RESEARCH ARTICLE

Large-scale regional model biases in the extratropical North Atlantic storm track and impacts on downstream precipitation

Marie Pontoppidan[1] | Erik W. Kolstad[1] | Stefan P. Sobolowski[1] | Asgeir Sorteberg[2] |
Changhai Liu[3] | Roy Rasmussen[3]

[1]NORCE Norwegian Research Centre,
Bjerknes Centre for Climate Research,
Bergen, Norway

[2]Geophysical Institute, University of
Bergen, Bjerknes Centre for Climate
Research, Bergen, Norway

[3]National Center for Atmospheric Research,
Boulder, Colorado

Correspondence
Marie Pontoppidan, NORCE Norwegian
Research Centre, Bjerknes Centre for
Climate Research, Bergen, Norway.
Email:
marie.pontoppidan@norceresearch.no

Abstract

Global climate models have circulation biases that the community aims to reduce, for instance through high-resolution dynamical downscaling. We used the Weather Research and Forecasting model (WRF) to downscale both ERA-Interim and a bias-corrected version of the Norwegian climate model NorESM1-M on a high-resolution grid. By varying the domain size, we investigated the influence of the driving data and highly resolved topography on the North Atlantic storm track and the precipitation in its exit region. In our largest domains, we found large-scale circulation and storm track biases similar to those seen in global models and with spatial patterns independent of the driving data. The biases in the smaller domains were more dependent on the quality of the driving data. Nevertheless, the biases had little effect on the simulated precipitation in Norway. Although the added value of downscaling was clear with respect to the global climate models, all the downscaled simulations showed similar precipitation frequencies and intensities. We posit that, because the precipitation is so strongly governed by the local topographic forcing, a correct storm track is less critical for the precipitation distribution.

KEYWORDS

North Atlantic storm tracks, precipitation, WRF model biases

1 | INTRODUCTION

Extratropical cyclones are one of the main drivers of the day-to-day variability of wintertime weather in northwestern Europe. The cyclones tend to follow a northeastward path over the North Atlantic, and often bring strong winds and precipitation (e.g., Woollings, 2010). The ability to model the tracks of these storms correctly is essential for assessing the impact of future weather extremes. However, modern global climate models (GCMs) have biases in their North Atlantic storm tracks. In general, the tracks are too zonally oriented (i.e., they do not have enough of a southwest–northeast tilt) and the

cyclones are too weak (Colle *et al.*, 2013; Zappa *et al.*, 2013). Dynamical downscaling has helped overcome biases related to unresolved local to regional processes, but the large-scale biases have, to some degree, persisted in the regional models (Seiler *et al.*, 2017; Poan *et al.*, 2018). A strategy where systematic large-scale biases are corrected prior to downscaling has been shown to lessen the negative effects of the biases in several regions (Xu and Yang, 2012; 2015; Bruyère *et al.*, 2014; Wang and Kotamarthi, 2015; Rocheta *et al.*, 2017). However, when the same strategy was applied to the Northeast Atlantic and western Norway, the storm track in the downscaled simulations was still biased (Pontoppidan *et al.*, 2018).

The influence of orography on the simulation of storm tracks has been investigated through simplified modeling frameworks and linear models (Smith, 1984; 1986; Broccoli and Manabe, 1992; Chang *et al.*, 2002; Chang, 2009). Brayshaw *et al.* (2009) showed that the upper-level tropospheric jet deflection induced by the Rocky Mountains enhances storm track development along the eastern coastline of the U.S. continent, and that the triangular shape of the North American continent contributes to this through the development of a cold pool over the northeastern land surface, which strengthens the meridional temperature gradient and low-level baroclinicity. Corroborating results were obtained by Pithan *et al.* (2016), who found that the typical too-zonal storm track in climate models closely aligns with the outcome of running a climate model with the parametrized low-level orographic drag switched off, and by Wilson *et al.* (2009), who linked the influence of the Rocky Mountains with the tilt of the North Atlantic storm track in an idealized model. Some recent studies suggest that models can reproduce some of the orographic interactions correctly (Sobolowski *et al.*, 2007; Berckmans *et al.*, 2013; Hoskins and Woollings, 2015; Rasmussen and Houze, 2016), and that a higher grid resolution indeed leads to improved representation of storm tracks and their associated downstream precipitation (Willison *et al.*, 2013; 2015; Booth *et al.*, 2018). To ensure that nonlinear orographic interactions, their downstream effects on the North Atlantic storm track, and ultimately the individual storm impact in northwestern Europe are included explicitly, it might make sense to define a large high-resolution domain where the Rocky Mountains are included. This would require a very large domain in the regional model, as investigated in "big-brother" experiments. In these, a reference simulation is performed in a larger outer domain (the "big brother") to enable validation against a "perfect" simulation. Such experiments have been performed for very large domains covering the contiguous U.S. region (Leduc and Laprise, 2009; Diaconescu and Laprise, 2013), but have yet to be investigated for the North Atlantic region. Previous big-brother experiments in smaller domains have shown that, when the quality of the driving data is good, the skill in reproducing large scales decreases with increasing domain size. However, when the driving data is coarse, the representation of the circulation improves with a larger domain (Denis *et al.*, 2002). In the North Atlantic region (the area of interest for this study), Køltzow *et al.* (2008) argued that a larger domain reduces large-scale biases, although their large domain was considerably smaller than the domain used here and their model grid spacing coarser.

While big-brother experiments are useful for diagnosing the biases that arise from the nesting procedure, the model bias itself stays unaddressed, because it is also present in the outer domain, which is used as validation data in the big-brother setup. A recent case study of three cold fronts in southeastern Europe showed that not only the size but also the position of the boundaries in the domains had great influence on model biases (Lamraoui *et al.*, 2018). The main questions addressed in this article are the following.

1. When a very large (10,200 km × 6,800 km) high-resolution domain is used, can the Weather Research and Forecasting model (WRF) reduce the North Atlantic storm track bias?
2. Does the model skill depend on the driving data?
3. What is the impact of the driving data and the domain size on the simulation of precipitation in northwestern Europe?

We seek answers to these questions by dynamically downscaling ERA-Interim (Dee *et al.*, 2011) and an IPCC-class earth system model (NorESM1-M: Bentsen *et al.*, 2012; Iversen *et al.*, 2013) over a domain that includes the Rocky Mountains, the North Atlantic Ocean, and northwestern Europe (see Figure 1a).

2 | METHODS

The simulations were performed with WRF model version 3.8.1 (Skamarock *et al.*, 2008), with driving data from ERA-Interim (Dee *et al.*, 2011) and a bias-corrected (Pontoppidan *et al.*, 2018) transient model run of the NorESM1-M (realization r1i1p1), with a grid spacing of 1.9° × 2.5° (Bentsen *et al.*, 2012; Iversen *et al.*, 2013). To mimic the resolution of a GCM in the driving data, we used a coarse 2° version of ERA-Interim to force most of the simulations. Because observations have been assimilated into ERA-Interim prior to the interpolation, this does not exclude the finer scales completely, but the aim was to reduce the effect as much as possible. The control simulation (ENUDG) was driven by the original 0.75° ERA-Interim. As shown in Table 1, the simulations all used the Thompson microphysical scheme (Thompson *et al.*, 2008), the Rapid Radiative Transfer Model for General Circulation Models (RRTMG) long- and shortwave radiation schemes (Iacono *et al.*, 2008), the Yonsei University scheme planetary boundary-layer scheme (Hong *et al.*, 2006), the Noah multi-physics land surface model (Niu *et al.*, 2011), and the Tiedtke cumulus scheme (Tiedtke, 1989; Zhang *et al.*, 2011). The simulations were run with a horizontal grid spacing of 20 km. There were 40 vertical levels, going up to 10 hPa. In the control simulation (ENUDG), spectral nudging was applied to maintain the flow as close as possible to actual conditions. To allow for the interior flow to develop freely in the remaining six simulations, nudging was not used there.

We ran several configurations for each simulation period: (a) the full domain including North America, to allow WRF to simulate interactions with the high-resolution topography
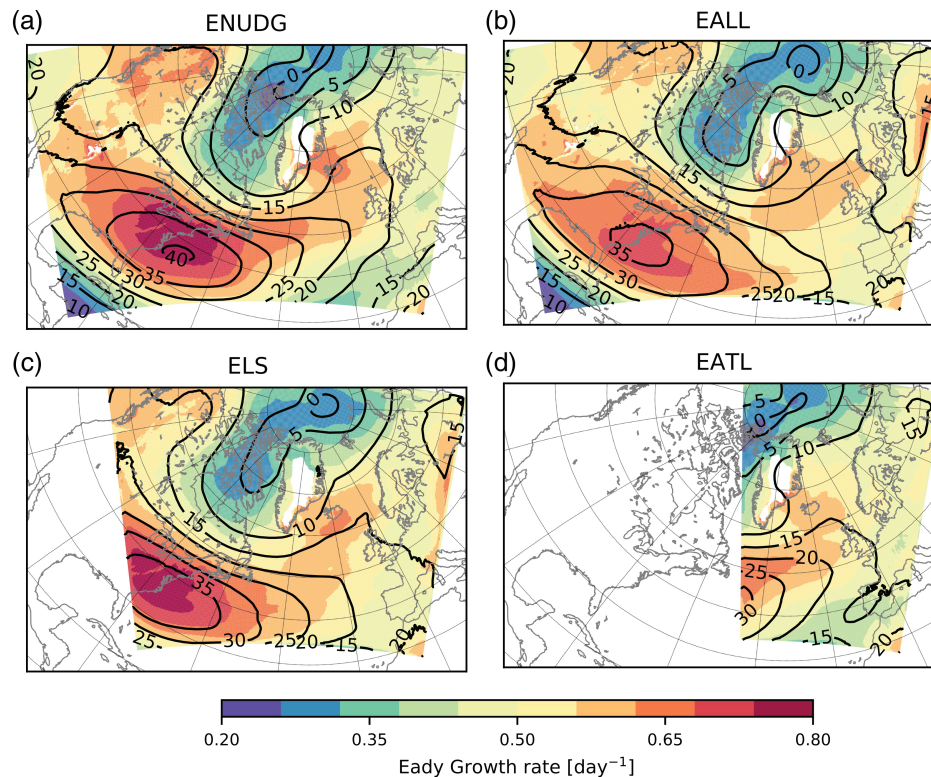
FIGURE 1 Eady growth rate at 500 hPa in shading and 300-hPa zonal wind in contours, for (a) the ENUDG simulation, (b) the EALL simulations, (c) the ELS simulation, and (d) the EATL simulation

