Received: 12 October 2020 | Revised: 9 December 2020 | Accepted: 5 January 2021

DOI: 10.1002/qj.3974

Quarterly Journal of the Royal Meteorological Society RMetS

RESEARCH ARTICLE

The representation of winter Northern Hemisphere atmospheric blocking in ECMWF seasonal prediction systems

Paolo Davini[1] | **Antje Weisheimer**[2,3] | **Magdalena Balmaseda**[2] | **Stephanie J. Johnson**[2] | **Franco Molteni**[2] | **Christopher D. Roberts**[2] | **Retish Senan**[2] | **Timothy N. Stockdale**[2]

[1]Istituto di Scienze dell'Atmosfera e del Clima, Consiglio Nazionale delle Ricerche, Torino, Italy

[2]European Centre for Medium-Range Weather Forecasts, Reading, UK

[3]Department of Physics, University of Oxford, National Centre for Atmospheric Science, Oxford, UK

Correspondence
P. Davini, CNR-ISAC, Corso Fiume 4, Torino 10133, Italy.
Email: p.davini@isac.cnr.it

Abstract

The simulation and prediction of winter Northern Hemisphere atmospheric blocking in the seasonal prediction systems from the European Centre for Medium-Range Weather Forecasts (ECMWF) is analysed. Blocking statistics from the operational November-initialised seasonal hindcasts are evaluated in three generations of models: System3, System4, and System5 (SEAS5). Improvements in the climatological representation of blocking are observed in the most recent model configurations, with reduced bias over North Pacific and Greenland. Minor progress is seen over the European sector, where SEAS5 still underestimates the observed blocking frequency. SEAS5 blocking interannual variability is underestimated too and is proportional to the climatological frequency, highlighting that a negative bias in the blocking frequency implies an underestimation of the interannual variance. SEAS5 predictive skill and signal-to-noise ratio remain low, but interesting positive results are found over Western and Central Europe. Improved forecasts with reduced ensemble spread are obtained during El Niño years, especially at low latitudes. Complementary experiments show that the statistics of blocking are improved following atmospheric and oceanic resolution increase. Conversely, they remain largely insensitive to coupled model sea-surface temperature (SST) errors. On the other hand, the implementation of stochastic parameterisations tends to displace blocking activity equatorward. Finally, by comparing seasonal hindcasts with climate runs using the same model, we highlight that the largest contributors to the chronic underestimation of blocking are persistent errors in the atmospheric model. It is also shown that SST errors have a larger impact on blocking bias in climate runs than in seasonal runs, and that increased ocean model resolution contributes to improved blocking more effectively in climate runs. Seasonal

forecasts can thus be considered a suitable test-bed for model development targeting blocking improvement in climate models.

KEYWORDS

atmospheric blocking, atmospheric dynamics, climate models, seasonal forecast, SEAS5

1 | INTRODUCTION

Atmospheric blocking is one of the most investigated weather patterns in current climate science. It occurs in the midlatitudes, typically at the exit of Atlantic and Pacific storm tracks, where the jet stream weakens (Tyrlis and Hoskins, 2008; Tibaldi and Molteni, 2018; Woollings *et al.*, 2018). It can be described as a quasistationary equivalent-barotropic low-vorticity system which can persist for several days, sometimes for weeks, blocking or diverting the eastward path of synoptic cyclones (Rex, 1950). Given its persistence, blocking has a substantial impact on the local weather, and it is associated with both cold spells in winter (Buehler *et al.*, 2011; Sillmann *et al.*, 2011) and heat waves in summer (Pfahl and Wernli, 2012; Schaller *et al.*, 2018).

Because of its impact on human activities, it is important to investigate the presence of possible sources of blocking predictability and to determine the extent to which blocking frequency is influenced by internal or forced climate variability. Several studies have addressed the prediction of blocking in medium and extended-range weather forecasts, as an initial value problem. Blocking onset has been shown to be scarcely predictable, due to the chaotic nature of the midlatitude flow (Pelly and Hoskins, 2003; Mauritsen and Källén, 2004; Matsueda, 2009). Other studies have addressed the ability of global climate models (GCMs) to represent the statistics of blocking, showing that GCMs still underestimate the average blocking frequency, especially over the European sector (Scaife *et al.*, 2010; Masato *et al.*, 2013; Davini and D'Andrea, 2016; 2020; Schiemann *et al.*, 2020). The origin of such underestimation has often been linked to errors in the mean state that affect Rossby-wave propagation, which are usually responsible for the onset of atmospheric blocking events (Scaife *et al.*, 2010; Davini and D'Andrea, 2016).

On the other hand, less attention has been given to the ability of seasonal forecasting systems to represent and predict the statistics of blocking events. Seasonal forecasts lie at the interface of medium-range forecasts and climate prediction problems: they are sensitive to both the initialisation of the slow-varying components of the Earth system (e.g., land, sea-ice, ocean, stratosphere) and the prescription of climate boundary conditions (e.g., greenhouse gases, volcanic aerosols, land-surface types). However, a fundamental difference between climate simulations and seasonal forecasts is the extent to which the model systematic error is fully developed. This ultimately affects the magnitude of the mean biases and the interaction between mean state and variability (Roberts *et al.*, 2020). Given that model errors are significant in the extratropics (e.g., Johnson *et al.*, 2019), the seasonal predictability of midlatitude flow is a particularly challenging problem.

Despite the recent improvements in forecasting the North Atlantic Oscillation (NAO) on interannual time-scales (Dunstone *et al.*, 2016; Weisheimer *et al.*, 2019), which is intimately related to blocking at high latitudes (Woollings *et al.*, 2008), many authors report low or negligible year-to-year skill of atmospheric blocking (e.g., Prodhomme *et al.*, 2016). Athanasiadis *et al.* (2014; 2020) investigated the winter NAO and atmospheric blocking in seasonal (UKMO and CMCC) and decadal (CESM) prediction systems. While they find promising results when looking at the NAO and Greenland Blocking, the skill remains low for blocking over Central Europe.

Although some encouraging results have been obtained in recent years (Matsueda, 2011; Davini and D'Andrea, 2020), improving blocking in weather and climate simulations remains a challenging task: reduced biases have been obtained following atmospheric horizontal resolution refinement (Jung *et al.*, 2012; Davini and D'Andrea, 2020) and when an oceanic model with reduced North Atlantic sea-surface temperature (SST) bias is used (Scaife *et al.*, 2010). However, results are model- (Schiemann *et al.*, 2017) and time-scale-dependent (Roberts *et al.*, 2020) and conclusions are often limited by the large midlatitude natural variability (Hartung *et al.*, 2017). On top of that, improvements can be achieved for the wrong reasons (Davini *et al.*, 2017), so that no unique recipe for "healing" blocking biases can be provided yet.

The work presented here documents the representation of atmospheric blocking in three European Centre for Medium-Range Weather Forecasts (ECMWF) seasonal prediction systems, namely System3, System4, and System5 (SEAS5). To interpret the evolution of blocking in

these systems further, a set of seven sensitivity experiments will be analysed to assess the sensitivity of blocking to (a) increased atmospheric and oceanic horizontal resolution, (b) ocean/sea-ice coupling, and (c) stochastic atmospheric parameterisations. A focused analysis on blocking properties of SEAS5, ECMWF's current operational forecasting system, will be then carried out: the skill, the interannual variance, and the signal-to-noise ratio will be analysed in order to unveil both limitations and encouraging aspects of the current seasonal prediction system. In the last section, a comparison with a set of climate runs using the same model version as SEAS5, which follow the High Resolution Model Intercomparison Project (HighResMIP) protocol (Haarsma *et al.*, 2016), will be carried out in order to assess the relevance of model initialisation and the impact of SST biases.

2 | DATA AND METHODS

Several hindcasts from a range of ECMWF seasonal prediction systems have been evaluated. November starting dates have been used to study the Northern Hemisphere winter variability in the December, January, and February season (DJF). Three operational prediction systems were used: System3 (S3, Stockdale *et al.*, 2011), System4 (S4, Molteni *et al.*, 2011), and SEAS5 (S5, Johnson *et al.*, 2019). All configurations are based on successive versions of the same atmospheric model, the Integrated Forecast System (IFS): S3 uses cy31r1, S4 uses cy36r4, and S5 uses cy43r1. The S3 ocean component is based on the Hamburg Ocean Primitive Equation (HOPE) ocean model (Wolff *et al.*, 1997), while S4 and S5 use progressive versions of the Nucleus for European Modelling of the Ocean (NEMO) model. The common hindcast period, which includes November starting dates, from 1981–2011 (extending up to March 2012) has been used for the analyses based on all three systems. If the analysis is focused only on S5, November starting dates from 1981–2019 (extending up to March 2020) are used.

In order to explore the model sensitivity in different configurations, seven sensitivity experiments are considered: two hindcast experiments based on S4 and five based on S5. These include different atmospheric and oceanic horizontal resolutions, atmosphere-only hindcasts, and simulations where the default stochastic physics parameterisations were switched off. The number of available ensemble members ranges from 25 up to 51. Moreover, six noninitialised climate runs from the ECMWF contribution to the European Union funded PRIMAVERA (PRocess-based climate sIMulation: AdVances in high-resolution modelling and European climate Risk Assessments) project,[1] part of the HighResMIP protocol (Haarsma *et al.*, 2016) have been considered (from 1980–2014). These runs are based on the same model components as S5 (i.e., IFS cy43r1 and NEMO 3.4.1), although some modifications in the atmosphere/ocean resolution, forcing, and tuning have been introduced (Roberts *et al.*, 2018).

A full description of the experiments considered, including the model resolution, the time range, and the ensemble size of each model configuration, is provided in Table 1.

Data from the ECMWF ERA5 reanalysis (Hersbach *et al.*, 2020) are used as a reference. Similar results have been obtained when using the ECMWF ERA-Interim Reanalysis (Dee *et al.*, 2011). Before any computation, all data are interpolated on a common $2.5° \times 2.5°$ grid with a bilinear remapping method.

Atmospheric blocking is then detected using the Davini *et al.* (2012) blocking index. This is an Eulerian blocking index based on the reversal of the meridional gradient of the daily geopotential height at 500 hPa ($Z500$), following the classic definition by Tibaldi and Molteni (1990). However, the index is extended into a two-dimensional form between 30°N and 75°N (following Scherrer *et al.*, 2006). Two meridional gradients of geopotential height are defined:

$$GHGS(\lambda_0, \phi_0) = \frac{Z500(\lambda_0, \phi_0) - Z500(\lambda_0, \phi_S)}{\phi_0 - \phi_S}, \quad (1)$$

$$GHGN(\lambda_0, \phi_0) = \frac{Z500(\lambda_0, \phi_N) - Z500(\lambda_0, \phi_0)}{\phi_N - \phi_0}, \quad (2)$$

and ϕ_0 ranges from 30°N to 75°N while λ_0 ranges from 0° to 360°, $\phi_S = \phi_0 - 15°$, and $\phi_N = \phi_0 + 15°$. Instantaneous Blocking is thus identified when

$$GHGS(\lambda_0, \phi_0) > 0, \qquad GHGN(\lambda_0, \phi_0) < -10 \ \text{m/°lat}. \quad (3)$$

Further constraints have been applied to instantaneous blocking. Firstly, Large-Scale Blocking is defined when Instantaneous Blocking is extended for at least 15° of continuous longitude. Secondly, a Large-Scale Blocking Event is defined for each grid point when Large-Scale Blocking is occurring within 5° longitude (2 grid points) and 2.5° latitude (1 grid point) of it. Finally, a Blocking Event at a certain grid point is defined when a Large-Scale Blocking Event lasts for at least 5 days. Those constraints ensure that Blocking Events have a significant longitudinal extension and are persistent and quasistationary. The percentage of days per season in which Blocked Events occur

[1] https://www.climateurope.eu/primavera/

TABLE 1 Details of the simulations analysed

Experiment	Resolution	Years	Ens. size	Description
S3	T_L159L62 HOPE1L29	1960–2011	40	System3 (operational 2007–2011)
S4	T_L255L91 ORCA1Z42	1981–2017	51	System4 (operational 2011–2017)
S4-ObsSST	T_L255L91	1981–2013	51	System4 with prescribed observed SSTs
S4-noStoch	T_L255L91 ORCA1Z42	1981–2013	51	System4 without stochastic physics in the atmosphere
S5	T_{co}319L91 ORCA025Z75	1981–2019	51	SEAS5 (operational since 2017)
S5-ObsSST	T_{co}319L91	1981–2018	25	Atmosphere-only SEAS5 with prescribed observed SSTs
S5-noStoch	T_{co}319L91 ORCA025Z75	1981–2014	51	SEAS5 without stochastic physics in the atmosphere
S5-LR	T_{co}199L91 ORCA1Z42	1981–2014	25	Low-resolution SEAS5
S5-LR-ObsSST	T_{co}199L91	1981–2016	25	Atmosphere-only low-resolution SEAS5 with prescribed observed SSTs
S5-LR-noStoch	T_{co}199L91 ORCA1Z42	1981–2014	25	Low-resolution SEAS5 without stochastic physics in the atmosphere
PRIM-LALO	T_{co}199L91 ORCA1Z42	1980–2014	8	PRIMAVERA low-resolution for atmosphere and ocean
PRIM-LAHO	T_{co}199L91 ORCA025Z75	1981–2014	3	PRIMAVERA low-resolution atmosphere and high-resolution ocean
PRIM-HALO	T_{co}399L91 ORCA1Z42	1980–2014	1	PRIMAVERA high-resolution atmosphere and low-resolution ocean
PRIM-HAHO	T_{co}399L91 ORCA025Z75	1980–2014	6	PRIMAVERA high-resolution for atmosphere and ocean
PRIM-LA-ObsSST	T_{co}199L91	1980–2014	8	PRIMAVERA low-resolution atmosphere-only with prescribed observed SSTs
PRIM-HA-ObsSST	T_{co}399L91	1980–2014	6	PRIMAVERA high-resolution atmosphere-only with prescribed observed SSTs

(i.e., blocked days) defines the blocking frequency climatology. On top of Blocking Events frequency, it is possible to estimate for each grid point the Blocking Events average duration and the number of 5-day long Blocking Events, that is, the number of Blocking Events onsets. A complete description of the blocking climatology and of the blocking detection scheme may be found in Davini *et al.* (2012). The whole set of blocking data has been produced with the Mid-Latitude Evaluation System (MiLES) R suite (Davini, 2019). The code is freely available on GitHub[2] and provides several blocking definitions and diagnostics.

