

Assessing mid-latitude dynamics in extreme event attribution systems

Daniel Mitchell¹ · Paolo Davini² · Ben Harvey³ · Neil Massey¹ · Karsten Haustein¹ · Tim Woollings⁴ · Richard Jones^{5,1} · Fredi Otto¹ · Benoit Guillod¹ · Sarah Sparrow⁶ · David Wallom⁶ · Myles Allen^{1,4}

Received: 22 January 2016 / Accepted: 2 August 2016 / Published online: 23 August 2016
© The Author(s) 2016. This article is published with open access at Springerlink.com

Abstract Atmospheric modes of variability relevant for extreme temperature and precipitation events are evaluated in models currently being used for extreme event attribution. A 100 member initial condition ensemble of the global circulation model HadAM3P is compared with both the multi-model ensemble from the Coupled Model Inter-comparison Project, Phase 5 (CMIP5) and the CMIP5 atmosphere-only counterparts (AMIP5). The use of HadAM3P allows for huge ensembles to be computed relatively fast, thereby providing unique insights into the dynamics of extremes. The analysis focuses on mid Northern Latitudes (primarily Europe) during winter, and is compared with ERA-Interim reanalysis. The tri-modal Atlantic eddy-driven jet distribution is remarkably well captured in HadAM3P, but not so in the CMIP5 or AMIP5 multi-model mean, although individual models fare better. The well known underestimation of blocking in the Atlantic region is apparent in CMIP5 and AMIP5, and also, to a lesser extent, in HadAM3P. Pacific blocking features are well produced in all modeling initiatives. Blocking duration is biased towards models reproducing too many short-lived events in all three modelling systems. Associated storm

tracks are too zonal over the Atlantic in the CMIP5 and AMIP5 ensembles, but better simulated in HadAM3P with the exception of being too weak over Western Europe. In all cases, the CMIP5 and AMIP5 performances were almost identical, suggesting that the biases in atmospheric modes considered here are not strongly coupled to SSTs, and perhaps other model characteristics such as resolution are more important. For event attribution studies, it is recommended that rather than taking statistics over the entire CMIP5 or AMIP5 available models, only models capable of producing the relevant dynamical phenomena be employed.

Keywords Event attribution · Dynamics · Mid-latitudes · Extreme

1 Introduction

Attributing the changing probability of specific extreme events to human induced climate change is an emerging field. The ability to quantify human influences on events plays a vital role in academic and societal understanding of climate change impacts. It is recognised not only in international climate change initiatives (IPCC, AR5) but also in studies addressing wider atmospheric interactions (Hoskins and Woollings 2015).

While there are multiple methods used for event attribution (Uhe et al. 2016, in review), two of the most widely used are distributed computing under the Weather@home framework (WAH) (Massey et al. 2014) and analysis of the CMIP5 ensemble (Lewis and Karoly 2013). CMIP5 employs the most up-to-date climate models currently available, whereas the WAH methodology employs a climate model that was state-of-the-art 15 years ago. The use of an older model is important, because the climate

✉ Daniel Mitchell
mitchell@atm.ox.ac.uk

¹ Environmental Change Institute, Oxford University, Oxford, UK

² Laboratoire de Météorologie Dynamique/IPSL, Ecole Normale Supérieure, Paris, France

³ NCAS, Department of Meteorology, University of Reading, Reading, UK

⁴ AOPP, Oxford University, Oxford, UK

⁵ Met Office Hadley Centre, Exeter, UK

⁶ OeRC, Oxford University, Oxford, UK

experiments are run on volunteer personal computers, rather than conventional supercomputers, and so computing power and memory is at a premium. However, using an older model may have disadvantages, and if the model cannot capture the event of interest for the correct reasons this must be reflected in the conclusions of the study.

“All models are wrong, but some models are useful” (Box 1987) and it is up to the scientific community to address how useful the model is for event attribution studies. Extreme events can be impacted through direct thermodynamical changes due to increased CO₂, which are relatively well captured by our latest GCMs (Palmer 2013), or indirect changes from the projection of thermodynamics on dynamical modes. It is well known that the latest coupled atmosphere-ocean models do not reproduce aspects of the relevant dynamical regimes for extreme events, for instance jet latitude variability, and European winter blocking (Davini et al. 2012; Zappa et al. 2013; Anstey et al. 2013; Harvey et al. 2014). Therefore the use of an older model for event attribution, as in WAH, could be called into question. This is especially true as the WAH model poorly resolves a number of key atmospheric regions, including the boundary layer which is important for land-surface feedbacks (Jaeger and Seneviratne 2011), and the stratosphere, which is important for winter mid-latitude circulation patterns (Mitchell et al. 2013). The comparison is not as clear-cut as suggested, because the WAH methodology employs prescribed SSTs, whereas the latest CMIP5 models have a coupled ocean. SSTs are known to affect surface baroclinicity and heat fluxes (Inatsu and Hoskins 2004), so it is reasonable to assume they may impact onto certain dynamical modes. Indeed, studies with the UK Met Office Hadley centre models (HadGEM) show evidence that SST biases can lead to atmosphere biases (Keeley et al. 2012; Scaife et al. 2011).

In this study we address the adequacy of event attribution systems directly, by comparing the simulation of relevant dynamical modes in the CMIP5 model-ensemble and the WAH initial condition ensemble. A systematic comparison of all diagnostics with the atmosphere-only equivalents of CMIP5 (AMIP5) is also provided.

The paper is structured as follows. Section 2 describes the data and methods used for the diagnostics. CMIP5, AMIP5 and HadAM3P are evaluated in Sect. 3. The implications for event attribution studies are discussed in Sect. 4, and the analysis is summarized in Sect. 5.

2 Data and methodology

2.1 Models

We make use of the 20 CMIP5 models (Taylor et al. 2012) with the relevant data available to calculate the diagnostics

needed for this analysis (see subsequent sections). They are, along with their AMIP5 counterparts; bcc-csm1-1, BNU-ESM, CanESM2, CCSM4, CMCC-CM, CMCC-CMS, CMCC-CESM, CNRM-CM5, EC-Earth, FGOALS-g2, GFDL-CM3, GFDL-ESM2M, HadGEM2-CC, IPSL-CM5A-LR, IPSL-CM5A-MR, MIROC5, MPI-ESM-LR, MRI-CGCM3, MRI-ESM1 and NorESM1-M. So as not to give more weight to any model, we take only the first ensemble member of each. For the AMIP5 simulations, the SSTs are taken from observations derived by merging data from the HadISST1 data set and the National Oceanic and Atmospheric Administration (NOAA) weekly optimum interpolation (OI) SST analysis (Hurrell et al. 2008).

To fully understand extreme weather events a large number of ensemble members must be simulated (Sippel et al. 2015). The WAH project allows for distributed computer power to run very large numbers of ensemble members (Massey et al. 2014) by allowing anyone to run climate simulations on their home computer, essentially turning PCs into climate simulators. The WAH setup currently employs the atmosphere only N96 model with 19 levels in the vertical (L19), HadAM3P, and the science of extreme climate events has seen major advances through this system (Stainforth et al. 2005; Pall et al. 2011). The use of the HadAM versions of the model is required, because the models were built to be very memory efficient and so can be run on current day home PCs (as is the nature of the WAH project). An initial condition perturbation of each single year within the period 1985–2010 is performed 100 times. The 1985–2010 period is used as it is the period covered by the SST driving data (Massey et al. 2014). The starting conditions for each year are from perturbations of a single simulation that was run continuously over the period 1984–2010, thereby allowing for adequate spin up simulations of 1-year.

All model forcing data for the WAH model are described in Massey et al. (2014), and are largely the same as those recommended for the CMIP5 initiative.

The analysis is compared against the most recent European Centre for Medium-range Weather Forecast (ECMWF) ERA-I reanalysis (Dee et al. 2011), and interpolated to map onto the HadAM3P grid unless otherwise specified.