TABLE 1 Overview of the WRF configuration in the simulations

WRF configuration	
Grid spacing	20 km
Vertical levels	40
Model top	10 hPa
Microphysics	Thompson
Radiation	RRTMG
Planetary boundary layer	YSU
Land surface model	Noah MP
Cumulus	Tiedtke

Abbreviations: YSU, Yonsei University scheme; MP, multi-physics.

TABLE 2 Overview of the different experiments and their driving data. ENUDG: the spectrally **NUDG**ed simulation forced with **ER**A-Interim, EALL: the full **ALL** domain forced with **E**RA-Interim, NALL: the full domain forced with **N**orESM1-M, ELS and NLS: the domains including the **L**and–**S**ea contrast forced with ERA-Interim and NorESM1-M, respectively, and EATL and NATL: the domain covering the **ATL**antic forced with ERA-Interim and NorESM1-M, respectively

	NorESM1-M	**ERAI 2°**	**ERAI 0.75°**
Spectrally nudged			ENUDG
ALL	NALL	EALL	
LS	NLS	ELS	
ATL	NATL	EATL	

(extensions: ALL, see domain in Figure 1b); (b) a reduced domain, but still including the contiguous U.S. east coast, with the aim of excluding explicit resolution of the orographic effects of the Rocky Mountains (extensions: LS, see domain in Figure 1c); (c) an even more reduced domain, excluding the contiguous United States entirely, with the aim of excluding the effects of both the Rockies and the land–sea contrast along the U.S. east coast (extensions: ATL, see domain in Figure 1d). In addition to using ERA-Interim as driving data, we also performed simulations that were forced with the bias-corrected NorESM1-M; for details, see Pontoppidan *et al.* (2018). The configurations are summarized in Table 2. Apart from the domain size, all the WRF simulations used identical settings and the parametrization schemes listed in Table 1.

The interannual variability in the reanalysis does not align with the variability in the GCM. In an attempt to select a representative subset, we chose six winters from the reanalysis period based on the phase of the North Atlantic Oscillation (NAO). The NAO index correlates well with wintertime precipitation in northwestern Europe in general and our focus region in western Norway in particular (Hanssen-Bauer, 2005). We selected the two most NAO-positive winters, the two most NAO-negative winters, and the two winters for which the NAO index was closest to zero. We then used the same strategy to select six winters from the GCM simulations. Although the two sets of winters cannot be compared one against the other, they can potentially shed light on systematic behaviors and biases in the regional model.

TABLE 3 Overview of winter seasons selected for simulation

NAO phase	ERA-Interim	NorESM1-M
NAO positive	1988–1989	1982–1983
	1994–1995	1995–1996
NAO neutral	1987–1988	1985–1986
	2003–2004	1996–1997
NAO negative	1995–1996	1988–1989
	2009–2010	1999–2000

The selected ERA-Interim winter seasons were 1988–1989 and 1994–1995 (NAO-positive), 1995–1996 and 2009–2010 (NAO-negative), and 1987–1988 and 2003–2004 (NAO-neutral). For the simulations forced with the bias-corrected NorESM1-M data, we selected the winter seasons in 1982–1983 and 1995–1996 (NAO-positive), 1988–1989 and 1999–2000 (NAO-negative), and 1985–1986 and 1996–1997 (NAO-neutral), as summarized in Table 3. All the simulations were conducted for a period of four months from November to February, using November as a spin-up period.

To analyze the representation of baroclinic development associated with the North Atlantic storm track, we used the Eady growth rate (e.g., Simmonds and Lim, 2009), calculated using the following equation:

$$egr = 0.3098 * \frac{|f * \frac{du}{dz}|}{N}, \quad (1)$$

where

$$N = \sqrt{\frac{g}{\Theta}\frac{d\Theta}{dz}},$$

better known as the Brunt–Väisäla frequency, f is the Coriolis parameter, g is the gravity and z is the height of the vertical levels of u (the zonal wind) and Θ (the potential temperature). The Eady growth rate was calculated at every time step for the 500-hPa level, using instantaneous values of Θ and u at 300 and 700 hPa, and averaged over the entire period, as suggested by Simmonds and Lim (2009). In addition, the upper-level jet over the North Atlantic Ocean was examined based on the 300-hPa zonal wind.

A comparison of spatial patterns was done using the Pearson spatial pattern correlation (SPC). It correlates two maps of values at their corresponding locations. In this study, we correlated the spatial mean value of the Eady growth rate in the nudged simulation with the remaining simulations inside their common domains, and similarly for the mean values of 300-hPa zonal wind. We also calculated the mean absolute deviation (MAD).

The precipitation comparison was performed for 10 selected observational stations from the national meteorological network in Norway. The choice of stations was made to obtain the largest amount of data available during the six simulated winter periods. The data were collected at an

hourly resolution and aggregated to three-hourly time periods to be compatible with the temporal resolution of the model data. We used the Perkins skill score (Perkins *et al.*, 2007) for validation. The Perkins skill score calculates the cumulative minimum value of two dataset distributions for each bin, quantifying the fraction of overlap. The equation is as follows:

$$PSS = \sum_{i=1}^{n} \min(D1_i, D2_i), \quad (2)$$

where $D1_i$ and $D2_i$ are the two datasets binned in n bins. A number close to one suggests a good fit, whereas zero indicates no fit.

3 | RESULTS

3.1 | Large-scale biases

First, we focus on the large-scale phenomenon with the largest influence on the winter climate in northwestern Europe–the North Atlantic storm track. Figure 1a shows the Eady growth rate in shading and the 300-hPa zonal wind in contours for ENUDG. The simulation reproduces the southwest to northeast tilt in the storm track, as well as the tilt in the upper-level jet. The Eady growth rate is highest along the U.S. east coast, where the cyclogenesis is also at its largest. It is persistent across the Atlantic basin, over the Norwegian Sea, and extends into the Barents Sea. The upper-level jet maximum is colocated with the Eady growth rate maximum, and the jet extends across the Atlantic basin. To confirm the effect of the nudging, we also ran a simulation with 2° ERA-Interim data on the boundaries and spectral nudging. Since it showed similar results to the ENUDG simulation, we will only refer to the ENUDG simulation henceforth.

Figure 1b–d shows the same variables, but for the remaining ERA-Interim-driven simulations. The magnitudes of the Eady growth rate maxima vary considerably, as do the tilts and the maxima of the upper-level jet.

