupled simulations (hist-1950), comparing LR models to their HR counterparts, revealed notable differences in various aspects. Among these, an improvement in wintertime blocking frequency near the stormtrack exit region was found with increasing resolution (Fig. 1b and Fig. S1). This improvement is evident also in Schiemann et al. (2020) who examined the same multi-model ensemble. Blocking in this broad area is referred to as European blocking [Masato et al. (2013)], it relates to anticyclonic Rossby wave breaking [Masato et al. (2012); Weijenborg et al. (2012); Davini et al. (2012); Messori et al. (2019)] and is dynamically consistent with a poleward deviation of the eddy-driven jet and the stormtrack [Booth et al. (2012)]. Thus, the northern-jet regime [Woollings et al. (2010a)] is inextricably linked to blocking in this area. Before continuing, it should be mentioned that adding a 5-day persistence criterion in the blocking detection leads to very similar results, yet with lower blocking frequencies (Fig. S2).

On the other hand, most PRIMAVERA models exhibit a reduction in wintertime SST bias in the central North Atlantic with increasing resolution (Fig. 1a). Inspired by earlier work of Namias (1964) and Scaife et al. (2011), we explore the hypothesis that the reductions of both biases

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-21-0515.1.

(i.e. in SST and blocking) are linked to each other. In this respect, it is worth mentioning that these SST biases are at least twice as large as the observed interannual standard deviation in that area, and so are the respective biases in the meridional SST gradient within the indicated area (Figs. 1a, 3a). Before proceeding with the analysis of North Atlantic midlatitude SST biases, it should be noted that remote climatological SST biases, for example in the tropical Pacific, may also have an influence on North Atlantic midlatitude blocking and atmospheric circulation regimes. A thorough investigation in this direction is out of the scope of this study, though it is noted that the most robust mitigation of SST biases with increasing resolution was found to be (Fig. S7) the one discussed above (Fig. 1a), while other improvements in global SST biases were found to be model dependent.

We start by analysing how each model responded to the resolution increase with regard to the SST biases in the central North Atlantic. Figure 2 shows that even though the bias reduction pattern is quite different from model to model, the most significant improvements occurred in the models that increased not only the atmospheric but also the oceanic resolution (Table 1). The SST bias changes seen in the central North Atlantic generally correspond to a reduction in a cold SST bias that most LR models have. In fact, similar cold biases have been a long-standing issue in past-generation climate models [e.g. Wang et al. (2014)]. There is evidence [Roberts et al. (2020); Storkey et al. (2018); Marzocchi et al. (2015)] that these cold biases are endogenous to the respective oceanic model components and are reduced with increasing resolution, likely thanks to better resolving the oceanic mesoscale. Further evidence for the oceanic origin of these SST biases comes from the examination of the respective biases in surface turbulent heat fluxes (sensible and latent) presented in Fig. 5; if the cold SST biases were caused by excessive heat-loss to the atmosphere, the bias in surface fluxes would have the opposite sign (positive rather than negative). Finally, the oceanic origin of these SST biases is corroborated by examining historical ocean-only simulations forced

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-21-0515.1.

with observed atmospheric fields. Such simulations were available for 3 out of the 13 different ocean model configurations (Table 1) and were found to exhibit cold SST biases in the central North Atlantic (not shown) comparable to the respective coupled historical simulations. Similar cold biases at the central North Atlantic are common in forced ocean-only simulations [Tsujino et al. (2020)] and are found to be reduced with increasing resolution [Chassignet et al. (2020)].

A possible causal relationship between the reduction in the cold SST biases and the improvements in European blocking frequency (Fig. 1) may involve a modification of near-surface atmospheric baroclinicity by the underlying SST gradients. Such an influence has been clearly demonstrated in previous studies [Nakamura et al. (2008); Nakamura and Yamane (2009); Hotta and Nakamura (2011)] in reference to sharp SST fronts, but may also apply here. The idea is that differential surface heat fluxes across SST fronts have a key role in maintaining low-level baroclinicity, which, in turn, may affect blocking through modifying baroclinic eddy activity (genesis and growth); thus SST biases may cause blocking biases. To assess this hypothesis, first the climatological meridional SST gradient for each model was plotted against the respective air-temperature gradient at 850 hPa (Fig. 3a), representing low-level baroclinicity. All meridional gradients are considered positive equatorward. While the air-temperature gradients have been spatially averaged within the area indicated in Fig. 1a, the SST gradients are averaged zonally along the respective SST front identified at each longitude through the meridional location of the maximum SST gradient. To aid the understanding of the SST gradient differences, it is noted that the respective biases result from the differential displacement of the isotherms due to the cold SST biases at the northern part of the indicated area, which is evident in Fig. 5a. Very interestingly, all models show clear improvements with increasing resolution (round markers to diamond markers) in the representation of both gradients. Moreover, the distribution of the examined 15 model configurations in the two-dimensional space (shown in Fig. 3a) indicates a strong relationship between the climatological

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-21-0515.1.

SST gradient and low-level baroclinicity. In addition, such an approximately linear relationship is also evident for the respective 10-year averages (Fig. 3b) indicating that decadal variations in the SST gradients at the central North Atlantic modulate the respective variability in low-level baroclinicity. The fact that the cloud of points corresponding to the observations (black dots) is more steeply inclined and more linear suggests that low-level baroclinicity in the models is less sensitive to changes in the underlying SST gradients than in the observations. We hypothesize that this under-represented sensitivity may be related to the signal-to-noise deficit exhibited by climate models, something that will be investigated in a follow up study.

Low-level baroclinicity is expected to affect baroclinic activity and subsequently the eddy-driven jet and blocking downstream. Figure 4 allows the comparison of European blocking frequencies, the above-examined SST gradients and the occurrence of the northern-jet regime for all models and the observations. Again, two salient facts emerge: (*i*) all models exhibit significant improvements with increasing resolution in all of the above three aspects, and (*ii*) the climatological values of European blocking frequency and North-jet occurrence relate strongly to each other. It is also worth noting that extrapolating / interpolating the line segments connecting the round markers (LR model configurations) to the diamond markers (HR model configurations) until meeting the observed SST gradient (Fig. 4a) would still leave all models with a deficit in European blocking. This suggests that beyond the cold SST bias in the central North Atlantic there are also other causes for the underestimation of European blocking. Considering the CMCC-CM2 model (strongest bias in European blocking), this model suffers from a comparably strong bias (deficit) also in its atmosphere-only forced historical simulation (termed "highresSST-present" in HighResMIP protocol), indicating that in this climate model, blocking strongly depends on the atmospheric model component. Finally, it should be mentioned again that adding a 5-day persistence criterion

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-21-0515.1.

to the blocking detection leads to analogous results, yet with blocking frequencies lowered by a factor of about 3 (Fig. S2).

This analysis is conducted using a single realization for each model configuration. To assess the uncertainty associated with internal variability, as discussed in section 2, mean climate differences between single realizations were assessed for some of the simulations available in multiple realizations. The examined SST biases were found to differ very little, while biases in European blocking and the respective differences between realisations were found to be relatively small, so that the multi-model results discussed in this section do not differ significantly when different model realisations are used (Fig. S3).

Even though this study highlights the role of increasing the oceanic model resolution, it should be mentioned that part of the demonstrated improvements in European blocking and the North Atlantic jet in the coupled historical simulations are likely due to the accompanying increase in atmospheric resolution (see Table 1). An indication for this is found in the fact that improvements in European blocking that are similar to —yet significantly weaker than— those shown in Fig. 1b are found comparing the respective forced historical atmosphere-only simulations ("highresSST-present"). This result is shown in the supplemental material (Fig. S6) and is consistent with the findings of Schiemann et al. (2020) who analyzed the same set of simulations. Nevertheless, atmosphere-only simulations are deprived of air–sea coupling, and thus these findings are not 'transferable' to the coupled simulations. To disentangle the role of oceanic and atmospheric resolution increase, a suitable approach would require comparing coupled simulations in which the oceanic and the atmospheric resolutions are increased separately so as to identify the influence of each individual realm in a fair way (i.e. with coupling always active). While it would be best to do so following a multi-model approach, the HighResMIP protocol adopted in PRIMAVERA did not require a resolution increase realm-by-realm. The only available examples of such simulations are those

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-21-0515.1.

performed with the ECMWF-IFS and HadGEM3-GC31 models (Table 1). Focusing on these simulations in Fig. 3 and Fig. 4, one can see very clearly that the most significant improvements in European blocking and the representation of the northern-jet regime are brought by increasing the oceanic resolution (and not the atmospheric resolution) from a typically coarse resolution (1°) to an eddy-permitting resolution (0.25°), and it is exactly this kind of resolution increase that also reduces most effectively the cold SST bias in the central North Atlantic.

b. The causal chain from SST biases to blocking biases

Having documented the improvements in European blocking occurring with increasing model resolution, the assumed causal relationship between SST biases and biases in blocking and the jet is further investigated. For this a number of diagnostics (climatologies and biases) are compared between the LOW-RES and the HIGH-RES multi-model ensembles (Table 1). At this point the switch from the LR / HR to the LOW / HIGH paired sub-ensembles is dictated by the unavailability of some data (AWI-CM) and the choice of giving equal weight to all models. Figure 5 (panels a, b) shows the respective SST model biases (shading) and full fields (dashed contours) as well as the observed winter climatology (solid contours). Biases in the meridional SST gradient are evident by the more densely packed isotherms (LOW-RES compared to HadISST2, in Fig. 5a) in the area indicated by the rectangular frame [15–45°W, 40–50°N]. The prominent cold bias in LOW-RES around (38°W, 48°N), where the Gulf Stream ends and the North Atlantic Current begins, represents a well-known and long-standing issue in ocean modeling. In some models [e.g., for NEMO see: Storkey et al. (2018) and Marzocchi et al. (2015)] this cold bias has been ameliorated, as in this study, by increasing horizontal resolution.

Then, panels c, d of Fig. 5 show the LOW-RES and HIGH-RES biases in turbulent (sensible and latent) surface heat fluxes. As mentioned earlier, the matching sign between the SST biases

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-21-0515.1.

and the biases in surface fluxes points to an oceanic driving of the latter, meaning that the colder SSTs lead to reduced surface heat fluxes as expected by the bulk formulas used in the respective parameterization. As a direct consequence, the biases of these two fields exhibit similar spatial patterns. It is important to note that on the western side of the examined area, the magnitude of the bias in surface fluxes ($>150\,\mathrm{W\,m^{-2}}$) is almost as large as the respective climatology ($\sim 200\,\mathrm{W\,m^{-2}}$), thus modifying profoundly the diabatic heating in the overlying atmospheric boundary layer. Recent studies [Croci-Maspoli and Davies (2009); Pfahl et al. (2015); Sheldon et al. (2017); Steinfeld and Pfahl (2019)] discuss the significance of sensible and latent heating for the formation of blocking, while Yamamoto et al. (2020) also show a high concentration of air-parcels experiencing diabatic heating in North Atlantic blockings that originate from this area.

The increased static stability (Fig. S4) resulting from the biases in surface heat fluxes drastically inhibits vertical motion (Fig. S4) and precipitation in the same area. Indeed, Fig. 6 (panels a, b) reveals significant biases in precipitation, arguably a combination of reduced moisture content (due to the co-located deficit in latent heat flux) and weakened vertical motion due to increased stability (deficit in sensible heat flux). The LOW-RES precipitation bias represents approximately 30–40% of the climatological value at the western side of the examined area. This strong bias is also a proxy for reduced diabatic heating in the interior of the troposphere, which has a role for blocking as discussed in the previous paragraph. Finally, there are also significant SST biases at the Gulf Stream separation area, but these improve little with increasing resolution and differ substantially from model to model.

Beyond the impact that biases in diabatic heating may have on blocking, another effect is found to be important. Panels e, f of Fig. 5 show that with increasing resolution the biases in low-level baroclinicity are strongly reduced in the examined area. The negative bias in HIGH-RES to the west of the examined area was found to originate primarily from a single model (see next paragraph).

In the area of interest, the notable positive bias in baroclinicity in LOW-RES appears to be caused by the positive bias in the underlying meridional SST gradient, while the downstream extension of the former is consistent with advection by the prevailing westerlies. It is evident that this bias effectively extends the zone of maximum climatological low-level baroclinicity (cyclogenesis area) to the east (also see Fig. 7, panels a, b). This way, storms undergo deeper development, and their impact on the mean flow (westerly acceleration) strengthens and extends farther zonally. Linking the biases in low-level baroclinicity with the SST biases is in agreement with the changes shown in Fig. 3a, as well as with those shown in Fig. 5 (panels c, d, surface heating) and Fig. 6 (panels a, b, heating in the interior).

Figure 7 provides further evidence for the above-mentioned effect. Biases in the maximum Eady growth rate (EGR) at 775 hPa are shown for LOW-RES and HIGH-RES. Broadly mirroring the biases in the temperature gradients, LOW-RES exhibits a strong localized positive bias that is drastically reduced in HIGH-RES. This bias ($\sim$0.1 day^{-1}) corresponds to about 20% of the respective climatological values on the western side of the examined area. The negative EGR bias in HIGH-RES seen farther upstream has been assessed and was found to originate primarily from a single model (the case of CNRM-CM6-1-HR and its biases farther upstream is discussed in more detail in the next section).

Figure 7 also shows (panels c–f) biases in the E-vector components, the divergence of which measures the impact of synoptic transient eddies on the westerly flow [Hoskins et al. (1983); Novak et al. (2015)]. The vertical component of the three-dimensional E-vector is proportional to the meridional eddy heat flux $v'T'$, here evaluated at 850 hPa (panels c–d). Given that winds become significantly reduced approaching the surface, the value of this eddy covariance is indicative of its vertical divergence between this pressure level and the surface. Thus, in LOW-RES, the localized positive bias in the examined area is indicative of an anomalous eastward acceleration sustaining a

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-21-0515.1.

positive bias in the low-level jet, as can be seen in Fig. 6. Similarly, the divergence of the horizontal E-vector components (Fig. 7, panels e, f) indicates an eastward acceleration aloft.