In the last part of the article, the relation between SSTs and atmospheric blocking is investigated, using monthly SSTs from each integration. The Hadley Centre Sea Ice and Sea Surface Temperature data set (HadISST: Rayner *et al.*, 2003) is used as a reference.

[2]https://github.com/oloapinivad/MiLES

3 | RESULTS

3.1 | Climatology and sensitivity analysis (1981–2011)

The climatological representation of winter blocking over the common period (DJF 1981–2011) in the ERA5 Reanalysis is displayed in Figure 1a. This shows an absolute maximum over the North Pacific (about 20% of blocked days per season) and two secondary maxima over Greenland and Europe (9–12% of blocked days).

From a dynamical point of view, blocking occurring over Greenland and the North Pacific is quite different from the blocking over the European sector, since the former is characterised by cyclonic wave breaking, while the latter is associated with anticyclonic wave breaking (Davini *et al.*, 2012). Their impact on the mesoscale and synoptic circulation is also different: European

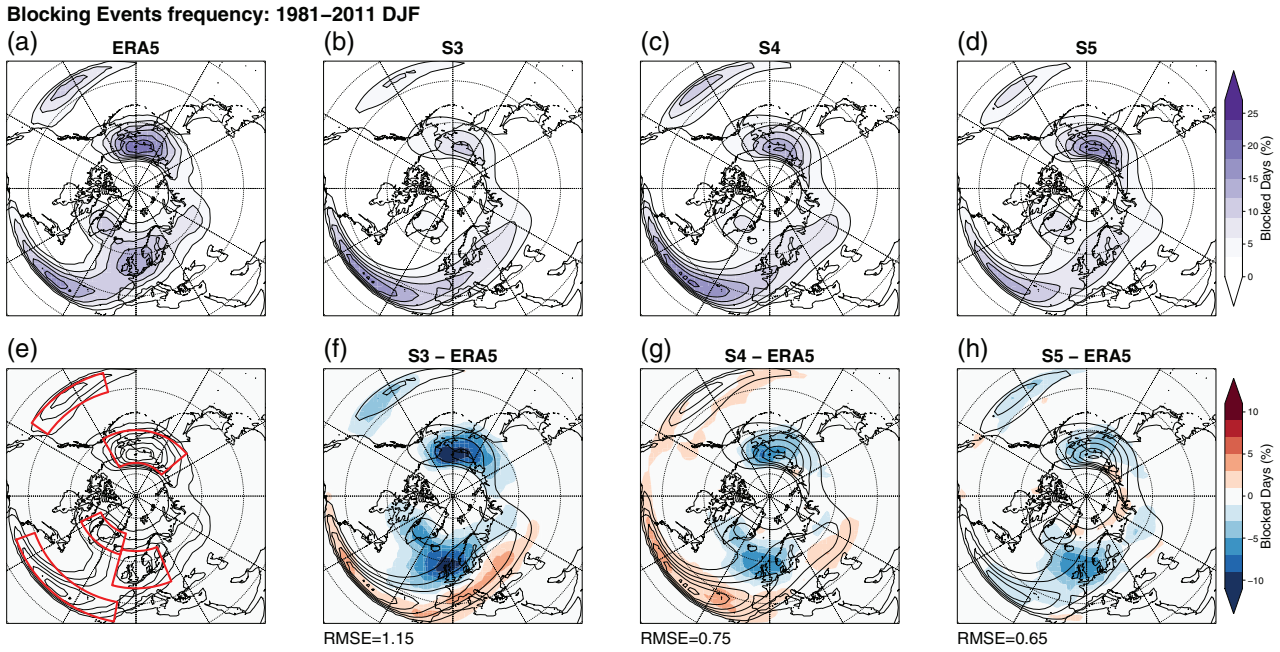

FIGURE 1 1981–2011 DJF Blocking Events climatological frequency for (a) ERA5, (b) System3, (c) System4, and (d) SEAS5, highlighted by both contours and shadings. (f–h) The same as (b–d), but with shadings showing their biases against ERA5. Contours are drawn each 3%. Solid boxes in (e) identify the Pacific LLB, Atlantic LLB, North Pacific, Greenland, and Central Europe sectors. At the bottom left of panels (f–h), the RMSE against ERA5 is reported

blocking—and its extension further downstream over Western Russia, usually identified as Ural blocking (Wang *et al.*, 2010; Luo *et al.*, 2016)—represents the most archetypal blocking event, which is able to literally block storm paths. In contrast, Greenland and North Pacific blocking, by virtue of its location—several degrees north from the jet stream—is not able to obstruct the passage of synoptic disturbances in a similar way; for these reasons, such conditions are often defined as high-latitude blocking (e.g., Berrisford *et al.*, 2007). However, their synoptic relevance has been recognised, considering that blocking over Greenland is bound intimately with the negative phase of the NAO (Woollings *et al.*, 2008).

Two other relative maxima at lower latitudes are detected, between 30°N and 40°N: one over the Central North Atlantic—approximately over the Azores—and the other over the Central North Pacific. These are unusual regions for blocking and they cannot be considered as actual blocking events: rather, they represent minor Rossby-wave breaking events, which are unable to perturb the jet stream effectively, and for such reasons they have been classified as low-latitude blocking (LLB: Davini *et al.*, 2012). Even if their impact on the weather is less relevant (over the Atlantic they are usually associated with a zonal flow and a positive NAO phase), Pacific and Atlantic LLBs are, however, interesting, since they provide information on where anticyclonic wave breaking is occurring on the equatorward flank of the Pacific and Atlantic jet streams.

The three above-described blocking maxima (North Pacific, Greenland, and Central Europe), together with the two LLB sectors (Atlantic LLB and Pacific LLB), defines the five sectors shown in red in Figure 1e. These sectors will be used hereafter to summarise the blocking performance.

Figure 1b,c,d reports the blocking climatology for the three ECMWF systems S3, S4, and S5, while Figure 1f,g,h shows the corresponding biases with respect to ERA5. The three systems show similar biases, all characterised by the absence of a relative maximum over the European continent. This leads to the well-known negative bias over Europe, a common feature of almost all climate models (Masato *et al.*, 2013; Davini and D'Andrea, 2020). However, an evident improvement is visible, with both S4 and S5 showing a considerably reduced bias over Europe compared with S3. The underestimation of North Pacific blocking—a model error visible in all successive seasonal forecasting systems—has also been alleviated with S5. Progress is also appreciated in the overactive Atlantic LLB region. In both S3 and S4, low-latitude overactivity is pervasive in the Atlantic sector, while in S5 the low-latitude blocking bias is much smaller and slightly negative over both the Pacific and the Atlantic sectors.

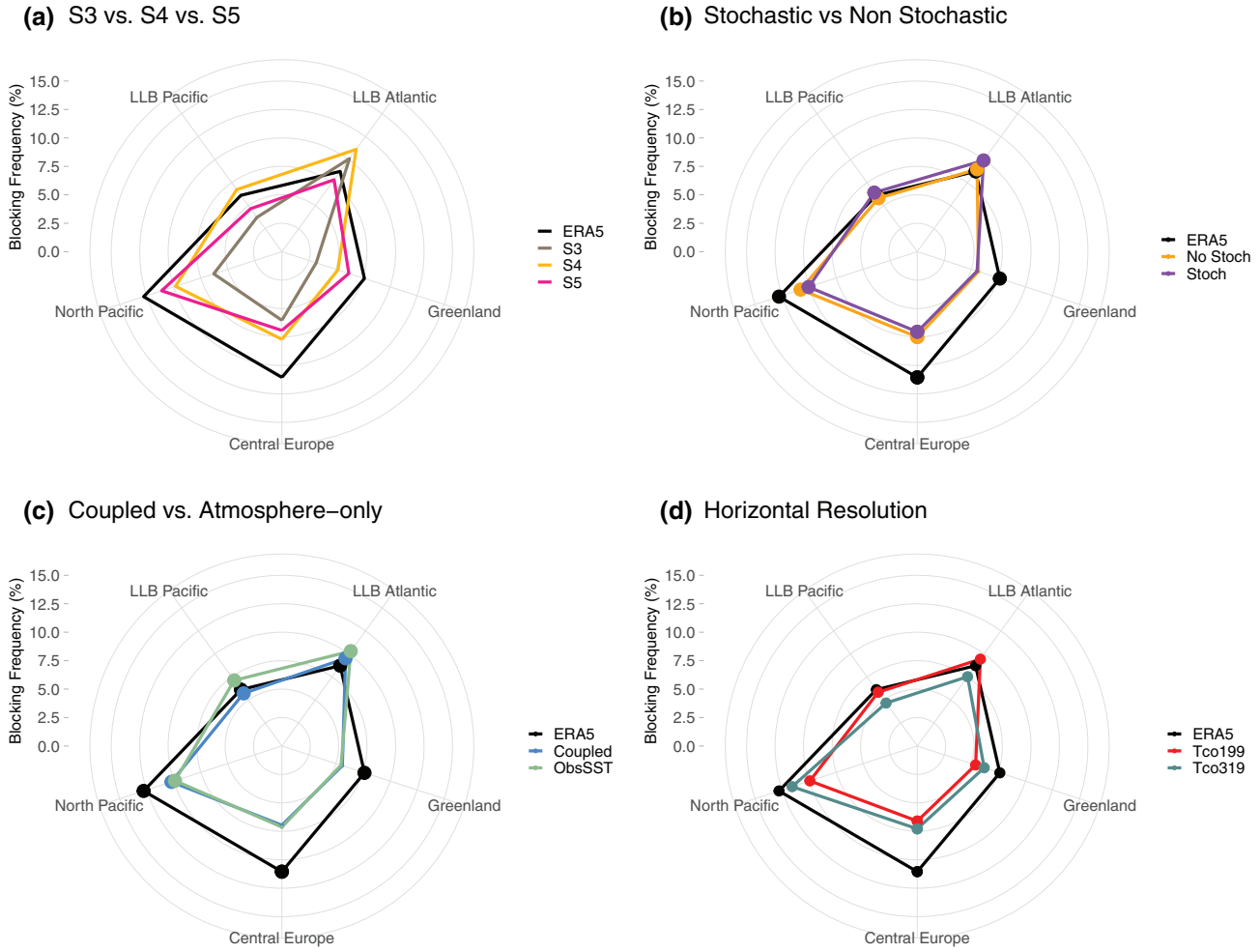

FIGURE 2 Radar chart showing the 1981–2011 DJF Blocking Events frequency of the five main regions: (a) S3, S4 and S5, (b) sensitivity to stochastic physics, (c) role of oceanic coupling, and (d) sensitivity to horizontal resolution. The black curve shows the verification data from ERA5. In (b,c,d), dots indicate where the difference between the two model configurations is significant with a Welch *t*-test at 5% level

Overall, S5 shows considerable improvements with respect to previous seasonal forecasting systems: the absence of a positive frequency bias in Atlantic low-latitude blocking is an unprecedented feature in the ECMWF seasonal systems. Furthermore, S5 exhibits the smallest biases over the North Pacific and Greenland sectors. Most importantly, the south-eastward displacement of Atlantic blocking activity seen in S3 and S4 (identified by the dipole over the North Sea and Central/Eastern Europe, a common bias in global climate models usually associated with a too-strong Atlantic jet stream) has been reduced in S5. This picture is confirmed by the root-mean-squared error (RMSE), reported in the bottom left of Figure 1f,g,h: it shows a decrease from 1.15% (for S3) to 0.75% (for S4) up to 0.65% (for S5).