The differences between the simulations driven by 2° ERA-Interim (EALL, ELS, and EATL) and the nudged simulation driven by 0.75° data (ENUDG) are shown in Figure 2. To assess the magnitude of the differences, we also calculated the standard deviation of ERA-Interim over a 30-year winter period and marked the areas where the magnitude of the differences was larger than one standard deviation. The EATL domain in Figure 2e,f shows good agreement with the ENUDG simulation with a low MAD and a high SPC value. However, the performance deteriorates gradually as the domain size is increased to include the eastern part of the contiguous United States (Figure 2c,d) and further to include the Rockies (Figure 2a,b). Compared with the ENUDG simulation, the Eady growth rate is significantly weaker in the storm track entry region and weaker along most of the storm
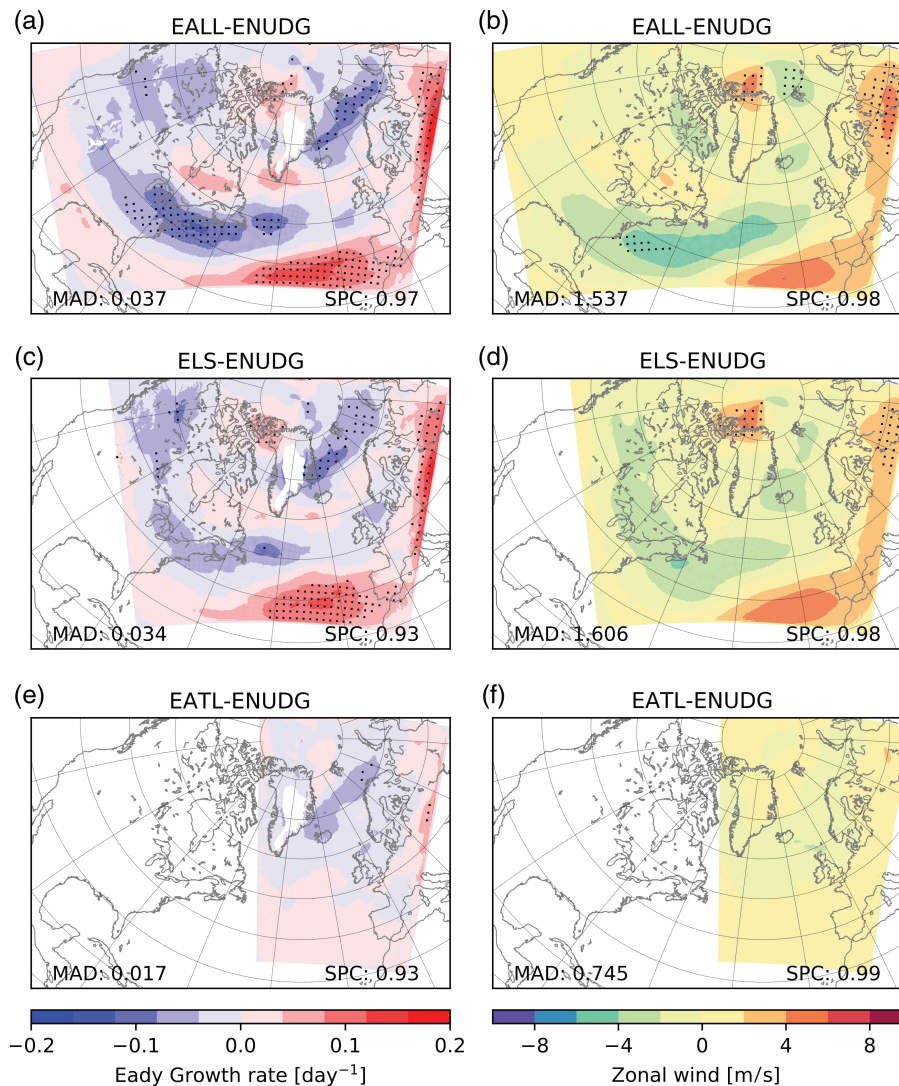
FIGURE 2 Biases in the simulations: (a) difference in 500-hPa Eady growth rate between the EALL and ENUDG simulations, (b) difference in 300-hPa zonal wind between the EALL and ENUDG simulations. (c) and (d), as (a) and (b), but for the difference between the ELS and ENUDG simulations. (e) and (f), as (a) and (b), but for the difference between the EATL and ENUDG simulations. Mean absolute deviation (MAD) is printed in the lower left corner, and the spatial pattern correlation (SPC) between mean maps of the two compared simulations is printed in the lower right corner. The dots indicate differences larger than one standard deviation of 30 years (DJF 1980–2010) data from ERA-Interim

track over the ocean. This increasing bias is quantified by a doubling of the MAD in the ALL domain, but the SPC is maintained (increased) to 0.93 (0.97) for the LS (ALL) domain. Also, the upper-level jet is too zonally oriented in ELS and EALL and too weak at its maximum. Again, the ALL domain MAD value is twice as large as in the ATL domain. The jet is generally wider in both simulations than the ENUDG jet, but the very high spatial correlations in all domains confirm the fairly similar patterns. It is worth noting that the SPC is generally higher for the 300-hPa zonal winds than for the 500-hPa Eady growth rate.

3.2 | Linkage to GCM simulations

As a correct representation of the storm track is important for understanding climate projections in Norway, we produced

a similar set of simulations, based on the NAO state, with exactly the same configurations, but driven by bias-corrected NorESM1-M data.

The results for the model runs driven by 1.9° × 2.5° NorESM1-M data are shown in Figure 3, which shows the differences with respect to ENUDG. Comparing Figure 2a,b with Figure 3a,b, we find some similarities. In both sets of simulations, we find a lower Eady growth rate in the entry region. This is probably causing the lower storm activity in this region of the storm track (the jet is also weaker in the same area). The biases in the upper-level jet, and in particular along the storm track, appear to be consistent circulation biases endemic to the regional model. The biases are clearly dependent on our domain selection and the inclusion of the Rockies does not reduce this circulation bias. With regard to the LS domain, the Eady growth-rate biases are fairly similar,
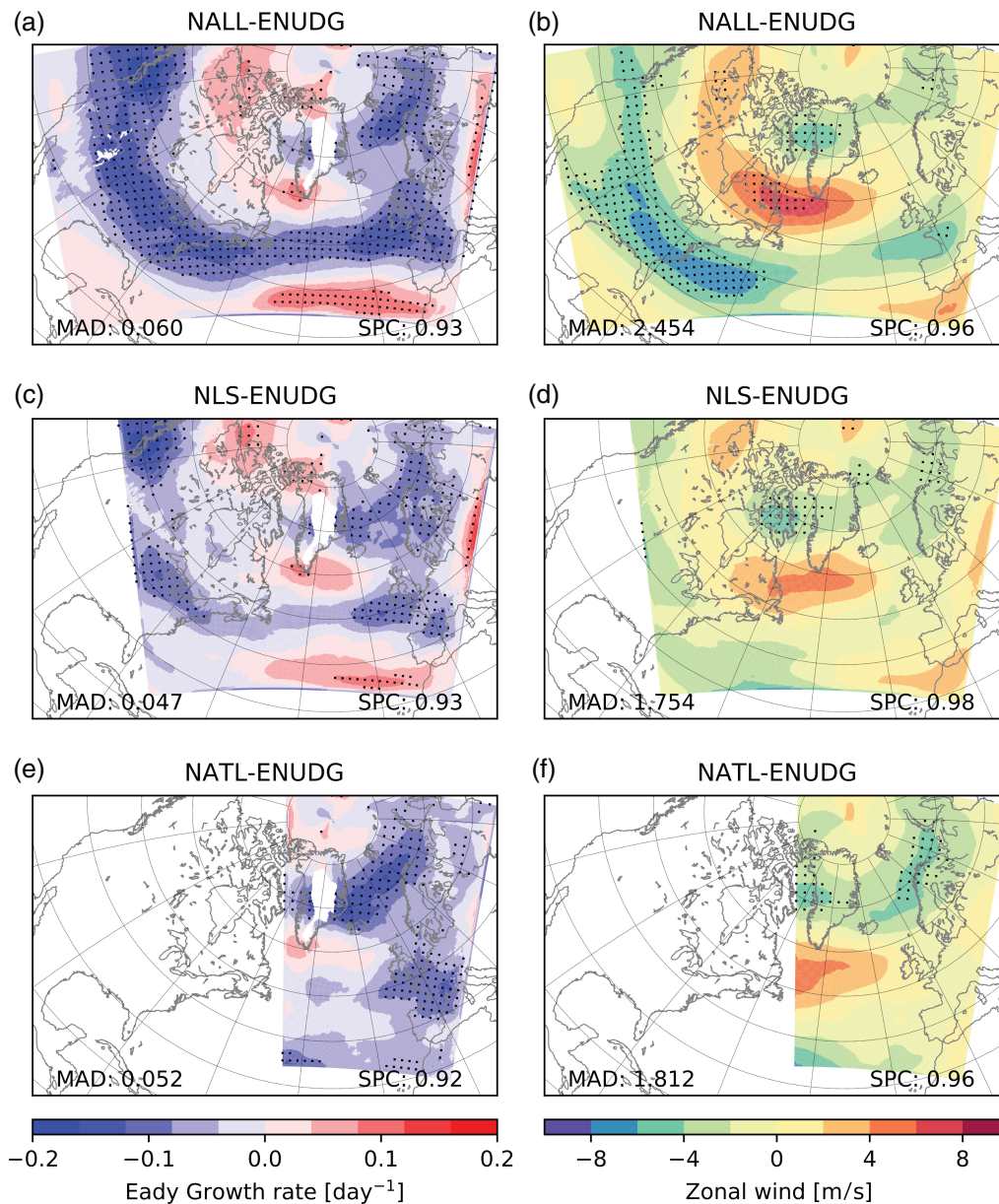
FIGURE 3 Similar to Figure 2, but for differences between the NorESM1-M driven simulations and the ENUDG simulation. Mean absolute deviation (MAD) is printed in the lower left corner, and the spatial pattern correlation (SPC) between mean maps of the two compared simulations is printed in the lower right corner. The dots indicate differences larger than one standard deviation of 30 years (DJF 1980–2010) data from ERA-Interim

but differences emerge for the upper-level wind. These differences arise mainly because the upper-level jet is wider in the NorESM1-M simulations, extending further north towards the tip of Greenland. When the NATL domain is used, the storm track in the Norwegian Sea is underestimated. In that small domain, we also find the largest differences between the simulations driven by ERA-Interim and the ones driven by NorESM1-M. SPCs vary little and are all above 0.92 for the Eady growth rate and 0.96 for the upper-level wind. The MAD again mainly increases with increased domain size, but with a lower magnitude than for the simulations driven by ERA-Interim. We also note that, although the areas exceeding one standard deviation of the 30-year ERA-Interim data

are colocated in our experiments driven by ERA-Interim and the experiments driven by NorESM1-M, the area is larger in the latter set of experiments.