To conclude the chain of arguments, we next discuss how the above-mentioned biases in the eddy-driven jet forcing relate to the biases in European blocking. Figure 6 shows the resulting bias in the wintertime eddy-driven jet (zonal wind at 850 hPa) and blocking. In the eastern part of the North Atlantic, LOW-RES, comparing to HIGH-RES, exhibits a stronger positive bias in the low-level zonal winds between 40–55°N and a stronger negative bias in blocking frequency between 50–65°N. These two aspects are consistent from a dynamical viewpoint as higher climatological blocking frequency at one grid point is expected to be accompanied by a climatologically weaker jet at the equatorward flank of the blocking anticyclones, about 10–15 degrees to the south [e.g., see Fig. 5 in Athanasiadis et al. (2014)]. Furthermore, regarding variability in observational data, when the eddy-driven jet strengthens between 40–55°N, corresponding to the negative polarity of the Eastern Atlantic Pattern [Wallace and Gutzler (1981)] and the A2 pattern of Athanasiadis et al. (2010) (see their Fig. 6) and to an increased occurrence of the central-jet regime [Woollings et al. (2010a)], this change is accompanied by a weakening of the eddy-driven jet to the north and of the subtropical jet to the south. Even though this study deals with biases and not with observed low-frequency variability, similar spatial connections to the ones discussed above are evident here indicating that these biases are related to changes in the frequency of occurrence of well-known patterns of variability, as suggested by Kushnir et al. (2002) for the atmospheric response to extratropical SST forcing.

More specifically, to relate the jet and blocking biases seen in LOW-RES, which are significantly reduced in HIGH-RES, it must be noted that blocking near the stormtrack exit relates to anticyclonic wave breaking accompanied by a poleward shifted jet and stormtrack (northern-jet regime). On the contrary, when the eddy-driven jet tends to occupy more frequently its zonal state, to be stronger

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-21-0515.1.

and to extend farther downstream (as it does in LOW-RES comparing to the observations due to the stronger SST gradients associated with the cold SST biases), this clearly disfavors the formation of European blocking that would imply opposite jet anomalies. In the absence of a general theory for blocking, another supporting argument relates to considering blocking in the stormtrack exit area as the result of dissipating extratropical cyclones [Shutts (1983); Illari and Marshall (1983); Shutts (1986)] so that when the latter become revitalized in that area by the increased near-surface baroclinicity in LOW-RES, the eddy-forcing of blocking anticyclones is reduced.

Orlanski (1998), referring to the North Pacific stormtrack in November as opposed to mid-winter, and in La Niña as opposed to El Niño years, underlines the importance of the eastward extended low-level baroclinicity in reducing the eddy-induced poleward deflection of the stormtrack. Here, this effect appears to be active in the North Atlantic due to the influence of the cold SST biases (stronger in LOW-RES) that effectively extends the zone of intense low-level baroclinicity farther to the east (Fig. 5e). This way, the eddies are revitalized and their shape characteristics at the Eastern North Atlantic no longer favours the poleward deflection of the stormtrack associated with European blocking.

Recent studies on blocking that examine nonlinear dynamics [Nakamura et al. (1997); Naoe and Matsuda (2002)] and the role of Rossby wave propagation along the midlatitude jet acting as a wave-guide [Nakamura and Huang (2018)] indicate the importance of changes in the basic state flow (jet speed) and upstream transient-wave-activity for modulating blocking frequency. In this light, our results, appear to be consistent by linking negative biases in European blocking frequency with increased eddy forcing (jet acceleration) caused by the identified cold SST biases. Thus, as a continuation to this study, it would be worth investigating also the modification of stationary eddies and local wave-activity in relation to blocking biases.

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-21-0515.1.

4. SST biases at the GSE area and the atmospheric circulation

a. Demonstrating the relationship to blocking and the jet

So far we discussed the important model improvements in European blocking and the representation of the eddy-driven jet near the stormtrack exit that came with increasing resolution (HIGH-RES vs LOW-RES) thanks to a notable reduction in SST biases in the central North Atlantic. In view of these results, it is logical to question how the models responded to increasing resolution farther upstream, and more specifically at the Gulf Stream extension (GSE) SST front, which has been shown to be important for the existence of the North Atlantic stormtrack [Hoskins and Valdes (1990); Minobe et al. (2008); Brayshaw et al. (2011)] and its variability linked to jet-latitude (southern-jet vs central-jet regimes) [O'Reilly et al. (2016, 2017)] and blocking over Greenland [Joyce et al. (2019); Athanasiadis et al. (2020)], hereafter also referred to as high-latitude blocking (HLB). Even though for this region models do not exhibit consistent improvements with increasing resolution across the multi-model ensemble, some dependencies were noted between SST biases and the climatological frequencies of HLB and the southern-jet regime, which are discussed below.

Figure 8 presents the climatological frequency of occurrence of the southern-jet regime (jet latitude $< 38°$N) in wintertime versus the respective climatological meridional SST gradient in the area [25–60°W, 35–45°N], indicated in Fig. 9. A salient, seemingly linear, relationship is seen for the examined 15 models suggestive of a physical relationship (weaker SST gradients linked to more frequent southern-jet occurrence). It should be mentioned that although the exact positioning of the markers in the plot exhibits some sensitivity to the choice of the GSE area and the latitude threshold used to define the southern-jet regime, the alignment of the markers is a rather robust feature and still holds without the outliers (CNRM-CM6-1-HR and the CMCC models). Given that the SST front is more sharply defined between 50–70°W, the latter longitudinal range may

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-21-0515.1.

appear as a reasonable choice instead of the used range (25–60°W). The selected area, however, encompasses SST biases from all models, and was found to lead to a clearer statistical relationship with the jet. It should be mentioned that there is some overlap between this area and the one used in the previous section in relation to European blocking (Fig. 1a). Yet, one should expect that HLB and European blocking, as well as the frequencies of the respective jet regimes, are not influenced by the SSTs only in mutually exclusive areas but also respond —rather differently— to common SST features. Some evidence for this can be found in Brayshaw et al. (2011) and Lee et al. (2018).

In Fig. 8b, a similar scatter plot for the frequency of HLB [60–70°N, 15–50°W] and the same SST gradients is shown. Finding here a similar relationship to the one discussed above is expected as Woollings et al. (2008); Athanasiadis et al. (2014) and Madonna and Woollings (2017), among other studies, have shown that an equatorward displaced eddy-driven jet tends to concur with Greenland blocking. The distribution of the models in the vertical direction in the two panels of Fig. 8 suggests that the link between Greenland blocking and the southern-jet regime holds also for climatologies.

Regarding the interpretation of this apparent relationship, in contrast to the results discussed in the previous section linking biases in European blocking to SST biases in the central North Atlantic, here it would be wrong to consider the atmospheric biases solely as a passive response to the SST biases. This is supported by two arguments: (*i*) the SST biases at the GSE area are more strongly influenced by coupled feedbacks and thus are less endogenous to the respective oceanic model components, as indicated, for instance, by finding the SST biases in two forced ocean-only simulations (with the CMCC-CM2-HR4 and CNRM-CM6-1-HR models, not shown) to be similar but considerably weaker than in the coupled simulations analyzed here, and (*ii*) in some cases the respective atmosphere-only simulations forced with observed SSTs exhibit similar —yet weaker— atmospheric biases (as for the CMCC-CM2-HR4 model, not shown) indicating

that the atmospheric biases in the respective coupled simulations are, at least in part, caused by the atmospheric model components. Therefore, we interpret this relationship as indicative of a positive coupled feedback: on the one hand, overly strong SST gradients in the broader GSE area tend to disfavor HLB blocking and to force a poleward displacement and intensification of the eddy-driven jet and, thus, a less frequent southern-jet occurrence. O'Reilly et al. (2017) demonstrate a similar behavior for the eddy-driven jet assessing its response to smoothing the GSE SST gradients. On the other hand, these same atmospheric circulation anomalies, representative of the positive NAO polarity, tend to increase the SST gradient in the GSE area by the anomalous fluxes accompanying the changes in near-surface winds [Deser et al. (2010)] and also force a stronger subtropical gyre via the increased wind-stress curl, leading to an intensified meridional heat transport by the Gulf Stream and, eventually, a sharper temperature contrast across the GSE SST front.

While there seems to be a small improvement with increasing resolution in all examined aspects (Fig. 8), models still tend to have overly strong SST gradients in the broader GSE area. It is conceivable that the excessively strong westerlies downstream linked to the under-representation of European blocking (Fig. 6, mainly for LOW-RES models) are contributing to this problem. Identifying the origins of the associated SST biases (apparent also for HIGH-RES models) goes beyond the scope of the present study.

b. Understanding the ocean-to-atmosphere forcing

Having discussed the apparent relationship between the climatological GSE SST gradient and the respective frequencies of HLB and the southern-jet regime, interpreted as a manifestation of a positive coupled feedback, this subsection is dedicated to better understanding one side of the hypothesized two-way interaction, namely the influence of SST biases along the GSE on the atmospheric circulation. This general topic has been the subject of careful past studies following

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-21-0515.1.

a single-model approach, such as Keeley et al. (2012) and Lee et al. (2018), which provided motivation for the present analysis.

To address this topic we use the two most extreme outliers, namely the historical coupled simulations by the CMCC-CM2-HR4 and CNRM-CM6-1-HR models that exhibit, respectively, the most positive and most negative biases in the GSE SST gradient (Fig. 8). In a simplified view, letting aside coupled feedbacks and biases endogenous to the atmospheric model components, these outliers may be seen as two extreme sensitivity experiments conducted with different models and loosely opposite forcings, considering the respective biases in the GSE SST gradient as prescribed oceanic forcings to the atmosphere. Nevertheless, bearing in mind the possibility of these two models not being representative of the behavior of the multi-model ensemble, additional analysis (discussed further down) has been conducted using the full ensemble. Figure 9 shows the SST biases for these two models (panels a, b) as well as the respective meridional gradients (panels c, d). In the area of interest, which has been chosen considering all examined models to best demonstrate the previously discussed relationship, the two outliers exhibit roughly opposite SST biases leading to biases of opposite sign in the respective meridional gradients.

In the same figure (panels e, f), the corresponding biases in low-level baroclinicity are displayed. As shown by Nakamura et al. (2004, 2008); Nakamura and Yamane (2009) and Hotta and Nakamura (2011), among other studies, the sharp SST fronts along western boundary currents with the associated large meridional gradients in diabatic heating anchor and fuel above them zones of maximum low-level baroclinicity. Hotta and Nakamura (2011) emphasized the dominance of sensible heat fluxes in restoring low-level baroclinicity, whilst Papritz and Spengler (2015) and Hoskins and Valdes (1990) underlined the role of latent heating (occurring in the interior thanks to the ascending eddy motions) in maintaining baroclinicity against the eroding action of baroclinic eddy fluxes. In both cases, however, it is the meridional gradient of the diabatic heating that

creates baroclinicity. Thus, the biases in low-level baroclinicity seen here are thought to be ultimately forced by the underlying SST biases. The latter also affects low-level static stability and, together with the changes in baroclinicity, modifies the maximum Eady growth rate (Fig. 9, panels g, h). The low-level baroclinicity and EGR biases for the two outliers (CMCC-CM2-HR4 and CNRM-CM6-1-HR) involve features of opposite sign indicating meridional shifts in the respective zones of maximum climatological values (poleward for CMCC-CM2-HR4 and equatorward for CNRM-CM6-1-HR).

Ultimately, baroclinic activity is impacted [Stone (1978)] and consequently, in combination with the sharper air-temperature meridional gradient, the eddy heat fluxes and the horizontal E-vector components are modified in each of the models (Fig. 10). As for the SST gradients and low-level baroclinicity, the eddy heat flux and E-vector biases for the two outliers involve features indicating meridional shifts in opposite directions (poleward for CMCC-CM2-HR4 and equatorward for CNRM-CM6-1-HR) resulting in homo-directional shifts in the eddy-driven jet, as seen in the bottom panels (e, f) of Fig. 10.

One may question whether the outliers analyzed above are representative of the whole ensemble in terms of the physical processes underlying the statistical relationships evident in Fig. 8. To address this point and to add robustness to the analysis presented above, the multi-model ensemble has been partitioned into two sub-ensembles following the climatological SST gradient of the models in the area [25–60°W, 35–45°N] (Fig. S8a; above-median and below-median groups), and some differences between the respective sub-ensemble means have been examined (Fig. S8, panels b–e). The results are consistent with the conclusions drawn from the analysis of the two outliers and support the hypothesis of a positive feedback between biases in the two realms. This is also supported by the fact that the zonal wind differences between the same sub-ensembles of atmosphere-only forced simulations (highresSST-present; not shown) exhibit the same dipole

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-21-0515.1.

pattern as in Fig. S8c, yet with significantly weaker anomalies. Besides, this analysis has been repeated excluding the three obvious outliers (CMCC-CM2-HR4, CMCC-CM2-VHR4 and CNRM-CM6-1-HR) obtaining consistent results, albeit with weaker differences (not shown). An in-depth dynamical analysis of these biases is beyond the scope of this study and would require, also, a complete set of ocean-only forced historical simulations (forced with observed wind-stresses and heat fluxes) performed with the same models so as to investigate the interplay between biases endogenous to the atmospheric model components and those endogenous to the corresponding oceanic model components. In fact, to assess more precisely the role of coupling the two realms, in a future assessment it would also be useful to have single-realm historical simulations forced with data coming from the corresponding simulations of the other realm forced with observations.

