Improvement in the decadal prediction skill of the northern hemisphere extra-tropical winter circulation through increased model resolution

Mareike Schuster[1], Jens Grieger[1], Andy Richling[1], Thomas Schartner[2], Sebastian Illing[1], Christopher Kadow[1], Wolfgang A. Müller[4], Holger Pohlmann[3,4], Stephan Pfahl[1], and Uwe Ulbrich[1]

[1]Freie Universität Berlin, Institut für Meteorologie, Carl-Heinrich-Becker Weg 6-10, 12165 Berlin
[2]Deutscher Wetterdienst, Güterfelder Damm 87-91, 14532 Stahnsdorf
[3]Deutscher Wetterdienst, Bernhard-Nocht-Straße 76, 20359 Hamburg
[4]Max-Planck-Institut für Meteorologie, Bundesstraße 53, 20146 Hamburg

Correspondence: Mareike Schuster (mareike.schuster@met.fu-berlin.de)

Abstract. In this study the latest version of the MiKlip decadal hindcast system is analyzed and the effect of different horizontal and vertical resolutions on the prediction skill of the northern hemisphere extra-tropical atmospheric circulation is assessed. Four metrics - the stormtrack, blocking frequencies, cyclone frequencies and windstorm frequencies - are analyzed with respect to the anomaly correlation of their winter averages. The model bias and hindcast skill are evaluated in both, a lower resolution version (LR, atm: T63L47, ocean: 1.5° L40) and a higher resolution version (HR, atm: T127L95, ocean: 0.4° L40) of the MPI-ESM system, for the lead years 2-5 using initializations between 1978 and 2012. While the LR version shows common shortcomings of lower resolution climate models, e.g. a too zonal stormtrack and a negative bias of blocking frequencies over the eastern North Atlantic and Europe, the HR version works against these biases. As a result, a functional chain of significantly improved decadal prediction skill between all four metrics is found with the increase of the spatial resolution. While the stormtrack, is significantly improved primarily over the main source region of synoptic activity - the North Atlantic Current, the other extra-tropical measures experience a significant improvement downstream thereof. Thus, the skill of the cyclone frequencies is significantly improved over the central North Atlantic and Northern Europe, the skill of the blocking frequencies is significantly improved over the Mediterranean, Scandinavia and Eastern Europe and the skill of the windstorms is significantly improved over Newfoundland and Central Europe. Not only is the skill improved with the increase in resolution, but the HR system itself exhibits significant skill over large areas of the North Atlantic and European sector for all four circulation metrics. These results are particularly promising regarding the high socio-economic impact of European winter windstorms and blocking situations.

1   Introduction

The extra-tropical circulation plays an important role for extremes in weather and climate, as its variability determines i.a. the frequency of extreme cyclones, embedded storm fields and phases of blocked flow. The consequences include extremes in temperature, precipitation/drought and wind speed, often accompanied by immense damage and harm. Therefore, the societal

demand for reliable near term climate predictions of such features - also to support political, economical and administrative decision making - is perpetually growing.

Decadal Climate prediction itself is an active research field in climate sciences. Different research groups around the globe aim at the development of skillful prediction systems (Boer et al., 2016). Retrospective forecasts, termed hindcasts, are used to assess the ability of the model systems to predict climate variability on inter-annual to decadal time scales. Initialized from observation-based data and run for a period of 10-30 years, decadal climate predictions combine forecast elements from weather and seasonal forecast sectors (initial conditions) as well as from long-term climate projections (boundary conditions). To date, different designs of decadal prediction systems are prevalent. They either consist of a multi-member single-model suite, such as the UK MetOffice's Decadal Prediction System 'DePreSys' (Smith et al., 2007) and the German 'MiKlip system' (Marotzke et al., 2016) - or they are based upon a multi-model suite, as used e.g. within the 5th Coupled Model Intercomparison Project (CMIP5; Taylor et al. 2012). Several multi-model studies, using the CMIP5 decadal prediction suite, come to the conclusion that there exists significant prediction skill on decadal time scales (e.g. Kim et al., 2012; Doblas-Reyes et al., 2013; Meehl et al., 2014). Results from these studies have also been included in the 5th assessment report (AR) of the Intergovernmental Panel on Climate Change (IPCC; Kirtman et al. 2013). Currently, in preparation for the 6th AR of the IPCC and CMIP6, improved decadal prediction systems are developed (Eyring et al., 2016; Boer et al., 2016; Kushnir et al., 2019).

With respect to atmospheric quantities, various versions of the 'MiKlip system', based on the Max-Planck Institute Earth System model (MPI-ESM), show decadal prediction skill mainly in terms of global and regional temperature indices Müller et al. (2012); Pohlmann et al. (2013); Kröger et al. (2017). The skill for precipitation, due to its complex and partly small-scale nature, is however regionally confined and prevalently limited to lead year 1 (Kadow et al., 2016). One result of MiKlip ('Mittelfristige Klimaprognosen', German for medium-term climate predictions) of practical relevance is the skillful prediction of $10\,\mathrm{m}$ wind speed and wind energy over Central Europe, that exists irrespective of the ocean initialization technique (Moemken et al., 2016) and for wind speeds of different quantiles upward from 75% (Haas et al., 2016).

Kruschke et al. (2014) and Kruschke et al. (2016) analyzed the forecast skill of Northern Hemisphere cyclone and windstorm frequencies in the 'MiKlip system', respectively. They found probabilistic decadal forecast skill for cyclone frequencies in some areas over the Northern Hemisphere ocean basins, mainly the European North Sea and the central Pacific, with the sub-sample of strong cyclones exhibiting generally higher skill than the complete sample of all cyclones (Kruschke et al., 2014). Using a parametric bias adjustment approach, Kruschke et al. (2016) found windstorm frequencies for winters 2–5 and winters 2–9 to be skillful over large parts of the Northern Hemisphere when compared against climatological forecasts. Another study using the MPI-ESM decadal prediction system demonstrated that the decadal prediction skill for surface temperature and cyclone frequencies can be significantly improved by replacing each ensemble member's ocean state with the ensemble mean ocean state at regular intervals during the forecast period (Kadow et al., 2017). More studies using the MPI-ESM hindcasts are collected in a special issue of the Meteorologische Zeitschrift about the validation of the MiKlip system in its first phase (Kaspar et al., 2016).

With respect to the North Atlantic and European domain, in many climate models of lower resolution a cold sea surface temperature (SST) bias south of Greenland is present, due to a displacement of the North Atlantic Current or a too weak over-turning circulation (Park et al., 2016; Scaife et al., 2011; Wang et al., 2014). This common bias in the North Atlantic Current, is associated with a too zonal stormtrack, stronger geopotential height gradients in the mid-latitudes, increased westerlies and

5 reduced blocking frequencies over Europe. It has been found in many studies, that the atmospheric dynamics benefit not only from a coupling of the atmosphere and ocean but also from an increased model resolution (e.g. Shaffrey et al., 2009; Jung et al., 2012; Dawson et al., 2013). Scaife et al. (2011), for example, demonstrated that the increase in resolution in both the atmospheric and oceanic model components results in a functional chain of improvements, as they found a reduced SST bias in the higher resolution model which in turn lead to a better representation of westerly winds and blocking frequencies.

10 With an atmospheric resolution of T63L47 (about 1.875° horizontal grid spacing) and an oceanic resolution of 1.5°L40 the MPI-ESM-LR decadal prediction system, applied in the first phase of MiKlip, has a rather moderate spatial resolution. Mean-while, studies using higher resolution forecast systems are available, for instance Monerie et al. (2017) using 0.5° grid spacing in the atmosphere and Robson et al. (2018) using ∼0.9°. However, systematic analyses of the actual effect of the increase in resolution on the hindcast performance on the decadal scale are rare. Pohlmann et al. (2013) found for the hindcasts of mixed

15 resolution (MPI-ESM-MR) that an increase in vertical (atmosphere: T63L95) and horizontal resolution (ocean: 0.4°L40) com-pared to MPI-ESM-LR improves the tropical Pacific surface temperature predictions in the lead years 2-5 and leads to a good representation of the quasi-biennial oscillation (QBO), which remains in alignment with observations well beyond the first 12 months after initialization. Apart from that, the mixed resolution shows only modest benefit for the hindcast skill (Marotzke et al., 2016).

In this study, for the first time an analysis of the direct impact of the model resolution on the skill of decadal climate predic-tions of dynamical variables under otherwise unchanged model settings (parametrization and initialization) is performed. We evaluate the MiKlip hindcasts performed with the latest version of the Max-Planck Institute Earth System model with higher resolution (MPI-ESM-HR, Müller et al. 2018), which will contribute to CMIP6, and compare its decadal forecast skill to that

25 of a previous lower-resolution version (MPI-ESM-LR). While many studies analyzing the skill of decadal forecast systems tend to focus on basic atmospheric variables such as the surface temperature and precipitation (e.g. Smith et al., 2007; Keenly-side et al., 2008; Goddard et al., 2013; Kadow et al., 2016; Monerie et al., 2018; Xin et al., 2018), we emphasize the role of dynamical processes and therefore analyze a set of variables representing the extra-tropical winter dynamics - the stormtrack, blocking, cyclones and windstorms.

30

We introduce the MPI-ESM prediction system in Sect. 2.1, as well as the skill measure used to assess the hindcast quality in this study. In Sect. 2.2 we describe the different circulation quantities in detail and present their climatology in the ERA-Interim reanalysis. The model climatology and biases are discussed in Sect. 3.1. Finally, the prediction skill of the winter circulation is evaluated in Sect. 3.2 with a focus on the North Atlantic and European region, before we conclude our results in Sect. 4.