The results of the Blocking Events frequency from S3, S4, and S5 are summarised by the radar charts of Figure 2, which also shows results from the complementary sensitivity experiments. The vertices of the radar chart correspond to the climatological blocking frequencies over the five sectors shown in red in Figure 1e and the distance from the centre indicates the blocking frequency. The sensitivity experiments address the impact of stochastic physics (Figure 2b), ocean–atmosphere coupling (Figure 2c), and horizontal resolution (Figure 2d).

Figure 2a summarises the findings of Figure 1 as discussed above: the better performance of S4 and S5 with respect to S3 emerges clearly, as well as the S5 improvements over the North Pacific, Greenland, and the two LLB sectors. The best results over the European sector are obtained by S4, even though the underestimation of blocking is still large (about 30–40% of the ERA5 values).

In order to represent some of the inevitable uncertainties due to unresolved subgrid-scale variability in the atmosphere, the IFS model applies stochastic, flow-dependent perturbations to the tendencies of its prognostic variables, known as "stochastic physics parameterisations" (Buizza

et al., 1999; Weisheimer *et al.*, 2014). Figure 2b investigates the sensitivity of the different seasonal prediction systems to the activation of these stochastic physics parameterisation schemes: the yellow curve is obtained from all runs where stochastic physics in the atmosphere has been switched off (S5-noStoch, S4-noStoch, S5-LR-noStoch), while the purple curve includes all the corresponding runs with stochastic physics (i.e., S5, S4, S5-LR). The differences are very small, but significant in four out of five sectors: the current implementation of stochastic physics exacerbates the overestimation of Atlantic LLB in favour of a further decrease of European Blocking. Similarly, stochastic physics is associated with a moderate increase of Pacific LLB paired with a weak reduction of North Pacific blocking. This suggests that the introduction of the stochastic physics parameterisations tends to displace blocking activity slightly equatorward, so that it may need further development to affect the blocking representation positively.

Figure 2c analyses the importance of an interactive ocean: blue lines show the blocking frequency from the atmosphere–ocean coupled runs (S5, S4, S5-LR), while green lines show the blocking frequency averaged among a set of corresponding atmosphere-only simulations (S5-ObsSST, S4-ObsSST, S5-LR-ObsSST). Similarly to what was found for stochastic physics simulations, only small differences emerge: hindcasts with prescribed observed SSTs show a significant tendency to overestimate Atlantic and Pacific LLB and perform slightly worse than coupled ones over the North Pacific. These differences are very small and seem to suggest that prescribed observed SSTs do not lead to any clear improvements for blocking—they actually deteriorate the simulation slightly. This is consistent with previous literature interpretations (Prodhomme *et al.*, 2016; Roberts *et al.*, 2020): with the winter season being too close to the November initialisation date, the oceanic biases have not fully developed yet and so the largest contribution to the systematic errors of blocking comes from the atmosphere. We could then speculate that the slight improvement seen in coupled runs over the North Pacific may be produced by the presence of interactive atmosphere–ocean processes which are missing in atmosphere-only runs, as for instance noticed by Davini and D'Andrea (2016, see their figure 3).

Finally, Figure 2d assesses the role of atmospheric and ocean horizontal resolution, comparing the experiments run at $T_{co}199$-ORCA1 (~50 km in the atmosphere and ~100 km in the ocean, S5-LR and S5-LR-noStoch) with the ones run at $T_{co}319$-ORCA025 (~32 km in the atmosphere and ~25 km in the ocean, S5 and S5-noStoch). The increased horizontal resolution produces a considerable improvement, with reduced bias in midlatitude and high-latitude blocking sectors, but a larger error in the LLB sectors: this suggests that increasing the model resolution displaces blocking activity poleward, in an opposite way to that achieved by implementing the stochastic physics parameterisations (Figure 2b). Considering that, as shown in Figure 2c, prescribed SSTs have a negligible impact on blocking representation, we speculate that most of the improvement comes from the increase in atmospheric resolution. This hypothesis is supported by the better performance of S5-ObsSST versus S5-LR-ObsSST in the five sectors (S5-ObsSST RMSE=0.64% versus S5-LR-ObsSST RMSE=0.71%). It could thus be possible that a finer atmospheric grid—associated with a more resolved mean orography (Jung *et al.*, 2012; Davini *et al.*, 2017)—favours a poleward displacement of the jet stream, especially over the Atlantic: this could increase the frequency over Central Europe and decrease it over the Atlantic LLB sector. Overall, these findings are in agreement with what was found for other GCMs (e.g., Davini and D'Andrea, 2020), where a finer horizontal atmospheric grid provides an improvement of the mean state and of the atmospheric blocking simulation over Europe.

3.2 | Duration and number of onsets (1981–2011)

Another interesting feature worth analysing is the duration of Blocking Events and its sensitivity in the different model configurations. The scatter plots of Figure 3 compare the number of Blocking Events onsets and the Blocking Events average duration in the five main sectors: the product of these two quantities is an indicator of the Blocking Events frequency shown in Figures 1 and 2. Small duration differences are seen in the different sectors, with values ranging between 6.7 and 7.3 days. The duration over the European sectors is, on average, underestimated by 0.5 days, while the number of onsets is underestimated by 25%. The differences in both duration and number of onsets among the different sensitivity experiments for a given system are quite small compared with differences between systems, so we conclude that neither stochastic physics nor coupling helps significantly to increase the blocking duration: intrinsic atmospheric model errors seem again to be accountable for the observed biases. On the other hand, increasing atmospheric and oceanic horizontal resolution slightly increases the duration of high-latitude blocking (North Pacific and Greenland sectors), and slightly reduces the blocking persistence in the Atlantic and Pacific LLB sectors (in line with what is seen for blocking frequencies, Figure 2d). Figure 3 also shows that a proportionality exists between duration and number

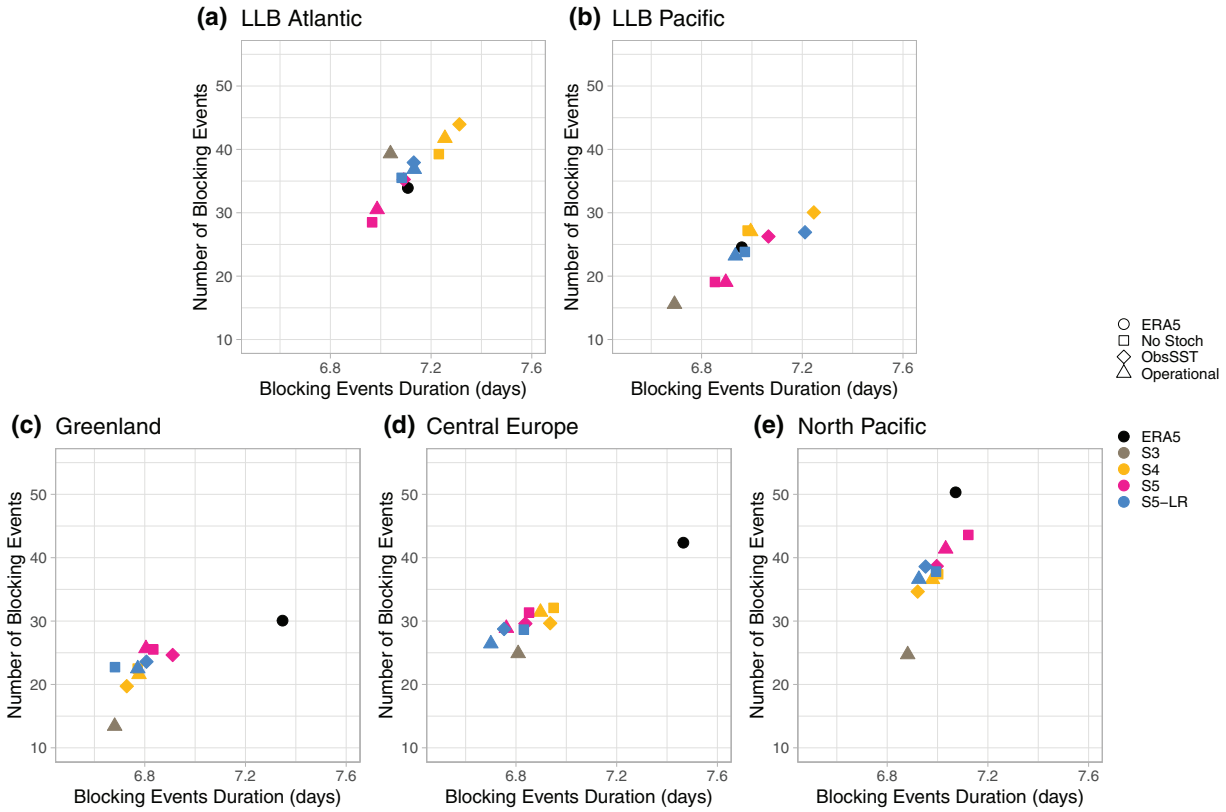

FIGURE 3 Scatter plot showing the ensemble mean for the number of Blocking Events onsets and the Blocking Events average duration during the 1981–2011 DJF period for different experiments over (a) Atlantic LLB, (b) Pacific LLB, (c) Greenland, (d) Central Europe, and (e) North Pacific sectors

of onsets, so that on average a longer duration implies more blocking onset too (or vice versa).

It is possible to investigate further the role of duration in determining the blocking frequency bias, decomposing the blocking frequency bias (Δf) in terms of "onsets" bias and "duration" bias. However, given the 5-day minimum threshold that has been introduced in the Blocking Events definition, the number of Blocking Events onsets N measured at day 5 is directly influenced by the persistence of the blocking anomaly in the prior days: for example, if the mechanism that maintains blocking is less effective in models than in observations, the number of onsets detected at day 5 will also be reduced.

Therefore, for the following analysis, the total number of events without any temporal filtering is considered, that is, the Large Scale Blocking Events—see Section 2—so that blocking situations lasting fewer than five days are also considered.

The bias in frequency could be written as

$$\Delta f = N_{\mathrm{m}} \cdot d_{\mathrm{m}} - N_0 \cdot d_0, \tag{4}$$

where N_{m} is the modelled number of onsets, d_{m} the modelled average duration, N_0 the observed number of onsets, and d_0 the observed average duration. This can be rearranged as

$$\Delta f = (N_{\mathrm{m}} - N_0) \cdot d_{\mathrm{m}} + (d_{\mathrm{m}} - d_0) \cdot N_{\mathrm{m}} - (N_{\mathrm{m}} - N_0) \cdot (d_{\mathrm{m}} - d_0), \tag{5}$$

where $(N_{\mathrm{m}} - N_0) \cdot d_{\mathrm{m}}$ represents the bias due to missing blocking onsets, $(d_{\mathrm{m}} - d_0) \cdot N_{\mathrm{m}}$ is the bias due to underestimated blocking duration, and $-(N_{\mathrm{m}} - N_0) \cdot (d_{\mathrm{m}} - d_0)$ is the bias due to the compound effect of both onsets and duration.

The results are reported in Figure 4, for all experiments and all sectors. It should be borne in mind that this decomposition does not reflect by construction what is shown in Figure 3, since here the 5-day filtering has been removed.