2.2 Jet latitude index

We follow the methodology used in Woollings et al. (2010) and the modifications employed by Anstey et al. (2013) to identify the lower level eddy-driven jet. Daily-mean zonal wind at 850 hPa is used, and is smoothed by calculating a 5-day running mean. This is then zonally averaged over the 0–60W region for the Atlantic basin. The location of the maximum of this quantity between 15N and 75N is defined as the jet latitude.

2.3 Blocking diagnostic

We use the bi-dimensional blocking definition of Davini et al. (2012), which is an extension on the uni-dimensional definition by Tibaldi and Molteni (1990). The method is based on reversal of the meridional gradient of 500 hPa daily geopotential height. Data are initially interpolated on a common 2.5×2.5 degrees grid. Then, for every grid point of coordinates (λ_0, Φ_0) we define:

$$GHGS(\lambda_0, \Phi_0) = \frac{Z500(\lambda_0, \Phi_0) - Z500(\lambda_0, \Phi_S)}{\Phi_0 - \Phi_S}, \quad (1)$$

$$GHGN(\lambda_0, \Phi_0) = \frac{Z500(\lambda_0, \Phi_N) - Z500(\lambda_0, \Phi_0)}{\Phi_N - \Phi_0} \quad (2)$$

where Φ_0 ranges from 30°N to 75°N and λ_0 ranges from 0° to 360° . $\Phi_S = \Phi_0 - 15^\circ$, $\Phi_N = \Phi_0 + 15^\circ$.

Therefore an instantaneous blocking (IB) is identified if:

$$GHGS(\lambda_0, \Phi_0) > 0 \quad GHGN(\lambda_0, \Phi_0) < -10 \text{ m}/^\circ\text{lat} \quad (3)$$

If the IB is larger than 15° of longitude, the diagnostic is deemed to be of sufficient spatial scale to be considered large scale blocking (i.e. larger than the Rossby deformation radius). Finally, a Blocking Event (or hereafter simply block) is defined at each grid box if, within a box 2.5° latitude and 5° longitude either side of the grid point, IB persists for at least 5 days. Further details on the blocking detection scheme may be found in ref Davini et al. (2012).

2.4 Storm track calculation

The measure of storm activity used here is the standard deviation of highpass-filtered daily-mean mean sea level pressure, providing a measure of the synoptic scale activity (Hoskins and Hodges 2002). The filter used is the daily difference filter of Chang et al. (2012), which admits most power in the 2–8 day band. All data is regridded onto a common n32 Gaussian grid prior to this calculation to allow a fair comparison. This is the diagnostic used in ref Harvey et al. (2014) for CMIP5 models, and therefore allows for a direct comparison.

3 Results

The study of dynamical modes from a super-ensemble framework is a new direction for extreme event attribution, as traditionally it is the statistics of climate that are analysed. Here, we explicitly consider leading modes of atmospheric variability that have an impact on extreme events at mid-Northern latitudes.

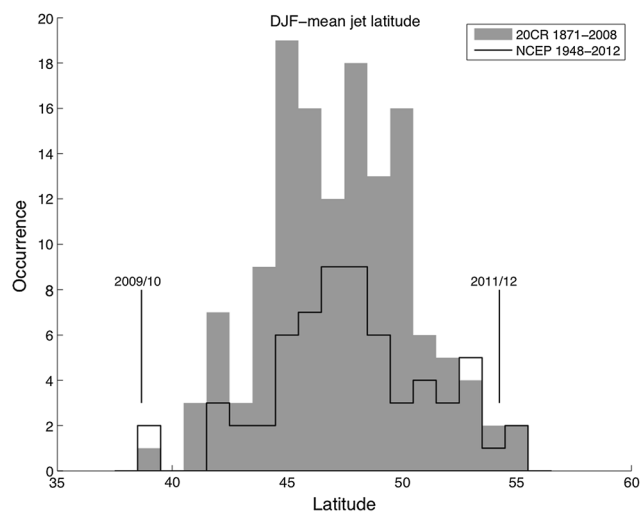


Fig. 1 Winter-averaged jet latitude for two different reanalyses spanning from 1871–2012. Two particularly extreme winters over Europe are also marked, 2009/2010 and 2011/2012

3.1 Jet stream

Most of the meteorological and climate patterns in the extra tropics are associated with the jet streams, which are stronger over ocean basins and so can lead to large model biases in those regions. A persistent jet location in winter can lead to extreme weather. For instance, Fig. 1 shows distributions of the average winter (DJF) location of the Atlantic Eddy driven jet.¹ The extremely persistent southerly jet throughout winter led to the extremely cold winter of 2009/2010 over Europe due to cold air being advected equatorward from high latitudes. In contrast, the extremely persistent northerly jet position of 2011/2012 led to a very dry winter, particularly over the south of England, because less moisture is picked up over the Atlantic. These winters are marked on Fig. 1 and are clearly at either extreme of the distribution.

On daily timescales the jet location also gives rise to extreme weather, and (Woollings et al. 2010) showed that the latitudinal position of the eddy-driven component of the jet in the Atlantic had a tri-modal structure, with each of the tri-modal regions possibly giving rise to different extreme weather patterns. This structure is poorly reproduced in both CMIP-3 models (Hannachi et al. 2013) and CMIP5 models (Anstey et al. 2013), although there is evidence that the latest Hadley Centre model (HadGEM3) is able to capture this (Williams et al. 2015). For CMIP5, in both the Atlantic and Pacific basins, the time-mean jet latitude was

¹ Note we use 20CR and NCEP reanalysis in this figure to build a large jet latitude sample.

1° too equatorward, on average (Anstey et al. 2013). Here, we use exactly the same metric as in these studies to calculate the Jet Latitude Index (JLI, see Sect. 2).

HadAM3P reproduces the distribution of the Atlantic Jet Latitude remarkably well (Fig. 2a). The tri-modal structure observed in reanalysis is clearly represented in all model ensemble members, with the jet located predominantly in its central position, then in its northern position, then in its southern position. The peak of these modes are located at 45N, 58N and 37N, respectively, all of which

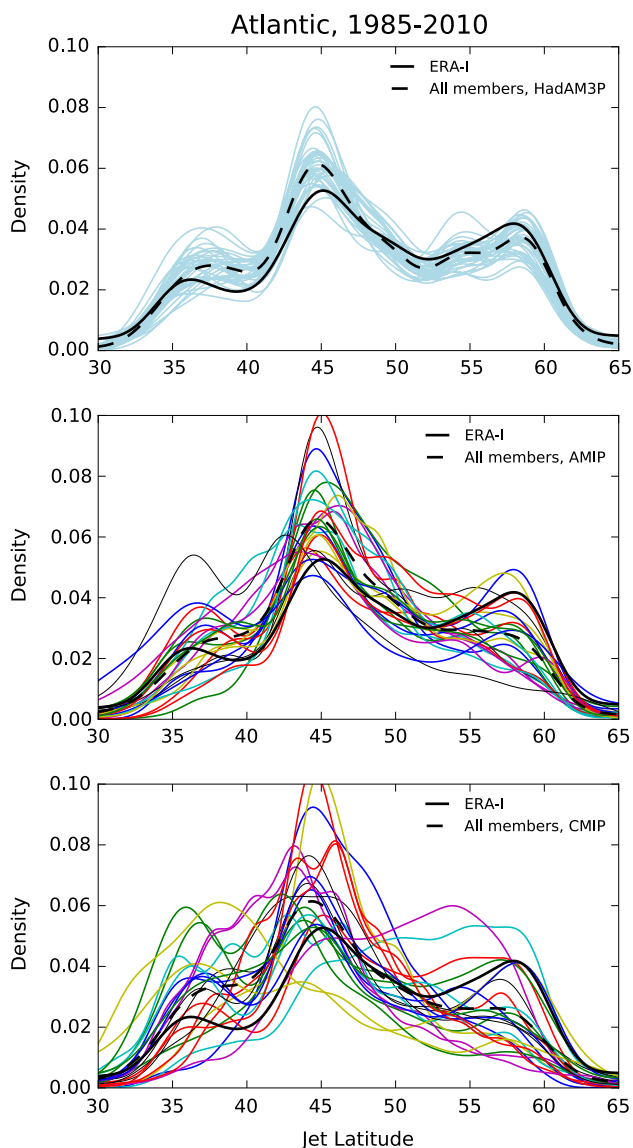


Fig. 2 Winter jet latitude location over the Atlantic for (top) HadAM3P, (middle) AMIP and (bottom) CMIP. The solid black line shows a Kernel estimate of ERA-I. The dashed black line shows the same but for the 3 different modelling initiatives. Thin lines show individual model simulations. In the top panel they are all ensemble members of HadAM3P, in the middle panel they are individual AMIP models, and in the bottom panel they are individual CMIP models