2 Data and methodology

The extra-tropical circulation in the Northern Hemisphere is most active during the winter season, with a stronger jet stream in the upper-troposphere and numerous strong cyclones developing in the mid-latitude baroclinic areas, favored by strong horizontal temperature contrasts resulting from relatively warm ocean currents near the surface and cold polar air masses.

5 Storms that strike the European continent at this time of the year are often powerful and damaging. We will therefore focus on the winter circulation and evaluate averages of the stormtrack and blocking, cyclone and windstorm frequencies from October through March. The stormtrack describes the variability of baroclinic waves on synoptic time-scales in the extra-tropics. These baroclinic waves are a combination of two contributing components, i.e. anti-cyclonic and cyclonic anomalies, which we will analyze in terms of blocking frequencies on the one hand and extra-tropical cyclone and windstorm frequencies on the other

10 hand.

To assess the model bias and to compute the prediction skill of the different variables in the decadal hindcasts, a reference, i.e. an observational data set is needed. However, there exists no gridded observational data set of the dynamical variables that we aim at. Instead we make use of a reanalysis product and derive the circulation quantities for the winters 1979/80 to 2016/17 from the ERA-Interim reanalysis (Dee et al., 2011), created by the European Centre for Medium-Range Weather Forecasts

15 (ECMWF), with a horizontal resolution of T255 ($\sim 0.75°$) on 60 levels and a top of the atmosphere at 0.1 hPa.

2.1 Forecast system and skill measures

The two decadal forecast systems that we compare are both based on the Earth System Model of the Max-Planck-Institute for Meteorology (MPI-ESM) version 1.2, which is a coupled atmosphere ocean model and consists of the atmospheric component ECHAM6.3 and the oceanic component MPI-OM1.6.2.

20 The lower resolution of MiKlip's pre-operational decadal prediction system (MPI-ESM-LR, termed LR hereafter) has an atmospheric horizontal resolution of T63 (1.875°) and 47 levels, with the top of the atmosphere at 0.01 hPa (Mauritsen et al., 2019). The ocean component is run with 1.5° L40. A general skill assessment of decadal predictions performed with the LR system can be found in Polkova et al. (2019). The higher resolution version (MPI-ESM-HR, termed HR hereafter) uses T127 (0.9375°) and 95 vertical levels for the atmosphere, and 0.4° L40 for the ocean (Müller et al., 2018). The HR version

25 therefore has a finer grid in both the atmosphere and the ocean components. For this analysis, both systems use the CMIP5 external forcing with respect to greenhouse gases and aerosols (for details see Giorgetta et al. 2013). Both systems are full-field initialized in the atmosphere, using ERA-40 (Uppala et al., 2005) and ERA-Interim (Dee et al., 2011); and anomaly-initialized in the ocean, using ORA-S4 (Balmaseda et al., 2013) and sea-ice concentration from the National Snow and Ice Data Center (NSIDC). The initialization procedure is identical to the one used for MiKlip's Baseline 1 system and is described in more

30 detail in Pohlmann et al. (2013). The LR system consists in total of 10 ensemble members, initialized annually between 1960 and 2016, with each initialization covering one decade. The integration period for each of the initializations spans 10 years. However, since the HR system - with an otherwise identical hindcast setup - consists of only 5 members, and to guarantee a fair comparison between the two forecast systems we only evaluate the first 5 members of LR as well.
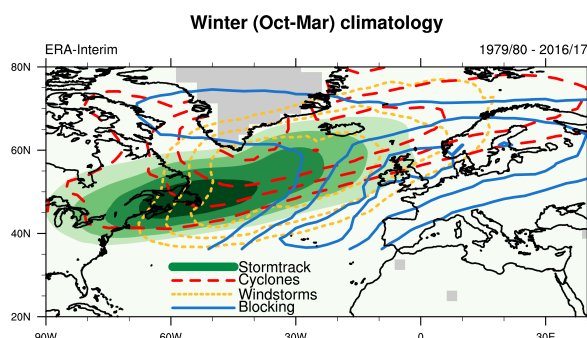
Figure 1. Climatology of the winter average (Oct-Mar) of different circulation quantities in the ERA-Interim reanalysis for the period 1979/80-2016/17. The stormtrack, i.e. the standard deviation of the 500 hPa geopotential height anomaly is shown in m (45-60 by 5). The fraction of blocked days is shown in % (4-8 by 2). The cyclone frequency (120-180 by 20) and windstorm frequency (25-30 by 2.5) are shown in number of tracks within a radius of 1000 km. Grey masked areas denote grid points with an orography larger than 1500 m, which have been omitted for cyclone identification.

To determine the skill of the two forecast systems, we analyze the winters 2-5 after initialization following the Decadal Climate Prediction Project (DCPP, Boer et al. 2016) protocol. We focus on the temporal variability by analyzing the anomaly correlation. Therefore, we employ the decadal climate prediction evaluation software of the MiKlip project (Illing et al., 2014). To match the period covered by the ERA-Interim reanalysis, we do not use the full set of initializations but instead use the

5   decadal hindcast experiments that are initialized between 1978 (winter 2: 1979/80, winter 5: 1982/83) and 2012 (winter 2: 2013/14, winter 5: 2016/17) in LR and HR. In total we therefore analyze 700 October-to-March winter seasons (5 members x 35 initializations x 4 lead years) per forecast system. The skill of each of the forecast systems (LR, HR) is first evaluated against the reanalysis data, i.e. the anomaly correlation between the respective hindcast and ERA-Interim is determined. Then, the two systems are compared against each other, i.e. the difference of the aforementioned correlations between the two forecast

10   systems is computed. To determine the significance of the correlation, the time series of reanalysis-hindcast pairs is resampled with replacement 1000 times (block bootstrap taking auto-correlation into account).

2.2   Circulation metrics

Stormtrack

The extra-tropical stormtrack is derived from the bandpass filtered variability of the geopotential height field at 500 hPa in

15   the window of 2.5 to 6 days - an Eulerian approach following Blackmon et al. (1976). Its long term winter average (October through March) is displayed in Fig. 1 for the North Atlantic and European region and the period 1979/80-2016/17 based on the ERA-Interim reanalysis. The North Atlantic stormtrack is visible in green shades, with its maximum of 60m located over the western North Atlantic and Newfoundland and a typical north-eastward tilt.

Blocking

The second synoptic scale feature that we analyze is atmospheric blocking. Here, a slightly modified version of the 2-dimensional blocking index of Scherrer et al. (2006), based on gradients in the daily 500 hPa geopotential height field, is used to identify instantaneously blocked grid points. In contrast to Scherrer et al. (2006), where a blocking area is defined in
5 between the blocking high and the associated low, here the position of detected blocked grid points is shifted north by 7.5° to correspond better with the anticyclonic part of a blocking situation. To account for large-scale and persistent blocking anticyclones between 35°N and 80°N, an adapted tracking algorithm for blocking regimes, similar to the approach by Barnes et al. (2012), is applied. With this tracking method, we only select contiguously blocked regions with a minimum zonal and meridional extension of ∼15° and an area of at least 1.5 x 10^6 km^2 lasting for a minimum of 4 days. A possible shifting,
10 merging and splitting of blocking areas in time is considered by adopting a blocking overlap area criterion of 750.000 km^2 between two consecutive days and a maximum distance between blocking centers of 1000 km. The climatology of the mean winter blocking frequency is displayed in blue isolines in Fig. 1. Its maximum of 8% blocked days stretches from the Azores to Scotland. A second region of increased blocking frequencies is found between Greenland and Iceland.

15 **Cyclones**

To identify and track extra-tropical cyclones we apply an objective Lagrangian feature tracking algorithm developed by Murray and Simmonds (1991) to the mean sea level pressure. Maxima of the Laplacian of the mean sea level pressure are identified and, if a minimum in the pressure field itself can(not) be detected in the vicinity, a closed (open) cyclone is identified. The system is then tracked in time, at 6-hourly time steps, using predicted locations for the successive time step, and probabilities
20 for the assignment of the systems in the consecutive time steps. The measure we ultimately use for our evaluation is the cyclone frequency, i.e. the number of cyclone tracks that pass within a radius of 1000 km of the respective grid point on a 2.5° x 2.5° grid. As the extrapolation of pressure to sea level can be erroneous over high terrain, cyclones are not identified at grid points where the orography is higher than 1500 m. In this study only cyclones are taken into account that are strong (Laplacian > 0.7 hPa (deg. lat.)$^{-2}$) and closed at least once during their lifetime and that last longer than a day. The winter average of the
25 cyclone frequency is displayed in Fig. 1 as red dashed contours. Its maximum is located at the southern tip of Greenland with 180 cyclones and a band of enhanced cyclone frequencies is located downstream of the stormtrack maximum with a similar southwest-northeast tilt.

Windstorms

30 Yet another objective Lagrangian tracking scheme is used to derive the frequency of extra-tropical windstorms (Leckebusch et al., 2008; Kruschke, 2014). This method is based on the exceedance of the local 98th percentile of the near-surface wind speed to define contiguous fields of strong wind. Percentiles are calculated for each hindcast and the reanalysis individually using 6 hourly data of the whole year between 1981 and 2010. For the hindcasts the percentiles of the uninitialized counterparts are used as done by Kruschke et al. (2016). Windstorms are identified if the area of wind exceedance above the percentile is
35 larger than 150.000 km^2 and if the feature is trackable for at least 18 hours. Tracking is done by means of a nearest neighbour

approach. Windstorm tracks are further used to calculate windstorm frequencies, which are computed identically to those of the cyclone frequencies. The yellow dotted contours in Fig. 1 represent the average winter windstorm frequency. It is nicely illustrated that the corresponding windstorm field is usually located to the south of the cyclone center, where the pressure gradients and thus geostrophic wind velocities are typically largest. Its maximum of 30 windstorms is located also downstream
5   of the stormtrack maximum, but slightly shifted southward compared to the cyclone frequencies.

