Deconstructing the climate change response of the Northern Hemisphere wintertime storm tracks

Article

Accepted Version

Climate Dynamics manuscript No.
(will be inserted by the editor)

Deconstructing the climate change response of the Northern Hemisphere wintertime storm tracks

B. J. Harvey · L. C. Shaffrey · T. J. Woollings

Received: date / Accepted: date

Abstract There are large uncertainties in the circulation response of the atmosphere to climate change. One manifestation of this is the substantial spread in projections for the extratropical storm tracks made by different state-of-the-art climate models. In this study we perform a series of sensitivity experiments, with the atmosphere component of a single climate model, in order to identify the causes of the differences between storm track responses in different models. In particular, the Northern Hemisphere wintertime storm tracks in the CMIP3 multi-model ensemble are considered. A number of potential physical drivers of storm track change are identified and their influence on the storm tracks is assessed. The experimental design aims to perturb the different physical drivers independently, by magnitudes representative of the range of values present in the CMIP3 model runs, and this is achieved via perturbations to the sea surface temperature and the sea-ice concentration forcing fields. We ask the question: can the spread of

B. J. Harvey and L. C. Shaffrey

National Centre for Atmospheric Science, Department of Meteorology, University of Reading, Reading, UK.

E-mail: b.j.harvey@reading.ac.uk

T. J. Woollings

Atmospheric, Oceanic and Planetary Physics, Department of Physics, University of Oxford, Oxford, UK.

projections for the extratropical storm tracks present in the CMIP3 models be accounted for in a simple way by any of the identified drivers? The results suggest that, whilst the changes in the upper-tropospheric equator-to-pole temperature difference have an influence on the storm track response to climate change, the large spread of projections for the extratropical storm track present in the northern North Atlantic in particular is more strongly associated with changes in the lower-tropospheric equator-to-pole temperature difference.

Keywords Storm tracks · Climate change · CMIP3 · Drivers of change · Polar amplification

1 Introduction

Extratropical cyclones are an important component of the climate system. They play a substantial role in the general circulation of the atmosphere through the transport of heat, momentum and moisture. They also play a role in the societal impacts of weather in the mid-latitudes through their associated precipitation, clouds and surface winds. It is therefore important to understand how the extratropical storm tracks, and the properties of extratropical cyclones, may change in the future.

Many studies have addressed this by either (i) analysing projections of the storm tracks simulated by one or more of the latest state-of-the-art climate models, to quantify and understand the range of changes that may take place, or (ii) by performing specific climate model experiments to understand individual mechanisms that might drive the changes. Over recent years the consensus has arisen that the zonal-mean properties of the storm tracks, and their associated eddy-driven surface winds, will move poleward under the projected greenhouse gas forcing (Yin 2005; Solomon et al. 2007). The precise mechanism causing such a change, however, is still under debate (Chen et al. 2008; Rivière 2011; Kidston et al. 2011). Recent studies (Woollings & Blackburn 2012; Barnes & Polvani 2013) have shown that the poleward shift of the eddy-driven jets in response to anthropogenic greenhouse gas forcing is remarkably robust in different

models, ocean basins and seasons, with the notable exception of the North Atlantic region in winter. In that case the multi-model mean response is small in comparison to the spread between different model responses. Other studies have investigated the responses of the storm tracks directly in large multi-model ensembles (Ulbrich et al. 2008; Harvey et al. 2012; Zappa et al. 2013b; Woollings et al. 2012; Harvey et al. 2013) and found less of a consensus. For the North Atlantic wintertime storm track in particular, these studies suggest that, rather than a poleward shift, there will be an intensification of its southern flank, effectively extending the storm track zonally towards Europe. However, there is a large spread in the projections of different models (e.g. Laîné et al. 2009).