are within 1° of the reanalysis estimates of jet latitude calculated over the same period. The spread in ensembles captures the observed distribution, especially given the observed tri-modal structure has a measure of decadal variability in it Woollings et al. (2014). In comparison to the AMIP5 and CMIP5 jet latitude distributions (Fig. 2b, c), HadAM3P performs particularly favourably. This is interesting because some studies have claimed that in order to adequately capture this type of regime behaviour, very high resolution models are needed (Dawson and Palmer 2014). HadAM3P has a resolution of N96, and there are numerous examples of higher resolution CMIP5 models that do not simulate this regime behaviour. The reason for why HadAM3P reproduces the tri-modal structure of the jet so well would require a more in-depth analysis, but it could be linked with balancing the various components of drag from different schemes, such as sub-grid orography and gravity waves.

The multi-model mean (MMM) of the AMIP5 and CMIP5 models shows very similar distributions to each other, both of which hint at the tri-modal structure observed in reanalysis. However, this tri-modality is really an averaging effect, with few of the CMIP5 and AMIP5 models simulating all three jet locations accurately in terms of occupancy and latitude. This result was also found in Anstey et al. (2013), where they discuss this in more detail.

3.2 Blocking

The eddy driven jet is important in its own right for extreme weather Santos et al. (2013), but it can also give rise to preferred regimes of blocking (over the Atlantic in this instance; Woollings et al. 2010; Davini et al. 2012). We first consider the mean biases in blocking frequency. The filled contours in the first three panels of Fig. 3 show the climatological winter (DJF) blocking frequency for HadAM3P, AMIP5 and CMIP5, respectively. The unfilled contours show the same quantity but for ERA-I (note that the blocking characteristics are very similar for other reanalyses (Barnes et al. 2014)). The bottom three panels show the biases. The spatial structure of blocking is well captured in these diagnostics, although clearly all three modeling initiatives have too little blocking in the European and Pacific sectors. Individual CMIP5/AMIP5 models show very different spatial structures, but the multi-model mean presented here compares favorably with reanalysis, a result also found in Anstey et al. (2013). In that sense, the HadAM3P ensemble fares particularly well, as it is a single model, and all blocking features are present, although there are also examples of CMIP5 models where blocking biases are lower (e.g. MIROC5; not shown). One likely cause for the European blocking bias is the resolution of orography in the models, which (Berckmans et al. 2013) showed to be

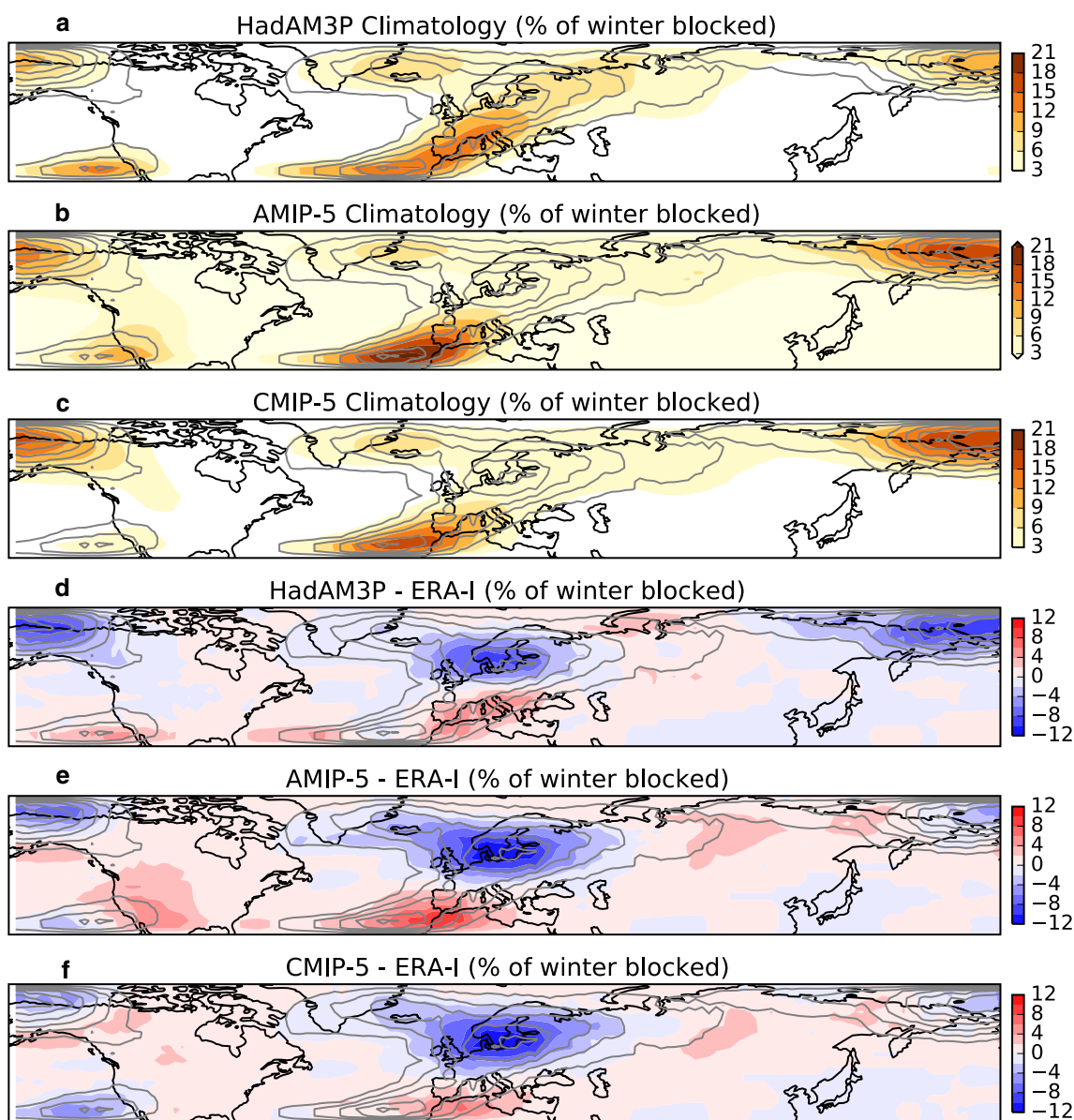


Fig. 3 The climatology of winter blocking frequency from **a** HadAM3P, **b** the AMIP5 ensemble and **c** the CMIP5 ensemble. The bias in winter blocking frequency from **d** HadAM3P, **e** AMIP5 and **f** CMIP5 with respect to ERA-I. *Grey contours* show the climatol-

ogy of ERA-I blocking frequency and are identical in all panels, with contour spacing at 3 % intervals. The climatology is defined over the 1985–2010 period

important. Note that a good reproduction of the JLI does not necessarily imply a good reproduction of blocking (Davini and Cagnazzo 2014).

The magnitude of the biases in HadAM3P are reduced compared with CMIP5 and AMIP5. At its peak, over Europe, ERA-I shows that between 10 and 15 % of the wintertime is blocked, but the frequency in HadAM3P is about half this, and in CMIP5/AMIP5, the frequency is only about a third of this. Over Greenland, the HadAM3P ensemble simulates blocking reasonably well, but there is still a slight negative bias in blocking for the CMIP5 and

AMIP5 ensemble means. Greenland blocking is associated with the southern jet regime (Woollings et al. 2010), which is better captured in the HadAM3P ensemble than the CMIP5 ensemble, and therefore agrees well with the biases presented here. Over the Pacific, all modeling initiatives capture the structure of the blocking well, but they are all biased negative, with HadAM3P not performing as well as the CMIP5 or AMIP5 multi model means.