In the LLB sectors, both Atlantic and Pacific, the errors in blocking frequency are small and they are characterised by an overestimation of the number of onsets and a slight underestimation of the duration. For midlatitude and high-latitude blocking, the impact of underestimated duration is much more evident, accounting for about 80–100 blocked days (i.e., about 3% of blocked days) in

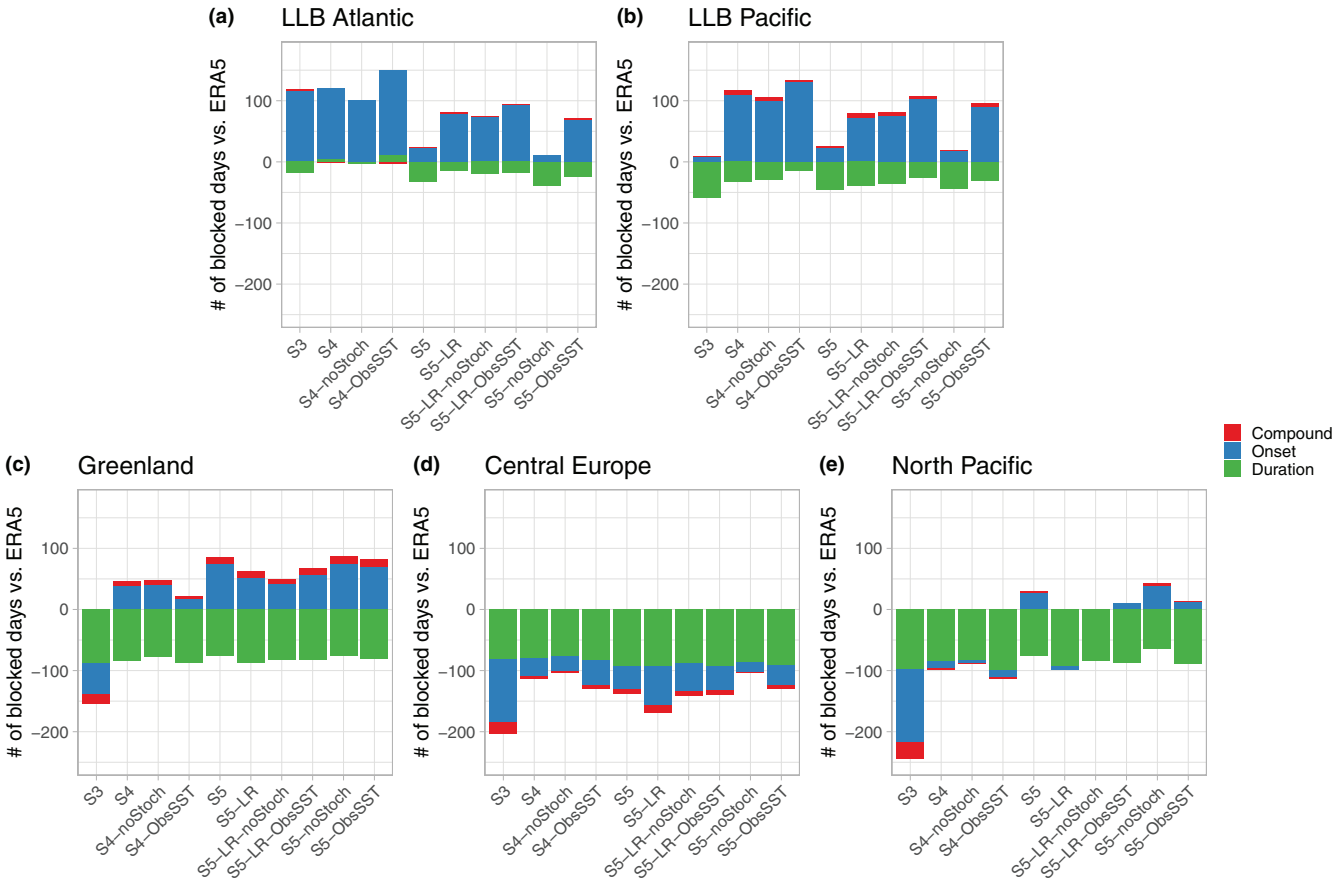

FIGURE 4 Linear decomposition of the Blocking Events frequency bias (versus ERA5) following Equation 5 into a part due to missing blocking onsets (blue), a part due to underestimated blocking duration (green), and their compound effect (red), over (a) Atlantic LLB, (b) Pacific LLB, (c) Greenland, (d) Central Europe, and (e) North Pacific sectors. Details on the linear decomposition are found in the text

all sectors. However, while over the North Pacific negligible bias is observed in blocking onsets, Central Europe is characterised by a negative bias and Greenland by a positive bias. Overall, one key piece of information can be summarised from this figure: both the different model configurations and the different blocking sectors show weak sensitivity to the duration of events (on average, too short in the model), showing that changes in blocking frequency can mainly be explained by changes in the number of onsets.

3.3 | Predictive skill of SEAS5 (1981–2019)

We will now focus on the performance of S5, which is the current operational seasonal prediction system at ECMWF, in forecasting interannual blocking variations. Figure 5a reports the pointwise Pearson's correlation between the seasonally averaged DJF blocking frequency of S5 and ERA5 reanalysis over the 1981–2019 hindcast period as an indicator of the S5 predictive skill. Atmospheric blocking is known for being scarcely predictable at a seasonal time-scale (Athanasiadis *et al.*, 2014), so that the result from Figure 5a can be interpreted as moderately positive. A few areas emerge as being significant at the 5% level with a two-tailed *t*-test (stippled area in Figure 5a), for example over parts of Western and Central Europe. However, these are paired with low or even negative skill over Northern Europe, so that, when averaging over the region that defines the European sector (as identified in Figure 1e), the result is a nonsignificant correlation skill of 0.18. Even in those regions that attain significant skill, the correlation values never exceed 0.4: indeed, a slight change of the time period chosen for the analysis can easily affect the findings of Figure 5a (e.g., the European sector skill is −0.1 over the 1981–2012 time window). Therefore, considerable caution is needed when discussing atmospheric blocking skill results.

Interestingly, significant skill is obtained for low-latitude blocking over the Atlantic LLB sector, suggesting a high skill for anticyclonic wave-breaking activity in this region: this is likely due to the fact that the Rossby-wave dynamics, which affects blocking, is more closely linked to tropical SSTs, which are represented

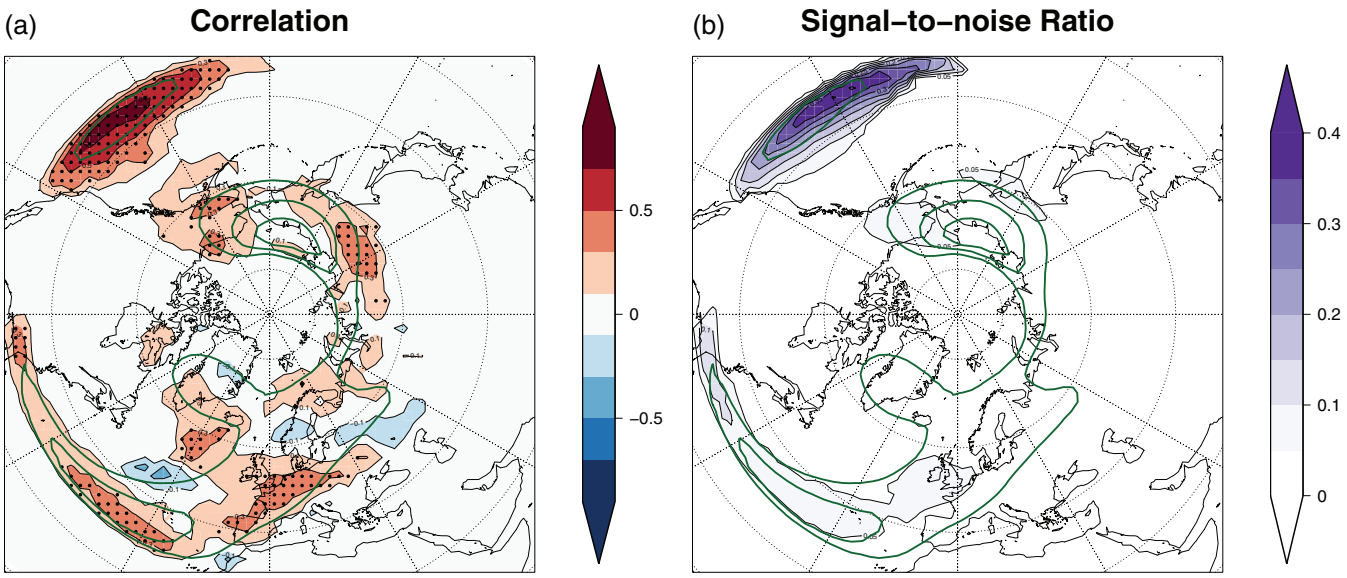

FIGURE 5 Predictability of System 5 DJF 1981–2019 Blocking Events: (a) interannual correlation skill against ERA5 Reanalysis and (b) signal-to-noise ratio. Stippling in (a) indicates the 5% significance level of being different from zero based on a two-tailed *t*-test. Green contours show the S5 blocking climatology frequency and are drawn every 5%. Values are only plotted when the Blocking Events frequency exceeds 2%

reasonably by the model and usually have a larger skill. Even larger correlations are achieved in Pacific LLB sector, with values attaining 0.8: the proximity of tropical Pacific convection is a good indicator of the connection between the Rossby-wave sources there and the wave-breaking activity in the Central Pacific. This region is also close to the area of El Niño Southern Oscillation (ENSO) direct influence, which can give tropic-wide warming, thus providing skill in subtropical temperature gradients that can affect Rossby-wave propagation and breaking. Conversely, the skill in high-latitude blocking regions (Greenland and North Pacific) is not significant, with the exception of two regions to the east and west of the North Pacific maximum, suggesting difficulties in forecasting the interannual variations of cyclonic wave breaking at high latitudes.

The signal-to-noise ratio *SNR* is shown in Figure 5b. This is computed as

$$SNR = \sigma^2_{\text{signal}}/\sigma^2_{\text{noise}} = \sigma^2_{\text{signal}}/(\sigma^2_{\text{total}} - \sigma^2_{\text{signal}}), \qquad (6)$$

where σ^2_{signal} is the interannual variance associated with the signal (i.e., the interannual variance of the ensemble mean), σ^2_{noise} is the interannual variance associated with the noise, and σ^2_{total} is the total interannual variance (i.e., considering all the ensemble members as part of the same experiment). The results over the Atlantic basin are quite low, with values everywhere below 10%. Over the Pacific, there is a clear distinction between the tropically driven Pacific LLB region and the North Pacific area, where the midlatitude noise dominates over the signal. Overall, the low predictability and signal-to-noise ratio shown in Figure 5 for most of the blocking regions is expected, given the dynamics of the midlatitudes, and it demonstrates further how difficult seasonal forecasts in this region are (e.g., Johnson *et al.*, 2019). The ratio of predictable components (RPC: Eade *et al.*, 2014), when the skill is positive, is generally confined between 1 and 2.5 (not shown), suggesting a moderate underconfidence of blocking in S5. However, considering that the skill for forecasts of the seasonal mean of 500-hPa geopotential height over Europe is low (Weisheimer *et al.*, 2019, see also the ECMWF website[3]), the regions of positive significant skill seen in Figure 5a are rather encouraging.

Figure 6 investigates the interannual blocking variability, focusing on the interannual variance. Here it is shown that regions of large interannual variance in ERA5 (Figure 6a) map well on to regions characterised by strong blocking activity. Considering the biases seen in Figure 1, it is not surprising that the underestimation of the interannual variance in S5 (Figure 6b,c) is particularly strong over the British Isles, Western Europe, and the Ural Blocking region. Indeed, in those regions, the S5 interannual variance is almost half the observed one. In order to highlight

[3]https://www.ecmwf.int/en/forecasts/charts/catalogue/seasonal_system5_anomaly_correlation_z500?facets=Range,Long%20(Months)%3BType,Verification%3BParameters,500%20hPa%20geopotential&time=2017110100,744,2017120200