To investigate the origins of the biases further, we separated the biases in the NALL simulation into biases potentially arising from the GCM, calculated as the difference between NorESM1-M and ERA-Interim, and biases potentially arising from the regional model, calculated as differences between NALL and the driving NorESM1-M. Figure 4 shows these biases. Note that the separation is somewhat artificial, since the GCM only influences the simulation through the lateral boundaries and the sea-surface temperatures. In the driving NorESM1-M, the 500-hPa Eady growth rate is generally too
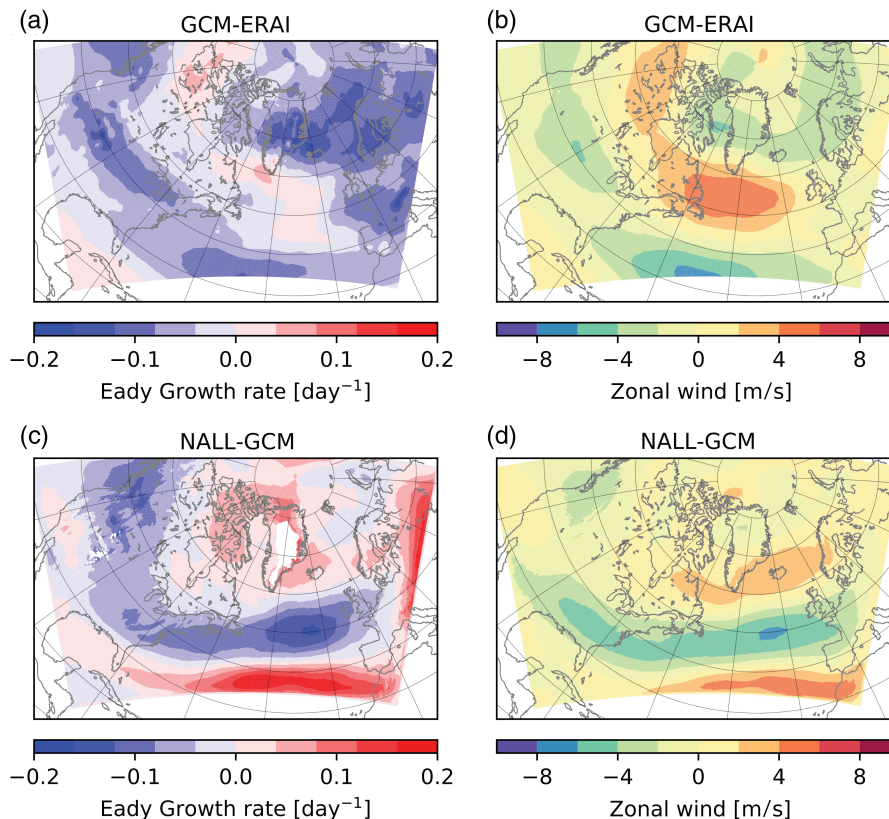
FIGURE 4 Estimated biases for the driving GCM (NorESM1-M) in (a) 500-hPa Eady growth rate and (b) 300-hPa zonal wind, and estimated model biases from the regional model in (c) 500-hPa Eady growth rate and (d) 300-hPa zonal wind

low in the entire domain (Figure 4a), and the regional model enhances this bias in the storm track region across the Atlantic Ocean (Figure 4c). Compared with Figure 3a, it is clear that the regional model bias is the main reason for the total bias of the Eady growth rate. Note also the similarities between Figures 4c and 2a, which also represent the regional model bias, but from the simulations driven by ERA-Interim. For the upper-level zonal wind, Figure 4b reveals a positive bias south of Greenland, which suggests a too-wide jet in the GCM. The regional model bias in Figure 4d has a similar spatial pattern to the Eady growth rate bias in Figure 4c, and dominates over the GCM bias in the entry region of the storm track, while the GCM bias dominates across the Atlantic basin.

Large differences in the storm track might be expected to affect the simulated precipitation, but, as shown in Figure 5, the differences in accumulated DJF precipitation occur mainly in the entrance region of the storm track. Despite these rather large biases over the western Atlantic, the differences over the eastern Atlantic are minor and seem little affected by either domain choice or driving data.

3.3 | Effects on precipitation in western Europe

Whilst the accumulated winter precipitation showed small differences in the eastern Atlantic, there might be larger differences in the daily variability. We therefore investigated the frequency and intensity of the precipitation at our highest available temporal resolution: three-hourly. We characterize a wet period as a three-hourly period with more than 0.1 mm, which is the lowest measurable value from the observational stations.

Figure 6 shows the frequency of precipitation periods. The three-hourly precipitation frequency at 10 stations in Norway is shown in Figure 6a. Note that two observational stations in southeastern Norway are quite close to each other, but both were included in the analysis to avoid a reduction of available data. A comparison of the observed values and the simulated values, interpolated from the nearest four grid points, is shown as the percentage of the observational values in the lower right corner of each panel. The simulated values are shown as color-filled circles. In the raw model data, an overestimation of the wet frequency is evident (Figure 6b,c); at the stations, the model values are more than 140% of the observed values. As the seven downscaled simulations show very similar results in the background precipitation, we only show the ALL domains (Figure 6d–f). The percentage of the observed values is shown in Table 4 and reveals a 40–50% reduction of the bias in the downscaled simulations.

For the intensity of precipitation during the wet periods, we turn to Figure 7. Again the downscaled simulations are all fairly similar, varying between 113 and 125% of the
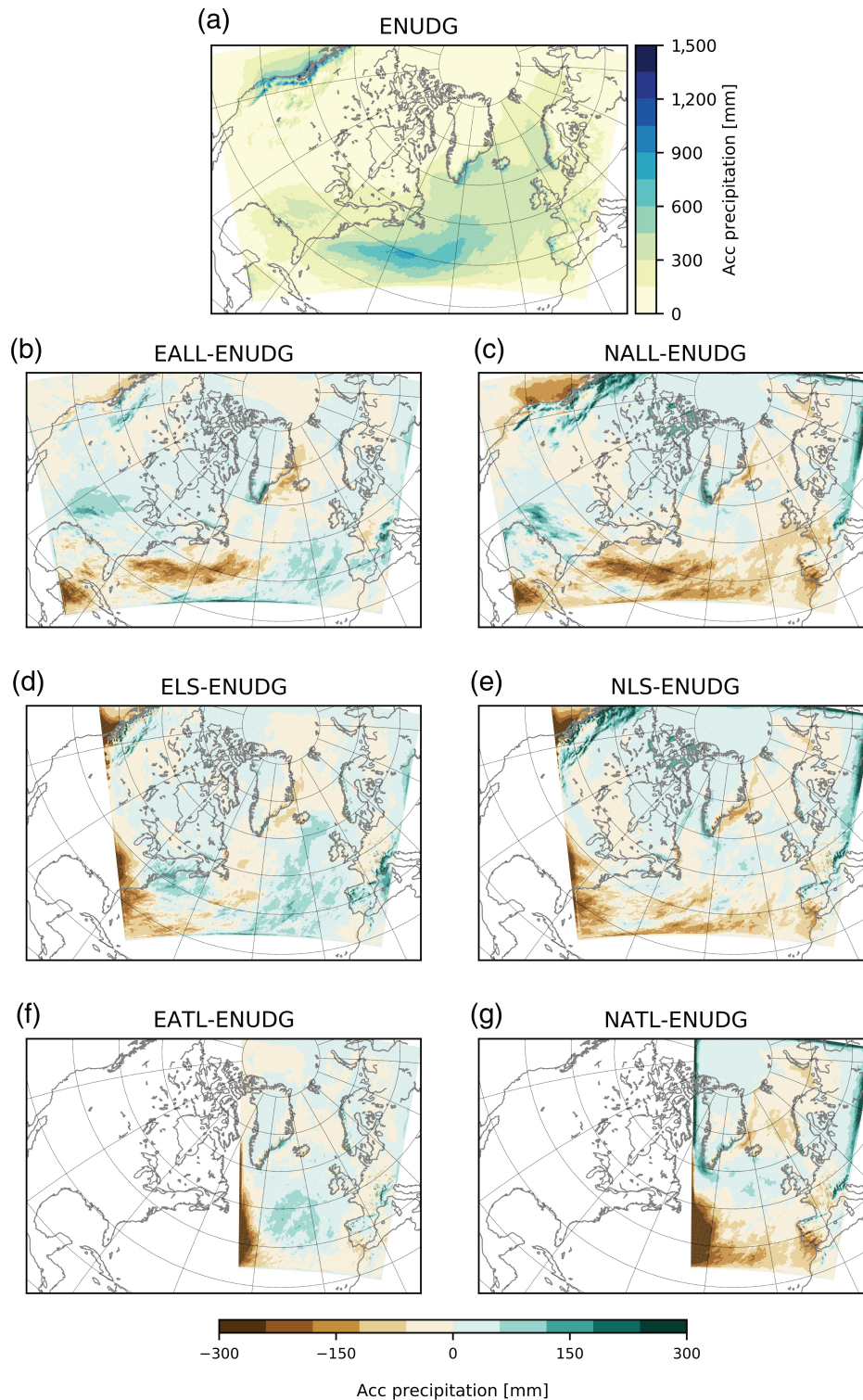
FIGURE 5 DJF accumulated precipitation: (a) the ENUDG simulation; (b), (d), and (f) difference between EALL, ELS, and EATL and the ENUDG simulation; (c), (e), and (g) difference between NALL, NLS, and NATL and the ENUDG simulation

observations (not all are shown), as listed in Table 4. The raw ERA-Interim field (Figure 7b) has too high intensity, 188% of the observations, whereas the original NorESM1-M data (Figure 7c) clearly have too low intensity during the wet periods, with only about 60% of the observational values, especially along the mountainous coastline.