To understand more fully the variability of blocking in the individual models and ensembles, the ensemble spread of blocking frequency at 60N is compared with

ERA-I (Fig. 4) shows the percentage of winter blocked at all longitudes for the climatology (1985–2010) of ERA-I (green) and each modeling initiative (blue), and is identical to a slice at 60N in Fig. 3. The inter-annual spread in ERA-I (green dashed), shows that extreme winters can lead to 30 % of the season blocked over Europe, with the peak occurring at the Greenwich Meridian. Clearly high levels of blocking are also observed over the Greenland and Pacific sectors. Over all these sectors HadAM3P captures the variability of blocking well, and certainly better than the mean alone suggests, but the distributions are more heavy-tailed in HadAM3P than ERA-I. For instance, at the Greenwich Meridian, the ERA-I data are approximately Gaussian, but the HadAM3P data are skewed (not shown).

Over the European sector, the peak blocking frequency in HadAM3P is around 5 ° eastward of ERA-I, which is small compared with an order of magnitude larger bias

(50 °) in the AMIP5 and CMIP5 models (Fig. 4 panels b and c). Clearly the AMIP5 and CMIP5 models have a real issue in capturing European blocking. At the Greenwich Meridian, only the most extreme years in AMIP5 and CMIP5 data have a similar frequency of blocking as the mean of ERA-I (15 % of the winter blocked). Conversely, over the Eastern Pacific, the CMIP5 and AMIP5 models produce too high a frequency of blocking. In general, climate models tend to produce a maximum of Euro-Atlantic blocking farther east, over Western Russia. This is related to jet dynamics and it is clearly visible in Fig. 4.

The analysis so far has concentrated on blocking frequency, but the duration of blocking events can be just as important. Blocking duration has been given far less attention in the literature than blocking frequency, primarily because sample sizes are not large enough, hence a multi-thousand ensemble member setup like WAH is

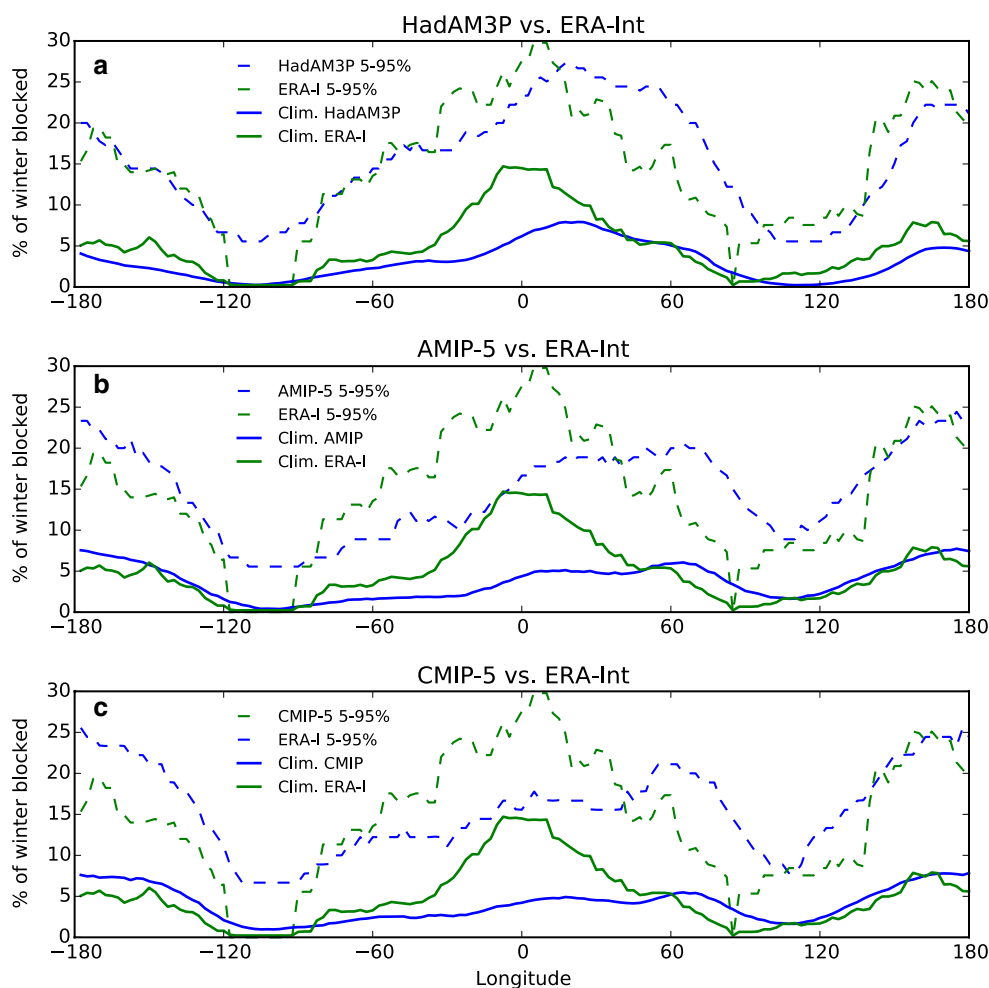


Fig. 4 Mid-latitude blocking frequency expressed as a percentage of winter blocked at 60N over all longitudes. *Blue dashed lines* show the 95 % range in inter-annual, inter-ensemble variability for (top) HadAM3P, (middle) AMIP5 and (bottom) CMIP5 (see legend). *Green*

dashed lines show the same but for ERA-I. Note the 5 % range is not plotted as it is zero in both cases (i.e. over 5 % of winters have no blocking). The *solid lines* show the winter climatology

ideal for this analysis. Figures 5 and 6 show the blocking duration for the (a) European sector, (b) Greenland sector and (c) Pacific sector. The most extreme blocking events can last up to 3 weeks in the reanalysis over Europe and the Pacific, and up to 2.5 weeks over Greenland. There are clear examples in HadAM3P, AMIP5 and CMIP5 where this length duration of events is captured, although there is a general underestimate of events with extremely long durations over Europe. Dunn-Sigouin and Son (2013) found a similar result for the annual mean blocking duration in CMIP5 models. HadAM3P clearly does as well as CMIP5 and AMIP5 on average, but all models show a tendency to have too many short events and not enough long events.

3.3 North Atlantic Oscillation

In the Atlantic, the positive and negative phases of the North Atlantic Oscillation (NAO) are equally as important

as blocking modes for synoptic variability (Dawson et al. 2012; Dawson and Palmer 2014). As we are comparing the variability of the NAO between different data sets, we choose not to use empirical orthogonal functions (EOFs), and we choose not to normalize it. Instead we calculate the NAO index as an area-weighted average over Iceland minus an area-weighted average over the Azores (Stephenson et al. 2006). This has the disadvantage that if the centers of action of the NAO are not the same across data sets, or if they are non-stationary in time (Lu and Greatbatch 2002), the magnitude of the dipole may not be well captured. However, the advantage is that the variability can be compared between different models, and give a physically meaningful interpretation of the pressure gradient over the Atlantic. The first EOF in the CMIP5 models also leads to inconsistent modes of variability with reanalyses (Davini and Cagnazzo 2014).