The software that was used to compute all the extra-tropical circulation quantities, as well as the evaluation procedure were implemented as separate plug-ins into the MiKlip Central Evaluation System (https://www-miklip.dkrz.de) - based on the Free Evaluation System Framework (Freva, Kadow et al. in preparation) - by their developers and authors of this paper. The single plug-ins and their documentation can be found under https://www-miklip.dkrz.de/plugins - plus the respective suffix
10 /stormtrack/detail/ (Stormtrack); /blocking_2d/detail/ (Blocking); /zykpak/detail/ (Cyclones); /wtrack/detail/ (Windstorms) and /murcss/detail/ (skill analysis).

3   Results

3.1   Model bias

Before we analyze the prediction skill, we will first evaluate the ensemble mean climatology and the model bias, in order to
15 assess the model's capability to represent the four atmospheric circulation features. For this, we only take into account those seasons that will be used for the skill analysis, i.e. the winters 2-5 of each of the 35 initializations (1978-2012) and 2x 5 members (LR and HR). To compute the model bias, we consider the entire reanalysis data set, i.e. winters from 1979/80 to 2016/17.

In Fig. 2 the model bias, compared to ERA-Interim, for the stormtrack and blocking frequency is displayed in colored shades
20 and the respective model climatology is shown in grey contours, for both the LR and HR ensemble mean. The grey contour levels are the same as for the ERA-Interim climatology in Fig. 1. The LR system shows the typical North Atlantic stormtrack along 45°N, with a maximum over the western part of the basin, however rather zonally aligned (Fig. 2a). Since the observed stormtrack is tilted, from south-west to north-east (see Fig. 1), this results in a negative bias (-10m) at higher latitudes and a positive bias (+8m) at lower latitudes in the LR prediction system. This bias can partly be corrected with the increase in
25 the model resolution, as the HR system increases the stormtrack activity where there is a negative bias in LR and vice versa (Fig. 2c), however this effect is strongest at the northern side of the stormtrack, as also seen in Müller et al. (2018) (their Fig. 10). In HR the North Atlantic stormtrack is more tilted, and therefore closer to observations (Fig. 2b). Not only does it extend further north in the higher resolution system, but it also extends further downstream, towards Central and Eastern Europe, and therefore reduces the negative bias over the North Sea and Scandinavia that is present in LR. The bias in HR is reduced at both,
30 northern and southern, flanks of the Atlantic stormtrack, however it is still present (-7m and +7m).

The blocking frequency shows a strong negative bias (-3%) in the LR system, just north of its climatological maximum, i.e. over a band stretching from the central North Atlantic and Great Britain towards the Baltic Sea, and a positive bias (+1.5%)
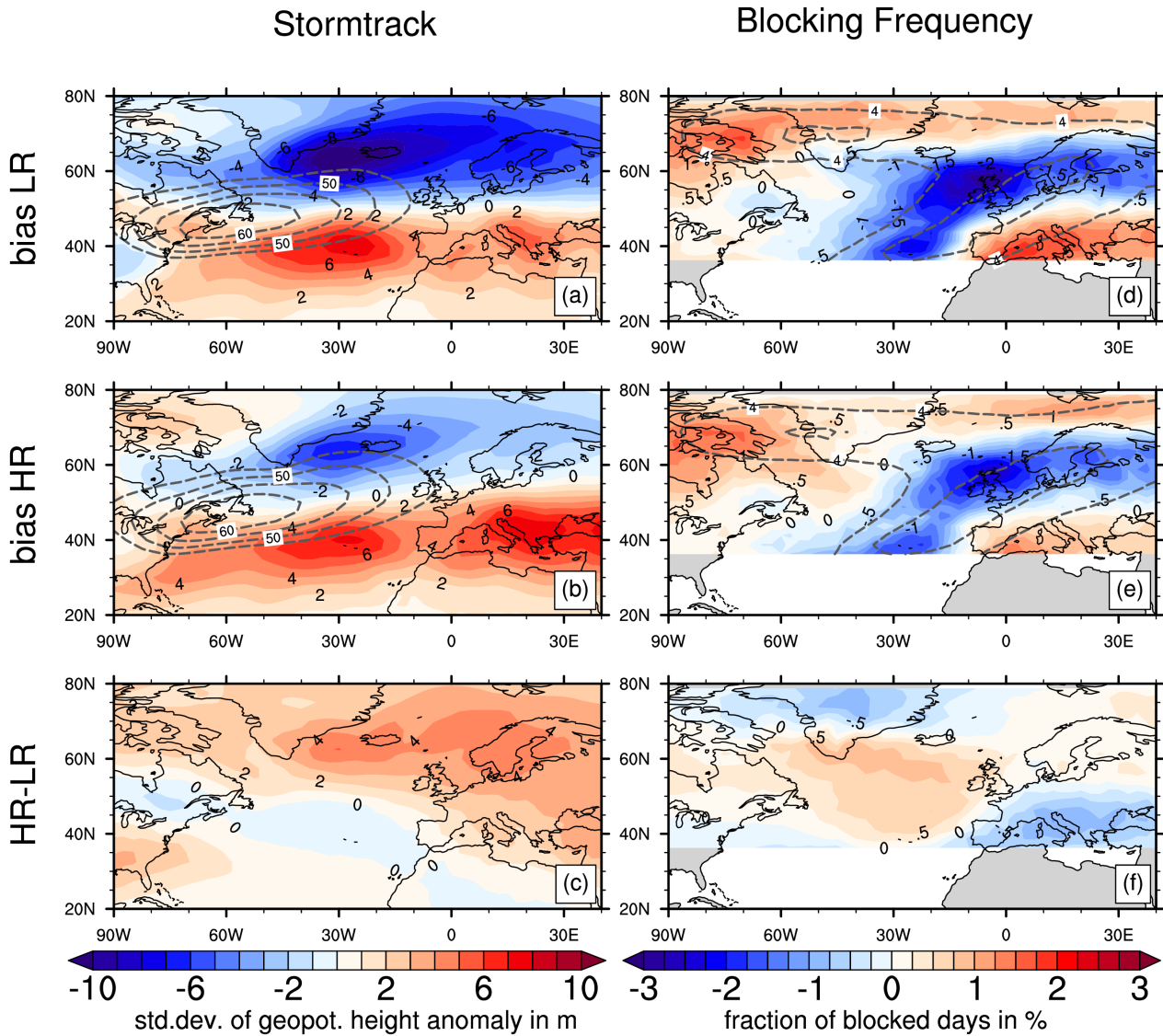
Figure 2. Ensemble mean model bias (shading) and model climatology (dashed contours) of the respective circulation quantity in LR (top row), HR (middle row) and the difference between HR and LR (bottom row). The circulation quantities displayed are the stormtrack (left) and the blocking frequency (right). Initializations from the period 1978-2012 are used for 5 members of each, LR and HR, and the ensemble mean is computed from lead-time averages over the hindcast winters 2-5 (Oct-Mar). In ERA-Interim the winters between 1979/80 and 2016/17 are used. The grey contours, i.e. ensemble mean climatology, have the same levels as in Fig. 1 - 45-60 by 5m for the stormtrack and 4-8 by 2 % for the blocking frequency.

over the Mediterranean (Fig. 2d). Fig. 2f nicely illustrates that again the HR prediction system counters these shortcomings of the LR system and reduces the bias in the right places, but the effect is rather marginal for this quantity. Though weaker, the
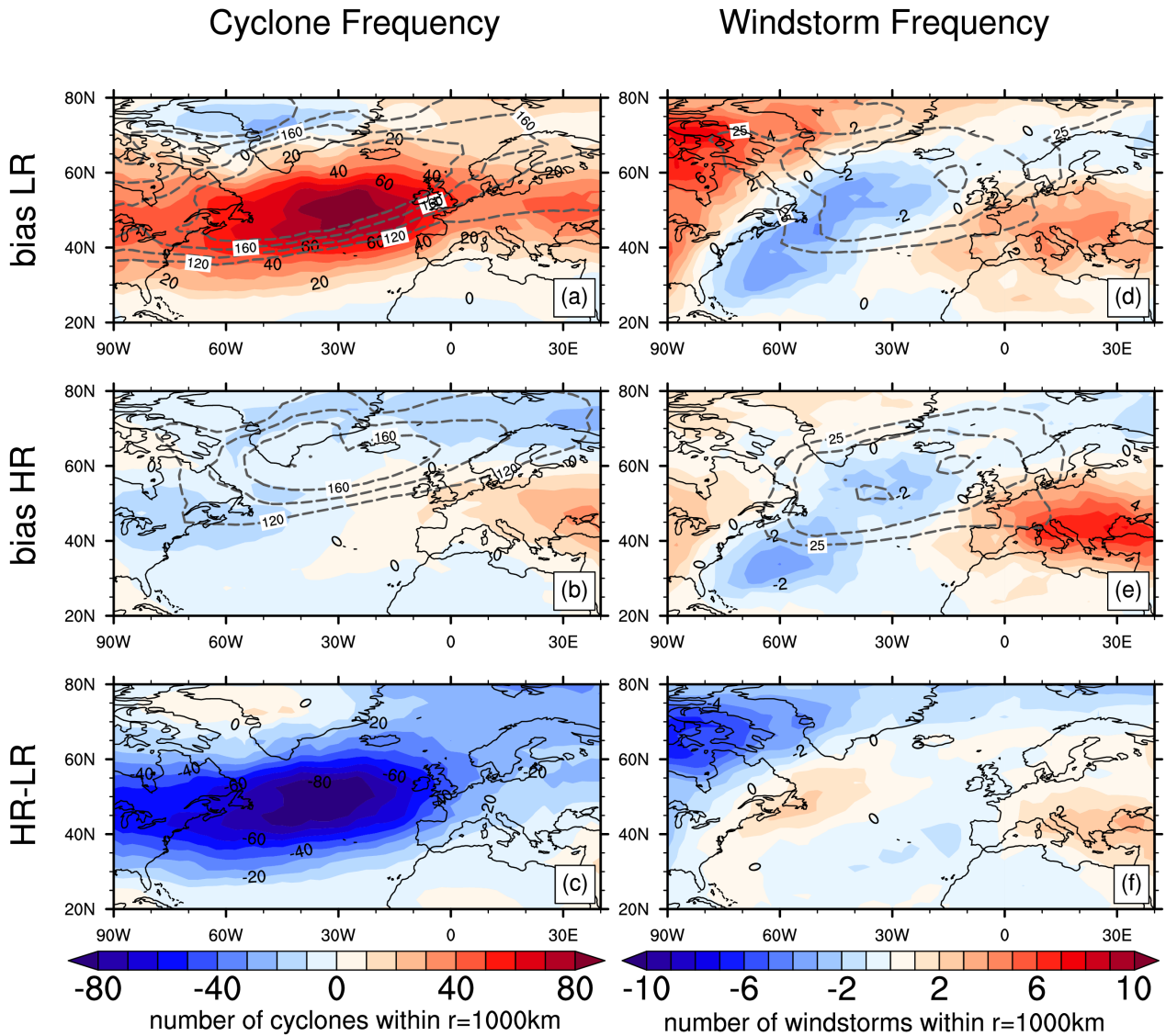
Figure 3. Same as Fig. 2 but for the cyclone frequency (left) and the windstorm frequency (right). The grey contours, i.e. ensemble mean climatology, have the same levels as in Fig. 1 - 120-180 by 20 cyclones for the cyclone frequency and 25-30 by 2.5 storms for the windstorm frequency.