The added value of the downscaling is demonstrated further when we compare the raw ERA-Interim and NorESM1-M fields with observations in Figure 8, where 100% means perfect agreement with the observations at the station. We expect some discrepancies with the observations, as the model values are based on four grid points, whereas the
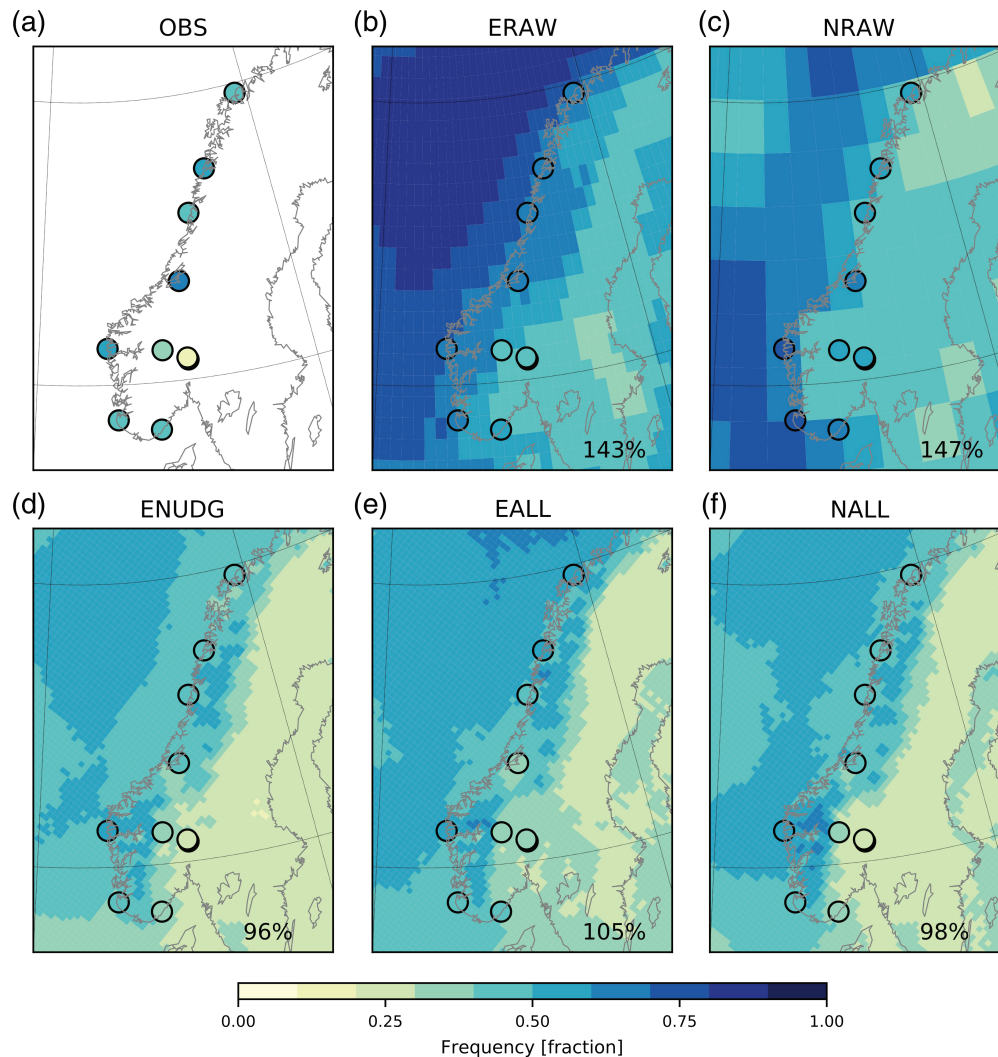
FIGURE 6 Fraction of wet 3-hr periods in (a) the observations, (b) the raw ERA-Interim, (c) the raw NorESM1-M, (d) the ENUDG simulation, (e) the EALL simulation, and (f) the NALL simulation. The values inside the subfigures denote the mean percentage of observation values at the position of the stations (interpolated from the four nearest grid points) in the respective simulations. A wet period is defined as a 3-hr period with more than 0.1 mm precipitation. Circles show the simulated fraction at the station position, integrated from the four nearest grid points, except in (a), which shows the observed values

observation value is valid for a single location only. At the 10 stations, the raw NorESM1-M data generally have too many wet periods compared with the observations, but the intensity is too low during those periods. This improves considerably in the downscaled runs. Even the ERA-Interim data, which have both too high frequency and too high intensity, agree with the observations better when downscaled. The Perkins skill score (Perkins *et al.*, 2007), calculated on the wet-day probability density function, increases slightly from 0.86 for the original NorESM1-M to a range of 0.87–0.91 for the downscaled simulations (a value of 1 means that the distributions are identical), as shown in Table 4. The comparison of results with observational data is slightly sensitive to the definition of a wet period; however, the conclusions remain the same regardless of the cutoff value. Crucially, the choice of domain or driving data does not have a large impact on the results.

4 | DISCUSSION

Our work was motivated by the idea that the North Atlantic storm track would be better represented by WRF in a large high-resolution model domain than in a smaller domain, due to more realistic topographic interactions in mountainous regions. Déqué *et al.* (2007) showed that the largest biases in regional models, albeit for an earlier generation of models and at coarser resolutions, were related to the driving GCMs, not the regional models themselves. Noguer *et al.* (1998) came to a similar conclusion. Also, big-brother experiments identified biases when using coarse-resolution data on the boundaries, but the biases arising purely from the nesting procedure were reduced when the domain size was increased; this has been shown for both Canada (Diaconescu *et al.*, 2007; Diaconescu and Laprise, 2013) and Norway (Køltzow *et al.*, 2008).
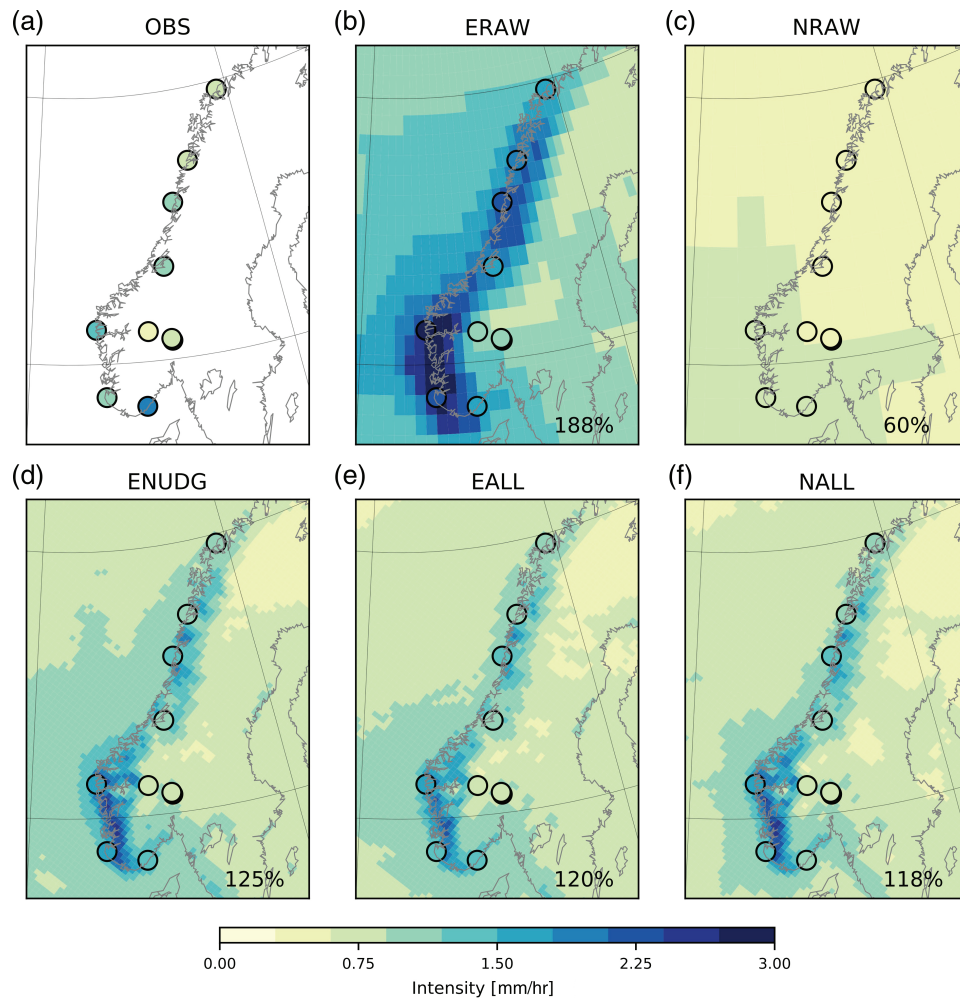
FIGURE 7 Similar to Figure 6, but for the mean intensity during wet periods. The values inside the subfigures denote the mean percentage of observation values at the position of the stations (interpolated from the four nearest grid points) in the respective simulations. Circles show the simulated intensity at the station position, integrated from the four nearest grid points, except in (a), which shows the observed values

TABLE 4 Score metrics for the simulated precipitation, interpolated from the four nearest grid points. Fraction and intensity are a percentage of the observational values, the rightmost column is the Perkins skill score (PSS)

	Fraction [%]	Intensity [%]	PSS
ERAW	143	188	0.77
NRAW	147	60	0.86
ENUDG	96	125	0.87
EALL	105	120	0.87
ELS	107	117	0.90
EATL	106	115	0.90
NALL	98	118	0.89
NLS	103	115	0.91
NATL	99	113	0.91

Of the model runs driven by ERA-Interim on the boundaries, we found that the one with the smallest domain (EATL) has the most accurate storm track and cyclogenesis representation over the eastern part of the North Atlantic. In the model runs where the WRF domain is expanded further west,

the domains are large enough to allow WRF to develop its own large-scale circulation. The result is a biased performance, with regards to all of storm track, storm magnitude, and upper-level zonal winds. A cyclone-tracking analysis confirmed these biases (not shown). It is of note that the inclusion of the Rocky Mountains in the EALL run does not reduce these circulation biases. This result, in particular, requires further investigation.