Figure 7 (top) shows the distributions of daily-mean winter NAO for HadAM3P, ERA-I and (left) AMIP5 and (right) CMIP5. The thin lines show individual members of

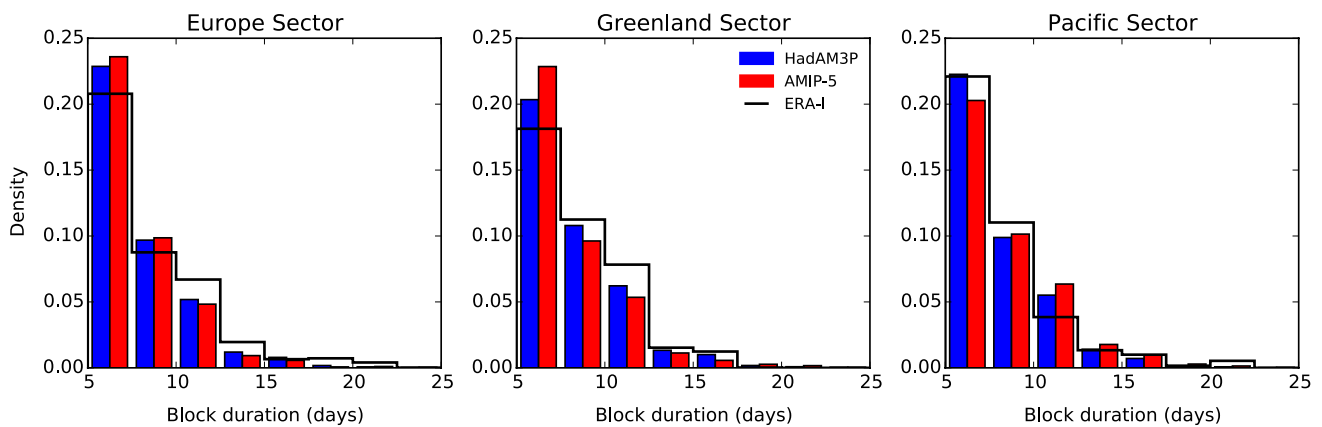


Fig. 5 Duration of winter blocking events over the period 1985–2010 in (black) ERA-I, (blue) HadAM3P and (red) AMIP5. Blocking is divided into the three principal regions of activity; Europe (15W–25E, Greenland (15W–60W) and the Pacific (150W–150E)

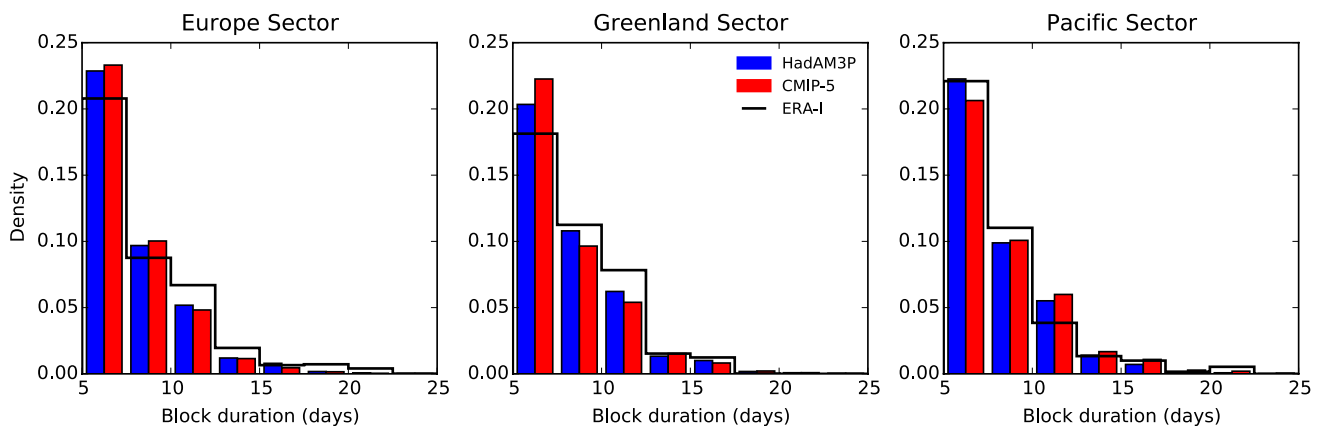


Fig. 6 As in previous figure but for CMIP5

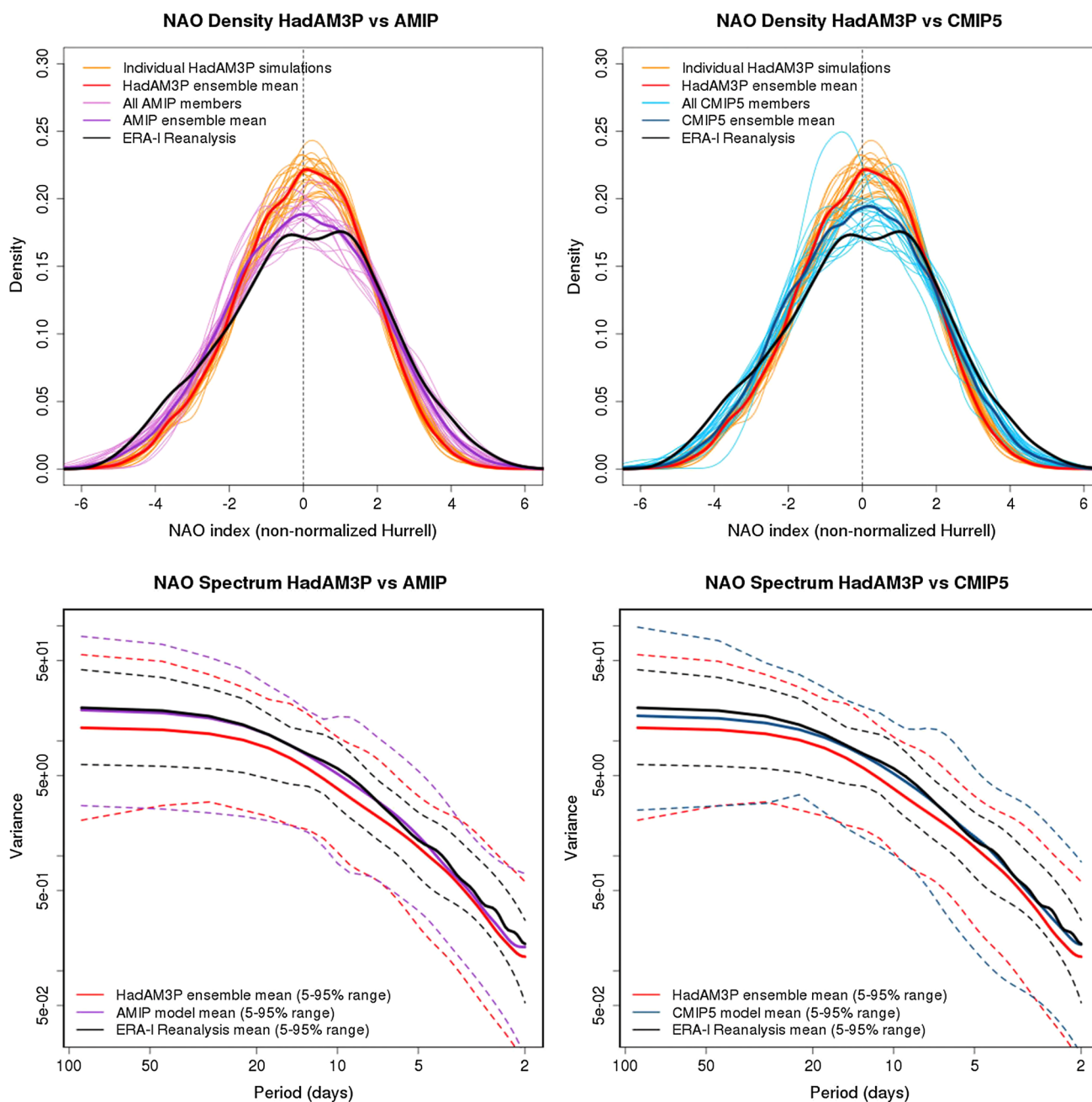


Fig. 7 Daily NAO variability during winter (DJF) expressed as (*top*) a PDF, and (*bottom*) a power spectral density. (*left*) A comparison of HadAM3P with AMIP5, and (*right*) a comparison of HadAM3P with

CMIP5. See legend for descriptions of *coloured lines*. The 5–95 % spread is plotted for the power spectra

the ensemble. Clearly most models fail to capture the most extreme negative and positive NAO events (more extreme than ± 3 hPa), suggesting some absence or poor representation of a physical process during winter. It is likely that models fail to capture these extreme events for different reasons (e.g. the presence or not of well resolved stratospheric connections, Mitchell et al. 2013; Seviour et al. 2016). However, it does seem that the NAO power across

all timescales is well captured by HadAM3P, AMIP5 and CMIP5 (Fig. 7, bottom).