5. Discussion and conclusions

a. A synopsis of the study

The aim of this study was twofold. First, it aimed to assess likely causal relationships between typical SST biases exhibited by current generation climate models in the midlatitude North Atlantic and other, similarly typical biases exhibited by the same models in the representation of North Atlantic eddy-driven jet and blocking. Second, the study aimed to evaluate improvements in the above-mentioned biases that occur with increasing model resolution. The multi-model HighResMIP coupled historical simulations conducted for PRIMAVERA (*hist-1950*) provided an ideal, designed on purpose, experimental framework for this analysis. The role of SST biases in two regions was investigated: (*i*) the beginning of the NAC, and (*ii*) the broader GSE area. The main findings of the study are outlined below:

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-21-0515.1.

- It was shown (Fig. 1) that increasing model resolution, primarily in the ocean (from 1° to 0.25°), led to a significant reduction in a typical wintertime cold SST bias at the beginning of the NAC, which is common among most of the examined models. This improves model biases in the adjacent meridional SST gradient, which in turn, is shown to affect low-level baroclinicity (Figs. 3, 5) and the maximum EGR (Fig. 7) in the central North Atlantic, directly upstream to the stormtrack exit. Thus, the reduction of this cold bias ultimately leads (for each model and for the multi-model mean) to a significant reduction of the respective deficit in European blocking frequency and northern-jet occurrence (Figs. 4, 6). This is mediated via an associated reduction in eddy–mean flow interaction biases (Fig. 7).

- It was shown (Fig. 3) that for each model the magnitude of the time-mean meridional SST gradient at the central North Atlantic [15–45°W, 40–50°N] determines the time-mean low-level baroclinicity in that area, and also, the decadal variations of the former determine the respective variations of the latter.

- Finally, regarding SST biases at the beginning of the NAC, it was shown that the severe cold biases exhibited by most low-resolution (LR) models lead to equally severe biases in surface turbulent heat fluxes (Fig. 5), which, in turn, strongly affect stability, vertical motion and precipitation in that area (Figs. S4, 6). Possible dynamical implications of these biases in diabatic heating are discussed below.

- Then, it was shown (Fig. 8) that across the multi-model ensemble there is a strong relationship between biases in the maximum meridional SST gradient along the GSE and farther downstream (referring to the area indicated in Fig. 9) and biases in high-latitude blocking and the southern-jet occurrence. The interpretation of this statistical relationship is discussed further down.

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-21-0515.1.

- Regarding climatological SST biases in the GSE area, increasing model resolution does not seem to lead to systematic improvements (Figs. 1, 2, 8) across the PRIMAVERA multi-model ensemble.

- It was discussed that claiming a uni-directional causal link between the above-mentioned SST biases (along the GSE and farther downstream) and impacts on the jet would be questionable; nevertheless, by comparing basic diagnostics for the most positive and most negative outliers across the multi-model ensemble (in terms of the bias in the maximum meridional SST gradient in the above-mentioned area), it was shown that the expected influences by the oceanic biases are in the same direction as the exhibited atmospheric biases (Figs. 9, 10). This is supported by the associated multi-model differences (Fig. S8).

b. Reflections on the main findings

This study proposes a causal relationship between North Atlantic SST biases and atmospheric circulation biases, considering the former as endogenous to the respective oceanic model components (a multi-model ensemble) effectively acting as forcing to the atmosphere. Previous studies [e.g. Keeley et al. (2012) and Lee et al. (2018)] have put forward similar arguments; nevertheless a direct comparison of results has little meaning as the latter depend on the exact pattern of the SST biases, which is model dependent.

In addition, regarding mechanisms proposed to explain the response, standard stormtrack dynamics have been invoked including eddy–mean flow interaction (diagnosed via the use of E-vectors) and a modification of the respective baroclinic activity caused by SST gradients. Masato et al. (2016) also based their analysis on eddy–mean flow interaction, and although they do not directly assess SST biases but biases in the jet, they comment: "these biases are weaker when the atmosphere model is forced with observed SSTs, suggesting that either it is vital to have the correct

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-21-0515.1.

SST distribution or ocean–atmosphere coupling plays a key role in the biases." In fact, we see no reason for these possibilities to be mutually exclusive.

The SST biases at the beginning of the NAC may contribute to biases in European blocking also via other physical processes. As discussed in Section 3, the severe deficit in diabatic heating caused by the cold SST bias in that area is thought to hinder blocking maintenance according to recent studies [Yamamoto et al. (2020); Steinfeld and Pfahl (2019)] that trace back diabatically warmed air-parcels from inside blockings to the marine boundary layer in that same area. As the present study focuses on instantaneous blocking, in a future study it would be interesting to assess the impact of such SST biases also on blocking persistence.

Since near-surface diabatic heating is an effective PV source [Hoskins et al. (1985); Ambaum and Athanasiadis (2007)], the severe negative biases in heat fluxes (sensible heat at the surface and latent heat in the near-surface interior) may cause a negative bias in PV. This may have two effects: (i) along with the increased stability, inhibit cyclone development in that area, thus modifying the shape of the North Atlantic stormtrack, reducing its poleward tilt, and (ii) through the respectively induced anticyclonic circulation have a similar effect on the climatological jet. These effects have not been assessed quantitatively, yet they appear plausible.

Regarding the robust relationship found between decadal variations in the meridional SST gradient and in low-level baroclinicity directly above (Fig. 3b), it is particularly interesting to note that the cloud of markers corresponding to the observations is more steeply inclined indicating considerably higher sensitivity to the underlying SST gradient compared to all models. We consider this to be a significant finding, likely contributing to the signal-to-noise problem [Scaife and Smith (2018)]. Even though models are expected to improve in this respect with increasing resolution, no such indication is evident here. Careful point-by-point analysis is needed to assess such improvements.

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-21-0515.1.

If the causal chain from SST gradients to low-level baroclinicity, baroclinic activity and finally European blocking is real, one may anticipate that predictable SST anomalies in the central North Atlantic may be a source of predictability for European blocking, too. Preliminary results based on ERA5 reanalysis and CMIP6 decadal hindcasts (ongoing analysis to be published separately) support this hypothesis. At this point it is worth noting that Namias (1964) proposed that changes in the central North Atlantic SST gradient may have favored the anomalously frequent and persistent blocking episodes of the notorious 1963 winter.

The apparently linear relationship found between the maximum meridional SST gradient in the broader GSE area and the frequency of occurrence of the southern-jet regime and HLB (Fig. 8) should be interpreted with caution, considering, for instance, likely positive coupled feedbacks between biases endogenous to the atmospheric and to the oceanic model components. In other words, although the importance of the Gulf Stream SST front for the stormtrack and the jet is clear [Hoskins and Valdes (1990); Minobe et al. (2008); Sampe et al. (2010); Brayshaw et al. (2011); Omrani et al. (2019)], the interaction between atmospheric and oceanic biases is two-way such that, especially for the Gulf Stream area, the atmosphere is not passively responding to the SST biases. HLB and the southern-jet regime relate strongly and directly to the NAO variability, and the latter has been shown to impact SSTs and the oceanic circulation, including the intensity and the meridional position of the Gulf Stream SST front [Taylor and Stephens (1998); Frankignoul et al. (2001); Eden and Jung (2001); de Coëtlogon et al. (2006); Joyce et al. (2009); McCarthy et al. (2018)]. Therefore, the above-mentioned (seemingly linear) statistical relationship is thought to reflect a two-way interaction between the respective oceanic and atmospheric anomalies (here climatological mean biases) that positively feed back on each other, thus tending to amplify the coupled response along a specific direction in the two-dimensional space.

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-21-0515.1.

The important point of why increasing model resolution brings the observed improvements (reduction of SST, blocking and jet biases) has been left to be discussed last. The most significant improvements in European blocking and in the representation of the jet latitude variability (Fig. 4) occurred in the models that also increased the oceanic resolution, thanks to a reduction in the typical cold SST bias in the central North Atlantic. Therefore, there is a clear indication that increasingly resolving the oceanic mesoscale (with model resolutions comparable to 0.25°) immediately benefits ocean representation. Further increases in oceanic resolution are expected to lead to additional improvements to both model components. At this point, it should be noted that increasing the oceanic resolution does not bring only improvements but also, in same specific aspects, deteriorates the biases. A notable example is the warm SST bias in the Labrador Sea (Fig. 5b) that appears or worsens when the oceanic resolution is increased (this occurs to 4 out of 5 models, see Table 1, specifically to the CNRM-CM6-1, EC-Earth3P, HadGEM3-GC31 and ECMWF-IFS models). While this issue is being investigated by the modeling community, Moreno-Chamarro et al. (2022) comment, citing Koenigk et al. (2022), that this warm bias might be related to excessive oceanic deep mixing in the coupled models using the NEMO ocean model at 0.25°–0.5° resolution.

Even if the oceanic resolution was found to be key, referring specifically to the above-discussed improvements in European blocking and the eddy-driven jet, the importance of atmospheric resolution should not be undervalued. The presented analysis focused on coupled historical simulations; nevertheless, assessing also the respective multi-model atmosphere-only forced historical simulations produced in PRIMAVERA (*highresSST-present*) showed that increasing the atmospheric resolution alone also brought noticeable improvements in the representation of North Atlantic eddy-driven jet (Fig. S5, note the wider and more realistic distributions for HIGH-RES) and blocking [Schiemann et al. (2020) and Fig. S6]. This is in agreement with numerous recent studies

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-21-0515.1.

highlighting the importance of high atmospheric resolution [e.g., Willison et al. (2013); Anstey et al. (2013); Smirnov et al. (2015); Parfitt et al. (2016); Schiemann et al. (2017, 2020)].

Acknowledgments. PJA and AB acknowledge funding from the PRIMAVERA project, funded by the European Union's Horizon 2020 programme under Grant Agreement 641727. MR acknowledges support form the Joint UK BEIS/Defra Met Office Hadley Centre Climate Programme (grant no. GA01101). N-E. Omrani was supported by: (1) Bjerknes Climate Prediction Unit with funding from the Trond Mohn Foundation (grant BFS2018TMT01), (2) the RCN funded ROADMAP project (316618) under a joint JPI Climate and JPI Ocean call. PA acknowledges funding also from the Italian Ministry of Education, University and Research (MIUR) through the JPI Oceans & JPI Climate "Next Generation Climate Science in Europe for Oceans" - ROADMAP Project- (D.M. 593/2016). Finally, the authors would like to thank three anonymous reviewers for their comments and suggestions that greatly improved the original manuscript.

Data availability statement. All data from PRIMAVERA HighResMIP simulations used in this study are available from the Earth System Grid Federation (ESGF). ERA5, ERA-40 and ERA-Interim reanalyses data are available from the European Centre for Medium Range Weather Forecasting (ECMWF). The HadISST2 high-resolution SST data used (1950–2014) are also available from the ESGF (https://doi.org/10.22033/ESGF/input4MIPs.1221). The OAFlux data used in this study are available from the Research Data Archive of the National Center for Atmospheric Research (NCAR, https://rda.ucar.edu/datasets/ds260.1).

References

Ambaum, M. H. P., and P. J. Athanasiadis, 2007: The response of a uniform horizontal temperature gradient to heating. *J. Atmos. Sci.*, **64**, 3708–3716, doi:https://doi.org/10.1175/JAS4038.1.

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-21-0515.1.

Anstey, J. A., and Coauthors, 2013: Multi-model analysis of Northern Hemisphere winter blocking: Model biases and the role of resolution. *J. Geophys. Res. Atmos.*, **118**, 3956–3971, doi:doi: 10.1002/jgrd.50231.

Athanasiadis, P. J., J. M. Wallace, and J. J. Wettstein, 2010: Patterns of wintertime jet stream variability and their relation to the storm tracks. *J. Atmos. Sci.*, **67**, 1361–1381, doi:https: //journals.ametsoc.org/view/journals/atsc/67/5/2009jas3270.1.xml.

Athanasiadis, P. J., S. Yeager, Y.-O. Kwon, A. Bellucci, D. W. Smith, and S. Tibaldi, 2020: Decadal predictability of North Atlantic blocking and the NAO. *npj Clim Atmos Sci*, **3**, 20, doi:https://doi.org/10.1038/s41612-020-0120-6.

Athanasiadis, P. J., and Coauthors, 2014: The representation of atmospheric blocking and the associated low-frequency variability in two seasonal prediction systems. *J. Climate*, **27**, 9082–9100, doi:https://doi.org/10.1175/JCLI-D-14-00291.1.

Berckmans, J., T. Woollings, M. Demory, P. Vidale, and M. Roberts, 2013: Atmospheric blocking in a high resolution climate model: influences of mean state, orography and eddy forcing. *Atmos. Sci. Lett.*, **14**, 34–40, doi:https://doi.org/10.1002/asl2.412.

Booth, J. F., E. Dunn-Sigouin, and S. Pfahl, 2012: The relationship between extratropical cyclone steering and blocking along the North American East coast. *Geophys. Res. Lett.*, **44**, 11 976–11 984, doi:https://doi.org/10.1002/2017GL075941.

Brayshaw, D. J., B. Hoskins, and M. Blackburn, 2011: The basic ingredients of the North Atlantic storm track. Part II: sea surface temperatures. *J. Atmos. Sci.*, **68**, 1784–1805, doi:https://doi.org/ 10.1175/2011JAS3674.1.

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-21-0515.1.

Chassignet, E., and Coauthors, 2020: Impact of horizontal resolution on global ocean-sea-ice model simulations based on the experimental protocols of the Ocean Model Intercomparison Project phase 2 (OMIP-2). *Geosci. Model Dev. Disc.*, 1–58, doi:https://doi.org/10.5194/gmd-2019-374.

Cherchi, A., and Coauthors, 2019: Global mean climate and main patterns of variability in the CMCC-CM2 coupled model. *Journal of Advances in Modeling Earth Systems*, **11**, 185–209, doi:{https://doi.org/10.1029/2018MS001369}.