We suggest that, when reanalysis data are used to drive the model at the boundaries, it is better to have a smaller domain to make sure that the synoptic scales are strongly constrained by the reanalysis. This does not appear to be the case when biased GCM data are used to drive the regional model. Even though the NorESM1-M data that we used had been bias-corrected prior to driving WRF, the storm-track-related biases are considerably larger in NATL than in EATL. This was expected, but we had hoped that using a larger domain with high horizontal grid spacing (20 km), and thereby a more reliable interaction between the orography and the atmosphere, would enable WRF to produce a more accurate
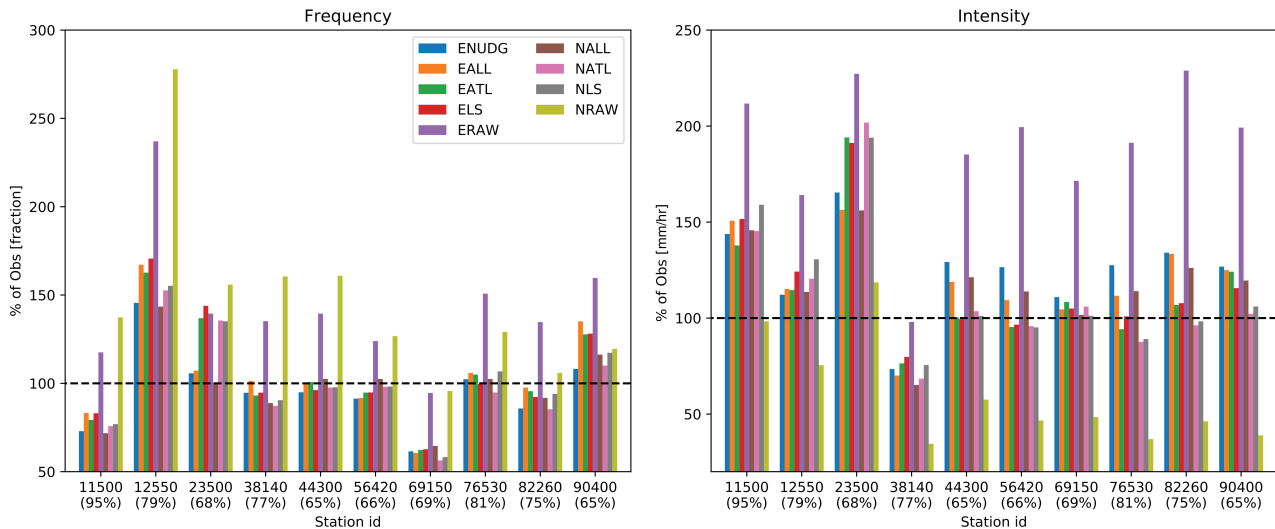
FIGURE 8 Precipitation from 10 selected observational stations and the simulations interpolated from the four nearest grid points and given as a percentage of observations for (a) the fraction of wet periods and (b) the mean intensity during wet periods. The percentage below the station ID shows the percentage of observational data available at the station

storm track. However, the NALL model run has biases that were similar to the ones in EALL. In fact, these biases are similar to well-known circulation biases in GCMs (Zappa *et al.*, 2013). This suggests that neither orographic interactions (dynamics) nor land–sea contrasts (thermodynamics) are simulated particularly well in WRF, even with a 20-km grid spacing. Some effort has been put into parametrization of low-level orographic blocking and gravity-wave drag in climate models, and this can reduce biases (Pithan *et al.*, 2016; van Niekerk *et al.*, 2016). However, in the WRF version used here (version 3.8.1), the combined low-level drag and gravity-wave parametrization scheme (Shin *et al.*, 2010) was not available for the map projection that we used. The newly released version 4.0.2 has this option, and it is activated by default. We therefore performed a simulation that was as similar as possible to EALL to test the influence of the gravity-wave drag parametrization, but we found no clear improvement with respect to the circulation biases (not shown). This could be because the subgrid orography parameter is not available at latitudes greater than 60°N, and/or the grid resolution of 20 km is sufficient to partly resolve the dominant wavelengths emanating from the Rocky Mountains, which could adversely affect the parametrized flow. Either way, WRF seems unreliable with regard to orographic drag and gravity-wave parametrization in high-latitude domains at present and further investigation is needed to clarify the role of this parametrization scheme in mitigating/exacerbating downstream circulation biases.

We summarize some possible reasons for the biases in the storm track representation. The ALL (EALL, and NALL) simulations struggle to simulate the eastern North American cold pool (not shown). This influences the position and strength of the upper-level jet and thereby the meridional temperature gradient across the contiguous U.S. east coast. A stronger temperature gradient is expected to lead to increased baroclinicity and storm development, and hence an increased Eady growth rate, while a weaker temperature gradient can lead to too zonally oriented a storm track. Our ENUDG simulation does indeed have the strongest meridional temperature gradient in the storm track entry region, as it is colder than the remaining simulations in the Canadian region and warmer in the troposphere above the southwestern part of the North Atlantic. The lower spatial correlation in the Eady growth rate compared with the upper-level winds suggests that the biases may result from the poor representation of small-scale thermodynamic processes that drive baroclinic instability, rather than the large-scale flow.

The lack of precipitation in the southern part of the storm track entry region in the EALL and NALL simulations is probably due to too little condensation aloft and not enough latent heating in the middle and upper troposphere. This could be part of a negative feedback connected to the reduced baroclinicity in the area. This idea is supported by an investigation of the 90th percentile of the 10-m wind speed (not shown). We found that the EALL and NALL simulations both underestimate the wind speeds along the U.S. east coast (with respect to ENUDG). This is likely because the storms are too weak, which is consistent with the precipitation (and latent heating) deficit.

Despite the circulation biases and precipitation biases over the western Atlantic, the effects on precipitation in northwestern Europe are surprisingly small. In fact, seasonal precipitation over land appears to be mostly unaffected by the domain size and the driving model. While this has been shown previously for domain size (Colin *et al.*, 2010; Diaconescu and Laprise, 2013) in a big-brother experimental setup, the effect of the driving model has not been clarified. Previous generations of regional models showed

biases dependent on the driving model in Europe (Noguer *et al.*, 1998; Déqué *et al.*, 2007) and Scandinavia (Jacob and Podzun, 1997), whereas newer big-brother studies have shown that the resolution of the driving model is less important for precipitation in Canada, including the Pacific Northwest region (Diaconescu and Laprise, 2013), which has a coastal topographic distribution similar to our study area. Køltzow *et al.* (2008) reached a similar conclusion for Norway, suggesting that the mountainous terrain at the end of the storm track was a dominating factor for local precipitation. Our study supports the latter conclusion; neither domain size nor driving model seem to have large influences on the seasonal precipitation. In western Norway, orographic enhancement is prominent during autumn and winter (Reuder *et al.*, 2007; Barstad and Caroletti, 2013; Pontoppidan *et al.*, 2017). We suggest that the local orographic influence on the simulated precipitation in our experiments is much larger than the effects of the domain size and the driving model.

Another reason for the similarities in the simulated precipitation may be linked to the source regions. Stohl *et al.* (2008) showed that midlatitude and subtropical regions are both important for wintertime precipitation in Norway. This suggests that our smallest domain is sufficient to resolve one of the main source regions, and that the source regions in general are well resolved in all of our simulations.

For climate projections, the added value of downscaling is important. As pointed out by, for example, Stephens *et al.* (2010), the global models produce precipitation far too often and with too low intensity. Our comparison of precipitation in WRF with the original NorESM1-M model output confirms this. At the 10 observational stations, the dynamical downscaling clearly adds value, as evidenced by the increased Perkins skill score. We also saw qualitative improvements, with regard to both the frequency and the intensity of the precipitation. The minor differences between the various simulations emphasize that the added value is largely a consequence of local interactions with topography, rather than upstream effects. The added value is consistent with numerous previous studies. For instance, Mayer *et al.* (2015) showed added value for downscaled simulations over Scandinavia; however, the improvement has also been shown to be negligible for the winter season over Europe (Kotlarski *et al.*, 2014; Glisan *et al.*, 2019).

Although the precipitation in northern Europe seems little affected by the circulation biases in the regional model, these biases influence other variables. Though wind was not the focus in this article, we emphasize that operations related to wind energy, fisheries, and shipping are highly dependent on a correct representation of synoptic-scale storms over the Atlantic, both on a day-to-day operational basis and for future planning. Therefore, more effort should be put into addressing GCM circulation biases.