3.4 Storm tracks

Extratropical cyclones propagate across the North Atlantic bringing extreme winds and rainfall to western Europe, however, the nature of the propagation is dependant on the

phase of the NAO. It is therefore insightful to assess the storm tracks, which are known to have large biases over the Atlantic in current generation climate models (Harvey et al. 2014). Figure 8 (left) shows the mean storm track magnitude for (top) HadAM3P, (middle) CMIP5 and (bottom) AMIP5 in the coloured contours, with the same quantity from ERA-I overlaid as gray line contours. The biases are shown in the right panels. The largest biases in HadAM3P (Fig. 8 top, left) are at high latitudes (poleward of 70N), where storm track density is low, and as such the biases are not so important. Note also that MSLP is biased to high

latitudes relative to wind. At mid-latitudes the storm track biases are small, especially over the Euro-Atlantic region. Crucially, the storm tracks are not too zonal, which is often the case in current generation climate models (Harvey et al. 2014; Zappa et al. 2013), including the AMIP5 and CMIP5 models analyzed here (Fig. 8 middle and bottom). For AMIP5 and CMIP5, perhaps the largest biases in storm track density are over the western coast of America and Canada, with the biases being larger in AMIP5 over CMIP5. For HadAM3P, there is no bias in this region, although there is a small negative bias over Eastern America.

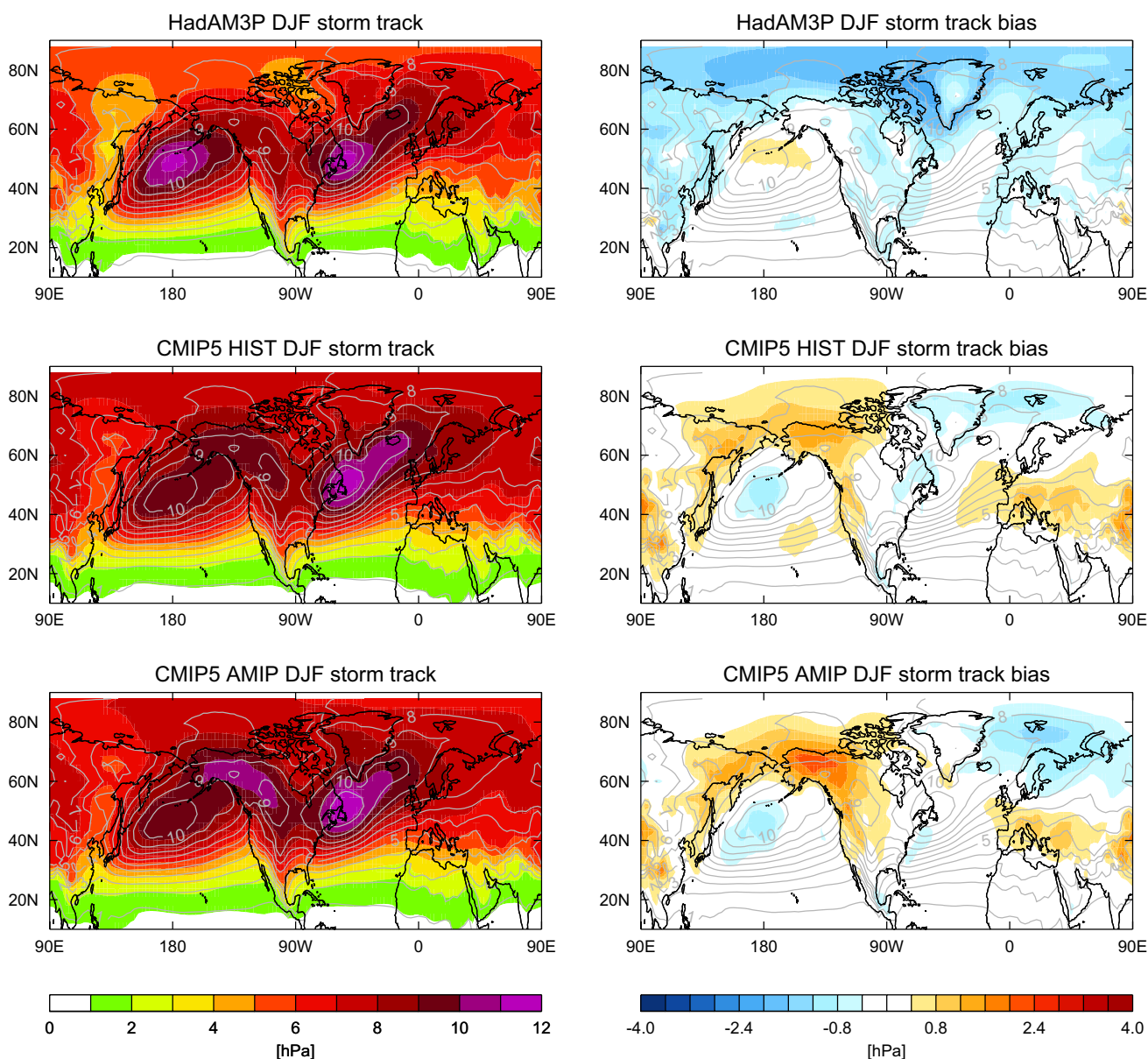


Fig. 8 (left) The mean DJF storm track for (top) HadAM3P, (middle) CMIP5 and (bottom) AMIP5. (right) The bias relative to ERA-Interim (years 1979–2013). Grey contours show the ERA-Interim

climatology. The storm track measure is the variance of time-filtered MSLP, where a daily-difference time filter is used, and the variances are converted to std dev for plotting (units: hPa)

4 Implications for event attribution studies

In the previous section the relevant dynamics relating to extreme weather events was assessed, and it is reasonable to ask, given the models' performances, what are the implications for past and future event attribution studies.

Often when changes in climate are considered the dynamical and the thermodynamical components are spoken of separately. Strictly this dichotomy is incorrect, because in general a change in dynamics can only be bought about by changes in thermodynamics² (yet the inverse is not true), so there is no such thing as a purely dynamical change. It is, however, sometimes convenient to think of changes as either dynamical, or thermodynamical, and this is especially true for regional climate change.

To take a simple theoretical example of the dynamics-to-thermodynamics interplay; for instance, as the troposphere warms, the depth of the troposphere, H , increases. This decreases the length scale of streamlines, L , for meridional displacements of stationary waves due to orography of depth h , according to the following relationship:

$$\Delta L \sim \frac{fh}{\beta} \cdot \frac{1}{\Delta H}, \quad (4)$$

where f is the Coriolis parameter and β is the beta parameter. So the thermodynamic expansion (ΔH) leads to a change in wave dynamic properties (ΔL). Hoskins and Woollings (2015) present similar theoretical arguments for changes in the wavelength and response to low-level heating of stationary waves. Given that for many dynamical examples the theory of how these may change under climate change is well established, it is natural that event attribution will use the theory and focus more and more on changes in extreme events from a dynamical perspective, especially for low signal-noise processes which may lead to large regional climate change (e.g. changes in blocking). For instance ref. Christidis and Stott (2015) attributed changes in Z500, ref. Mitchell et al. (2016) attributed changes in summer blocking, and ref. Schaller et al. (2016) attributed changes in regime occupancy.

In this study we have assessed whether models used in event attribution studies are capable of simulating relevant modes of variability for extreme events. Focusing on potential issues in attribution studies, notably there are still biases with models reproducing the mean frequency of winter blocking over Europe, but, reassuringly, HadAM3P has clear examples of ensemble members that are as extreme as the observed years. This is not the case for CMIP5, however, and it is recommended that event

attribution studies of winter blocking should exclude models that poorly represent blocking in this region, although one would worry about sample sizes for such an analysis. Likewise the duration of blocking events over Europe is slightly underestimated, which would have implications for extended cold snaps and as such health implications. Both HadAM3P and CMIP5 models have examples of long-duration events, however, without knowing explicitly what causes this bias, event attribution statements in this region are less reliable than in other regions associated with these phenomena.