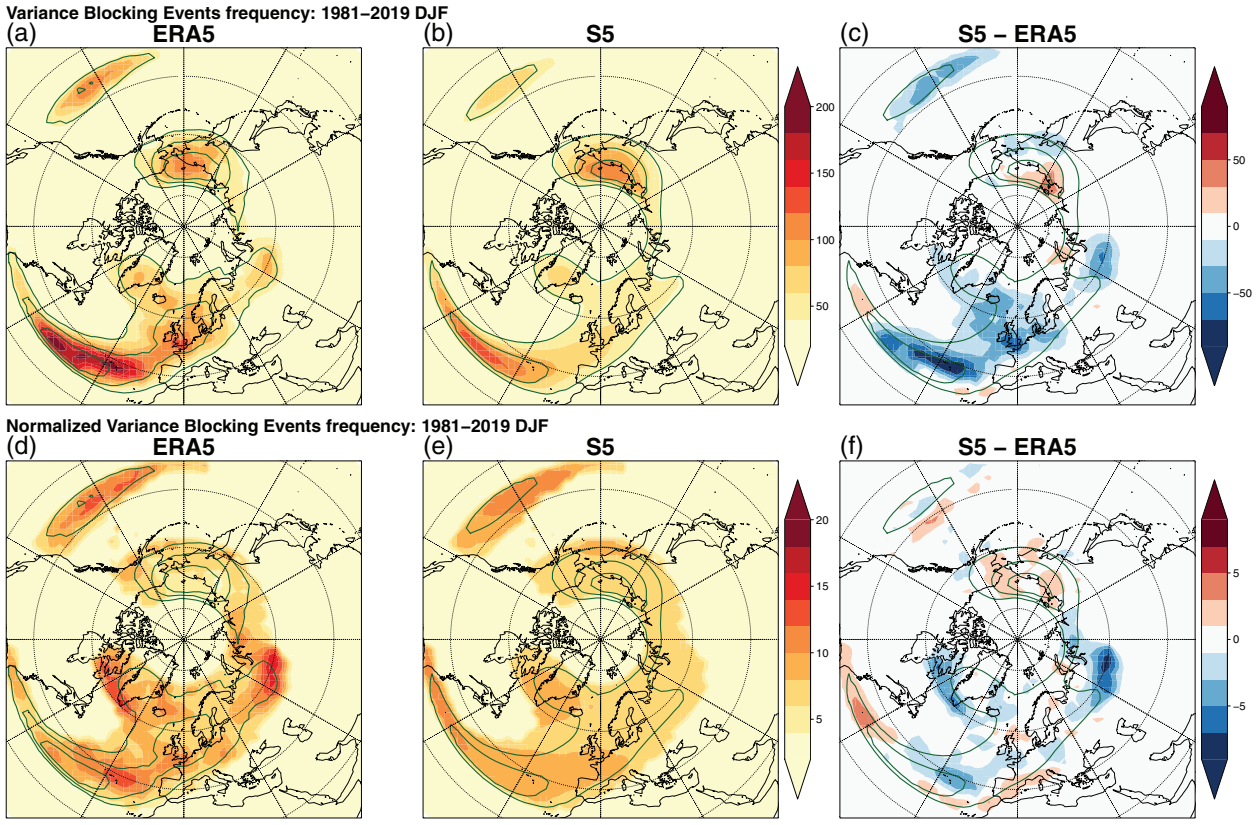

FIGURE 6 DJF 1981–2019 Interannual variance for (a) ERA5, (b) S5, and (c) S5—ERA5 differences. (d) Normalized interannual variance for ERA5, (e) the same for S5, and (f) S5—ERA5 differences. Green contours show the blocking climatology frequency for (a,d) ERA5 and (b,c,e,f) S5. Contours are drawn every 5%. Values are only plotted when the Blocking Events frequency exceeds 2%

regions where the blocking frequency variance is particularly large, the effect induced by the mean value has been removed: Figure 6d normalises the total variance of ERA5 blocking events by dividing by its mean frequency. In this way, it is possible to see that in ERA5 the largest "normalised variance" is now found over the Ural Blocking sector and over the Eastern subtropical Atlantic, often on the flanks of blocking regions, and perhaps associated with variability of the location of blocking activity. The S5 biases for normalised variance—obtained dividing the interannual variance by the climatological S5 blocking frequencies—in Figure 6e,f show that the largest errors occur over the Ural. Given that this region is characterised by few blocking events, it is hard to say if this bias is a robust feature or whether it is related to the sampling uncertainty in ERA5.

Overall, the fact that the spatial distribution of the normalised variance shows a uniform pattern further suggests the presence of a linear relationship between the interannual variance and the mean in blocking. Indeed, for grid points where the climatological values are larger than 2% blocked days, the Pearson's correlation between climatological frequencies and interannual variance is 0.84 for ERA5 and 0.96 for S5. This has an important consequence for blocking analysis, since it implies that a large climatological negative blocking frequency bias is associated with a negative bias in interannual variance, which may affect the predictability of blocking. In other words, it could be assumed that a correct interannual variability can be obtained only if the climatological frequency of blocking is properly simulated (or vice versa).

In order to conclude the analysis of the S5 predictive skill, Figure 7 shows the time series of blocking events plotted as a box plot for the five sectors defined in Figure 1e. The correlation is displayed at the top left of each panel and reflects the results already shown in Figure 5a, with significant results obtained only over the LLB sectors (0.34 for the Atlantic LLB, 0.76 for the Pacific LLB). Over the other three main midlatitude and high-latitude sectors, the skill in terms of Pearson's correlation is confirmed to be low: however, it should be noted that the SEAS5 overall skill for geopotential height over the Euro-Atlantic sector is not particularly high, considering for instance that NAO shows a skill of 0.43 (Johnson *et al.*, 2019). The ensemble spread is, overall, rather stable over the years, with little indication of flow dependence (as shown

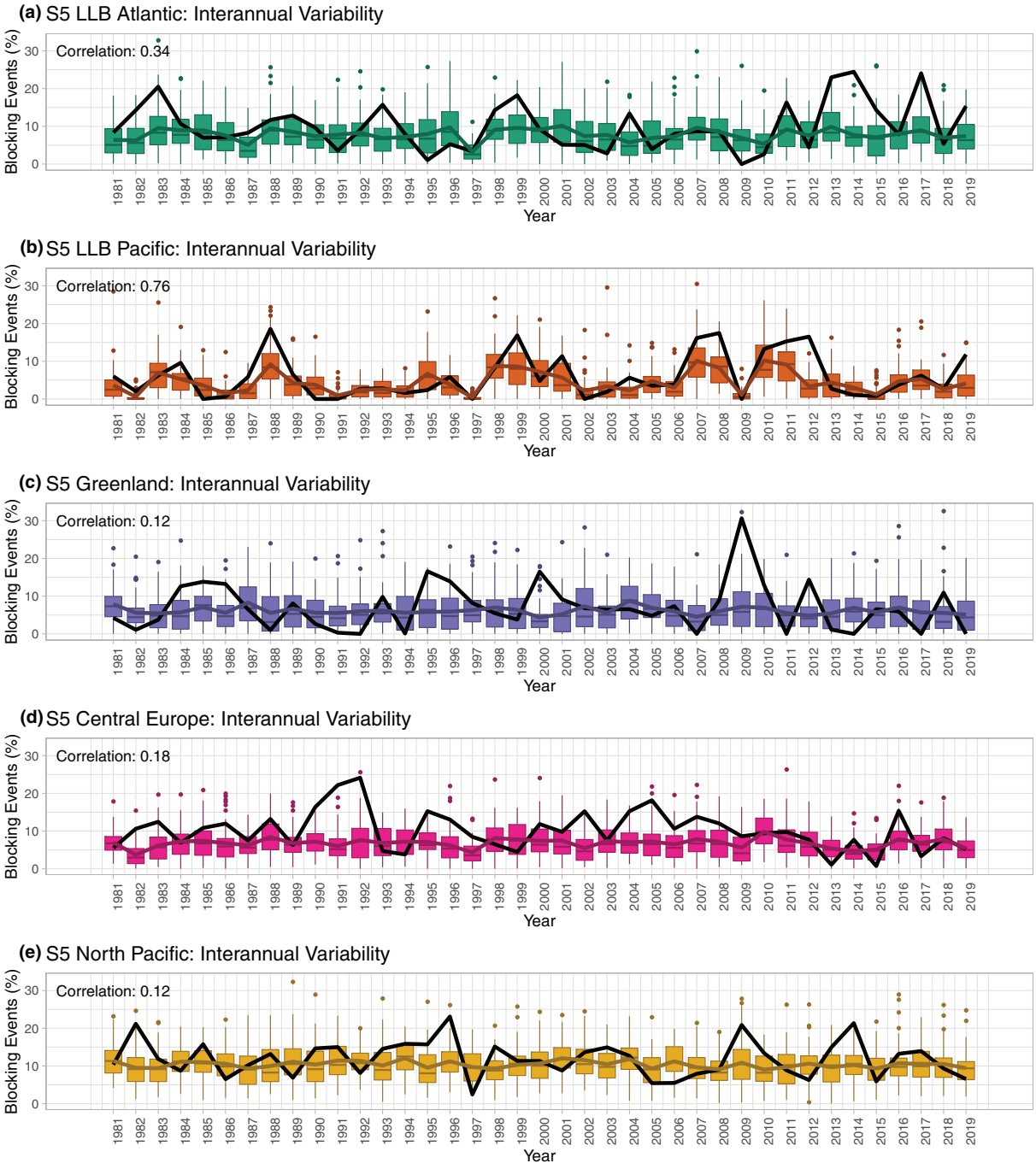

FIGURE 7 Box plot representing the seasonally averaged S5 DJF 1981–2019 Blocking Events time series over (a) Atlantic LLB, (b) Pacific LLB, (c) Greenland, (d) Central Europe, and (e) North Pacific sectors. Bold coloured lines represent the ensemble mean. The black line is the ERA5 reanalysis. The lower, central, and upper hinge show the first quartile, the median, and the third quartile, respectively. The upper (lower) whiskers extend from the third (first) quartile to the largest (smallest) value in the ensemble, but limited to an upper bound that is 1.5 times the interquartile range (i.e., the distance between the third and first quartiles). Dots show outliers. The correlation between the ensemble mean and reanalysis is shown in the upper left of each panel

by the box plot extension) for most of the sectors. The only clear exception is Pacific LLB, where a few starting dates show considerably reduced ensemble spread (e.g., 1982, 1991, 1997, 2009, 2015). Interestingly, the November 1997 starting date also shows a halved ensemble spread in the Atlantic LLB, and also partially in Central Europe. The fact that those coincide mostly with positive ENSO phases, with the 1997–1998 winter being characterised by the strongest positive ENSO event during the hindcast period, suggests the presence of a potential linkage

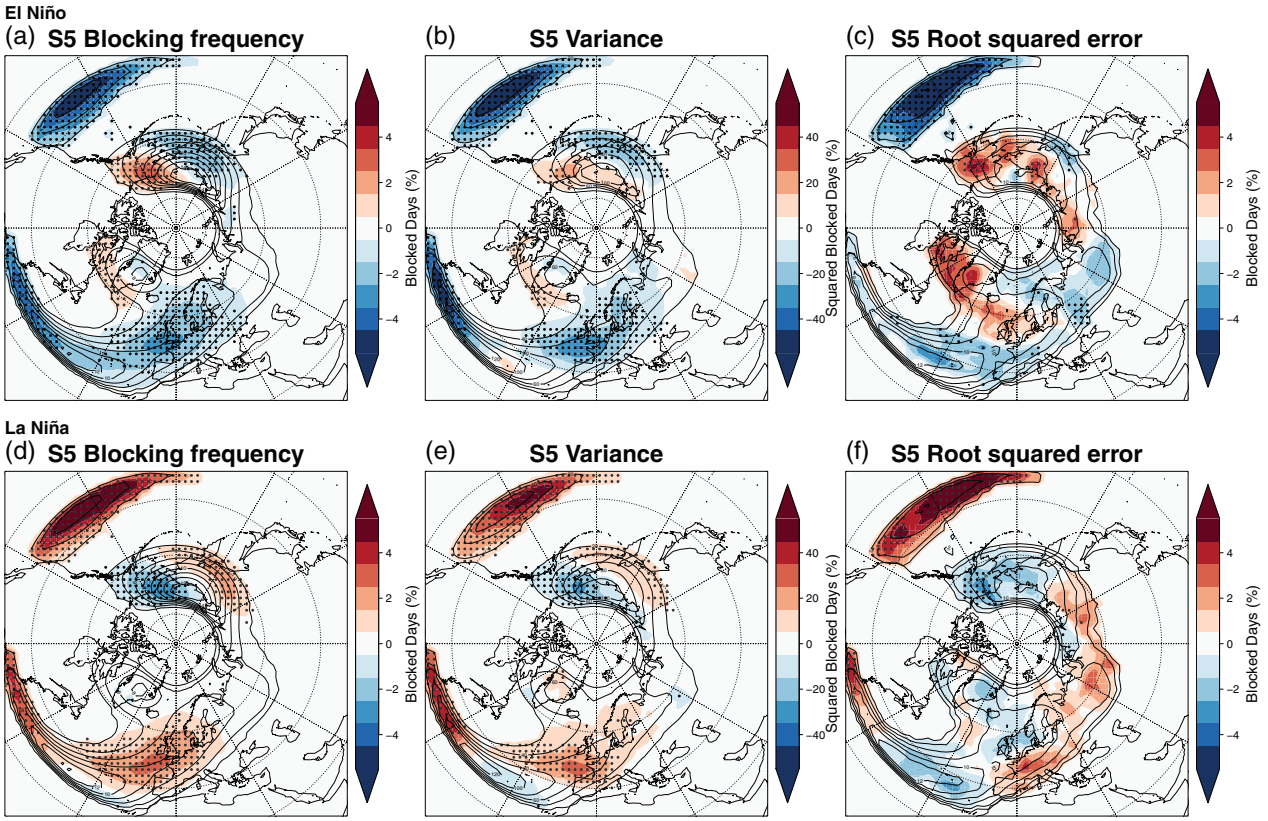

FIGURE 8 S5 DJF statistics of the blocking frequency anomalies (shading) during (a,b,c) El Niño years and (d,e,f) La Niña years, with respect to the value of the corresponding statistic over the full time window 1981–2019 (contours). (a) and (d) show the mean blocking frequency; (b) and (e) show the ensemble variance; (c) and (f) show the forecast mean-root-squared error. Stippling indicates the 5% significance level with a 1,000-trial bootstrap test

between the frequency of low-latitude anticyclonic wave breaking and ENSO: this is investigated in the following section.