Croci-Maspoli, M., and H. C. Davies, 2009: Key dynamical features of the 2005/06 European winter. *Mon. Wea. Rev.*, **137**, 664–678, doi:https://doi.org/10.1175/2008MWR2533.1.

Czaja, A., C. Frankignoul, S. Minobe, and B. Vannière, 2019: Simulating the midlatitude atmospheric circulation: what might we gain from high-resolution modeling of air-sea interactions? *Curr. Clim. Change Rep.*, **5**, 390–406, doi:https://doi.org/10.1007/s40641-019-00148-5.

Davini, P., C. Cagnazzo, S. Gualdi, and A. Navarra, 2012: Bidimensional diagnostics, variability, and trends of northern hemisphere blocking. *J. Climate*, **25**, 6496–6509, doi:https://journals.ametsoc.org/view/journals/clim/25/19/jcli-d-12-00032.1.xml.

Davini, P., and F. D'Andrea, 2016: Northern hemisphere atmospheric blocking representation in global climate models: Twenty years of improvements?,. *J. Climate*, **29**, 8823–8840, doi:https://doi.org/10.1175/JCLI-D-16-0242.1.

de Coëtlogon, G., C. Frankignoul, M. Bentsen, C. Delon, H. Haak, S. Masina, and A. Pardaens, 2006: Gulf stream variability in five Oceanic General Circulation Models. *Journal of Physical Oceanography*, **36**, 2119–2135, doi:https://doi.org/10.1175/JPO2963.1.

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-21-0515.1.

Dee, D. P., and co-authors, 2011: The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **137**, 553–597, doi:https://doi.org/10.1002/qj.828,http://doi.wiley.com/10.1002/qj.828.

Deser, C., M. A. Alexander, S.-P. Xie, and A. S. Phillips, 2010: Sea surface temperature variability: patterns and mechanisms. *Annual Review of Marine Science*, **2(1)**, 115–143, doi:10.1146/annurev-marine-120408-151453.

Drews, A., R. Greatbatch, H. Ding, M. Latif, and W. Park, 2015: The use of a flow field correction technique for alleviating the North Atlantic cold bias with application to the Kiel Climate Model. *Ocean Dynamics*, **65**, 1079–1093, doi:https://doi.org/10.1007/s10236-015-0853-7.

Eden, C., and T. Jung, 2001: North Atlantic interdecadal variability: oceanic response to the North Atlantic Oscillation (1865–1997). *J. Climate*, **14**, 676–691, doi:https://doi.org/10.1175/1520-0442(2001)014<3C0676:NAIVOR>3E2.0.CO;2.

Foussard, A., G. Lapeyre, and R. Plougonven, 2019: Storm track response to oceanic eddies in idealized atmospheric simulations. *J. Climate*, **32**, 445–463, doi:https://doi.org/10.1175/JCLI-D-18-0415.1.

Frankignoul, C., G. de Coëtlogon, T. M. Joyce, and S. Dong, 2001: Gulf stream variability and ocean–atmosphere interactions. *Journal of Physical Oceanography*, **31**, 3516–3529, doi:https://doi.org/10.1175/1520-0485(2002)031<3C3516:GSVAOA>3E2.0.CO;2.

Gutjahr, O., D. Putrasahan, K. Lohmann, J. H. Jungclaus, J. S. von Storch, N. Brüggemann, H. Haak, and A. St ossel, 2019: Max Planck Institute Earth System Model (MPI-ESM1.2) for the High-Resolution Model Intercomparison Project (HighResMIP). *Geosci. Model Dev.*, **12**, 3241–3281, doi:https://doi.org/10.5194/gmd-12-3241-2019.

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-21-0515.1.

Haarsma, R., J. García-Serrano, C. Prodhomme, O. Bellprat, P. Davini, and S. Drijfhout, 2019: Sensitivity of winter North Atlantic-European climate to resolved atmosphere and ocean dynamics. *Sci. Rep.*, **9**, 13 358, doi:https://doi.org/10.1038/s41598-019-49865-9.

Haarsma, R., and Coauthors, 2020: HighResMIP versions of EC-EARTH: EC-EARTH3P and EC-EARTH3P-HR. Description, model performance, data handling and validation. *Geoscientific Model Development Discussions*, 1–37, doi:https://doi.org/10.5194/gmd-2019-350.

Haarsma, R. J., and Coauthors, 2016: High Resolution Model Intercomparison Project (HighResMIP v1.0) for CMIP6. *Geosci. Model Dev.*, **9**, 4185–4208, doi:https://doi.org/10.5194/gmd-9-4185-2016.

Hersbach, H., B. Bell, and P. e. a. Berrisford, 2020: The ERA5 global reanalysis. *Quart. J. Roy. Meteor. Soc.*, **146**, 1999–2049, doi:https://doi.org/10.1002/qj.3803.

Hoskins, B. J., I. N. James, and G. H. White, 1983: The shape, propagation and mean-flow interaction of large-scale weather systems. *J. Atmos. Sci.*, **40**, 1595–1612, doi:https://journals.ametsoc.org/view/journals/atsc/40/7/1520-0469_1983_040_1595_tspamf_2_0_co_2.xml.

Hoskins, B. J., M. E. McIntyre, and A. W. Robertson, 1985: On the use and the significance of isentropic potential vorticity maps. *Quart. J. Roy. Meteor. Soc.*, **111**, 877–946, doi:https://doi.org/10.1002/qj.49711147002.

Hoskins, B. J., and P. J. Valdes, 1990: On the existence of storm-tracks. *J. Atmos. Sci.*, **47**, 1854–1864, doi:doi:10.1175/1520-0469(1990)0471854:OTEOST2.0.CO;2.

Hotta, D., and H. Nakamura, 2011: On the significance of the sensible heat supply from the ocean in the maintenance of the mean baroclinicity along storm tracks. *J. Climate*, **24**, 3377–3401, doi:https://journals.ametsoc.org/view/journals/clim/24/13/2010jcli3910.1.xml.

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-21-0515.1.

Illari, L., and J. C. Marshall, 1983: On the interpretation of eddy fluxes during a blocking episode. *J. Atmos. Sci.*, **40**, 2232–2242, doi:https://doi.org/10.1175/1520-0469(1983)040%3C2232:OTIOEF%3E2.0.CO;2.

Iqbal, W., W. Leung, and A. Hannachi, 2018: Analysis of the variability of the North Atlantic eddy-driven jet stream in CMIP5. *Climate Dyn.*, **51**, 235–247, doi:https://doi.org/10.1007/s00382-017-3917-1.

James, I. N., 1994: *Introduction to circulating atmospheres*. Cambridge atmospheric and space science series, Cambridge University Press.

Jiang, T., K. Evans, M. Branstetter, P. Caldwell, R. Neale, P. J. Rasch, Q. Tang, and S. Xie, 2019: Northern Hemisphere blocking in ∼25-km-resolution E3SM v0.3 atmosphere-land simulations. *Journal of Geophysical Research: Atmospheres*, **124**, 2465–2482, doi:https://doi.org/10.1029/2018JD028892.

Joyce, T. M., Y. Kwon, H. Seo, and C. C. Ummenhofer, 2019: Meridional Gulf Stream shifts can influence wintertime variability in the North Atlantic storm track and Greenland blocking. *Geophys. Res. Lett.*, **46**, 1702–1708, doi:https://doi.org/10.1029/2018GL081087.

Joyce, T. M., Y. Kwon, and L. Yu, 2009: On the Relationship between synoptic wintertime atmospheric variability and path shifts in the Gulf Stream and the Kuroshio Extension. *J. Climate*, **22**, 3177–3192, doi:https://doi.org/10.1175/2008JCLI2690.1.

Keeley, S., R. Sutton, and L. Shaffrey, 2012: The impact of North Atlantic sea surface temperature errors on the simulation of North Atlantic European region climate. *Quart. J. Roy. Meteor. Soc.*, **138**, 1774–1783, doi:https://doi.org/10.1002/qj.1912.

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-21-0515.1.

Kennedy, J., H. Titchner, N. Rayner, and M. Roberts, 2017: input4MIPs.MOHC.SSTsAndSeaIce.HighResMIP.MOHC-HadISST-2-2-0-0-0, version 20170505. *Earth System Grid Federation*, doi:https://doi.org/10.22033/ESGF/input4MIPs.1221, accessed 5 May 2017.

Koenigk, T., and Coauthors, 2022: Deep mixed ocean volume in the Labrador Sea in HighResMIP models. *Climate Dyn.*, **57**, 1895–1918, doi:https://doi.org/10.1007/s00382-021-05785-x.

Kushnir, Y., W. A. Robinson, I. Blade, N. M. J. Hall, S. Peng, and R. Sutton, 2002: Atmospheric GCM response to extratropical SST anomalies: synthesis and evaluation. *J. Climate*, **15**, 2233–2256, doi:https://doi.org/10.1175/1520-0442(2002)015%3C2233:AGRTES%3E2.0.CO;2.

Lee, R. W., T. J. Woollings, B. J. Hoskins, K. D. Williams, C. H. O'Reilly, and G. Masato, 2018: Impact of Gulf Stream SST biases on the global atmospheric circulation. *Climate Dyn.*, **51**, 3369–3387, doi:https://doi.org/10.1007/s00382-018-4083-9.

Madonna, L. C. G. C. M., E., and T. Woollings, 2017: The link between eddy-driven jet variability and weather regimes in the North Atlantic-European sector. *Quart. J. Roy. Meteor. Soc.*, **143**, 2960–2972, doi:doi:10.1002/qj.3155.

Marzocchi, A., J. J.-M. Hirschi, N. P. Holliday, S. A. Cunningham, A. T. Blaker, and A. C. Coward, 2015: The North Atlantic subpolar circulation in an eddy-resolving global ocean model. *Journal of Marine Systems*, **142**, 126–143, doi:https://doi.org/10.1016/j.jmarsys.2014.10.007.

Masato, G., B. J. Hoskins, and T. Woolings, 2012: Wave-breaking characteristics of midlatitude blocking. *Quart. J. Roy. Meteor. Soc.*, **138**, 1285–1296.

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-21-0515.1.

Masato, G., B. J. Hoskins, and T. Woollings, 2013: Winter and Summer Northern Hemisphere Blocking in CMIP5 Models. *J. Climate*, **26**, 7044–7059, doi:https://journals.ametsoc.org/view/journals/clim/26/18/jcli-d-12-00466.1.xml.

Masato, G., T. Woollings, K. Williams, B. Hoskins, and R. Lee, 2016: A regime analysis of Atlantic winter jet variability applied to evaluate HadGEM3-GC2. *Quart. J. Roy. Meteor. Soc.*, **142**, 3162–3170, doi:https://doi.org/10.1002/qj.2897.

McCarthy, G. D., T. M. M. Joyce, and S. A. Josey, 2018: Gulf Stream variability in the context of quasi-decadal and multidecadal Atlantic climate variability. *Geophys. Res. Lett.*, **45**, 11 257–11 264, doi:https://doi.org/10.1029/2018GL079336.

Messori, G., P. Davini, M. C. Alvarez-Castro, F. S. R. Pausata, P. Yiou, and R.Caballero, 2019: On the low-frequency variability of wintertime euro-atlantic planetary wave-breaking. *Climate Dyn.*, **52**, 2431–2450, doi:https://doi.org/10.1007/s00382-018-4373-2.

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:doi:10.1038/nature06690.

Moreno-Chamarro, E., L.-P. Caron, P. Ortega, L. T. S., and M. J. Roberts, 2021: Can we trust CMIP5/6 future projections of European winter precipitation? *Environ. Res. Lett.*, **16**, doi:https://doi.org/10.1088/1748-9326/abf28a.

Moreno-Chamarro, E., and Coauthors, 2022: Impact of increased resolution on long-standing biases in HighResMIP-PRIMAVERA climate models. *Geosci. Model Dev. Disc.*, **2022**, 1–33, doi:https://doi.org/10.5194/gmd-15-269-2022.

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-21-0515.1.

Nakamura, H., M. Nakamura, and J. L. Anderson, 1997: The role of high- and low-frequency dynamics in blocking formation. *Mon. Wea. Rev.*, **125**, 2074–2093, doi:https://doi.org/10.1175/1520-0493(1997)125%3C2074:TROHAL%3E2.0.CO;2.

Nakamura, H., T. Sampe, A. Goto, W. Ohfuchi, and S.-P. Xie, 2008: On the importance of midlatitude oceanic frontal zones for the mean state and dominant variability in the tropospheric circulation. *Geophys. Res. Lett.*, **35**, L15 709, doi:doi:10.1029/2008GL034010.

Nakamura, H., 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*, C. Wang, S.-P. Xie, and J. A. Carton, Eds., American Geophysical Union, 329–346, doi:https://doi.org/10.1029/147GM18.

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:https://journals.ametsoc.org/view/journals/clim/22/4/2008jcli2297.1.xml.

Nakamura, N., and C. S. Y. Huang, 2018: Atmospheric blocking as a traffic jam in the jet stream. *Science*, **361**, 42–47, doi:https://doi.org/10.1126/science.aat0721.

Namias, J., 1964: Seasonal persistence and recurrence of european blocking during 1958–1960. *Tellus*, **16**, 394–407, doi:https://doi.org/10.1111/j.2153-3490.1964.tb00176.x.

Naoe, H., and Y. Matsuda, 2002: Rossby wave propagation and blocking formation in realistic basic flows. *J. Meteor. Soc. Japan*, **80**, 717–731, doi:https://doi.org/10.2151/jmsj.80.717.

Novak, L., M. H. P. Ambaum, and R. Tailleux, 2015: The life cycle of the North Atlantic storm track. *J. Atmos. Sci.*, **72**, 821–833, doi:https://doi.org/10.1175/JAS-D-14-0082.1.

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-21-0515.1.