Likewise storm track biases were varied in CMIP5 and HadAM3P, and event attribution studies should be cautious when studying extreme precipitation events in these large bias areas. For instance, in HadAM3P there is a large negative bias in storm tracks at high latitudes (poleward of 75N). However, this is also a region of low storm track density, as well as being sparsely populated region of the world. As such, it is less likely that localised extreme events are relevant in that region.

Certain extreme events are linked to jet location and magnitude. For instance the 2013/2014 flooding in the UK was linked to a strengthened jet causing more persistent storms to track across the Atlantic (Huntingford et al. 2014). In HadAM3P the distribution of Atlantic jet location is very well simulated, implying that attribution statements about this type of event are more reliable. However, this is not the case for CMIP5 models in general, where the jet locations are not well captured, and also the Euro-Atlantic storm tracks are too zonal. As an example, Fig. 9 shows return period curves of zonal wind at 850 hPa over the North Atlantic for (blue) HadAM3P

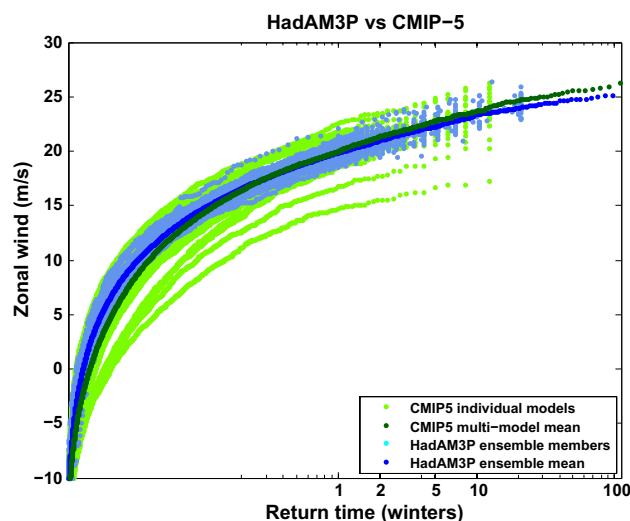


Fig. 9 Return time curves for daily winter zonal wind at 850 hPa over the Atlantic (0–60W, 55–60N) for the period 1985–2010. The comparison is between HadAM3P and CMIP5 models

² Counterexamples exist in palaeoclimate, such as the mechanical forcing of the ice sheet, but we are not aware of any in the anthropogenic context.

and (green) CMIP5 data. Such analyses are widely used in attribution studies by comparing the return period of a particular event under different forcing scenarios. Taking an example event of a return time of 1 winter, we can see that the two methods give different spreads in return values. For the HadAM3P method this corresponds to zonal winds of 18–22 ms^{-1} , but for the CMIP5 method it corresponds to 13–23 ms^{-1} . It is however reassuring to note that the multi-model mean of the HadAM3P and CMIP5 ensemble are very similar. Clearly there are at least two of the CMIP models which do not simulate zonal winds as extreme as the others, calling into question whether these models would be suitable for this analysis. Given that the HadAM3P model simulates the location and magnitude of the eddy driven jet well with respect to ERA-I reanalysis (Fig. 2), this is an example of where we may have more confidence in attribution with one event attribution system over another, and has particular implications for, e.g. wind storms impacting Western Europe. It is therefore recommended that models be selected before event attribution analyses are performed, based on how well they reproduce the required underlying mechanisms. The choice of what classes as 'good' is of course subjective, but model metrics have been developed and various decision flows established for model inter-comparison projects (Waugh and Eyring 2008; Knutti et al. 2010; Knutti 2010). The model selection employed by attribution studies needn't be any different from the model inter-comparison projects, although the model selection criteria should be based on extreme event-relevant metrics.

5 Summary

In this paper, relevant modes of atmospheric variability for winter extreme temperature and precipitation events have been assessed in models used for extreme event attribution studies (primarily HadAM3P and CMIP5 models). Clear implications for past and future event attribution have been laid out.

We have concentrated on winter mid-latitudes in the Northern Hemisphere, the focus of many event attribution studies. While much analysis has been performed on the atmospheric dynamical modes in CMIP5 models, very little has been performed in any systematic way for HadAM3P. As a by-product to this assessment, we have implicitly assessed:

1. The interplay between model uncertainty (from CMIP5) and initial condition uncertainty (from HadAM3P).
2. The constraint that SSTs have on atmospheric modes of variability.

Conclusions specific to the two ensembles are as follows:

HadAM3P As a whole, HadAM3P simulates the modes of variability considered here well, especially in comparison with CMIP5 models either in coupled or AMIP mode. Given this skill, and given the ability under WAH to reproduce super-ensembles ($\sim 10,000$), the potential to further understand the dynamics of extremes is particularly noteworthy.

The tri-modal structure of the Atlantic jet is remarkably well reproduced in HadAM3P, and is notoriously hard to capture in our latest generation models (Anstey et al. 2013). The variability of blocking frequency is well captured in all regions, with many examples of extreme years in HadAM3P of similar magnitude to the extreme years in ERA-I reanalysis. However, there is a bias in the mean blocking frequency over Europe, linked to too few high-duration blocking events in HadAM3P (although examples of extreme-duration blocking events are still found).

The storm track activity is well reproduced in the model, and the well known too zonal storm track issue in models is not apparent in HadAM3P. However, there are small negative biases over Western Europe and Eastern America.

CMIP5 The multi-model mean of CMIP5 fares less well in simulating the relevant atmospheric modes for extreme events than HadAM3P. The tri-modal structure of the Atlantic jet is not reproduced in many of the models, and even then only vaguely resembles the ERA-I calculated distribution. Extreme blocking events (in terms of duration of winter blocked) over the European sector are notably absent, with the top 5 % of blocking events from the CMIP5 models only matching the mean of ERA-I. Over the Pacific the CMIP5 models perform far better, and represent the variability and duration of winter blocking well (and more accurately than HadAM3P). Finally, on average, the CMIP5 models have too high a density of storm tracks over Western America and Canada, and too zonal a jet over the Euro-Atlantic region.

Our results show a need for event attribution studies to perform some initial 'suitability' tests on their chosen event attribution system. It may be that one system is more suitable than another, or it may be that no systems can capture the relevant dynamics of the event, in which case attribution may not be possible. For systems where multiple different models are used, such as the CMIP5 methodology, some models may be found to be inadequate, and should therefore be excluded from the analysis.

Acknowledgments DMM was supported by the NERC ACE-Africa project. RGJ was supported by the UK Joint Department for Energy and Climate Change (DECC), Department for Environment, Food and Rural Affairs (Defra) MOHC Climate Programme (GA01101). We would like to thank all of the volunteers who have donated their computing time to climateprediction.net and weather@home, and the Met Office Hadley Centre PRECIS team for their technical and