3.4 | Impact of ENSO in SEAS5 (1981–2019)

We composited blocking seasons depending on the ENSO Niño3.4 SST index (defined as the average SST in the 170°–120°W 5°S–5°N box) over the DJF 1981–2019 time window. El Niño (La Niña) years are then defined when the Niño3.4 SST index anomaly exceeds (falls below) 1 K (−1 K). A total of six El Niño and six La Niña winters are thus defined. The left column of Figure 8 shows the composite of positive and negative ENSO years for the blocking frequency. The impact of ENSO on blocking in S5 is quite robust and symmetric: El Niño leads to a moderate decrease of blocking frequency over Europe (−1.5%) and to a stronger decrease at low latitudes over both the Atlantic and the Pacific basin. A northeastward displacement is observed for North Pacific blocking, while no evident signal is seen over Greenland, suggesting a weak teleconnection between the NAO and ENSO in S5. Overall, this is in good agreement with observations (Renwick and Wallace, 1996; Barriopedro *et al.*, 2006; Tibaldi and Molteni, 2018). It is, however, important to recall that these anomalies are considerably smaller than the simulated interannual variations (e.g., the interannual standard deviation over Europe is 4.5%, i.e., three times larger than the ENSO signal). The central column of Figure 8 shows the ensemble variance anomalies during ENSO years: during El Niño events, the ensemble variance is notably reduced, especially at low latitudes. Conversely, in an almost symmetric fashion, La Niña events seem to increase the variability of wave breaking on the equatorward side of the jet, with a special impact on low-latitude blocking events. This would suggest that low blocking activity years are more predictable than high blocking activity years. For the small number of events (six), a meaningful correlation cannot be computed: however, the panels on the right of Figure 8 show the impact of ENSO on the S5 RMSE. These results are pointing in the same direction as seen for the blocking frequency and ensemble variance, with a reduced/increased RMSE for low-latitude blocking in the Pacific during El Niño/La Niña. A stronger response

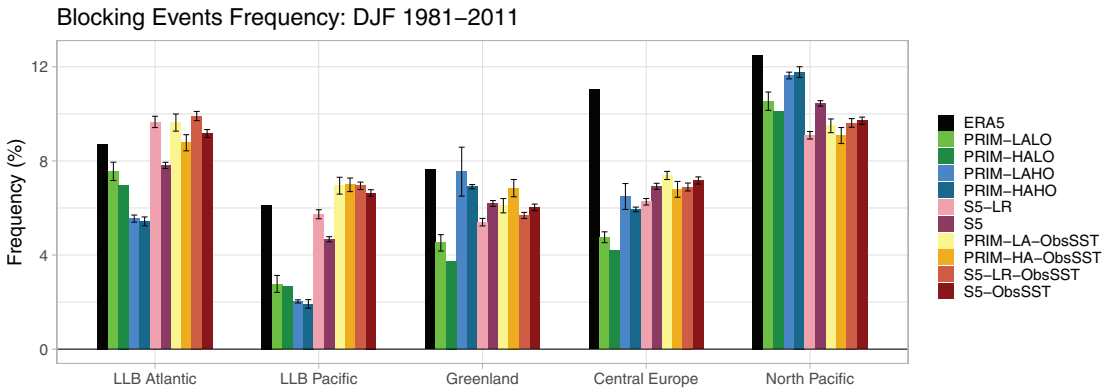

FIGURE 9 DJF 1981–2011 Blocking Events climatology averaged over the five different sectors shown in Figure 1e for ERA5, S5, S5-LR, S5-ObsSST, and S5-LR-ObsSST, and the PRIMAVERA climate runs. Lighter colours show low atmospheric resolution, darker colours high atmospheric resolution. Error bars represent the standard deviation of the mean (when more than one member is available)

is seen during El Niño years: however, the significance (assessed with 5% level with a 1,000-trial bootstrap) over the main regions of blocking is rarely achieved, with better results over the Pacific.

3.5 | Comparison with climate runs (1981–2011)

In the last part of this work, we compare a set of climatological runs (Roberts *et al.*, 2018) performed for the PRIMAVERA project using a setup similar to the one used for S5. The main difference from the S5 hindcasts is that the PRIMAVERA simulations are forced climate runs rather than initialised seasonal forecasts. PRIMAVERA runs are initialised from 1950 and continuously integrated up to 2014, while seasonal hindcasts are initialised every 1st of November from an observed state of the Earth system. This means that (a) PRIMAVERA runs are integrated on a longer time-scale, so that the ocean can reach its equilibrium state, possibly increasing its bias, and (b) PRIMAVERA simulations do not show a binary correspondence between the weather simulated in individual years and observations. On top of that, it is important to highlight that PRIMAVERA integrations are characterised by a limited number of ensemble members and that "high" resolution runs use the $T_{co}399$ grid rather than the $T_{co}319$ grid, as shown by Table 1.

Figure 9 shows the results from PRIMAVERA integrations, showing the averaged blocking frequency in the five main blocking sectors. As a reference, results from S5 and S5-LR—which are the hindcasts run with the same model—and the corresponding atmosphere-only runs (S5-ObsSST and S5-LR-ObsSST) are also shown. Interestingly, in almost all the PRIMAVERA configurations the difference between high and low atmospheric resolution seems to be marginal, showing a slight deterioration over most of the sectors (see green, blue, and yellow bars). Conversely, increasing the oceanic resolution brings a considerable increase for Greenland, Europe, and North Pacific sectors and a deterioration over the Atlantic and Pacific LLB sectors (compare blue bars and green bars), again suggesting a poleward displacement of blocking activity at higher resolution, as seen in Figure 2d.

While the improvement following oceanic grid refinement has already been seen in some specific models (e.g., Scaife *et al.*, 2010), the lack of impact of a finer atmospheric grid—or at most negligible differences—is contrary to the expectations from previous publications using the same atmospheric model (Jung *et al.*, 2012; Davini *et al.*, 2017). This is even more surprising, considering that for S5 the atmospheric resolution seemed to bring a moderate bias reduction (see Section 3.1).

Another relevant feature that can be easily spotted in Figure 9 is that benefits carried by the combination of high atmospheric and oceanic resolutions are much more relevant for climate runs than for seasonal runs. Indeed, great improvements in the Greenland, Central Europe, and North Pacific sectors are seen going from PRIM-LALO to PRIM-HAHO, while much smaller improvements are seen going from S5-LR to S5. Overall, the impact of ocean resolution is more noticeable in the PRIMAVERA integrations than in the seasonal runs, consistent with the results reported by Roberts *et al.* (2020).

Considering the large set of simulations available, it is also possible to investigate the impact of different model components and initialisation procedures in the representation of blocking.

1. By comparing atmosphere-only climate runs forced with prescribed SSTs (PRIM-LA-ObsSST and

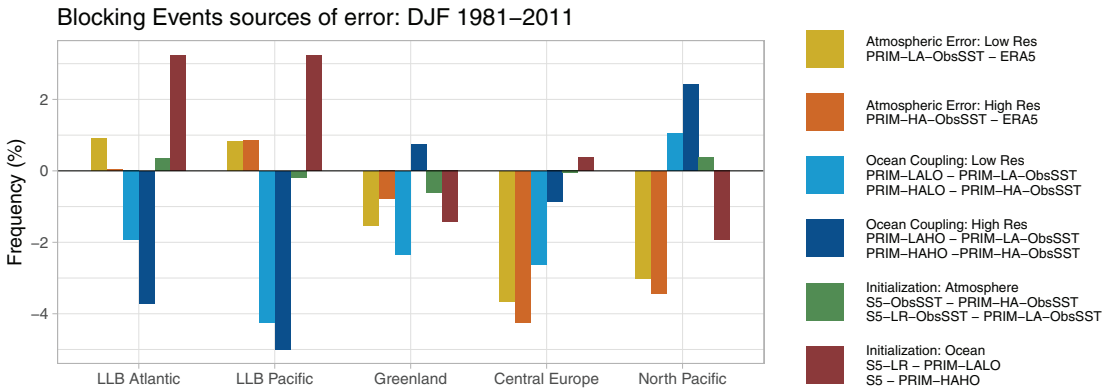

FIGURE 10 Blocking Events DJF 1981–2011 analysis of the sources of the blocking frequency error. For each sector, making use of the different experiments available, the different sources of blocking error have been investigated: the contribution of the intrinsic atmospheric model bias for (1) a low-resolution atmosphere ($T_{co}199$, yellow) and (2) a high resolution atmosphere ($T_{co}399$, orange), the impact of the coupling with (3) a low-resolution ocean (ORCA1, light blue) or (4) a high-resolution ocean (ORCA025, dark blue), (5) the impact of the November initialization of the atmosphere/land-surface (green), and (6) the impact of the November initialization of the ocean/sea-ice (red). Experiments used for each comparison are listed in the legend

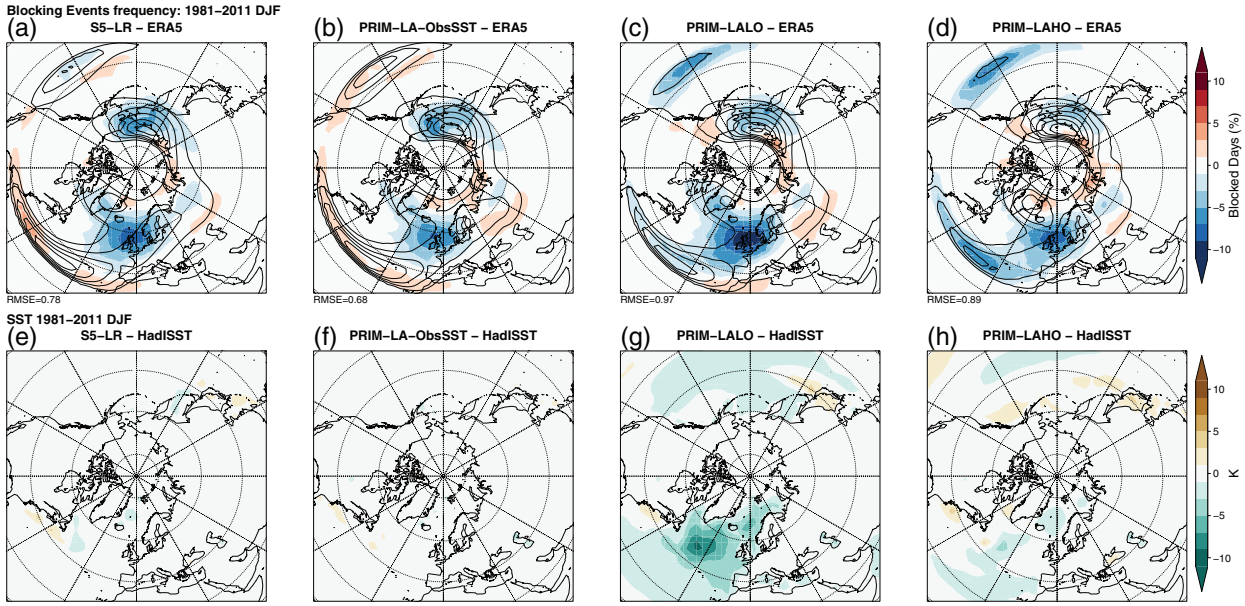

FIGURE 11 DJF 1981–2011 Blocking Events frequency for (a) S5-LR, (b) PRIM-LA-ObsSST, (c) PRIM-LALO, and (d) PRIM-LAHO (contours) and their bias against ERA5 reanalysis (shading). Contours are drawn each 3%. SST differences between (e) S5-LR, (f) PRIM-LA-ObsSST, (g) PRIM-LALO, and (h) PRIM-LAHO and the HadISST reanalysis

PRIM-HA-ObsSST) against observations, it is possible to estimate the intrinsic atmospheric model bias in blocking frequency at both low ($T_{co}199$) and high ($T_{co}399$) resolution.

2. By comparing coupled climate runs (PRIM-LALO, PRIM-LAHO, PRIM-HALO, and PRIM-HAHO) against atmosphere-only climate runs with prescribed SSTs (PRIM-LA-ObsSST and PRIM-HA-ObsSST), it is possible to assess the effect of coupling with a low (ORCA1) resolution or with a high (ORCA025) resolution oceanic model.
3. By comparing seasonal hindcasts with prescribed SSTs (S5-ObsSST and S5-LR-ObsSST) against atmosphere-only climate runs with prescribed SSTs (PRIM-LA-ObsSST and PRIM-HA-ObsSST), it is possible to assess the relevance that atmosphere and land-surface initialisation may have on blocking frequencies.