bias of the blocking frequency in HR is still considerable (-2.5% and +1% respectively). These findings are in line with the analysis of blocking in Müller et al. (2018) (their Fig. 12).

The climatology of the cyclone frequency with its maximum at the southern tip of Greenland, seen in Fig. 1, is also visible in

5   LR (Fig. 3a). In contrast to the stormtrack and blocking frequency, the cyclone frequency in LR does not exhibit a clear south-

ward shift compared to the reanalysis. Instead, in the low resolution system there are overall far too many cyclones present between 30°-70°N, but especially over the central North Atlantic, where a positive bias of up to +80 cyclones is found. Most impressively amongst all variables, this bias of the cyclone frequency is radically reduced and almost completely absent in the HR system (Fig. 3b). The numbers are reduced to a bias of -10 cyclones over the western North Atlantic and +10 cyclones

5 over Europe. The increase in horizontal and vertical resolution evidently eliminates many cyclone tracks in the MPI-ESM (Fig. 3c) over the entire North Atlantic domain and adjacent continents. This results in cyclone climatologies very close to those in ERA-Interim (Fig. 3b).

The windstorm frequency shows a slightly different behavior. There are too few windstorms (-3) present over the western

10 and central North Atlantic along the North Atlantic current, and too many windstorms over the continents - +3 over Europe and +5 over the US and Canada (Fig. 3d). Given that there are too many cyclones in LR, the negative windstorm bias over the Atlantic might seem contradictory, as windstorms are a consequence of strong cyclones. However, it should be highlighted that the cyclone tracking algorithm also detects cyclones in their weak phase, as long as they become strong at least once during their lifetime. Therefore, the positive cyclone bias is likely influenced by many weaker and/or short lived systems that are not

15 strong for a long enough time to develop a windstorm. The negative windstorm bias is therefore not contradictory. In fact it is in line with the too zonally oriented stormtrack (Fig. 2a,b), also resulting in too many windstorms over Central Europe and the Mediterranean and too few storms over Northern Europe (Fig. 3d,e). The increase of the model resolution yields an increase of windstorm frequency over the North Atlantic current (Fig. 3e) and a remarkable reduction over the Hudson Bay, implying that local temperature gradients along land-sea borders and related surface fluxes, are slightly better represented in HR. The

20 bias over South East Europe, however, is amplified. This leaves the higher resolution system with biases of -2 along the North Atlantic current and the central North Atlantic, and +6 over South East Europe (Fig. 3f).

While the exact location and magnitude of the extra-tropical circulation features over the North Atlantic and European region exhibits deviations from the observation, overall the MPI-ESM is capable to represent those dynamical variables. Also

25 in Müller et al. (2018), it is noted that although bias reductions from LR to HR are modest for the multitude of variables they analyzed, the dynamics of the atmosphere still benefit from the increase in resolution and make this model eligible for prediction studies. We therefore proceed to analyze the deterministic decadal prediction skill.

3.2 Prediction skill

The anomaly correlation between the stormtrack in the LR hindcast and ERA-Interim for the winters 2-5 after initialization is

30 shown in Fig. 4a. Although both, significant positive and negative correlations are equally valuable from a mathematical point of view, a significant negative correlation, i.e. a consistently opposite prediction of the observed climate variability is inconsistent with the physically-based model setup. We thus consider only significantly positive correlations as model prediction skill. The LR system shows skill for the stormtrack over the central North Atlantic, as well as over Canada, the Baffin Bay and the Barents Sea. However, southwestward of the climatological stormtrack maximum, over the North Atlantic Current, where
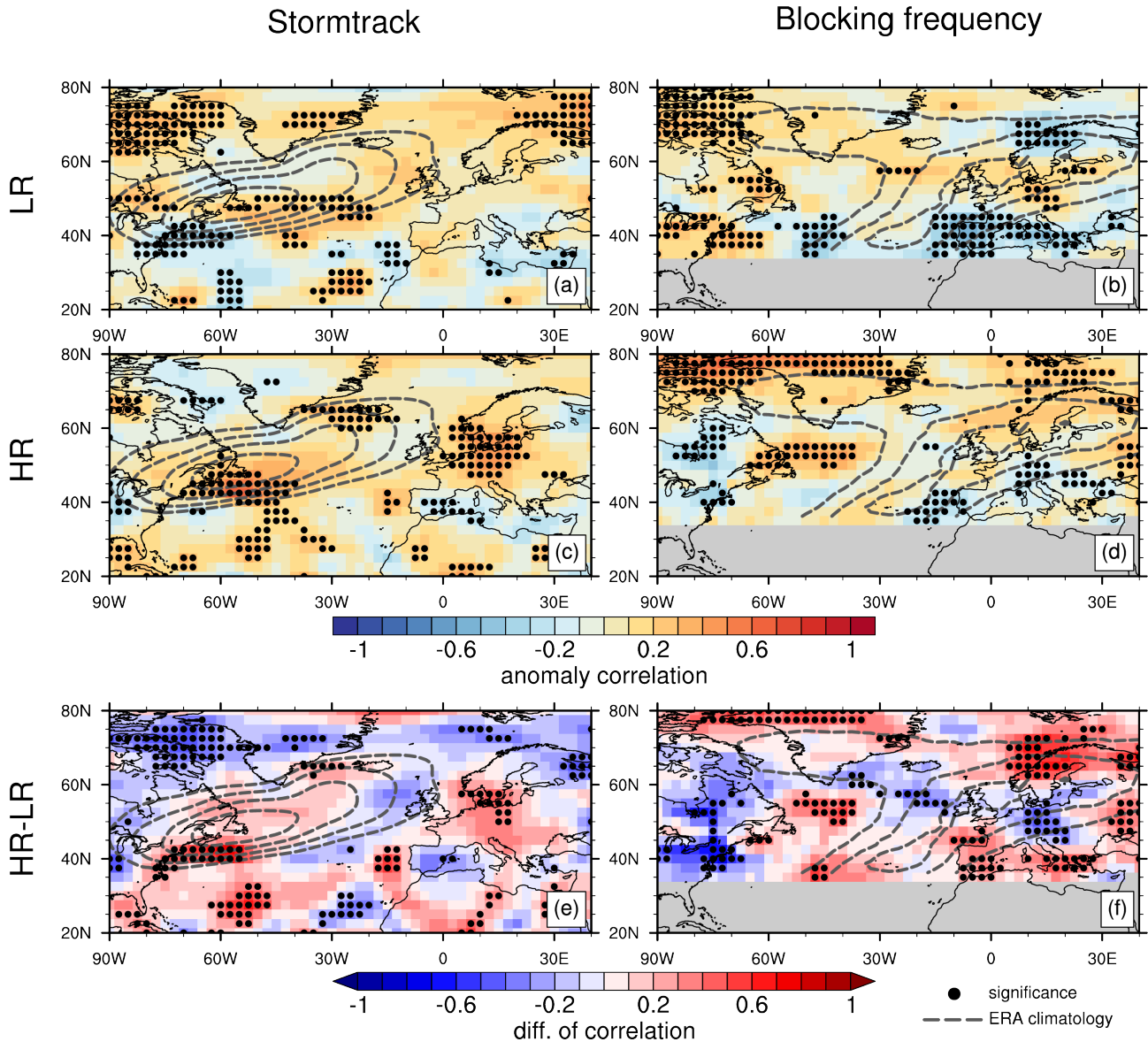
Figure 4. Anomaly correlation between the respective circulation quantity in ERA-Interim and LR (top row), between ERA-Interim and HR (middle row); and the difference between middle and top row (bottom row). The circulation quantities displayed are the stormtrack (left) and the blocking frequency (right). Initializations from the period 1978-2012 are used for both LR and HR and the correlation is computed for the winter (Oct-Mar) average of the hindcast winters 2-5. The dashed contours show the climatology of the circulation quantity in ERA-Interim (1979/80-2016/17) - as depicted in Fig. 1. The dots mark significance (1000 times resampling of reanalysis-hindcast time series).