5 | CONCLUDING REMARKS

In this study we have explored the large-scale biases that arise in large-domain, high-resolution WRF simulations. We speculate that the biases arise because WRF has fundamental problems in representing the interactions between orography and the atmospheric circulation as well as small-scale thermodynamic processes. Nevertheless, the impact of those biases on precipitation in western Norway was found to be minor, potentially due to the strong local orographic forcing. A fundamental question, which remains to be answered, is whether the precipitation is improved, but for the wrong reasons. How would an accurately simulated storm track influence the precipitation distribution in the area?

The future increase in computational resources will enable us to increase the resolution of global climate models and large efforts are currently under way to develop global variable- and high-resolution climate models further (e.g., Skamarock *et al.*, 2012). Despite the apparent insensitivity of regional precipitation to large-scale deficiencies, this needs to be confirmed in a multimodel ensemble setting. Also, large-scale circulation biases can affect many other near-surface variables that are critical for nature and society. Therefore, it is important to address and rectify large-scale circulation biases in order to obtain more credible assessments of future impacts of climate change.

ORCID

Marie Pontoppidan http://orcid.org/0000-0003-4023-6811
Erik W. Kolstad http://orcid.org/0000-0001-5394-9541
Stefan P. Sobolowski http://orcid.org/0000-0002-6422-4535

REFERENCES

Barstad, I. and Caroletti, G.N. (2013) Orographic precipitation across an island in southern Norway: model evaluation of time-step precipitation. *Quarterly Journal of the Royal Meteorological Society*, 139(675), 1555–1565. https://doi.org/10.1002/qj.2067.

Bentsen, M., Bethke, I., Debernard, J.B., Iversen, T., Kirkevåg, A., Seland, Ø., Drange, H., Roelandt, C., Seierstad, I.A., Hoose, C. and Kristjánsson, J.E. (2012) The Norwegian Earth System Model, NorESM1-M-Part 1: description and basic evaluation. *Geoscientific Model Development Discussions*, 5, 2843–2931. https://doi.org/10.5194/gmdd-5-2843-2012.

Berckmans, J., Woollings, T., Demory, M.E., Vidale, P.L. and Roberts, M. (2013) Atmospheric blocking in a high resolution climate model: influences of mean state, orography and eddy forcing. *Atmospheric Science Letters*, 14(1), 34–40. https://doi.org/10.1002/asl2.412.

Booth, J.F., Naud, C.M. and Willison, J. (2018) Evaluation of extratropical cyclone precipitation in the North Atlantic Basin: an analysis of ERA-Interim, WRF, and two CMIP5 models. *Journal of Climate*, 31(6), 2345–2360. https://doi.org/10.1175/JCLI-D-17-0308.1.

Brayshaw, D.J., Hoskins, B. and Blackburn, M. (2009) The basic ingredients of the North Atlantic storm track. Part I: land–sea contrast and orography. *Journal of the Atmospheric Sciences*, 66(9), 2539–2558. https://doi.org/10.1175/2009JAS3078.1.

Broccoli, A.J. and Manabe, S. (1992) The effects of orography on midlatitude Northern Hemisphere dry climates. *Journal of Climate*, 5(11), 1181–1201. https://doi.org/10.1175/1520-0442(1992)005<1181:TEOOOM>2.0.CO;2.

Bruyère, C.L., Done, J.M., Holland, G.J. and Fredrick, S. (2014) Bias corrections of global models for regional climate simulations of high-impact weather. *Climate Dynamics*, 43(7–8), 1847–1856. https://doi.org/10.1007/s00382-013-2011-6.

Chang, E.K.M. (2009) Diabatic and orographic forcing of northern winter stationary waves and storm tracks. *Journal of Climate*, 22(3), 670–688. https://doi.org/10.1175/2008JCLI2403.1.

Chang, E.K.M., Lee, S. and Swanson, K.L. (2002) Storm track dynamics. *Journal of Climate*, 15(16), 2163–2183. https://doi.org/10.1175/1520-0442(2002)015<02163:STD>2.0.CO;2.

Colin, J., Déqué, M., Radu, R. and Somot, S. (2010) Sensitivity study of heavy precipitation in Limited Area Model climate simulations: influence of the size of the domain and the use of the spectral nudging technique. *Tellus A*, 62(5), 591–604. https://doi.org/10.1111/j.1600-0870.2010.00467.x.

Colle, B.A., Zhang, Z., Lombardo, K.A., Chang, E., Liu, P. and Zhang, M. (2013) Historical evaluation and future prediction of eastern North American and Western Atlantic extratropical cyclones in the CMIP5 models during the cool season. *Journal of Climate*, 26(18), 6882–6903. https://doi.org/10.1175/JCLI-D-12-00498.1.

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., Hólm, E.V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A.P., Monge-Sanz, B.M., Morcrette, J.J., Park, B.K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, 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(656), 553–597. https://doi.org/10.1002/qj.828.

Denis, B., Laprise, R., Caya, D. and Côté, J. (2002) Downscaling ability of one-way nested regional climate models: the Big-Brother Experiment. *Climate Dynamics*, 18(8), 627–646. https://doi.org/10.1007/s00382-001-0201-0.

Déqué, M., Rowell, D.P., Lüthi, D., Giorgi, F., Christensen, J.H., Rockel, B., Jacob, D., Kjellström, E., De Castro, M. and Van Den Hurk, B. (2007) An intercomparison of regional climate simulations for Europe: assessing uncertainties in model projections. *Climatic Change*, 81(Supplement 1), 53–70. https://doi.org/10.1007/s10584-006-9228-x.

Diaconescu, E.P. and Laprise, R. (2013) Can added value be expected in RCM-simulated large scales?. *Climate Dynamics*, 41(7-8), 1769–1800. https://doi.org/10.1007/s00382-012-1649-9.

Diaconescu, E.P., Laprise, R. and Sushama, L. (2007) The impact of lateral boundary data errors on the simulated climate of a nested regional climate model. *Climate Dynamics*, 28(4), 333–350. https://doi.org/10.1007/s00382-006-0189-6.

Glisan, J.M., Jones, R., Lennard, C., Castillo Pérez, N.I., Lucas-Picher, P., Rinke, A., Solman, S. and Gutowski, W.J. (2019) A metrics-based analysis of seasonal daily precipitation and near-surface temperature within seven Coordinated Regional Climate Downscaling Experiment domains. *Atmospheric Science Letters*, 20(5), e897. https://doi.org/10.1002/asl.897.

Hanssen-Bauer, I. (2005) *Regional temperature and precipitation series for Norway: analyses of time-series updated to 2004*. Norwegian Meteorological Institute. Technical Report 15.

Hong, S.Y., Noh, Y. and Dudhia, J. (2006) A new vertical diffusion package with an explicit treatment of entrainment processes. *Monthly Weather Review*, 134(9), 2318–2341. https://doi.org/10.1175/MWR3199.1.

Hoskins, B. and Woollings, T. (2015) Persistent extratropical regimes and climate extremes. *Current Climate Change Reports*, 1(3), 115–124. https://doi.org/10.1007/s40641-015-0020-8.

Iacono, M.J., Delamere, J.S., Mlawer, E.J., Shephard, M.W., Clough, S.A. and Collins, W.D. (2008) Radiative forcing by long-lived greenhouse gases: calculations with the AER radiative transfer models. *Journal of Geophysical Research Atmospheres*, 113(13), 2–9. https://doi.org/10.1029/2008JD009944.

Iversen, T., Bentsen, M., Bethke, I., Debernard, J.B., Kirkevåg, A., Seland, Ø., Drange, H., Kristjánsson, J.E., Medhaug, I., Sand, M. and Seierstad, I.A. (2013) The Norwegian Earth System Model, NorESM1-M – Part 2: climate response and scenario projections. *Geoscientific Model Development Discussions*, 5(3), 2933–2998. https://doi.org/10.5194/gmdd-5-2933-2012.

Jacob, D. and Podzun, R. (1997) Sensitivity studies with the regional climate model REMO. *Meteorology and Atmospheric Physics*, 63(1–2), 119–129. https://doi.org/10.1007/BF01025368.

Køltzow, M., Iversen, T. and Haugen, J.E. (2008) Extended Big-Brother experiments: the role of lateral boundary data quality and size of integration domain in regional climate modelling. *Tellus A*, 60(3), 398–410. https://doi.org/10.1111/j.1600-0870.2008.00309.x.

Kotlarski, S., Keuler, K., Christensen, O.B., Colette, A., Déqué, M., Gobiet, A., Goergen, K., Jacob, D., Lüthi, D., Van Meijgaard, E., Nikulin, G., Schär, C., Teichmann, C., Vautard, R., Warrach-Sagi, K. and Wulfmeyer, V. (2014) Regional climate modeling on European scales: a joint standard evaluation of the EURO-CORDEX RCM ensemble. *Geoscientific Model Development Discussions*, 7(4), 1297–1333. https://doi.org/10.5194/gmd-7-1297-2014.

Lamraoui, F., Booth, J.F. and Naud, C.M. (2018) WRF hindcasts of cold front passages over the ARM Eastern North Atlantic site: A sensitivity study. *Monthly Weather Review*, 146(8), 2417–2432. https://doi.org/10.1175/mwr-d-17-0281.1.

Leduc, M. and Laprise, R. (2009) Regional climate model sensitivity to domain size. *Climate Dynamics*, 32(6), 833–854. https://doi.org/10.1007/s00382-008-0400-z.