4. By comparing coupled seasonal hindcasts (S5-LR and S5) against coupled climate runs (PRIM-LALO and PRIM-HAHO), it is possible to investigate the role of the ocean and sea-ice initialisation.

All this information is condensed in Figure 10, where a comparison amongst the different sources of Blocking Events error is presented. From here we can see the following.

- Over Europe and the North Pacific, blocking frequency appears to be little affected by various model features (e.g., resolution, interactive or prescribed ocean, initialisation of atmosphere, land or ocean) compared with the amplitude of the atmospheric model intrinsic bias in blocking frequency (see yellow and orange bars).
- As discussed above, PRIMAVERA integrations do not show evident improvements following an atmospheric horizontal resolution increase (compare yellow and orange bars). It is interesting to notice that atmospheric grid refinement leads to an improvement where the bias is relatively small (as for Greenland and Atlantic LLB, ∼1%) but a deterioration is produced where the bias is relatively large (Europe and North Pacific, ∼3%).
- Coupling the atmospheric model with an interactive ocean leads to deterioration of the representation of blocking over all sectors except the North Pacific (see blue bars), in agreement with previous works (e.g., Hartung *et al.*, 2017). This is particularly true for LLB sectors. Better performance for coupled GCMs over the North Pacific has already been observed by Davini and D'Andrea (2016) and this suggests that coupled processes may be more relevant in the North Pacific sector than elsewhere.
- Compared with low-resolution ocean, high-resolution ocean increases the blocking frequency over Greenland, Europe, and North Pacific sectors (compare light and dark blue bars), confirming the results from Roberts *et al.* (2020). Interestingly, a reduction of LLB frequencies is observed, showing once more the complementarity between LLB and mid/high-latitude blocking sectors.
- The November initialisation of the atmosphere and land surface has negligible benefits for the blocking frequency climatology (green bars). This is expected for the atmosphere, where the inherent chaotic nature of the flow erases any memory of the initial conditions after a few weeks. However, it is more surprising for the land surface: indeed, this means that features such as the snow cover initialisation have very little impact on the climatology of blocking.
- Conversely, November initialisation for the ocean and sea-ice has a large impact on blocking representation, except for Central Europe. What is particularly interesting is that this impact acts in the opposite direction to the effect of the coupling to the ocean (compare blue and red bars), in all the sectors analysed here. This suggests that the November initialisation of the ocean is able to keep the ocean model sufficiently far from its own attractor that it levels out the bias introduced by the coupling. In such a way, the seasonal hindcasts have a blocking climatology closer to their atmosphere-only counterparts. Indeed, the extent to which S5 results are closer to PRIM-HAHO or to S5-ObSST can be taken as a measure of the degree of development of the coupling-induced model biases affecting blocking frequencies. The largest differences are seen over the Atlantic and Pacific LLB sectors, where a correct initialisation seems to increase the blocking frequency, probably following their larger connection with tropical dynamics. A moderate decrease is seen over Greenland and the North Pacific, while negligible changes are found over Central Europe.

An intriguing finding from Figure 10 is the notable improvement obtained following the oceanic resolution refinement. Further insight can be gained from the analysis presented in Figure 11, where all low atmospheric resolution configurations ($T_{co}199$: similar results can be obtained looking at high-resolution models with $T_{co}399/T_{co}319$ resolution) are analysed, showing their blocking bias in the upper row and their SST bias in the lower row (against HadISST). We compare the low-resolution version of S5 (S5-LR) with low atmospheric resolution simulations from PRIMAVERA: with prescribed observed SSTs (PRIM-LA-ObsSST), with low resolution in the ocean (PRIM-LALO), and with high resolution in the ocean (PRIM-LAHO). These models, with the exception of the PRIM-LALO configuration, are characterised by a blocking bias of similar magnitude over the European sector.

Since coupled climate runs can drift freely over many years, it is not surprising to see that the PRIMAVERA runs are characterised by larger SST biases than the seasonal hindcasts (Figure 11, bottom row). For example, S5-LR, performed with a 1-degree ocean model, has an almost negligible SST bias (Figure 11e) compared with the coupled PRIMAVERA runs (PRIM-LAHO and PRIM-LALO). This is due to the ocean initialisation in seasonal forecasts, which makes S5-LR SST bias close to the atmosphere-only PRIM-LA-ObsSST (Figure 11f). Indeed, the S5-LR SST errors are smaller than those of PRIM-LAHO (Figure 11h), which is based on a four times finer oceanic mesh

with a quarter-degree resolution and has a bias of a few degrees in the subpolar gyre region. The differences are even more striking with PRIM-LALO (Figure 11g), which has a configuration very similar to S5-LR: here, there is a significant negative bias in the Atlantic sector of more than 10 °C. This potentially explains why, in PRIMAVERA climate runs, increasing the oceanic resolution provides a robust increase of blocking frequencies, but has a comparatively smaller impact on seasonal forecasts.

In summary, findings from Figure 11 suggest that the benefits of a high-resolution ocean may be more relevant in climate runs than in seasonal prediction. However, since it has been shown that a positive correlation exists between the magnitude of the bias of North Atlantic SSTs and the bias for European and Greenland blocking frequencies (Davini and D'Andrea, 2020), it could be possible that what really matters for blocking simulation in coupled GCMs is the presence of a reasonable oceanic mean state: of course, this could be obtained by increasing the horizontal resolution of the oceanic model, but other bias-reducing approaches could be optimal as well. It should, however, be recalled that the largest source of blocking frequency error is introduced by the atmospheric model component, so that reducing the oceanic model bias would be a second-tier goal.

4 | DISCUSSION AND CONCLUSIONS

This study has analysed the properties of atmospheric blocking in a set of ECMWF seasonal prediction systems. Three different generations, namely System3, System4, and SEAS5, have been compared. These three successive versions differ in many aspects, including atmospheric and oceanic resolution, different cycles of the atmospheric and ocean models, different details in the initialisation procedures, and number of ensemble members. While the "evolution" of the three different systems shows an improvement in the climatological representation of blocking, with reduced bias in almost all sectors, the substantial underestimation of blocking—especially over Central Europe—is still an open issue.

In addition to these three operational systems, the results from several controlled sensitivity experiments have been studied. This has been done in order to highlight potential benefits in blocking representation associated with increased atmospheric and oceanic resolution, with/without atmosphere–ocean coupling and with/without stochastic physics parameterisations. In agreement with previous works (e.g., Davini and D'Andrea, 2016; Hartung *et al.*, 2017), results show weak sensitivity to coupled model SST errors. On the other hand, minor effects by stochastic parameterisations—which seems to shift the blocking activity equatorward, increasing LLB at the expense of blocking at mid and high latitudes—is observed. A moderate improvement of blocking frequencies can be found when increasing resolution—a common feature in climate models (e.g., Davini and D'Andrea, 2020)—especially over the Greenland and North Pacific sectors, associated with a reduction of Atlantic and Pacific LLB. Overall, the three sensitivities analysed seems to operate in a similar fashion, that is, increasing/decreasing LLB to the detriment of mid/high-latitude blocking. This suggests that all model changes analysed here (i.e., horizontal resolution, prescribed/interactive SSTs, stochastic parameterisations) interact in a similar way with the model mean state (which affects Rossby-wave propagation), leading to meridional displacements in blocking frequencies.

It is shown that the blocking frequency bias can be decomposed into a part associated with a duration error and a part associated with an error in the number of onsets (where blocking onsets can be interpreted as successive Rossby-wave breaking events: e.g., Woollings *et al.* 2018). ECMWF seasonal prediction systems simulate blocking duration reasonably over the Atlantic and Pacific LLB sectors, while overestimating the number of blocking onsets. The situation for mid- and high-latitude sectors is more complicated: while in all three sectors the models underestimate duration considerably, the number of events is underestimated over Central Europe, simulated well over the North Pacific, and overestimated over Greenland. Furthermore, the fact that all the differences observed among the different configurations are caused mainly by changes in the number of onsets suggests that—especially for midlatitude and high-latitude blocking—a considerable part of the blocking frequency bias could be eliminated if the processes controlling the blocking duration were better resolved by the models. Dedicated studies involving specific processes, as performed by Steinfeld *et al.* (2020) for diabatic heating, are therefore recommended.

The second part of the article focuses on the properties and prediction skills of SEAS5. Over Europe, Greenland, and the North Pacific, the signal-to-noise ratio is very low ($\sim$ 10%). The blocking interannual variability is underestimated by S5 almost everywhere, following the mean frequency bias. Although a deeper understanding of blocking interannual variability is beyond the goal of the current work, the positive correlation between mean and variance of Blocking Events may suggest the possibility that, at an interannual time-scale, Blocking Events are following a Poisson-like distribution. This will be the subject of a forthcoming study by the authors.

S5 exhibits some skill in Blocking Events over Western and Central Europe, although it is sensitive to the period of analysis: it is interesting to notice that an analogous positive skill was found for a similar region by Athanasiadis *et al.* (2014) for the UKMO seasonal forecasting system. Even if blocking skill remains low, over Europe it is notably better than that observed for geopotential height as a whole (Weisheimer *et al.*, 2019). More robust results are obtained for the tropically driven Atlantic and Pacific LLB sectors. Such regions are the most sensitive to ENSO, showing smaller mean values, reduced forecast errors, and reduced variance during El Niño events (and vice versa during La Niña years). Overall, the S5 response to ENSO is in agreement with observations (e.g., Tibaldi and Molteni, 2018).

One interesting consequence of (a) blocking variance being positively correlated with blocking climatological frequency, (b) climatological blocking frequencies being underestimated, and (c) European blocking skill being larger than geopotential height skill, is that S5 possibly underestimates the predictable signal over Central Europe. Indeed, it could be possible that the low but significant skill currently found for blocking is not diagnosed in the midlatitude geopotential height field because the blocking-induced signal is too small (due to its underestimated variance). This suggests that improving the blocking mean frequency may also improve the overall seasonal skill in the midlatitudes.

The final part of the work is dedicated to the comparison between the seasonal hindcasts and a set of climate simulations by ECMWF run within the HighResMIP protocol. Those simulations share the same atmospheric and oceanic models, but they are run in climate mode (i.e., without initialising the model every November). Overall, these simulations present larger blocking bias than seasonal hindcasts when run in coupled mode, but quite similar biases when run in atmosphere-only mode. By comparing the two sets of simulations, it is possible to show that the largest source of bias is the inherent atmospheric error, which is usually increased when coupling the atmosphere to an interactive ocean. While atmosphere and land-surface November initialisation have a negligible impact on blocking statistics—suggesting that features such as November snow cover may have small relevance in shaping blocking-frequency winter climatology—it is shown that ocean and sea-ice initialisation provides some benefits over the Atlantic and Pacific LLB, but seems only marginally relevant for blocking over Greenland and Central Europe. Furthermore, it is shown that oceanic high resolution is much more relevant in climate runs than in seasonal hindcasts. Considering also that it has been shown that high ocean resolution does not bring evident improvements to blocking frequencies on seasonal time-scales (Prodhomme *et al.*, 2016; Roberts *et al.*, 2020), it is reasonable to conclude that—at current atmospheric resolution—other effects associated with an eddy-permitting ocean model (such as improved air–sea fluxes) are of secondary importance compared with the reduced mean-state SST bias.

To conclude, the present work suggests that seasonal forecasts may be an interesting test-bed to address blocking bias in numerical models: they are better than atmosphere-only climate runs, since they include coupled processes—even if they show very similar biases—and they are better than coupled climate runs, since they show much smaller SST biases. Furthermore, given the limited impact that a high-resolution ocean has on blocking statistics in seasonal hindcasts, they can also be run at lower oceanic resolution with reduced computing cost. We therefore encourage more research on atmospheric blocking biases in GCMs, taking advantage of seasonal forecasts.

ACKNOWLEDGEMENTS

We thank two anonymous reviewers for helping us to improve the article. Special thanks are due to Fabio D'Andrea for the illuminating discussion on the blocking duration.

ORCID

Paolo Davini https://orcid.org/0000-0003-3389-7849
Antje Weisheimer https://orcid.org/0000-0002-7231-6974
Magdalena Balmaseda https://orcid.org/0000-0002-9611-8788
Stephanie J. Johnson https://orcid.org/0000-0002-0569-043X
Franco Molteni http://orcid.org/0000-0003-0651-4566
Christopher D. Roberts https://orcid.org/0000-0002-2958-6637
Retish Senan https://orcid.org/0000-0003-1949-1893
Timothy N. Stockdale https://orcid.org/0000-0002-7901-0337

REFERENCES

Athanasiadis, P.J., Bellucci, A., Hermanson, L., Scaife, A.A., MacLachlan, C., Arribas, A., Materia, S., Borrelli, A. and Gualdi, S. (2014) The representation of atmospheric blocking and the associated low-frequency variability in two seasonal prediction systems. *Journal of Climate*, 27, 9082–9100.