the meridional gradient of the stormtrack climatology is strongest, there is significant negative correlation. This lack of skill in that area is overcome when the resolution of the dynamical model is increased. In the HR system (Fig. 4c) there is a large
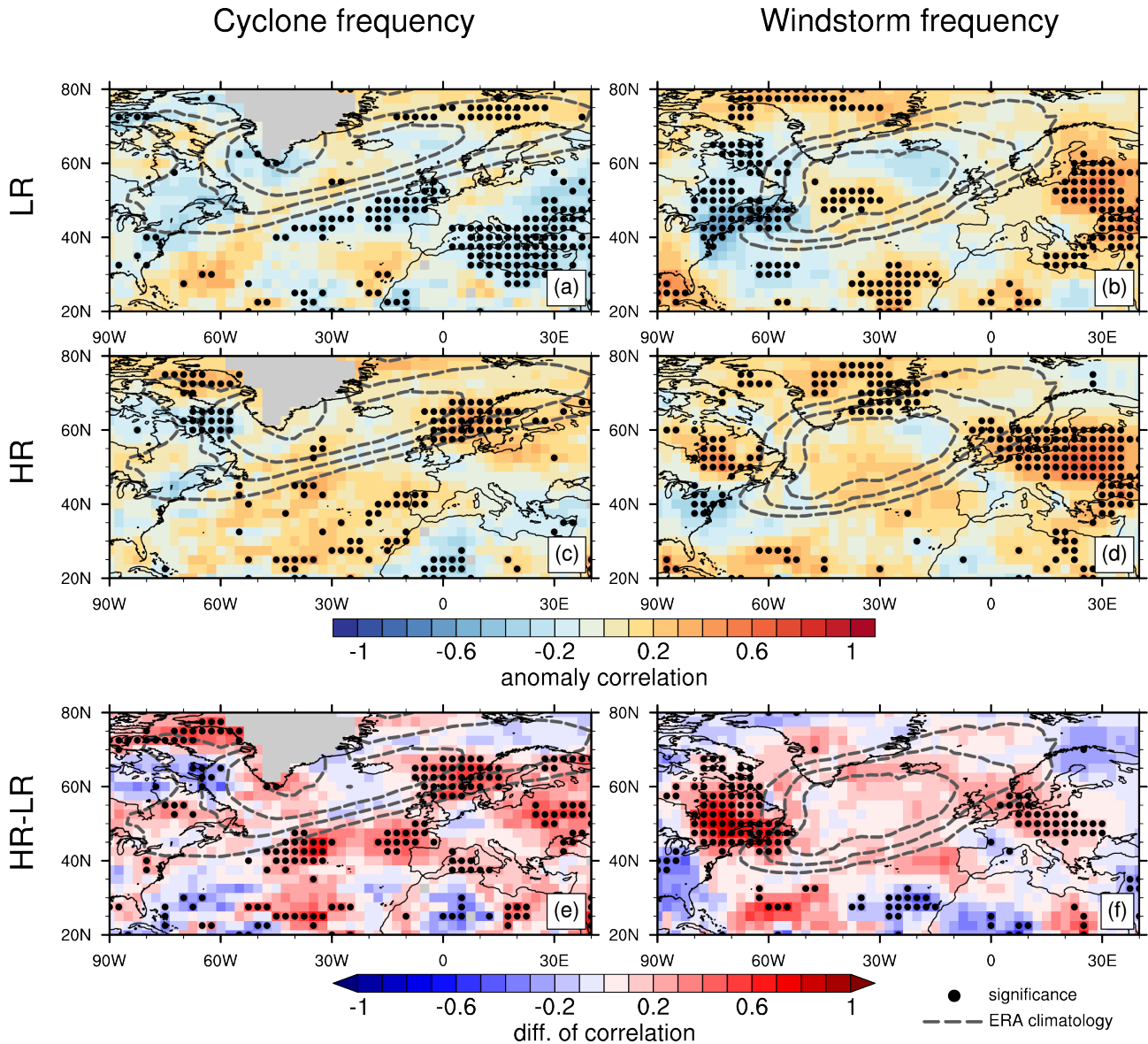
Figure 5. Same as Fig. 4 but for the cyclone frequency (left) and the windstorm frequency (right).

area of significant positive correlation over the North Atlantic Current and additionally over Iceland and Central Europe. The improvement from LR to HR, shown in Fig. 4e, is strongest over the North Atlantic Current and the tropical North Atlantic. Additionally, there are areas of significant skill improvement east of the Azores, west of Iceland and over the North and Baltic Seas. The improvement over the latter two regions is in line with the bias reductions found on the northern side and downstream

5   end of the stormtrack. Interestingly, the ensemble mean bias along the North Atlantic Current (Fig. 2c) is merely influenced by the increase in resolution, and yet appears to have a strong influence on inter-annual variability and prediction skill in that

region (Fig. 4e).

The anomaly correlation between LR and ERA-Interim for the winter blocking frequencies are illustrated in Fig. 4b. Similar to the stormtrack, there is skill over Canada. Although the correlation is positive in large areas over the North Atlantic and
5 Central Europe, it is only significant at a few of those grid points, e.g. south of Iceland and around the Baltic Sea. This changes in the HR system, larger areas around and downstream of Newfoundland and over Northern and Eastern Europe show skill for the winters 2-5 (Fig. 4d). Also, some areas of significantly negative correlation over the central North Atlantic around 40°N, the Mediterranean and Scandinavia, present in LR, are reduced in size or converted to positive correlation in HR. A significant improvement in correlation with respect to the blocking frequency is therefore found for several areas, such as east of New-
10 foundland and all around Europe, i.e. the Mediterranean, Eastern and Northern Europe (Fig. 4f) - except for Central Europe, which actually suffers from a significant decrease in correlation from LR to HR. The skill improvement around the Mediterranean and downstream of Newfoundland matches well with the bias reduction in those areas. The change in the anomaly correlation of the other regions cannot be directly explained by climatological changes.

15 For winter cyclone frequencies in the winters 2-5 in LR there is a small area of significant skill over the Arctic Ocean north of Scandinavia, however the rest of the domain is dominated by small or negative correlation (Fig. 5a). There are large regions with significantly negative correlation west of Great Britain and over the Mediterranean. Once again, with the increase in resolution in HR (Fig. 5c) this strongly improves and positive anomaly correlation bestrides the entire North Atlantic, and is skillful (significant) over a large contiguous area over the North Sea and Scandinavia and at scattered grid points over the
20 central North Atlantic. Thus, the skill for extra-tropical cyclone frequencies is significantly improved through the finer resolution in large areas over the central and eastern North Atlantic, the North Sea, Scandinavia and Eastern Europe (Fig. 5e). Those areas in which the skill is improved in HR coincide with the location of the maximum bias improvement and with the more accurately represented climatological cyclone frequencies on the downstream end along the European west coast. The analysis also reveals that not only the skill for the cyclone track frequency improves, but also for the cyclone genesis frequency, i.e.
25 the location where the cyclones form (not shown). There is significant skill improvement from LR to HR of the cyclogenesis frequency south of Greenland, over the entire eastern North Atlantic and over Northern Europe (not shown), indicating that not only the lifetime and pathway of existing maritime cyclones is improved but also the genesis of cyclones that form just off the European west coast and continental cyclones.

30 Prediction skill for the winter windstorms in the LR prediction system is present over the central North Atlantic, in the region of the maximum of the windstorm climatology, and over Eastern Europe (Fig. 5b). A large area of significant negative anomaly correlation is located around Newfoundland. It is remarkable that with the finer resolution the skill increases almost throughout the entire domain, i.e. it improves over the ocean but also and most strongly over continental areas. This effect is strongest and significant around Newfoundland and over Central and Eastern Europe (Fig. 5f). This matches the results for the
35 skill improvement of the cyclone frequencies in Fig. 5e, indicating that if the cyclone tracks are improved along the European

west coast, the downstream impact of the associated wind fields is also improved. Also, the skill improvement over Canada and Newfoundland is in line with the bias reduction of the ensemble mean windstorm climatology in this region. The HR system thus produces skillful windstorm predictions over large regions of the Northern Hemisphere, but most impressively over Central and Eastern Europe (Fig. 5d).

5   4   Discussion and Conclusion

This study has evaluated the changes in the deterministic decadal forecast skill of the atmospheric extra-tropical winter circulation in response to an increase in the horizontal and vertical resolution of the forecast system, under otherwise unchanged conditions (initialization technique, parametrization). Two hindcast sets initialized in the period 1978-2012, performed with the MiKlip pre-operational decadal prediction system, one of lower resolution (LR, atm. ~1.8°, ocean ~1.5°) and one of higher

10   resolution (HR, atm. ~0.9°, ocean ~0.4°), have been evaluated for the winters 2-5 after the initialization, using 5 members each. The forecast skill has been analyzed in terms of anomaly correlation for the stormtrack, blocking frequency, cyclone frequency and windstorm frequency. Additionally, the analysis of the ensemble mean model bias has provided additional insights into the modified atmospheric dynamics and into possible sources of improved forecast skill in the higher resolved system.

It has been demonstrated that with the increase in the horizontal and vertical resolution, the representation of the mid-latitude

15   dynamics in the MPI-ESM decadal prediction system is significantly improved. This applies to the ensemble mean climatology as well as the decadal prediction skill.

The stormtrack climatology in LR is represented too zonally and slightly shifted southward compared to the reanalysis, which is a well known weakness of many climate models and was discussed for CMIP5 models e.g. in Zappa et al. (2013). The increased model resolution counteracts this bias, leading to an extended and slightly more tilted stormtrack, but it cannot

20   fully compensate for it. Changes are strongest on the northward side of the stormtrack maximum. These results correspond to findings from Müller et al. (2018), who noted a reduced bias of the atmospheric jet stream position in the northern extra-tropics and a decrease of the storm track bias over the northern North Atlantic in HR.