scientific support for the development and application of weather@Home. PD acknowledges the funding from the European Union's Horizon 2020 research and innovation programme COGNAC under the European Union Marie Skłodowska-Curie Grant agreement No. 654942.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Anstey JA, Davini P, Gray LJ, Woollings TJ, Butchart N, Cagnazzo C, Christiansen B, Hardiman SC, Osprey SM, Yang S (2013) The ability of CMIP5 models to simulate north atlantic extratropical cyclones. *J Geophys Res Atmos* 118(10):3956
- Barnes EA, Dunn-Sigouin E, Masato G, Woollings T (2014) Exploring recent trends in Northern Hemisphere blocking. *Geophys Res Lett* 41(2):638
- Berkmans J, Woollings T, Demory ME, Vidale PL, Roberts M (2013) Atmospheric blocking in a high resolution climate model: influences of mean state, orography and eddy forcing. *Atmos Sci Lett* 14(1):34
- Chang EK, Guo Y, Xia X (2012) CMIP5 multimodel ensemble projection of storm track change under global warming. *J Geophys Res Atmos* 117(D23):D23118
- Christidis N, Stott PA (2015) Changes in the geopotential height at 500 hPa under the influence of external climatic forcings. *Geophys Res Lett* 42(24). doi:10.1002/2015GL066669
- Davini P, Cagnazzo C, Gualdi S, Navarra A (2012) Bidimensional diagnostics, variability, and trends of Northern Hemisphere blocking. *J Clim* 25(19):6496
- Davini P, Cagnazzo C (2014) On the misinterpretation of the North Atlantic Oscillation in CMIP5 models. *Clim Dyn* 43(5–6):1497
- Dawson A, Palmer T (2014) Simulating weather regimes: impact of model resolution and stochastic parameterization. *Clim Dyn* 44(7–8):2177
- Dawson A, Palmer T, Corti S (2012) Simulating regime structures in weather and climate prediction models. *Geophys Res Lett* 39(21):L21805
- Dee D, Uppala S, Simmons A, Berrisford P, Poli P, Kobayashi S, Andrae U, Balmaseda M, Balsamo G, Bauer P et al (2011) The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Q J R Meteorol Soc* 137(656):553
- Dunn-Sigouin E, Son SW (2013) Northern Hemisphere blocking frequency and duration in the CMIP5 models. *J Geophys Res Atmos* 118(3):1179
- Hannachi A, Barnes EA, Woollings T (2013) Behaviour of the winter North Atlantic eddy-driven jet stream in the CMIP3 integrations. *Clim Dyn* 41(3–4):995
- Harvey B, Shaffrey L, Woollings T (2014) Equator-to-pole temperature differences and the extra-tropical storm track responses of the CMIP5 climate models. *Clim Dyn* 43(5–6):1171
- Hoskins BJ, Hodges KI (2002) New perspectives on the Northern Hemisphere winter storm tracks. *J Atmos Sci* 59(6):1041
- Hoskins B, Woollings T (2015) Persistent extratropical regimes and climate extremes. *Curr Clim Change Rep* 1(3):115
- Huntingford C, Marsh T, Scaife AA, Kendon EJ, Hannaford J, Kay AL, Lockwood M, Prudhomme C, Reynard NS, Parry S et al (2014) Potential influences on the United Kingdom's floods of winter 2013/14. *Nat Clim Change* 4(9):769
- Hurrell J, Hack J, Shea D, Caron J, Rosinski J (2008) A new sea surface temperature and sea ice boundary dataset for the Community Atmosphere Model. *J Clim* 21(19):5145
- Inatsu M, Hoskins BJ (2004) The zonal asymmetry of the Southern Hemisphere winter storm track. *J Clim* 17(24):4882
- Jaeger E, Seneviratne S (2011) Impact of soil moisture-atmosphere coupling on European climate extremes and trends in a regional climate model. *Clim Dyn* 36(9–10):1919
- Keeley S, Sutton R, Shaffrey L (2012) The impact of North Atlantic sea surface temperature errors on the simulation of North Atlantic European region climate. *Q J R Meteorol Soc* 138(668):1774
- Knutti R, Furrer R, Tebaldi C, Cermak J, Meehl GA (2010) Challenges in combining projections from multiple climate models. *J Clim* 23(10):2739
- Knutti R (2010) The end of model democracy? *Clim Change* 102(3–4):395
- Lewis SC, Karoly DJ (2013) Anthropogenic contributions to Australia's record summer temperatures of 2013. *Geophys Res Lett* 40(14):3705
- Lu J, Greatbatch RJ (2002) The changing relationship between the NAO and northern hemisphere climate variability. *Geophys Res Lett* 29(7):52
- Massey N, Jones R, Otto F, Aina T, Wilson S, Murphy J, Hassell D, Yamazaki Y, Allen M et al (2014) weather@ home-development and validation of a very large ensemble modelling system for probabilistic event attribution. *Q J R Meteorol Soc* 141:1528–1545
- Mitchell D, Gray L, Anstey J, Baldwin M, Charlton-Perez A (2013) The influence of stratospheric vortex displacements and splits on surface climate. *J Clim* 26(8):2668–2682
- Mitchell D, Heaviside C, Vardoulakis S, Huntingford C, Masato G, Guillod BP, Frumhoff P, Bowery A, Wallom D, Allen M (2016) Attributing human mortality during extreme heat waves to anthropogenic climate change. *Environ Res Lett* 11(7):074006. <http://stacks.iop.org/1748-9326/11/i=7/a=074006>
- Pall P, Aina T, Stone DA, Stott PA, Nozawa T, Hilberts AG, Lohmann D, Allen MR (2011) Anthropogenic greenhouse gas contribution to flood risk in England and Wales in autumn 2000. *Nature* 470(7334):382
- Palmer T (2013) Climate extremes and the role of dynamics. *Proc Natl Acad Sci* 110(14):5281
- Santos JA, Woollings T, Pinto JG (2013) Are the winters 2010 and 2012 archetypes exhibiting extreme opposite behavior of the North Atlantic jet stream? *Mon Weather Rev* 141(10):3626
- Scaife AA, Copeley D, Gordon C, Harris C, Hinton T, Keeley S, O'Neill A, Roberts M, Williams K (2011) Improved Atlantic winter blocking in a climate model. *Geophys Res Lett* 38(23):L23703
- Schaller N, Kay AL, Lamb R, Massey NR, Van Oldenborgh GJ, Otto FE, Sparrow SN, Vautard R, Yiou P, Ashpole I et al (2016) Human influence on climate in the 2014 southern England winter floods and their impacts. *Nat Clim Change* 6:627–634
- Seviour WJM, Gray LJ, Mitchell DM (2016) Stratospheric polar vortex splits and displacements in the high-top CMIP5 climate models. *J Geophys Res Atmos* 121:1400–1413. doi:10.1002/2015JD024178
- Sippel S, Mitchell D, Black MT, Dittus AJ, Harrington L, Schaller N, Otto FE (2015) Combining large model ensembles with extreme value statistics to improve attribution statements of rare events. *Weather Clim Extrem* 9:25
- Stainforth DA, Aina T, Christensen C, Collins M, Faull N, Frame D, Kettleborough J, Knight S, Martin A, Murphy J et al (2005) Uncertainty in predictions of the climate response to rising levels of greenhouse gases. *Nature* 433(7024):403

- Stephenson D, Pavan V, Collins M, Junge M, Quadrelli R et al (2006) North Atlantic Oscillation response to transient greenhouse gas forcing and the impact on European winter climate: a CMIP2 multi-model assessment. *Clim Dyn* 27(4):401
- Taylor KE, Stouffer RJ, Meehl GA (2012) An overview of CMIP5 and the experiment design. *Bull Am Meteorol Soc* 93(4):485
- Tibaldi S, Molteni F (1990) On the operational predictability of blocking. *Tellus* 42A:343
- Uhe P, Otto FEL, Hausteil K, van Oldenborgh GJ, King A, Wallom D, Allen MR, Cullen H (2016) Comparison of Methods: Attributing the 2014 record European temperatures to human influences. *Geophys Res Lett*. doi:[10.1002/2016GL069568](https://doi.org/10.1002/2016GL069568)
- Waugh DW, Eyring V (2008) Quantitative performance metrics for stratospheric-resolving chemistry-climate models. *Atmos Chem Phys* 8(18):5699
- Williams K, Harris C, Bodas-Salcedo A, Camp J, Comer R, Copsey D, Fereday D, Graham T, Hill R, Hinton T et al (2015) The met office global coupled model 2.0 (GC2) configuration. *Geosci Model Dev Discuss* 8(1):521
- Woollings T, Charlton-Perez A, Ineson S, Marshall A, Masato G (2010) Associations between stratospheric variability and tropospheric blocking. *J Geophys Res* 115(D6):D06108
- Woollings T, Czuchnicki C, Franzke C (2014) Twentieth century North Atlantic jet variability. *Q J R Meteorol Soc* 140(680):783
- Zappa G, Shaffrey LC, Hodges KI (2013) The ability of CMIP5 models to simulate north atlantic extratropical cyclones. *J Clim* 26(15):5379