Omrani, N.-E., F. Ogawa, H. Nakamura, N. Keenlyside, S. W. L., and M. Katja, 2019: Key role of the ocean western boundary currents in shaping the Northern Hemisphere climate. **9**, 3014, doi:https://doi.org/10.1038/s41598-019-39392-y.

O'Reilly, C., S. Minobe, and A. Kuwano-Yoshida, 2016: The influence of the Gulf Stream on wintertime European blocking. *Climate Dyn.*, **47**, 1545–1567, doi:https://doi.org/10.1007/s00382-015-2919-0.

O'Reilly, C., S. Minobe, A. Kuwano-Yoshida, and T. Woollings, 2017: The Gulf Stream influence on wintertime North Atlantic jet variability. *Quart. J. Roy. Meteor. Soc.*, **143**, 173–183, doi:https://doi.org/10.1002/qj.2907.

Orlanski, I., 1998: Poleward deflection of storm tracks. *J. Atmos. Sci.*, **55**, 2577–2602, doi:https://doi.org/10.1175/1520-0469(1998)055%3C2577:PDOST%3E2.0.CO;2.

Papritz, L., and T. Spengler, 2015: Analysis of the slope of isentropic surfaces and its tendencies over the North Atlantic. *Quart. J. Roy. Meteor. Soc.*, **141**, 3226–3238, doi:https://doi.org/10.1002/qj.2605.

Parfitt, R., A. Czaja, S. Minobe, and A. Kuwano-Yoshida, 2016: The atmospheric frontal response to SST perturbations in the Gulf Stream region. *Geophys. Res. Lett.*, **43**, 2299–2306, doi:doi:10.1002/2016GL067723.

Pfahl, S., C. Schwierz, C. M. Grams, H. Wernli, M. Croci-Maspoli, C. M. Grams, and H. Wernli, 2015: Importance of latent heat release in ascending air streams for atmospheric blocking. *Nat. Geosci.*, **8**, 610–615, doi:https://doi.org/10.1038/NGEO2487.

Roberts, C. D., R. Senan, F. Molteni, S. Boussetta, M. Mayer, and S. P. E. Keeley, 2018: Climate model configurations of the ECMWF Integrated Forecasting System (ECMWF-IFS cy-

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-21-0515.1.

cle 43R1) for HighResMIP. *Geosci. Model Dev.*, **11**, 3681–3712, doi:https://doi.org/10.5194/gmd-11-3681-2018.

Roberts, C. D., F. Vitart, M. A. Balmaseda, and F. Molteni, 2020: The time-scale-dependent response of the wintertime North Atlantic to increased ocean model resolution in a coupled forecast model. *J. Climate*, **33**, 3663–3689, doi:https://journals.ametsoc.org/view/journals/clim/33/9/jcli-d-19-0235.1.xml.

Roberts, M. J., and Coauthors, 2019: Description of the resolution hierarchy of the global coupled HadGEM3-GC3.1 model as used in CMIP6 HighResMIP experiments. *Geosci. Model Dev.*, **12**, 4999–5028, doi:10.5194/gmd-12-4999-2019.

Rohrer, M., S. Brönnimann, O. Martius, C. C. Raible, and M. Wild, 2019: Decadal variations of blocking and storm tracks in centennial reanalyses. *Tellus*, **71**, 1–21, doi:doi:10.1080/16000870.2019.1586236.

Sampe, T., H. Nakamura, A. Goto, and W. Ohfuchi, 2010: Significance of a midlatitude SST frontal zone in the formation of a storm track and an eddy-rriven westerly jet. *J. Climate*, **23**, 1793–1814, doi:https://doi.org/10.1175/2009JCLI3163.1.

Scaife, A., and D. Smith, 2018: A signal-to-noise paradox in climate science. *npj Clim Atmos Sci*, **1**, 28, doi:https://doi.org/10.1038/s41612-018-0038-4.

Scaife, A. A., and Coauthors, 2011: Improved Atlantic winter blocking in a climate model. *Geophys. Res. Lett.*, **38**, n/a–n/a, doi:10.1029/2011GL049573.

Scherrer, S. C., M. Croci-Maspoli, C. Schwierz, and C. Appenzeller, 2006: Two-dimensional indices of atmospheric blocking and their statistical relationship with winter climate patterns in the Euro-Atlantic region. *Int. J. Climatol.*, **26**, 233–249, doi:https://doi.org/10.1002/joc.1250.

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-21-0515.1.

Schiemann, R., and Coauthors, 2017: The resolution sensitivity of Northern Hemisphere blocking in four 25-km Atmospheric Global Circulation Models. *J. Climate*, **30**, 337–358, doi:https://doi.org/10.1175/JCLI-D-16-0100.1.

Schiemann, R., and Coauthors, 2020: Northern Hemisphere blocking simulation in current climate models: evaluating progress from the Climate Model Intercomparison Project Phase 5 to 6 and sensitivity to resolution. *Wea. Clim. Dyn.*, **1**, 277–292, doi:https://doi.org/10.5194/wcd-1-277-2020.

Sein, D. V., and Coauthors, 2017: Ocean modeling on a mesh with resolution following the local Rossby radius. *J. Adv. Model Earth Syst.*, **9**, 2601–2614, doi:https://doi.org/10.1002/2017MS001099.

Sheldon, L., A. Czaja, B. Vanniére, C. Morcrette, B. Sohet, M. Casado, and D. Smith, 2017: A 'warm path' for Gulf Stream-troposphere interactions. *Tellus*, **69**, 1–15, doi:https://doi.org/10.1080/16000870.2017.1299397.

Shutts, G. J., 1983: The propagation of eddies in diffluent jetstreams: eddy vorticity forcing of 'blocking' flow fields. *Quart. J. Roy. Meteor. Soc.*, **109**, 737–761, doi:https://doi.org/10.1002/qj.49710946204.

Shutts, G. J., 1986: A case study of eddy forcing during an atlantic blocking episode. *Advances in Geophysics*, **29**, 135–162.

Siqueira, L., and B. P. Kirtman, 2016: Atlantic near-term climate variability and the role of a resolved Gulf Stream. *Geophys. Res. Lett.*, **43**, 3964–72, doi:https://doi.org/10.1002/2016gl068694.

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-21-0515.1.

Small, R., R. Tomas, and F. Bryan, 2014: Storm track response to ocean fronts in a global high-resolution climate model. *Climate Dyn.*, **43**, 805–828, doi:https://doi.org/10.1007/s00382-013-1980-9.

Smirnov, D., M. Newman, M. A. Alexander, Y.-O. Kwon, and C. Frankignoul, 2015: Investigating the local atmospheric response to a realistic shift in the Oyashio sea surface temperature front. *J. Climate*, **28**, 1126–1147, doi:https://doi.org/10.1175/JCLI-D-14-00285.1.

Steinfeld, D., and S. Pfahl, 2019: The role of latent heating in atmospheric blocking dynamics: a global climatology. *Climate Dyn.*, **53**, 6159–6180, doi:https://doi.org/10.1007/s00382-019-04919-6.

Stone, P. H., 1978: Baroclinic adjustment. *J. Atmos. Sci.*, **35**, 561–571, doi:https://doi.org/10.1175/1520-0469(1978)035%3C0561:BA%3E2.0.CO;2.

Storkey, D., and Coauthors, 2018: UK Global Ocean GO6 and GO7: a traceable hierarchy of model resolutions. *Geosci. Model Dev.*, **11**, 3187–3213, doi:https://doi.org/10.5194/gmd-11-3187-2018.

Taylor, A. H., and J. A. Stephens, 1998: The North Atlantic Oscillation and the latitude of the Gulf Stream. *TELLUS A: Dynamic Meteorology and Oceanography*, **50**, 134–142, doi:10.3402/tellusa.v50i1.14517.

Tsujino, H., and Coauthors, 2020: Evaluation of global ocean–sea-ice model simulations based on the experimental protocols of the Ocean Model Intercomparison Project phase 2 (OMIP-2). *Geosci. Model Dev.*, **13**, 3643–3708, doi:https://doi.org/10.5194/gmd-13-3643-2020.

Uppala, S. M., and co-authors, 2005: The ERA-40 re-analysis. *Quart. J. Roy. Meteor. Soc.*, **31**, 2961–3012, doi:https://doi.org/10.1256/qj.04.176,http://doi.wiley.com/10.1256/qj.04.176.

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-21-0515.1.

Voldoire, A., and Coauthors, 2019: Evaluation of CMIP6 deck experiments with CNRM-CM6-1. *J. Adv. Model Earth Syst.*, **11**, 2177–2213, doi:https://doi.org/10.1029/2019MS001683.

Wallace, J. M., and D. S. Gutzler, 1981: Teleconnections in the geopotential height field during the Northern Hemisphere winter. *Mon. Wea. Rev.*, **109**, 784–812, doi:https://doi.org/10.1175/1520-0493(1981)109<0784:TITGHF>2.0.CO;2.

Wang, C., L. Zhang, S.-K. Lee, L. Wu, and C. R. Mechoso, 2014: A global perspective on CMIP5 climate model biases. *Nature Cli. Change*, **4**, 201–205, doi:https://doi.org/10.1038/nclimate2118.

Weijenborg, C., H. de Vries, and R. Haarsma, 2012: On the direction of Rossby wave breaking in blocking. *Climate Dyn.*, **39**, 2823–2831, doi:https://doi.org/10.1007/s00382-012-1332-1.

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:https://doi.org/10.1175/JAS-D-12-0226.1.

Woollings, T., A. Hannachi, and B. J. Hoskins, 2010a: Variability of the North Atlantic eddy-driven jet stream. *Quart. J. Roy. Meteor. Soc.*, **136**, 856–868.

Woollings, T., B. Hoskins, M. Blackburn, D. Hassell, and K. Hodges, 2010b: Storm track sensitivity to sea surface temperature resolution in a regional atmosphere model. *Climate Dyn.*, **35**, 341–353, doi:https://doi.org/10.1007/s00382-009-0554-3.

Woollings, T., B. J. Hoskins, M. Blackburn, and P. Berrisford, 2008: A new Rossby wave-breaking interpetation of the North Atlantic Oscillation. *J. Atmos. Sci.*, **65**, 609–626, doi:https://doi.org/10.1175/2007JAS2347.1.

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-21-0515.1.

Yamamoto, A., M. Nonaka, P. Martineau, A. Yamazaki, Y.-O. Kwon, H. Nakamura, and B. Taguchi, 2020: Oceanic origins for wintertime Euro-Atlantic blocking. *Weather Clim. Dynam. Discuss.*, in review, doi:https://doi.org/10.5194/wcd-2020-39.

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

Zappa, G., L. C. Shaffrey, and K. I. Hodges, 2013: The ability of CMIP5 models to simulate North Atlantic extratropical cyclones. *J. Climate*, **26**, 5379–5396, doi:https://doi.org/10.1175/JCLI-D-12-00501.1.

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-21-0515.1.

LIST OF TABLES

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-21-0515.1.

1037 TABLE 1. PRIMAVERA models and simulations. Columns detail the model name,
1038 the atmosphere grid spacing at 50°N, nominal ocean grid spacing, and the respective
1039 ensemble.

No.	Model	Atm. grid (km)	Ocean grid (km)	LR / HR	LOW / HIGH
1	AWI-CM-1-1-LR	129	50	LR	–
2	AWI-CM-1-1-HR	67	25	HR	–
3	CMCC-CM2-HR4	64	25	LR	LOW-RES
4	CMCC-CM2-VHR4	18	25	HR	HIGH-RES
5	CNRM-CM6-1	142	100	LR	LOW-RES
6	CNRM-CM6-1-HR	50	25	HR	HIGH-RES
7	EC-Earth3P	71	100	LR	LOW-RES
8	EC-Earth3P-HR	36	25	HR	HIGH-RES
9	ECMWF-IFS-LR	50	100	LR	LOW-RES
10	ECMWF-IFS-MR	50	25	HR	HIGH-RES
11	ECMWF-IFS-HR	25	25	HR	–
12	HadGEM3-GC31-LL	135	100	LR	LOW-RES
13	HadGEM3-GC31-MM	60	25	HR	HIGH-RES
14	HadGEM3-GC31-HM	25	25	HR	–
15	HadGEM3-GC31-HH	25	8	HR	HIGH-RES
16	MPI-ESM1-2-HR	67	40	LR	LOW-RES
17	MPI-ESM1-2-XR	34	40	HR	HIGH-RES

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-21-0515.1.

LIST OF FIGURES

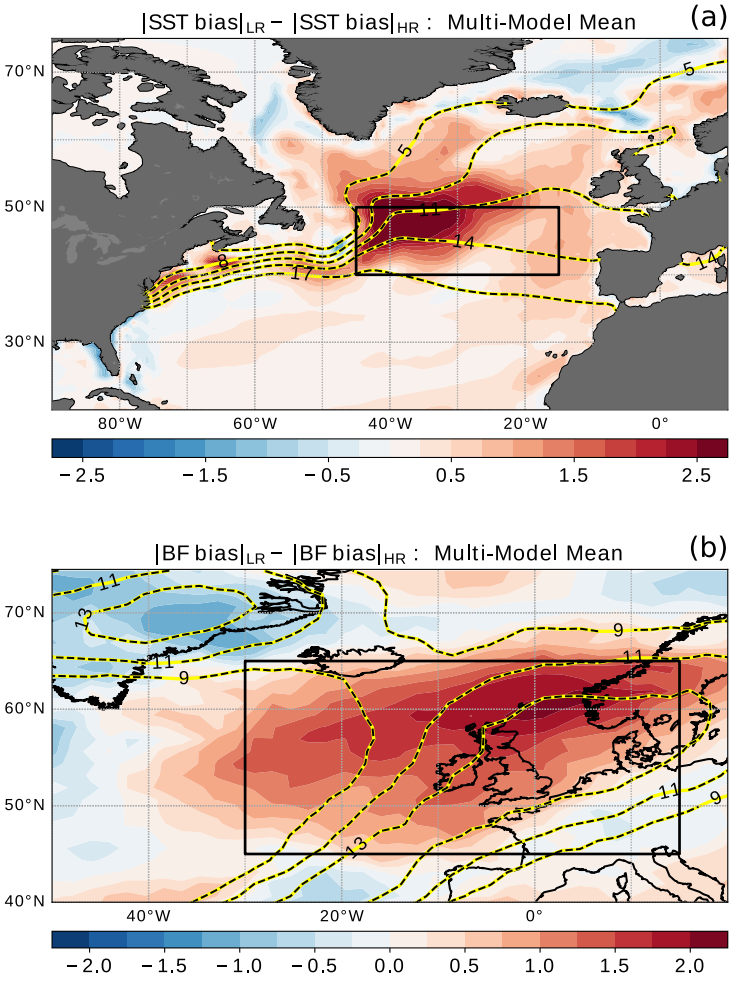


FIG. 1. Differences in wintertime absolute bias between "LR" and "HR" multi-model means (Table 1). Upper panel: SST biases (shading, in K) with the HadISST2 1950–2014 climatology in contours (°C). Lower panel: blocking frequency biases (shading, in % of blocked days) with the ERA-JOINT climatological blocking frequency in contours (% of blocked days). Details on the blocking detection method are provided in the text. The rectangular frames (selected areas) correspond to [15–45°W, 40–50°N] (upper panel) and [30°W–15°E, 45–65°N] (lower panel).