The slightly better representation of the stormtrack over the eastern North Atlantic is in line with a slightly improved blocking frequency, this means less dominance of westerlies along the European west coast and instead more influence by blocking

25   situations. However, amongst all variables, the blocking frequency is affected least by the increased resolution. The lack of blocking over the eastern North Atlantic and the southward shift of the blocking climatology in LR is only marginally modified by the higher resolution - effects are strongest over the Mediterranean and downstream the southern tip of Greenland. This bias pattern is common in climate models, and has been reported e.g. by Scaife et al. (2010) and also for the MPI-ESM by Müller et al. (2018).

30   Cyclone frequencies are overall too high in the entire domain in the LR version - a feature that has been found in previous studies with the MPI-ESM (Kruschke et al., 2014) and its predecessor ECHAM5 (Bengtsson et al., 2006). This cyclone bias is strongest over the eastern North Atlantic just west of Great Britain and is consistent with a negative sea-level pressure bias, i.e. systematically too low pressure values between 2 and 5 hPa, in the same region found in LR (Müller et al., 2018). Earlier,

we argued that strong cyclones are usually accompanied by windstorms and that because of the different bias patterns of these two variables, the intense bias seen in the LR cyclones is likely due to weak and moderate systems. This is in line with the evaluation of Kruschke et al. (2014), who actually showed that in LR the strong positive cyclone bias can mainly be attributed to weak and moderate systems, by illustrating a remarkably reduced bias over the North Atlantic and Europe when only intense

5 cyclones, i.e. the strongest 25% in terms of the Laplacian of the sea-level pressure, are considered. In contrast to the minor influence on blocking, the increase in model resolution has a powerful effect on cyclone frequencies and successfully manages to decrease the strong Atlantic cyclone frequency bias to a minimum, leaving its climatology to resemble that of the reanalysis very closely. This is also in line with the reduced sea-level pressure bias in HR found by Müller et al. (2018).

In the low resolution system the windstorm frequency is too small over the Atlantic Ocean and too high over land, a phe-

10 nomenon that has been reported previously, e.g. by Kruschke et al. (2016). The higher resolution seems to improve especially the strong windstorm biases along North American coastlines, i.e. the Hudson Bay, Newfoundland and Nova Scotia. Although the tilt of the windstorm track density over the North Atlantic is mended, the model still generates too many windstorms over Europe and the positive bias there is generally amplified.

15 With respect to the decadal prediction skill this analysis showed that the increased resolution of the MPI-ESM decadal prediction system significantly improves the anomaly correlation in crucial regions of the North Atlantic and Europe on lead times of 2-5 winters for all four extra-tropical circulation measures. Furthermore, the areas with improved forecast skill are key regions for the genesis of synoptic weather systems in the North Atlantic and for their impact on Europe. This is particularly evident for the stormtrack, for which a strong and significant skill improvement is found along the North Atlantic Current and

20 over Central Europe. Given the important role of surface heat fluxes and local SST gradients for the dynamics of the stormtrack (Brayshaw et al., 2011), these are likely sources of improved atmospheric variability in the HR system. The stronger tilt and downstream extension of the stormtrack climatology in HR results in improved and significant decadal forecast skill east of Greenland and in Central Europe.

Significant improvement in the anomaly correlation of the blocking frequency is found downstream of where the stormtrack

25 skill is improved, and in large patches all around Europe, except for Central Europe. The strongest effect of the bias reduction, found over the Mediterranean, coincides with skill improvement in that area, however since the bias reduction is generally marginal, a direct effect on the decadal prediction skill is not necessarily given.

The strong misrepresentation of the cyclone climatology in LR results in no decadal forecast skill throughout the North Atlantic and European domain. Thus the striking climatological bias reduction in HR also impacts the prediction skill, which

30 is improved throughout the entire domain (significantly along the outskirts of the cyclone frequency maximum) and results in skillful cyclone frequency predictions over Northern Europe. The improved representation of cyclones in this region may also be beneficial for the prediction of blocking over Scandinavia (where the skill in HR is significantly improved), as cyclones can contribute to downstream blocking formation through eddy vorticity forcing (Shutts, 1983) and diabatic processes (Pfahl et al., 2015). A more accurate representation of smaller-scale diabatic processes may also be a reason for the increased forecast skill

of cyclones at the southern flank of the main stormtrack, over the subtropical North Atlantic and the Mediterranean, as moist processes are thought to be particularly important for such subtropical systems (e.g. Davis, 2010).

In line with these skill improvements in the cyclone frequency, the skill for windstorms improves as well significantly over North-East and Central Europe, i.e. south of the cyclones' signal. This matches with the general south-eastward displacement of the maximum wind speeds relative to the cyclone center (Leckebusch et al., 2008). This indicates that the variability of strong North Atlantic cyclones traveling towards Scandinavia and leading to windstorms in North and Central Europe is much better captured by the high resolution decadal prediction system. Interestingly, the skill is not negatively affected in South-East Europe, although the climatological windstorm bias is amplified in HR in that region. On the other hand, the strong bias reduction over North America and Canada appears to impact the prediction skill, thus a significant skill improvement for windstorm frequency is found over Canada and parts of the North Atlantic Current. Although, the forecast skill for 10m wind speeds and wind energy output only differs slightly between different ocean initializations (Moemken et al., 2016), this study reveals that the increased resolution has a large impact on the hindcast skill of synoptic scale features, such as cyclones and windstorms.

Overall we demonstrated that there is a chain of decadal prediction skill improvement amongst the extra-tropical circulation metrics with the increase in model resolution, similar to the interrelations laid out in Scaife et al. (2011). Also, our results are in agreement with previous studies by Prodhomme et al. (2016) and Befort et al. (2019) who found skill improvements in different seasonal prediction systems for blocking, windstorm and cyclone frequencies when the model resolution is increased. Our study showed that there is a significant improvement of the stormtrack skill along the North Atlantic Current followed by downstream improvement of the cyclone frequency skill over the central North Atlantic and finally improved skill of the cyclone, windstorm and blocking frequencies over the impact area Europe. Additionally, not only does the prediction skill improve with a finer grid (HR vs. LR), the HR system itself offers significant deterministic forecast skill in large regions over the North Atlantic and Europe (HR vs. ERA-Interim). An important question remains, as to which physical processes form the basis of this detected decadal prediction skill of the different circulation variables, and should be explored in future research.

Acknowledgements. The authors would like to acknowledge funding from CoreLogic SARL Paris and from the Federal Ministry of Education and Research in Germany (BMBF) through the research program MiKlip II (FKZ: 01LP1519A, 01LP1519B, 01LP1520A) and the CMIP6-DICAD project (FKZ: 01LP1605D). We acknowledge support by the Open Access Publication Fund of the Freie Universität Berlin.

References

Balmaseda, M. A., Mogensen, K., and Weaver, A. T.: Evaluation of the ECMWF ocean reanalysis system ORAS4, Quarterly Journal of the Royal Meteorological Society, 139, 1132–1161, https://doi.org/10.1002/qj.2063, 2013.

Barnes, E. A., Slingo, J., and Woollings, T.: A methodology for the comparison of blocking climatologies across indices, models and climate scenarios, CLIMATE DYNAMICS, 38, 2467–2481, https://doi.org/10.1007/s00382-011-1243-6, 2012.

Befort, D. J., Wild, S., Knight, J. R., Lockwood, J. F., Thornton, H. E., Hermanson, L., Bett, P. E., Weisheimer, A., and Leckebusch, G. C.: Seasonal forecast skill for extratropical cyclones and windstorms, Quarterly Journal of the Royal Meteorological Society, 145, 92–104, https://doi.org/10.1002/qj.3406, 2019.

Bengtsson, L., Hodges, K. I., and Roeckner, E.: Storm tracks and climate change, Journal of Climate, 19, 3518–3543, https://doi.org/10.1175/JCLI3815.1, 2006.

Blackmon, M. L., Wallace, J., Lau, N., and Mullen, S. L.: An observational study of the Northern Hemisphere wintertime circulation, Journal of Atmospheric Sciences, 34, 1040–1053, https://doi.org/10.1175/1520-0469(1977)034<1040:AOSOTN>2.0.CO;2, 1976.

Boer, G. J., Smith, D. M., Cassou, C., Doblas-Reyes, F., Danabasoglu, G., Kirtman, B., Kushnir, Y., Kimoto, M., Meehl, G. A., Msadek, R., Müller, W. A., Taylor, K. E., Zwiers, F., Rixen, M., Ruprich-Robert, Y., and Eade, R.: The Decadal Climate Prediction Project (DCPP) contribution to CMIP6, GEOSCIENTIFIC MODEL DEVELOPMENT, 9, 3751–3777, https://doi.org/10.5194/gmd-9-3751-2016, 2016.

Brayshaw, D. J., Hoskins, B., and Blackburn, M.: The Basic Ingredients of the North Atlantic Storm Track. Part II: Sea Surface Temperatures, Journal of the Atmospheric Sciences, 68, 1784–1805, https://doi.org/10.1175/2011JAS3674.1, https://doi.org/10.1175/2011JAS3674.1, 2011.

Davis, C. A.: Simulations of Subtropical Cyclones in a Baroclinic Channel Model, Journal of the Atmospheric Sciences, 67, 2871–2892, https://doi.org/10.1175/2010JAS3411.1, 2010.

Dawson, A., Matthews, A. J., Stevens, D. P., Roberts, M. J., and Vidale, P. L.: Importance of oceanic resolution and mean state on the extra-tropical response to El Niño in a matrix of coupled models, Climate Dynamics, 41, 1439–1452, https://doi.org/10.1007/s00382-012-1518-6, https://doi.org/10.1007/s00382-012-1518-6, 2013.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., 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.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Quarterly Journal of the Royal Meteorological Society, 137, 553–597, https://doi.org/10.1002/qj.828, 2011.