Athanasiadis, P.J., Yeager, S., Kwon, Y.O., Bellucci, A., Smith, D.W. and Tibaldi, S. (2020) Decadal predictability of North Atlantic blocking and the NAO. *NPJ Climate and Atmospheric Science*, 3, 1–10.

Barriopedro, D., Garcia-Herrera, R., Lupo, A. and Hernandez, E. (2006) A climatology of Northern Hemisphere Blocking. *Journal of Climate*, 19, 1042–1063.

Berrisford, P., Hoskins, B. and Tyrlis, E. (2007) Blocking and Rossby wave breaking on the dynamical tropopause in the Southern Hemisphere. *Journal of the Atmospheric Science*, 64, 2881–2898.

Buehler, T., Raible, C. and Stocker, T. (2011) The relationship of winter season North Atlantic blocking frequencies to extreme cold or dry spells in the ERA-40. *Tellus A*, 63, 212–222.

Buizza, R., Milleer, M. and Palmer, T.N. (1999) Stochastic representation of model uncertainties in the ECMWF ensemble prediction system. *Quarterly Journal of the Royal Meteorological Society*, 125, 2887–2908.

Davini, P. (2019) MiLES-Mid latitude evaluation system. *Zenodo*, http://doi.org/10.5281/zenodo.2578139.

Davini, P., Cagnazzo, C., Gualdi, S. and Navarra, A. (2012) Bidimensional diagnostics, variability and trends of Northern Hemisphere blocking. *Journal of Climate*, 25, 6996–6509.

Davini, P., Corti, S., D'Andrea, F., Riviere, G. and von Hardenberg, J. (2017) Improved winter European atmospheric blocking frequencies in high-resolution global climate simulations. *Journal of Advances in Modeling Earth Systems*, 9, 2615–2634.

Davini, P. and D'Andrea, F. (2016) Northern hemisphere atmospheric blocking representation in global climate models: twenty years of improvements?. *Journal of Climate*, 29, 8823–8840.

Davini, P. and D'Andrea, F. (2020) From CMIP-3 to CMIP-6: Northern Hemisphere atmospheric blocking simulation in present and future climate. *Journal of Climate*, 33(23), 10021–10038.

Dee, D., Uppala, S., Simmons, A., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A.C.M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A.J., Haimberger, L., Healy, S.B., Hersbach, H., Holm, E.V., Isaksen, L., Kallberg, P., Koehler, M., Matricardi, M., McNally, A.P., Monge-Sanz, B.M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thepaut, J.-N. and Vitart, F. (2011) The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 137, 553–597.

Dunstone, N., Smith, D., Scaife, A., Hermanson, L., Eade, R., Robinson, N., Andrews, M. and Knight, J. (2016) Skilful predictions of the winter North Atlantic Oscillation one year ahead. *Nature Geoscience*, 9. 809

Eade, R., Smith, D., Scaife, A., Wallace, E., Dunstone, N., Hermanson, L. and Robinson, N. (2014) Do seasonal-to-decadal climate predictions underestimate the predictability of the real world?. *Geophysical Research Letters*, 41, 5620–5628.

Haarsma, R.J., Roberts, M.J., Vidale, P.L., Senior, C.A., Bellucci, A., Bao, Q., Chang, P., Corti, S., Fučkar, N.S., Guemas, V., von Hardenberg, J., Hazeleger, W., Kodama, C., Koenigk, T., Leung, L.R., Lu, J., Luo, J.J., Mao, J., Mizielinski, M.S., Mizuta, R., Nobre, P., Satoh, M., Scoccimarro, E., Semmler, T., Small, J. and von Storch, J.S. (2016) High resolution model intercomparison project (HighResMIP v1. 0) for CMIP6. *Geoscientific Model Development*, 9, 4185–4208.

Hartung, K., Svensson, G. and Kjellström, E. (2017) Resolution, physics and atmosphere–ocean interaction–how do they influence climate model representation of Euro-Atlantic atmospheric blocking?. *Tellus A: Dynamic Meteorology and Oceanography*, 69, 1406252. https://doi.org/10.1080/16000870.2017.1406252.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R.J., Holm, E., JaniskovÃą, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S. and Thepaut, J.N. (2020) The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146, 1999–2049.

Johnson, S.J., Stockdale, T.N., Ferranti, L., Balmaseda, M.A., Molteni, F., Magnusson, L., Tietsche, S., Decremer, D., Weisheimer, A., Balsamo, G., Keeley, S., Mogensen, K., Zuo, H. and Monge-Sanz, B. (2019) SEAS5: the new ECMWF seasonal forecast system. *Geoscientific Model Development*, 12, 1087–1117.

Jung, T., Miller, M., Palmer, T., Towers, P., Wedi, N., Achuthavarier, D., Adams, J., Altshuler, E., Cash, B., Kinter III, J., Marx, L., Stan, C. and Hodges, K.I. (2012) High-resolution global climate simulations with the ECMWF model in Project Athena: experimental design, model climate, and seasonal forecast skill. *Journal of Climate*, 25, 3155–3172.

Luo, D., Xiao, Y., Yao, Y., Dai, A., Simmonds, I. and Franzke, C.L. (2016) Impact of Ural blocking on winter warm Arctic–cold Eurasian anomalies. Part I: blocking-induced amplification. *Journal of Climate*, 29, 3925–3947.

Masato, G., Hoskins, B.J. and Woollings, T. (2013) Winter and summer Northern Hemisphere blocking in CMIP5 models. *Journal of Climate*, 26, 7044–7059.

Matsueda, M. (2009) Blocking predictability in operational medium-range ensemble forecasts. *SOLA*, 5, 113–116.

Matsueda, M. (2011) Predictability of Euro-Russian blocking in summer of 2010. *Geophysical Research Letters*, 38, L06801. https://doi.org/10.1029/2010GL046557

Mauritsen, T. and Källén, E. (2004) Blocking prediction in an ensemble forecasting system. *Tellus A: Dynamic Meteorology and Oceanography*, 56, 218–228.

Molteni, F., Stockdale, T., Balmaseda, M., Balsamo, G., Buizza, R., Ferranti, L., Magnusson, L., Mogensen, K., Palmer, T. and Vitart, F. (2011) *The New ECMWF Seasonal Forecast System (System 4)* Vol. 49. Reading, UK: European Centre for Medium-Range Weather Forecasts.

Pelly, J. and Hoskins, B. (2003) A new perspective on blocking. *Journal of the Atmospheric Science*, 60, 743–755.

Pfahl, S. and Wernli, H. (2012) Quantifying the relevance of atmospheric blocking for co-located temperature extremes in the northern hemisphere on (sub-) daily time scales. *Geophysical Research Letters*, 39, L12807. https://doi.org/10.1029/2012GL052261.

Prodhomme, C., Batté, L., Massonnet, F., Davini, P., Bellprat, O., Guemas, V. and Doblas-Reyes, F. (2016) Benefits of increasing the model resolution for the seasonal forecast quality in EC-Earth. *Journal of Climate*, 29, 9141–9162.

Rayner, N., Parker, D., Horton, E., Folland, C., Alexander, L., Rowell, D., Kent, E. and Kaplan, A. (2003) Global analyses of sea-surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *Journal of Geophysical Research*, 108. 4407

Renwick, J. and Wallace, J. (1996) Relationships between North Pacific wintertime blocking, El Niño, and the PNA pattern. *Monthly Weather Review*, 124, 2071–2076.

Rex, D. (1950) Blocking action in the middle troposphere and its effect upon regional climate: I. An aerological study of blocking action. *Tellus*, 2, 196–211.

Roberts, C., Vitart, F., Balmaseda, M. and Molteni, F. (2020) The time-scale-dependent response of the wintertime north atlantic to increased ocean model resolution in a coupled forecast model. *Journal of Climate*, 33, 3663–3689.

Roberts, C.D., Senan, R., Molteni, F., Boussetta, S., Mayer, M. and Keeley, S. (2018) Climate model configurations of the ECMWF integrated forecast system (ECMWF-IFS cycle 43r1) for HighResMIP. *Geoscientific Model Development Discussions*, 2018, 1–48.

Scaife, A., Woollings, T., Knight, J., Martin, G. and Hinton, T. (2010) Atmospheric blocking and mean biases in climate models. *Journal of Climate*, 23, 6143–6152.

Schaller, N., Sillmann, J., Anstey, J., Fischer, E.M., Grams, C.M. and Russo, S. (2018) Influence of blocking on Northern European and Western Russian heatwaves in large climate model ensembles. *Environmental Research Letters*, 13. 054015

Scherrer, S., Croci-Maspoli, M., Schwierz, C. and Appenzeller, C. (2006) Two-dimensional indices of atmospheric blocking and their statistical relationship with winter climate patterns in the Euro-Atlantic region. *International Journal of Climatology*, 26, 233–249.

Schiemann, R., Athanasiadis, P., Barriopedro, D., Doblas-Reyes, F., Lohmann, K., Roberts, M.J., Sein, D.V., Roberts, C.D., Terray, L. and Vidale, P.L. (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. *Weather and Climate Dynamics*, 1, 277–292.

Schiemann, R., Demory, M.E., Shaffrey, L.C., Strachan, J., Vidale, P.L., Mizielinski, M.S., Roberts, M.J., Matsueda, M., Wehner, M.F. and Jung, T. (2017) The resolution sensitivity of Northern Hemisphere blocking in four 25-km atmospheric global circulation models. *Journal of Climate*, 30, 337–358.

Sillmann, J., Croci-Maspoli, M., Kallache, M. and Katz, R.W. (2011) Extreme cold winter temperatures in Europe under the influence of North Atlantic Atmospheric Blocking. *Journal of Climate*, 24, 5899–5913.

Steinfeld, D., Boettcher, M., Forbes, R. and Pfahl, S. (2020) The sensitivity of atmospheric blocking to changes in upstream latent heating–numerical experiments. *Weather and Climate Dynamics Discussions*, 1, 405–426.

Stockdale, T.N., Anderson, D.L., Balmaseda, M.A., Doblas-Reyes, F., Ferranti, L., Mogensen, K., Palmer, T.N., Molteni, F. and Vitart, F. (2011) ECMWF seasonal forecast system 3 and its prediction of sea-surface temperature. *Climate Dynamics*, 37, 455–471.

Tibaldi, S. and Molteni, F. (1990) On the operational predictability of blocking. *Tellus*, 42A, 343–365.

Tibaldi, S. and Molteni, F. (2018). Atmospheric blocking in observation and models, *Oxford Research Encyclopedia of Climate Science*. Oxford, UK: Oxford University Press.

Tyrlis, E. and Hoskins, B. (2008) The morphology of Northern Hemisphere blocking. *Journal of Atmospheric Science*, 65, 1653–1665.

Wang, L., Chen, W., Zhou, W., Chan, J.C., Barriopedro, D. and Huang, R. (2010) Effect of the climate shift around mid 1970s on the relationship between wintertime Ural blocking circulation and East Asian climate. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, 30, 153–158.

Weisheimer, A., Corti, S., Palmer, T. and Vitart, F. (2014) Addressing model error through atmospheric stochastic physical parametrizations: impact on the coupled ECMWF seasonal forecasting system. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 372. 20130290

Weisheimer, A., Decremer, D., MacLeod, D., O'Reilly, C., Stockdale, T.N., Johnson, S. and Palmer, T.N. (2019) How confident are predictability estimates of the winter North Atlantic Oscillation?. *Quarterly Journal of the Royal Meteorological Society*, 145, 140–159.

Wolff, J.-O., Maier-Reimer, E. and Legutke, S. (1997). The hamburg ocean primitive equation model. World data center for climate (WDCC) at DKRZ. https://doi.org/10.2312/WDCC/DKRZ_Report_No13

Woollings, T., Barriopedro, D., Methven, J., Son, S.W., Martius, O., Harvey, B., Sillmann, J., Lupo, A.R. and Seneviratne, S. (2018) Blocking and its response to climate change. *Current Climate Change Reports*, 4, 287–300.

Woollings, T., Hoskins, B., Blackburn, M. and Berrisford, P. (2008) A new Rossby-wave breaking interpretation of the North Atlantic oscillation. *Journal of Atmospheric Science*, 65, 609–326.

How to cite this article: Davini P, Weisheimer A, Balmaseda M, *et al*. The representation of winter Northern Hemisphere atmospheric blocking in ECMWF seasonal prediction systems. *Q.J.R. Meteorol. Soc*. 2021;1–20. https://doi.org/10.1002/qj.3974