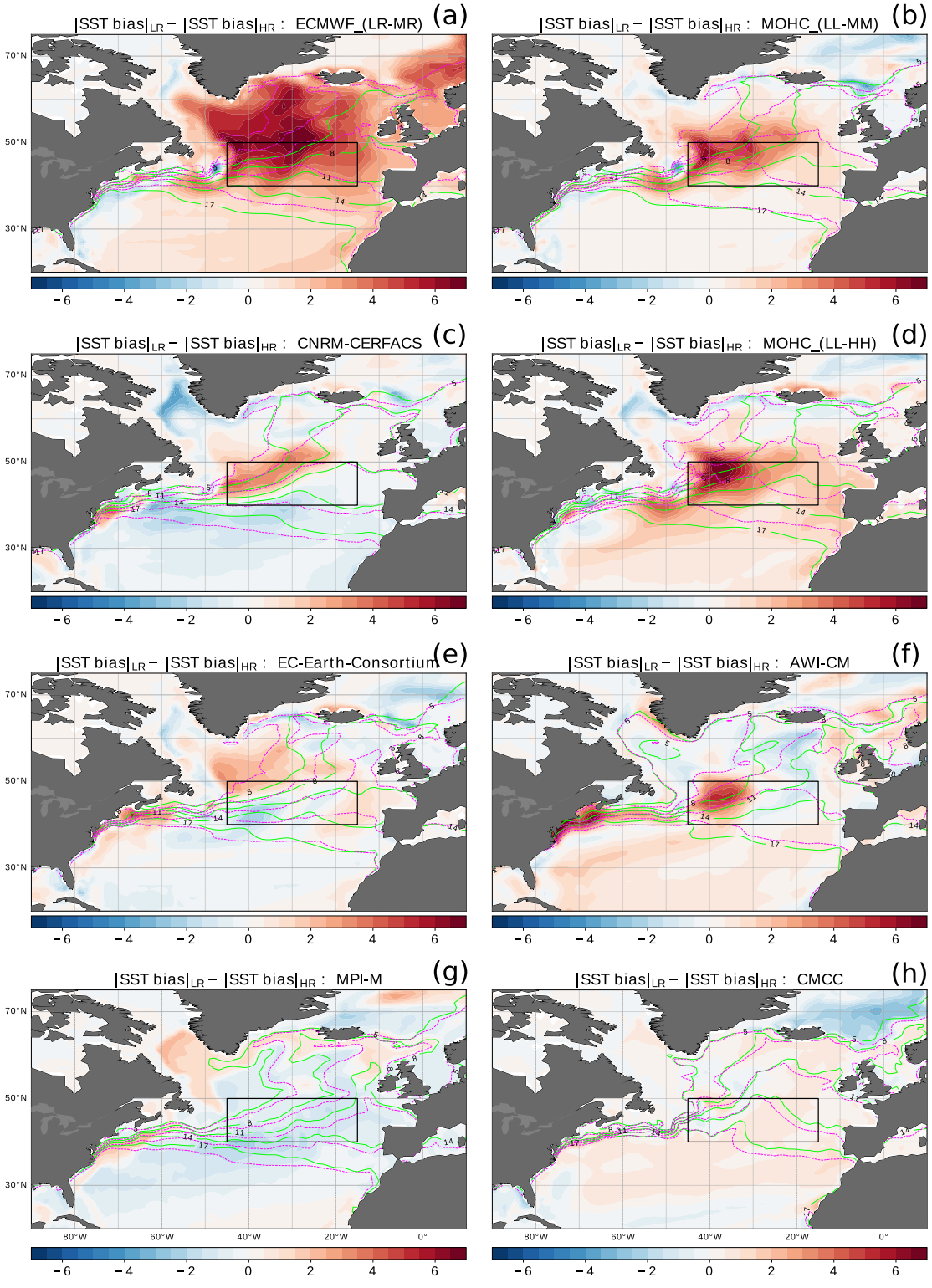
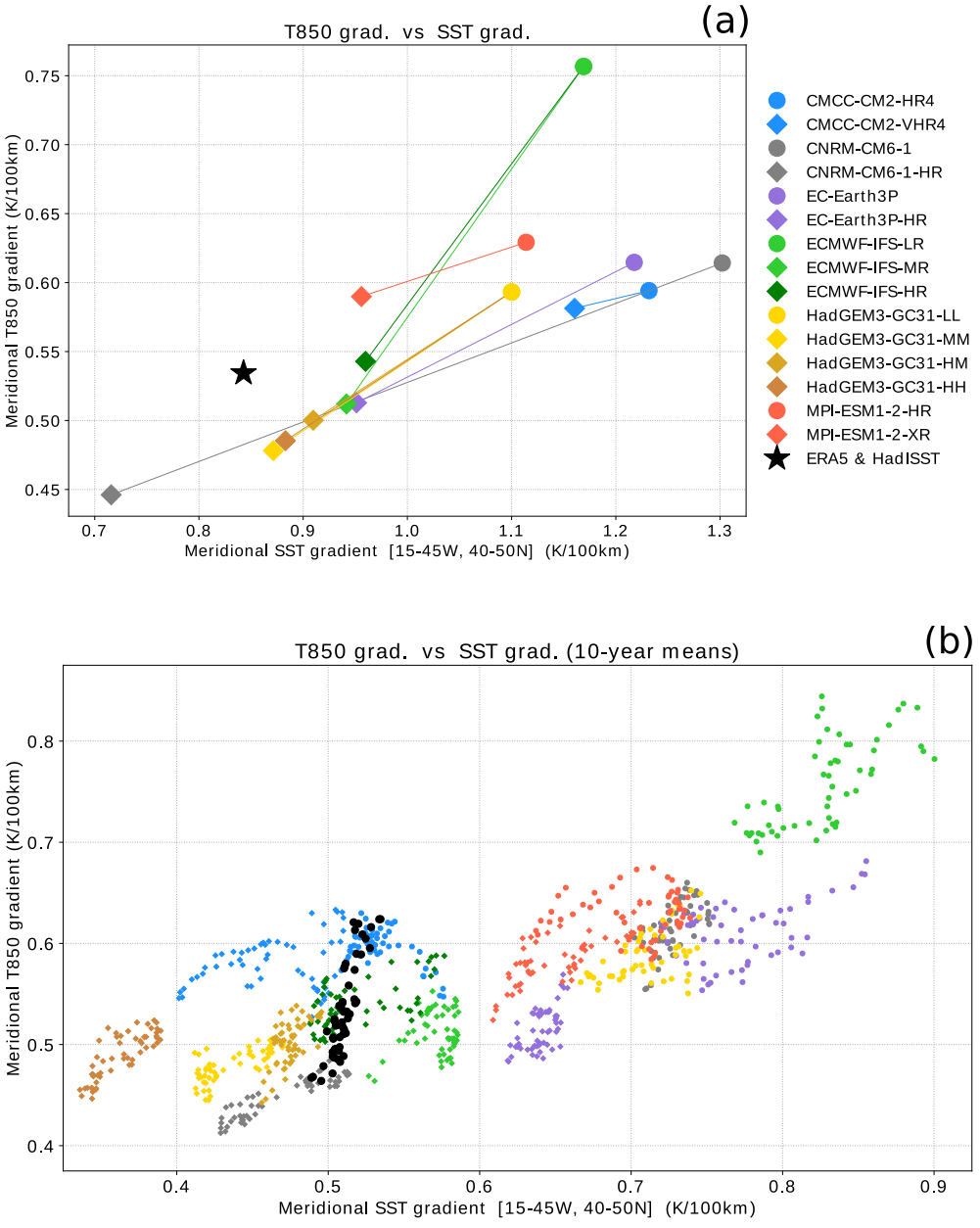


FIG. 2. Differences in absolute SST bias between "LR" and "HR" model pairs (Table 1). Climatological isotherms are shown in contours, solid green for "LR" models and dashed violet for "HR" models. The rectangular frame defines the area of interest (referred to as Central North Atlantic in the text) where the model biases modify the climatological meridional SST gradient. The bottom panels correspond to the models (CMCC and MPI-M) that increase only the atmospheric resolution. In (b) and (d), MOHC stands for *Met. Office Hadley Centre* and is used as a shortcut for the HadGEM3-GC31 model. All rectangular frames correspond to [15–45°W, 40–50°N], as in Fig. 1a.

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-21-0515.1.



1123 FIG. 3. Upper panel: climatological meridional gradient of air-temperature at 850 hPa, spatially
1124 averaged in the area [15–45°W, 40–50°N] as indicated by the frame in Fig. 1, versus the respective
1125 climatological meridional SST gradient (details in the text) in the same area for each model in
1126 Table 1 (except AWI-CM due to missing data). Units: $\mathrm{K}\,(100\,\mathrm{km})^{-1}$. Round markers for LR
1127 models, diamond markers for HR models and the black asterisk for the respective observations.
1128 Lower panel: as for the upper panel but showing values for 10-year winter means in the same
1129 historical period (1950–2014) and using simple spatial averages for both gradients (T850 and SST).
1130 Gradients are considered to be positive equatorward. Same color code as in the legend above.

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-21-0515.1.

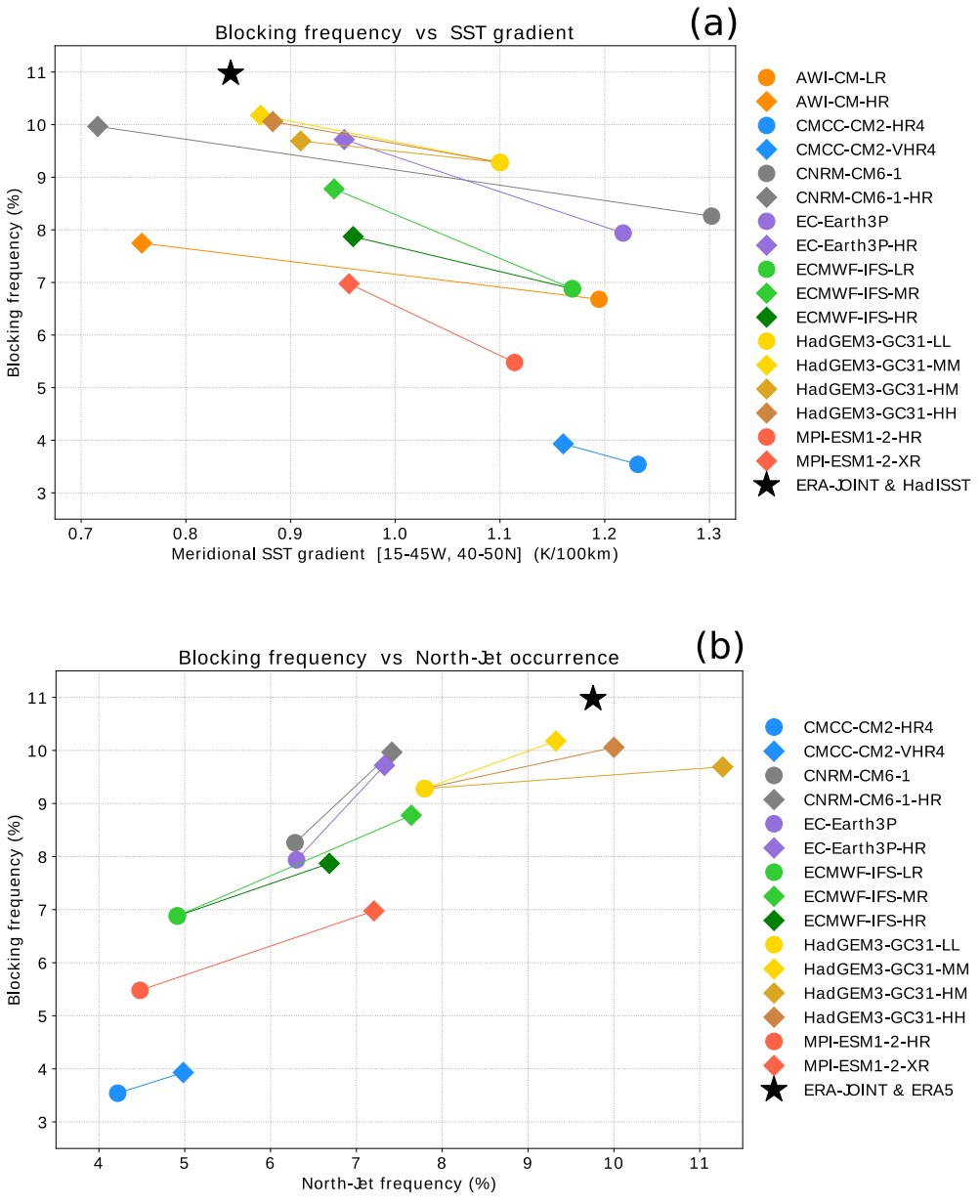


FIG. 4. Upper panel: climatological instantaneous blocking frequency, spatially averaged in the area [30°W–15°E, 45–65°N] as indicated by the frame in Fig. 1, versus the respective climatological meridional SST gradient (details in the text) in the area [15–45°W, 40–50°N] for each LR and HR model (Table 1). Units: % of blocked days and K (100 km)$^{-1}$. Round markers for LR models, diamond markers for HR models and the black asterisk for the respective observations. Lower panel: as for the upper panel but with the climatological frequency of the northern-jet regime (jet latitude > 60°N, details in the text) in the x-axis.