Doblas-Reyes, F. J., Andreu-Burillo, I., Chikamoto, Y., Garcia-Serrano, J., Guemas, V., Kimoto, M., Mochizuki, T., Rodrigues, L. R. L., and van Oldenborgh, G. J.: Initialized near-term regional climate change prediction, Nature Communications, 4, 1715, https://doi.org/10.1038/ncomms2704, 2013.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, Geoscientific Model Development, 9, 1937–1958, https://doi.org/10.5194/gmd-9-1937-2016, 2016.

Giorgetta, M. A., Jungclaus, J., Reick, C. H., Legutke, S., Bader, J., Boettinger, M., Brovkin, V., Crueger, T., Esch, M., Fieg, K., Glushak, K., Gayler, V., Haak, H., Hollweg, H.-D., Ilyina, T., Kinne, S., Kornblueh, L., Matei, D., Mauritsen, T., Mikolajewicz, U., Müller, W.,

Notz, D., Pithan, F., Raddatz, T., Rast, S., Redler, R., Roeckner, E., Schmidt, H., Schnur, R., Segschneider, J., Six, K. D., Stockhause, M., Timmreck, C., Wegner, J., Widmann, H., Wieners, K.-H., Claussen, M., Marotzke, J., and Stevens, B.: Climate and carbon cycle changes from 1850 to 2100 in MPI-ESM simulations for the Coupled Model Intercomparison Project phase 5, Journal of Advances In Modeling Earth Systems, 5, 572–597, https://doi.org/10.1002/jame.20038, 2013.

5 Goddard, L., Kumar, A., Solomon, A., Smith, D., Boer, G., Gonzalez, P., Kharin, V., Merryfield, W., Deser, C., Mason, S. J., Kirtman, B. P., Msadek, R., Sutton, R., Hawkins, E., Fricker, T., Hegerl, G., Ferro, C. A. T., Stephenson, D. B., Meehl, G. A., Stockdale, T., Burgman, R., Greene, A. M., Kushnir, Y., Newman, M., Carton, J., Fukumori, I., and Delworth, T.: A verification framework for interannual-to-decadal predictions experiments, Climate Dynamics, 40, 245–272, https://doi.org/10.1007/s00382-012-1481-2, 2013.

Haas, R., Reyers, M., and Pinto, J. G.: Decadal predictability of regional-scale peak winds over Europe using the Earth System Model
10 of the Max-Planck-Institute for Meteorology, Meteorologische Zeitschrift, 25, 739–752, https://doi.org/10.1127/metz/2015/0583, http://dx.doi.org/10.1127/metz/2015/0583, 2016.

Illing, S., Kadow, C., Oliver, K., and Cubasch, U.: MurCSS: A Tool for Standardized Evaluation of Decadal Hindcast Systems, Journal of Open Research Software, 2, 2014.

Jung, T., Miller, M. J., Palmer, T. N., Towers, P., Wedi, N., Achuthavarier, D., Adams, J. M., Altshuler, E. L., Cash, B. A., Kinter, J. L., Marx,
15 L., Stan, C., and Hodges, K. I.: 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, https://doi.org/10.1175/JCLI-D-11-00265.1, https://doi.org/10.1175/JCLI-D-11-00265.1, 2012.

Kadow, C., Illing, S., Kunst, O., Schartner, T., Grieger, J., Schuster, M., Richling, A., Kirchner, I., Rust, H., Cubasch, U., and Ulbrich, U.: Freva - Free Evaluation System Framework for Earth System Modeling, Geoscientific Model Development, in preparation.

20 Kadow, C., Illing, S., Kunst, O., Rust, H. W., Pohlmann, H., Müller, W. A., and Cubasch, U.: Evaluation of forecasts by accuracy and spread in the MiKlip decadal climate prediction system, METEOROLOGISCHE ZEITSCHRIFT, 25, 631–643, https://doi.org/10.1127/metz/2015/0639, 2016.

Kadow, C., Illing, S., Kroener, I., Ulbrich, U., and Cubasch, U.: Decadal climate predictions improved by ocean ensemble dispersion filtering, JOURNAL OF ADVANCES IN MODELING EARTH SYSTEMS, 9, 1138–1149, https://doi.org/10.1002/2016MS000787, 2017.

25 Kaspar, F., Rust, H. W., Ulbrich, U., and Becker, P.: Verification and process oriented validation of the MiKlip decadal prediction system, Meteorologische Zeitschrift, 25, 629–630, https://doi.org/10.1127/metz/2016/0831, http://dx.doi.org/10.1127/metz/2016/0831, 2016.

Keenlyside, N. S., Latif, M., Jungclaus, J., Kornblueh, L., and Roeckner, E.: Advancing decadal-scale climate prediction in the North Atlantic sector, NATURE, 453, 84–88, https://doi.org/10.1038/nature06921, 2008.

Kim, H. M., Webster, P. J., and Curry, J. A.: Evaluation of short-term climate change prediction in multi-model CMIP5 decadal hindcasts,
30 Geophysical Research Letters, 39, L10 701, https://doi.org/10.1029/2012GL051644, 2012.

Kirtman, B., Power, S., Adedoyin, J., Boer, G., Bojariu, R., Camilloni, I., Doblas-Reyes, F., Fiore, A., Kimoto, M., Meehl, G., Prather, M., Sarr, A., Schär, C., Sutton, R., van Oldenborgh, G., Vecchi, G., and Wang, H.: Near-term Climate Change: Projections and Predictability, book section 11, p. 953–1028, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, https://doi.org/10.1017/CBO9781107415324.023, www.climatechange2013.org, 2013.

35 Kröger, J., Pohlmann, H., Sienz, F., Marotzke, J., Baehr, J., Köhl, A., Modali, K., Polkova, I., Stammer, D., Vamborg, F. S. E., and Müller, W. A.: Full-field initialized decadal predictions with the MPI earth system model: an initial shock in the North Atlantic, Climate Dynamics, https://doi.org/10.1007/s00382-017-4030-1, 2017.

Kruschke, T.: Winter wind storms: Identication, verication of decadal predictions, and regionalization, Ph.D. thesis, Freie Universität Berlin, 2014.

Kruschke, T., Rust, H. W., Kadow, C., Leckebusch, G. C., and Ulbrich, U.: Evaluating decadal predictions of northern hemispheric cyclone frequencies, Tellus Series A-dynamic Meteorology and Oceanography, 66, 22 830, https://doi.org/10.3402/tellusa.v66.22830, 2014.

5 Kruschke, T., Rust, H. W., Kadow, C., Müller, W. A., Pohlmann, H., Leckebusch, G. C., and Ulbrich, U.: Probabilistic evaluation of decadal prediction skill regarding Northern Hemisphere winter storms, METEOROLOGISCHE ZEITSCHRIFT, 25, 721–738, https://doi.org/10.1127/metz/2015/0641, 2016.

Kushnir, Y., Scaife, A. A., Arritt, R., Balsamo, G., Boer, G., Doblas-Reyes, F., Hawkins, E., Kimoto, M., Kolli, R. K., Kumar, A., Matei, D., Matthes, K., Müller, W. A., O'Kane, T., Perlwitz, J., Power, S., Raphael, M., Shimpo, A., Smith, D., Tuma, M., and Wu, B.: Towards
10 operational predictions of the near-term climate, Nature Climate Change, 9, 94–101, https://doi.org/10.1038/s41558-018-0359-7, 2019.

Leckebusch, G. C., Renggli, D., and Ulbrich, U.: Development and application of an objective storm severity measure for the Northeast Atlantic region, Meteorologische Zeitschrift, 17, 575–587, https://doi.org/10.1127/0941-2948/2008/0323, http://dx.doi.org/1fv0.1127/0941-2948/2008/0323, 2008.

Marotzke, J., Müller, W. A., Vamborg, F. S. E., Becker, P., Cubasch, U., Feldmann, H., Kaspar, F., Kottmeier, C., Marini, C., Polkova, I.,
15 Proemmel, K., Rust, H. W., Stammer, D., Ulbrich, U., Kadow, C., Koehl, A., Kroeger, J., Kruschke, T., Pinto, J. G., Pohlmann, H., Reyers, M., Schroeder, M., Sienz, F., Timmreck, C., and Ziese, M.: MIKLIP A NATIONAL RESEARCH PROJECT ON DECADAL CLIMATE PREDICTION, BULLETIN OF THE AMERICAN METEOROLOGICAL SOCIETY, 97, 2379–2394, https://doi.org/10.1175/BAMS-D-15-00184.1, 2016.

Mauritsen, T., Bader, J., Becker, T., Behrens, J., Bittner, M., Brokopf, R., Brovkin, V., Claussen, M., Crueger, T., Esch, M., Fast, I., Fiedler,
20 S., Fläschner, D., Gayler, V., Giorgetta, M., Goll, D. S., Haak, H., Hagemann, S., Hedemann, C., Hohenegger, C., Ilyina, T., Jahns, T., Jimenéz-de-la Cuesta, D., Jungclaus, J., Kleinen, T., Kloster, S., Kracher, D., Kinne, S., Kleberg, D., Lasslop, G., Kornblueh, L., Marotzke, J., Matei, D., Meraner, K., Mikolajewicz, U., Modali, K., Möbis, B., Müller, W. A., Nabel, J. E. M. S., Nam, C. C. W., Notz, D., Nyawira, S.-S., Paulsen, H., Peters, K., Pincus, R., Pohlmann, H., Pongratz, J., Popp, M., Raddatz, T. J., Rast, S., Redler, R., Reick, C. H., Rohrschneider, T., Schemann, V., Schmidt, H., Schnur, R., Schulzweida, U., Six, K. D., Stein, L., Stemmler, I., Stevens, B., von
25 Storch, J.-S., Tian, F., Voigt, A., Vrese, P., Wieners, K.-H., Wilkenskjeld, S., Winkler, A., and Roeckner, E.: Developments in the MPI-M Earth System Model version 1.2 (MPI-ESM1.2) and Its Response to Increasing CO2, Journal of Advances in Modeling Earth Systems, 0, https://doi.org/10.1029/2018MS001400, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018MS001400, 2019.