Mayer, S., Maule, C.F., Sobolowski, S., Christensen, O.B., Danielsen Sørup, H.J., Sunyer, M.A., Arnbjerg-Nielsen, K. and Barstad, I. (2015) Identifying added value in high-resolution climate simulations over Scandinavia. *Tellus A*, 67(24941), 1–18. https://doi.org/10.3402/tellusa.v67.24941.

Niu, G.Y., Yang, Z.L., Mitchell, K.E., Chen, F., Ek, M.B., Barlage, M., Kumar, A., Manning, K., Niyogi, D., Rosero, E., Tewari, M. and Xia, Y. (2011) The community Noah land surface model with multiparametrization options (Noah-MP): 1. Model description and evaluation with local-scale measurements. *Journal of Geophysical Research*, 116(D12), D12 109. https://doi.org/10.1029/2010JD015139.

Noguer, M., Jones, R. and Murphy, J. (1998) Sources of systematic errors in the climatology of a regional climate model over Europe. *Climate Dynamics*, 14(10), 691–712. https://doi.org/10.1007/s003820050249.

Perkins, S.E., Pitman, A.J., Holbrook, N.J. and McAneney, J. (2007) Evaluation of the AR4 Climate Models' simulated daily maximum temperature, minimum temperature, and precipitation over Australia using probability density functions. *Journal of Climate*, 20(17), 4356–4376. https://doi.org/10.1175/JCLI4253.1.

Pithan, F., Shepherd, T.G., Zappa, G. and Sandu, I. (2016) Climate model biases in jet streams, blocking and storm tracks resulting from missing orographic drag. *Geophysical Research Letters*, 43(13), 7231–7240. https://doi.org/10.1002/2016GL069551.

Poan, E.D., Gachon, P., Laprise, R., Aider, R. and Dueymes, G. (2018) Investigating added value of regional climate modeling in North American winter storm track simulations. *Climate Dynamics*, 50(5–6), 1799–1818. https://doi.org/10.1007/s00382-017-3723-9.

Pontoppidan, M., Kolstad, E., Sobolowski, S. and King, M.P. (2018) Improving the reliability and added value of dynamical downscaling via correction of large-scale errors: a Norwegian perspective. *Journal of Geophysical Research: Atmospheres*, 123(21), 11,875–11,888. https://doi.org/10.1029/2018JD028372.

Pontoppidan, M., Reuder, J., Mayer, S. and Kolstad, E.W. (2017) Downscaling an intense precipitation event in complex terrain: the importance of high grid resolution. *Tellus A*, 69(1), 1271 561. https://doi.org/10.1080/16000870.2016.1271561.

Rasmussen, K.L. and Houze, R.A. (2016) Convective initiation near the Andes in subtropical South America. *Monthly Weather Review*, 144(6), 2351–2374. https://doi.org/10.1175/MWR-D-15-0058.1.

Reuder, J., Fagerlid, G.O., Barstad, I. and Sandvik, A. (2007) Stord Orographic Precipitation Experiment (STOPEX): an overview of phase I. *Advances in Geosciences*, 10, 17–23. https://doi.org/10.5194/adgeo-10-17-2007.

Rocheta, E., Evans, J.P. and Sharma, A. (2017) Can bias correction of regional climate model lateral boundary conditions improve low-frequency rainfall variability?. *Journal of Climate*, 30(24), 9785–9806. https://doi.org/10.1175/JCLI-D-16-0654.1.

Seiler, C., Zwiers, F.W., Hodges, K.I. and Scinocca, J.F. (2017) How does dynamical downscaling affect model biases and future projections of explosive extratropical cyclones along North America's

Atlantic coast?. *Climate Dynamics*, 50(1–2), 677–692. https://doi.org/10.1007/s00382-017-3634-9.

Shin, H.H., Hong, S.Y., Dudhia, J. and Kim, Y.J. (2010) Orography-induced gravity wave drag parametrization in the global WRF: implementation and sensitivity to shortwave radiation schemes. *Advances in Meteorology*, 2010, 1–8. https://doi.org/10.1155/2010/959014.

Simmonds, I. and Lim, E.P. (2009) Biases in the calculation of Southern Hemisphere mean baroclinic eddy growth rate. *Geophysical Research Letters*, 36(1), 1–5. https://doi.org/10.1029/2008GL036320.

Skamarock, W., Klemp, J., Dudhia, J., Gill, D., Barker, D., Duda, M., Huang, X.Y., Wang, W. and Powers, J. (2008) *A description of the advanced research WRF version 3*. National Center for Atmospheric Research. Technical Report: TN-475+STR.

Skamarock, W.C., Klemp, J.B., Duda, M.G., Fowler, L.D., Park, S.H. and Ringler, T.D. (2012) A multiscale nonhydrostatic atmospheric model using centroidal Voronoi tesselations and C-grid staggering. *Monthly Weather Review*, 140(9), 3090–3105. https://doi.org/10.1175/MWR-D-11-00215.1.

Smith, R.B. (1984) A theory of lee cyclogenesis. *Journal of the Atmospheric Sciences*, 41(7), 1159–1168. https://doi.org/10.1175/1520-0469(1984)041<1159:ATOLC>2.0.CO;2.

Smith, R.B. (1986) Further development of a theory of lee cyclogenesis. *Journal of the Atmospheric Sciences*, 43(15), 1582–1602. https://doi.org/10.1175/1520-0469(1986)043<1582:FDOATO>2.0.CO;2.

Sobolowski, S., Gong, G. and Ting, M. (2007) Northern Hemisphere winter climate variability: response to North American snow cover anomalies and orography. *Geophysical Research Letters*, 34(16), 2–6. https://doi.org/10.1029/2007GL030573.

Stephens, G.L., L'Ecuyer, T., Forbes, R., Gettlemen, A., Golaz, J.C., Bodas-Salcedo, A., Suzuki, K., Gabriel, P. and Haynes, J. (2010) Dreary state of precipitation in global models. *Journal of Geophysical Research Atmospheres*, 115(24), 1–14. https://doi.org/10.1029/2010JD014532.

Stohl, A., Forster, C. and Sodemann, H. (2008) Remote sources of water vapor forming precipitation on the Norwegian west coast at 60° – a tale of hurricanes and an atmospheric river. *Journal of Geophysical Research: Atmospheres*, 113(D05), 102. https://doi.org/10.1029/2007JD009006.

Thompson, G., Field, P.R., Rasmussen, R.M. and Hall, W.D. (2008) Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part II: implementation of a new snow parametrization. *Monthly Weather Review*, 136(12), 5095–5115. https://doi.org/10.1175/2008MWR2387.1.

Tiedtke, M. (1989) A comprehensive mass flux scheme for cumulus parametrization in large-scale models. *Monthly Weather Review*, 117(8), 1779–1800. https://doi.org/10.1175/1520-0493(1989)117<1779:ACMFSF>2.0.CO;2.

van Niekerk, A., Shepherd, T.G., Vosper, S.B. and Webster, S. (2016) Sensitivity of resolved and parametrized surface drag to changes in resolution and parametrization. *Quarterly Journal of the Royal Meteorological Society*, 142(699), 2300–2313. https://doi.org/10.1002/qj.2821.

Wang, J. and Kotamarthi, V.R. (2015) High-resolution dynamically downscaled projections of precipitation in the mid and late 21st century over North America. *Earth's Future*, 3(7), 268–288. https://doi.org/10.1002/2015EF000304.

Willison, J., Robinson, W.A. and Lackmann, G.M. (2013) The importance of resolving mesoscale latent heating in the North

Atlantic Storm Track. *Journal of the Atmospheric Sciences*, 70(7), 2234–2250. https://doi.org/10.1175/JAS-D-12-0226.1.

Willison, J., Robinson, W.A. and Lackmann, G.M. (2015) North Atlantic storm-track sensitivity to warming increases with model resolution. *Journal of Climate*, 28(11), 4513–4524. https://doi.org/10.1175/JCLI-D-14-00715.1.

Wilson, C., Sinha, B. and Williams, R.G. (2009) The effect of ocean dynamics and orography on atmospheric storm tracks. *Journal of Climate*, 22(13), 3689–3702. https://doi.org/10.1175/2009JCLI2651.1.

Woollings, T. (2010) Dynamical influences on European climate: an uncertain future. *Philosophical Transactions of the Royal Society A: Mathematical Physical and Engineering Sciences*, 368, 3733–3756. https://doi.org/10.1098/rsta.2010.0040.

Xu, Z. and Yang, Z.L. (2012) An improved dynamical downscaling method with GCM bias corrections and its validation with 30 years of climate simulations. *Journal of Climate*, 25(18), 6271–6286. https://doi.org/10.1175/JCLI-D-12-00005.1.

Xu, Z. and Yang, Z.L. (2015) A new dynamical downscaling approach with GCM bias corrections and spectral nudging. *Journal of Geophysical Research: Atmospheres*, 120(8), 3063–3084. https://doi.org/10.1002/2014JD022958.

Zappa, G., Shaffrey, L.C. and Hodges, K.I. (2013) The ability of CMIP5 Models to simulate North Atlantic extratropical cyclones. *Journal of Climate*, 26(15), 5379–5396. https://doi.org/10.1175/JCLI-D-12-00501.1.

Zhang, C., Wang, Y. and Hamilton, K. (2011) Improved representation of boundary layer clouds over the Southeast Pacific in ARW-WRF using a modified Tiedtke cumulus parametrization scheme. *Monthly Weather Review*, 139(11), 3489–3513. https://doi.org/10.1175/MWR-D-10-05091.1.