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-21-0515.1.

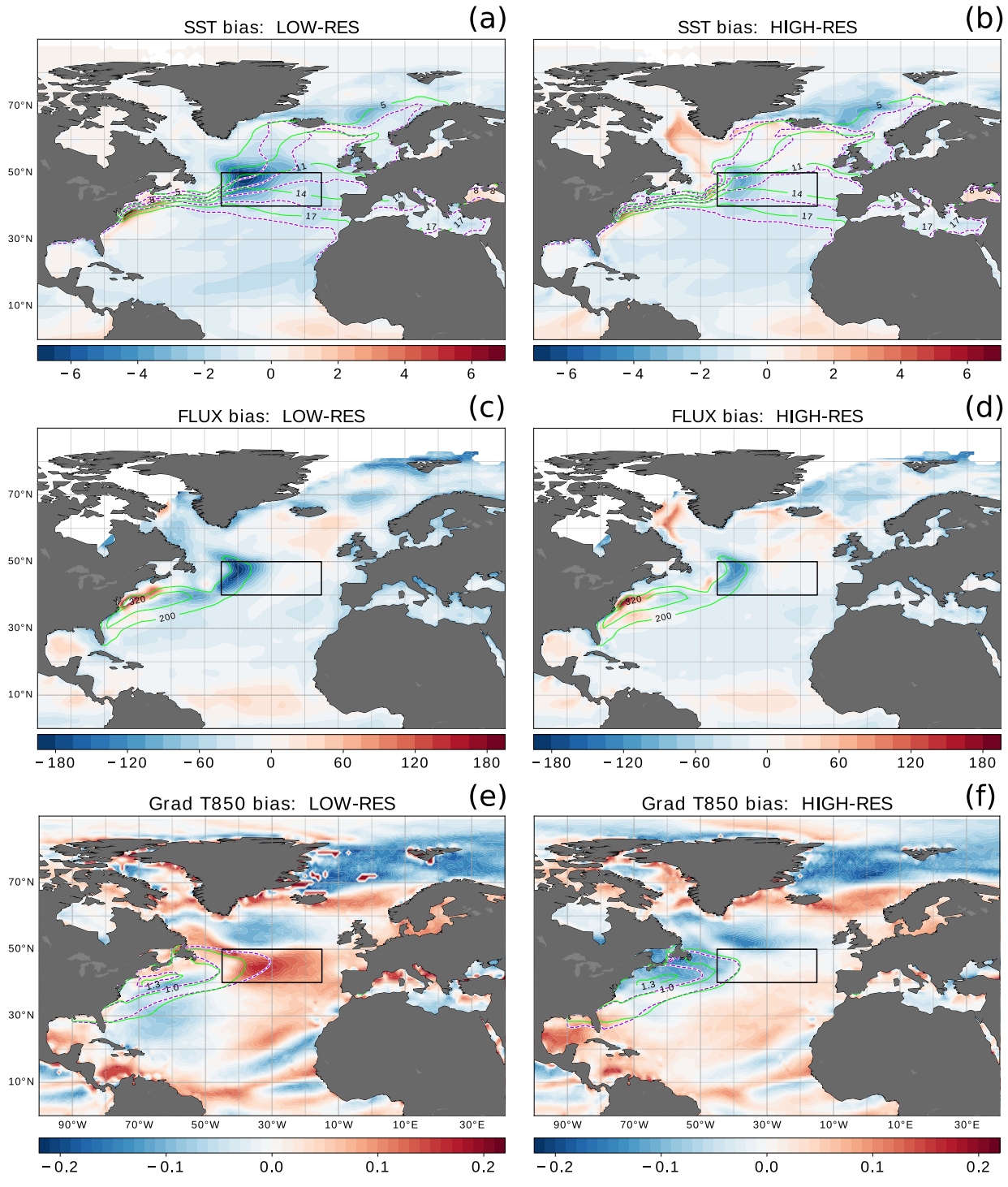


FIG. 5. Multi-model mean biases in wintertime for the LOW-RES (left) and HIGH-RES (right) sub-ensembles as indicated in Table 1. Upper panels: SST biases with respect to HadISST2, in K, middle panels: biases in surface turbulent heat fluxes (sensible plus latent, positive upward) in respect to OA-Flux, in W m^{-2}, lower panels: biases in the meridional gradient of air-temperature at 850 hPa in respect to ERA5 (gradients are considered to be positive equatorward, in K (100 km)$^{-1}$). Specific climatological contours are labeled for each variable, solid green for the observed and dashed violet for the respective multi-model mean climatology. All rectangular frames correspond to [15–45°W, 40–50°N], as in Fig. 2.

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-21-0515.1.

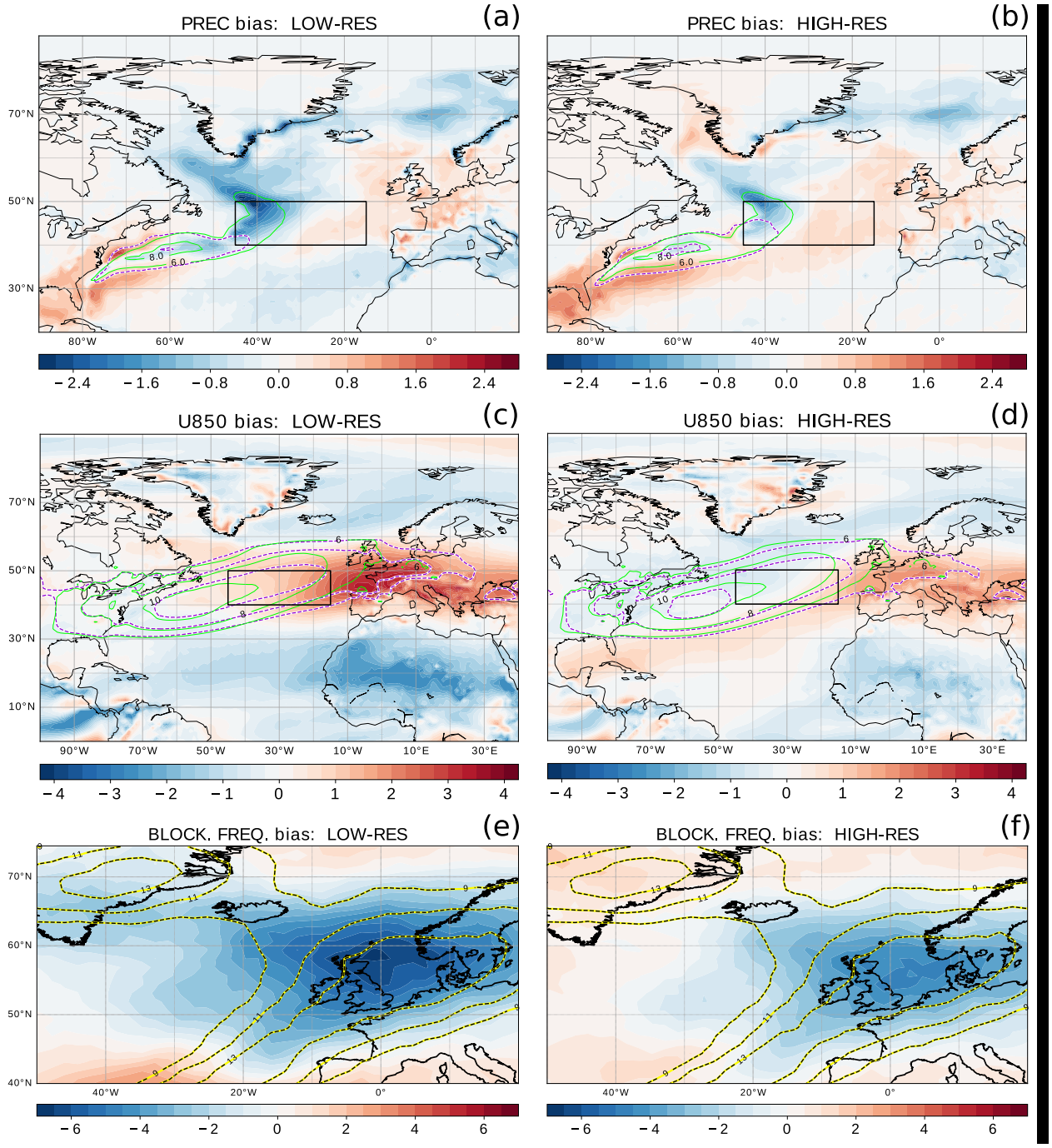


FIG. 6. As in Fig. 5 but for: (upper panels) biases in total precipitation, in mm day^{-1}, (middle panels) biases in zonal wind at 850 hPa, in m s^{-1}, and (lower panels) biases in blocking frequency in % of blocked days (details in the text); selected contours of the observed climatological blocking frequency are shown for reference. All rectangular frames correspond to [15–45°W, 40–50°N], as in Fig. 5.

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-21-0515.1.

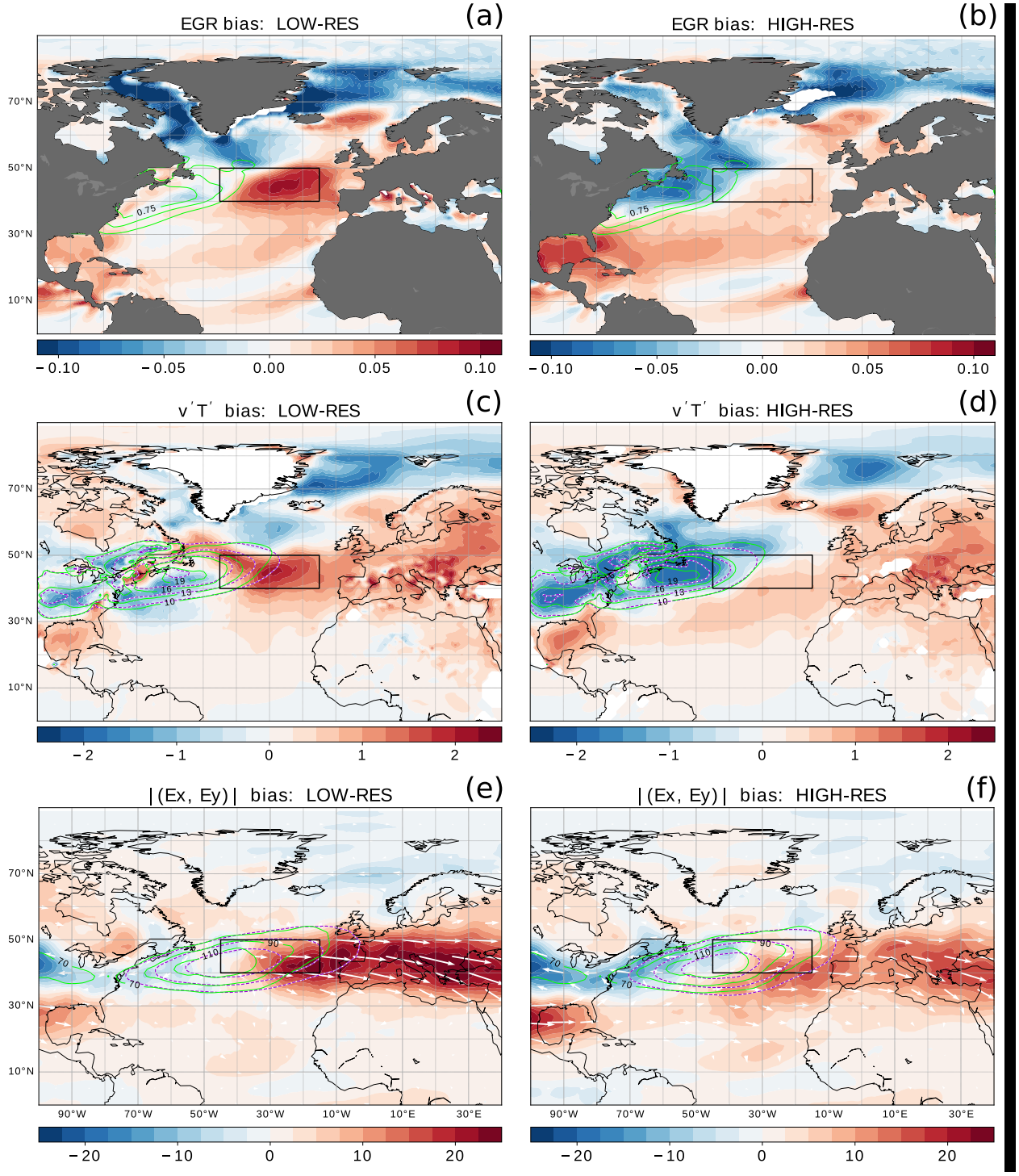
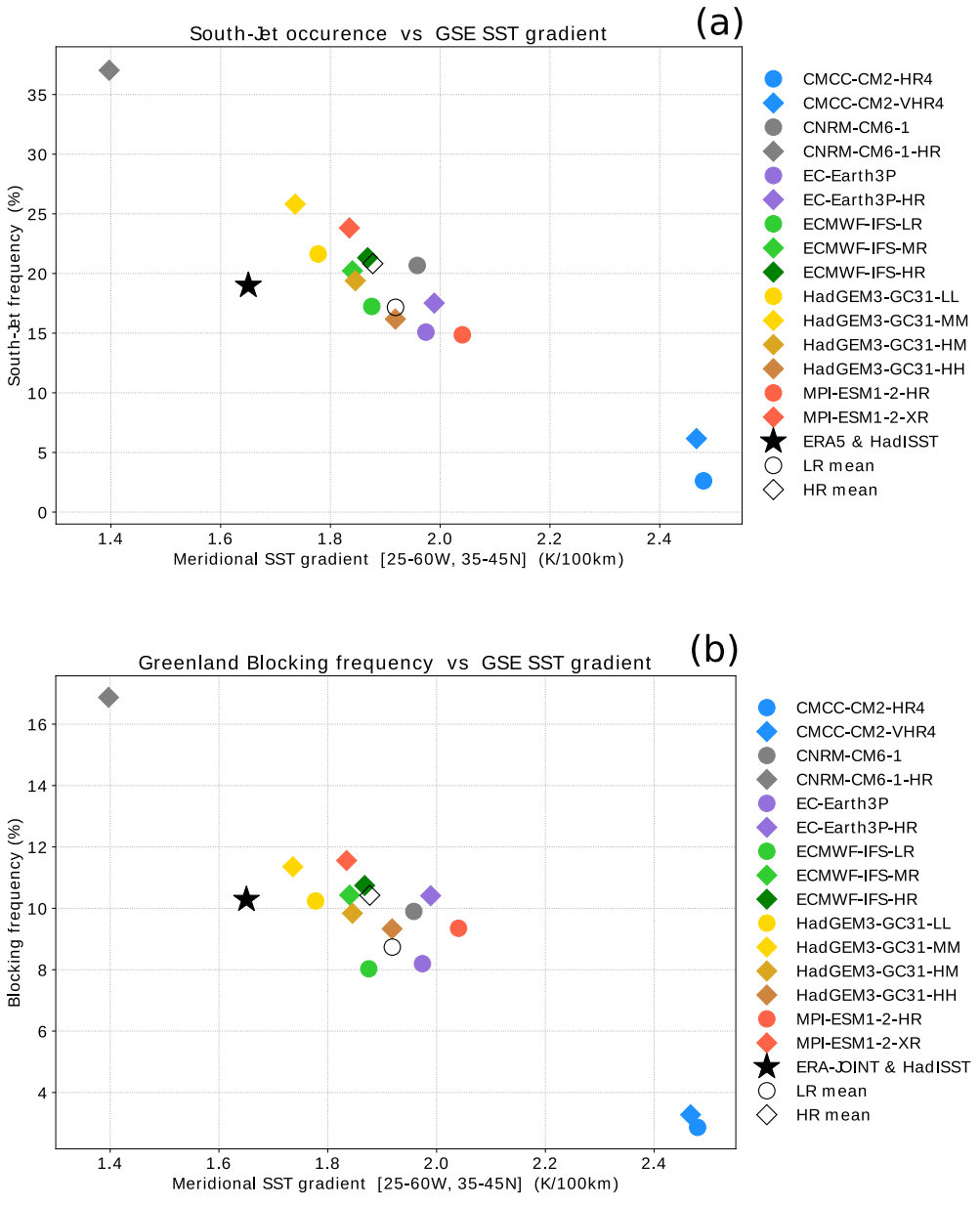