Meehl, G. A., Goddard, L., Boer, G., Burgman, R., Branstator, G., Cassou, C., Corti, S., Danabasoglu, G., Doblas-Reyes, F., Hawkins, E., Karspeck, A., Kimoto, M., Kumar, A., Matei, D., Mignot, J., Msadek, R., Navarra, A., Pohlmann, H., Rienecker, M., Rosati, T., Schneider,
30 E., Smith, D., Sutton, R., Teng, H., van Oldenborgh, G. J., Vecchi, G., and Yeager, S.: DECADAL CLIMATE PREDICTION An Update from the Trenches, BULLETIN OF THE AMERICAN METEOROLOGICAL SOCIETY, 95, 243–267, https://doi.org/10.1175/BAMS-D-12-00241.1, 2014.

Moemken, J., Reyers, M., Buldmann, B., and Pinto, J. G.: Decadal predictability of regional scale wind speed and wind energy potentials over Central Europe, TELLUS SERIES A-DYNAMIC METEOROLOGY AND OCEANOGRAPHY, 68,
35 https://doi.org/10.3402/tellusa.v68.29199, 2016.

Monerie, P.-A., Coquart, L., Maisonnave, É., Moine, M.-P., Terray, L., and Valcke, S.: Decadal prediction skill using a high-resolution climate model, Climate Dynamics, 49, 3527–3550, https://doi.org/10.1007/s00382-017-3528-x, https://doi.org/10.1007/s00382-017-3528-x, 2017.

Monerie, P. A., Robson, J., Dong, B. W., and Dunstone, N.: A role of the Atlantic Ocean in predicting summer surface air temperature over North East Asia?, Climate Dynamics, 51, 473–491, https://doi.org/10.1007/s00382-017-3935-z, 2018.

Murray, R. J. and Simmonds, I.: A numerical scheme for tracking cyclone centres from digital data. Part I: Development and operation of the scheme, Australian Meteorological Magazine, 39, 155–166, 1991.

5 Müller, W. A., Jungclaus, J. H., Mauritsen, T., Baehr, J., Bittner, M., Budich, R., Bunzel, F., Esch, M., Ghosh, R., Haak, H., Ilyina, T., Kleine, T., Kornblueh, L., Li, H., Modali, K., Notz, D., Pohlmann, H., Roeckner, E., Stemmler, I., Tian, F., and Marotzke, J.: A Higher-resolution Version of the Max Planck Institute Earth System Model (MPI-ESM1.2-HR), Journal of Advances in Modeling Earth Systems, 0, https://doi.org/10.1029/2017MS001217, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2017MS001217, 2018.

Müller, W. A., Baehr, J., Haak, H., Jungclaus, J. H., Kroeger, J., Matei, D., Notz, D., Pohlmann, H., von Storch, J. S., and Marotzke, J.:
10 Forecast skill of multi-year seasonal means in the decadal prediction system of the Max Planck Institute for Meteorology, GEOPHYSICAL RESEARCH LETTERS, 39, https://doi.org/10.1029/2012GL053326, 2012.

Park, T., Park, W., and Latif, M.: Correcting North Atlantic sea surface salinity biases in the Kiel Climate Model: influences on ocean circulation and Atlantic Multidecadal Variability, Climate Dynamics, 47, 2543–2560, https://doi.org/10.1007/s00382-016-2982-1, https://doi.org/10.1007/s00382-016-2982-1, 2016.

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

Pohlmann, H., Muller, W. A., Kulkarni, K., Kameswarrao, M., Matei, D., Vamborg, F. S. E., Kadow, C., Illing, S., and Marotzke, J.: Improved forecast skill in the tropics in the new MiKlip decadal climate predictions, Geophysical Research Letters, 40, 5798–5802, https://doi.org/10.1002/2013GL058051, 2013.

20 Polkova, I., Brune, S., Kadow, C., Romanova, V., Gollan, G., Baehr, J., Glowienka-Hense, R., Greatbatch, R. J., Hense, A., Illing, S., Köhl, A., Kröger, J., Müller, W. A., Pankatz, K., and Stammer, D.: Initialization and Ensemble Generation for Decadal Climate Predictions: A Comparison of Different Methods, Journal of Advances in Modeling Earth Systems, 11, 149–172, https://doi.org/10.1029/2018MS001439, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018MS001439, 2019.

Prodhomme, C., Batté, L., Massonnet, F., Davini, P., Bellprat, O., Guemas, V., and Doblas-Reyes, F. J.: Benefits of Increasing the Model
25 Resolution for the Seasonal Forecast Quality in EC-Earth, Journal of Climate, 29, 9141–9162, https://doi.org/10.1175/JCLI-D-16-0117.1, https://doi.org/10.1175/JCLI-D-16-0117.1, 2016.

Robson, J., Polo, I., Hodson, D. L. R., Stevens, D. P., and Shaffrey, L. C.: Decadal prediction of the North Atlantic subpolar gyre in the HiGEM high-resolution climate model, Climate Dynamics, 50, 921–937, https://doi.org/10.1007/s00382-017-3649-2, 2018.

Scaife, A. A., Woollings, T., Knight, J., Martin, G., and Hinton, T.: Atmospheric Blocking and Mean Biases in Climate Models, JOURNAL
30 OF CLIMATE, 23, 6143–6152, https://doi.org/10.1175/2010JCLI3728.1, 2010.

Scaife, A. A., Copsey, D., Gordon, C., Harris, C., Hinton, T., Keeley, S., O'Neill, A., Roberts, M., and Williams, K.: Improved Atlantic winter blocking in a climate model, GEOPHYSICAL RESEARCH LETTERS, 38, https://doi.org/10.1029/2011GL049573, 2011.

Scherrer, S. C., Croci-Maspoli, M., Schwierz, C., and Appenzeller, C.: 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,
35 https://doi.org/10.1002/joc.1250, 2006.

Shaffrey, L. C., Stevens, I., Norton, W. A., Roberts, M. J., Vidale, P. L., Harle, J. D., Jrrar, A., Stevens, D. P., Woodage, M. J., Demory, M. E., Donners, J., Clark, D. B., Clayton, A., Cole, J. W., Wilson, S. S., Connolley, W. M., Davies, T. M., Iwi, A. M., Johns, T. C., King, J. C., New, A. L., Slingo, J. M., Slingo, A., Steenman-Clark, L., and Martin, G. M.: U.K. HiGEM: The New U.K. High-Resolution Global Envi-

ronment Model - Model Description and Basic Evaluation, Journal of Climate, 22, 1861–1896, https://doi.org/10.1175/2008JCLI2508.1, https://doi.org/10.1175/2008JCLI2508.1, 2009.

Shutts, G. J.: The Propagation of Eddies In Diffluent Jetstreams - Eddy Vorticity Forcing of Blocking Flow-fields, Quarterly Journal of the Royal Meteorological Society, 109, 737–761, https://doi.org/10.1002/qj.49710946204, 1983.

5 Smith, D. M., Cusack, S., Colman, A. W., Folland, C. K., Harris, G. R., and Murphy, J. M.: Improved surface temperature prediction for the coming decade from a global climate model, SCIENCE, 317, 796–799, https://doi.org/10.1126/science.1139540, 2007.

Taylor, K. E., Stouffer, R. J., and Meehl, G. A.: An Overview of CMIP5 and the Experiment Design, Bulletin of the American Meteorological Society, 93, 485–498, https://doi.org/10.1175/BAMS-D-11-00094.1, https://doi.org/10.1175/BAMS-D-11-00094.1, 2012.

Uppala, S. M., Kallberg, P. W., Simmons, A. J., Andrae, U., Bechtold, V. D., Fiorino, M., Gibson, J. K., Haseler, J., Hernandez, A., Kelly,

10 G. A., Li, X., Onogi, K., Saarinen, S., Sokka, N., Allan, R. P., Andersson, E., Arpe, K., Balmaseda, M. A., Beljaars, A. C. M., Van De Berg, L., Bidlot, J., Bormann, N., Caires, S., Chevallier, F., Dethof, A., Dragosavac, M., Fisher, M., Fuentes, M., Hagemann, S., Holm, E., Hoskins, B. J., Isaksen, L., Janssen, P. A. E. M., Jenne, R., McNally, A. P., Mahfouf, J. F., Morcrette, J. J., Rayner, N. A., Saunders, R. W., Simon, P., Sterl, A., Trenberth, K. E., Untch, A., Vasiljevic, D., Viterbo, P., and Woollen, J.: The ERA-40 re-analysis, Quarterly Journal of the Royal Meteorological Society, 131, 2961–3012, https://doi.org/10.1256/qj.04.176, 2005.

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

Xin, X. G., Gao, F., Wei, M., Wu, T. W., Fang, Y. J., and Zhang, J.: Decadal prediction skill of BCC-CSM1.1 climate model in East Asia, International Journal of Climatology, 38, 584–592, https://doi.org/10.1002/joc.5195, 2018.

Zappa, G., Shaffrey, L. C., and Hodges, K. I.: The Ability of CMIP5 Models to Simulate North Atlantic Extratropical Cyclones, Journal of

20 Climate, 26, 5379–5396, https://doi.org/10.1175/JCLI-D-12-00501.1, 2013.