FIG. 7. As in Fig. 5 but for: (upper panels) biases in maximum Eady growth rate at 775 hPa, in day^{-1}, (middle panels) biases in baroclinic eddy fluxes ($v'T'$) at 850 hPa, in K m s^{-1}, and (lower panels) biases in the magnitude of the horizontal E-vector ($\overline{v'^2 - u'^2}$, $\overline{-u'v'}$) at 250 hPa, in m^2 s^{-2} with the arrows showing biases in the respective components. Primes denote high-pass filtering of daily data (details in the text). All rectangular frames correspond to [15–45°W, 40–50°N], as in Fig. 6.

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-21-0515.1.



1154 FIG. 8. Upper panel: climatological frequency (% of days) of the southern-jet regime (jet
1155 latitude $< 38°$N) in wintertime for each model versus the respective climatological meridional SST
1156 gradient in the area [25–60°W, 35–45°N] as indicated by the frames in Fig. 9 (details in the text).
1157 Round markers for "LR" models, diamond markers for "HR" models, and the black asterisk for
1158 the respective observations. Lower panel: as above but for high-latitude [10–55°W, 60–75°N]
1159 blocking frequency in the y-axis, in % of days (details in the text).

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-21-0515.1.

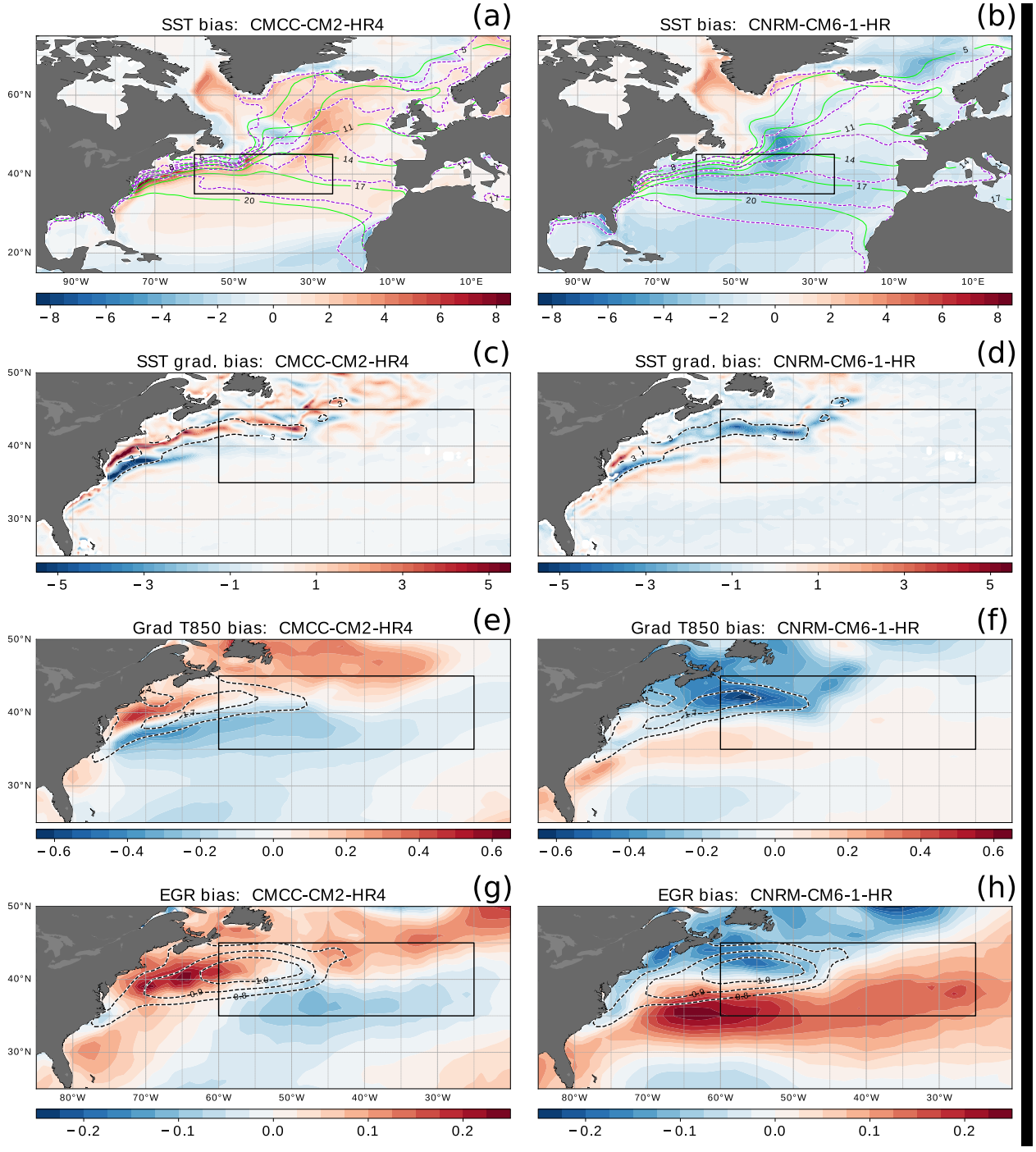


FIG. 9. Comparison of biases for the "outlier" models (CMCC-CM2-HR4 and CNRM-CM6-1-HR) seen in the previous figure. Top panels: SST biases, in K, with solid green contours for the observed climatology and dashed violet contours for the respective model climatology, 2nd row panels: biases in meridional SST gradient, in $\mathrm{K}\,(100\,\mathrm{km})^{-1}$, 3rd row panels: biases in meridional gradient of air-temperature at 850 hPa, in $\mathrm{K}\,(100\,\mathrm{km})$, bottom panels: biases in maximum Eady growth rate at about 885 hPa, in day^{-1}. Dashed black contours refer to the respective observed climatology in the same season (DJF) and period (1950–2014). All rectangular frames correspond to [25–60°W, 35–45°N].

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-21-0515.1.

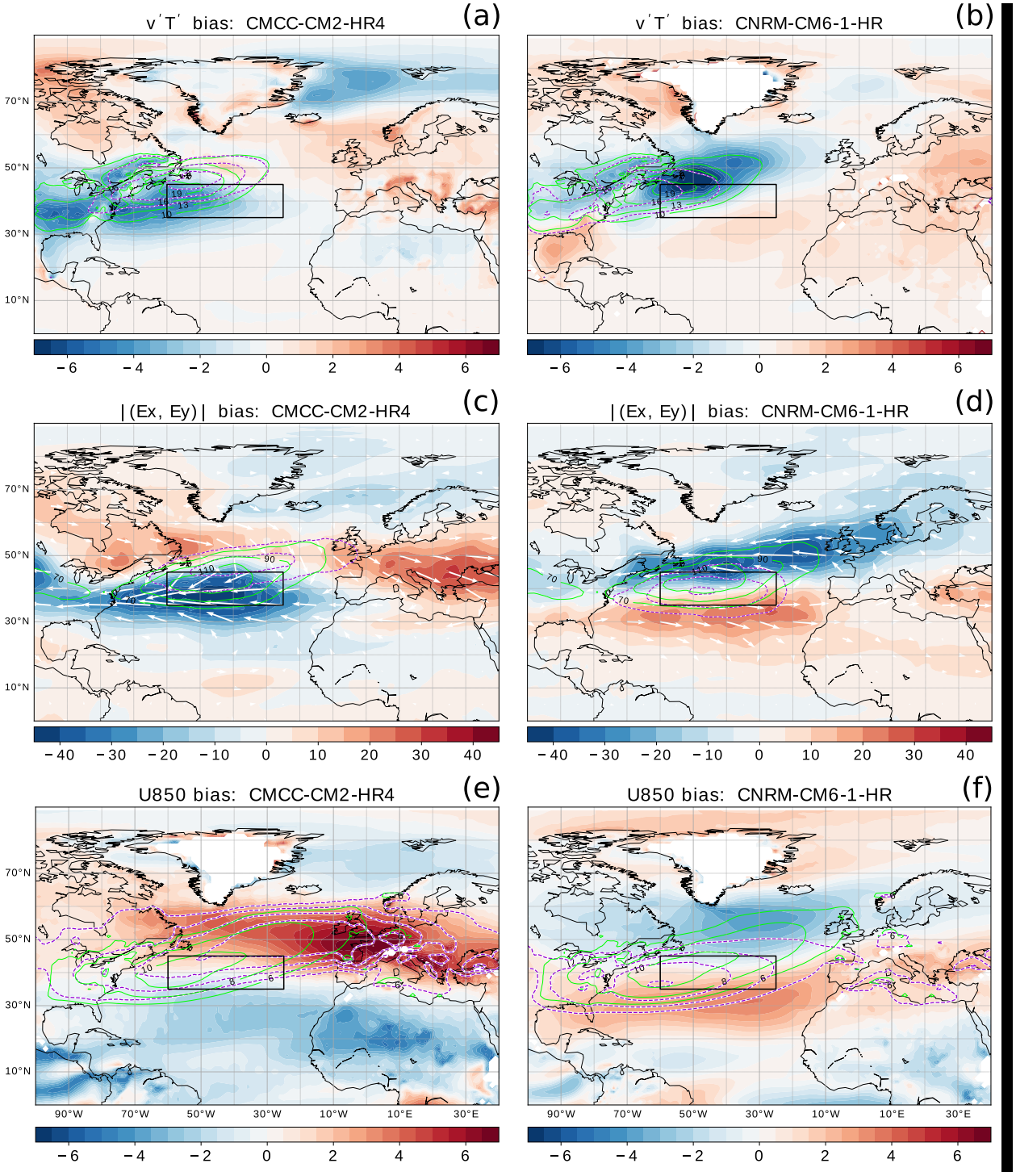


FIG. 10. As in Fig. 9 but showing: (top panels) biases in baroclinic eddy fluxes ($v'T'$) at 850 hPa, in K m s^{-1}, (middle panels) biases in the magnitude of the horizontal E-vector ($\overline{v'^2 - u'^2}$, $\overline{-u'v'}$) at 250 hPa, in m^2 s^{-2} with the arrows showing biases in the respective components, and (bottom panels) biases in zonal wind at 850 hPa, in m s^{-1}. As in Fig. 6, primes denote high-pass filtering of daily data (details in the text). All rectangular frames correspond to [25–60°W, 35–45°N], as in Fig. 9.

Accepted for publication in *Journal of Climate*. DOI 10.1175/JCLI-D-21-0515.1.