As an illustration of the problem, Figure 1 shows the multi-model mean climate change responses of several wintertime climate variables from the World Climate Research Programme's Coupled Model Intercomparison Project phase 3 (CMIP3) multi-model dataset. Here the climate change response is defined as the difference between 2081-2100 from the SRESA1B experiment and 1961-2000 from the 20C3M experiment. Panel **a** shows a measure of the storm activity, the standard deviation of the 2-6 day bandpass-filtered daily-mean mean sea level pressure (MSLP) field. This commonly-used diagnostic provides a simple assessment of synoptic-scale activity using only daily-mean MSLP data (Hoskins & Hodges 2002; Chang 2009), and will be referred to simply as the storm track in the following. The response of this measure of storm activity is in general agreement with the studies cited above. Here we note the key points relevant to later discussion: (i) there is a poleward shift and intensification of the multi-model mean North Pacific storm track in response to climate change, (ii) there is no corresponding poleward shift in the North Atlantic storm track, instead there is an intensification on its southern flank resulting in a zonal extension of the storm track towards Europe, and (iii) there is a large inter-model standard deviation of the responses, particularly in the North Atlantic region (Figure 1**b**). It is this large spread that we wish to understand.

On the global scale, the zonal-mean temperature response to increased GHGs consists of a warming throughout the troposphere with regions of enhanced warming in the tropics in the upper-troposphere and in the wintertime polar regions near the surface, as shown in Figure 1**c**. Many studies have suggested that both regions of warming may influence the storm tracks in different ways, for instance Harvey et al. (2013) show using CMIP5 data that a substantial fraction of the inter-model spread in wintertime storm track projections can be accounted for by a simple linear regression model based on differences in the equator-to-pole temperature difference at upper- and lower- tropospheric levels, with larger temperature differences at either level associated with stronger storm activity. A number of mechanisms have been suggested for how the warming influences the storm tracks in each case, including: changes in the baroclinicity, via changes in the meridional temperature gradient or the vertical static stability (Frierson 2006; Butler et al. 2010) or the height of the tropopause (Williams 2006; Lorenz & DeWeaver 2007) directly influencing the eddies, or changes in the mean circulation influencing the propagation of the eddies (Hartmann et al. 2000; Kidston et al. 2011) or their wave-breaking behaviour (Chen et al. 2008; Rivière 2011).

On a more local scale, Wilson et al. (2009) and Brayshaw et al. (2011) present evidence that local sea surface temperatures (SSTs) can influence the North Atlantic storm track, and Deser et al. (2000) and Magnusdottir et al. (2004) show that the distribution of Arctic sea ice concentration (SIC) is also important. Future changes in the SST and sea-ice distributions may be linked via changes in the Atlantic Meridional Overturning Circulation (AMOC, see Laîné et al. (2009); Brayshaw et al. (2009)), and indeed Woollings et al. (2012) show that there is a strong correlation between the responses of the AMOC and the North Atlantic storm track in the CMIP3 models. However, several questions remain: Is the inter-model spread in the responses of either the North Atlantic SSTs or the Arctic sea ice responsible for the large spread in the storm track responses?, and is the link between the responses of the AMOC and the North Atlantic storm track responses due to changes in North Atlantic SSTs, Arctic sea-ice, or both?

In this paper we attempt to deconstruct the climate change projections of the Northern Hemisphere storm track/eddy-driven jet system, and in particular aim to understand the causes of the inter-model differences in the CMIP3 multi-model dataset. In particular we ask the question: can the differences between the model projections be accounted for by differences in a small number of physical 'drivers' of the storm tracks? If so, which drivers are most important, and what is their impact on the storm tracks? This work is motivated by the studies of Woollings et al. (2012) and Harvey et al. (2013), which suggest respectively the AMOC and both the lower- and the upper-tropospheric large-scale equator-to-pole temperature differences as possible physical drivers of storm track change. These studies relied on inter-model correlations in multi-model ensembles and were therefore unable to infer any causality. In the present study, a series of atmosphere-only GCM experiments are presented, using a single model forced by perturbations to the SST and SIC fields. The perturbations are designed to isolate the effects of the drivers of storm track change, and reflect the spread of these drivers between the different CMIP3 models projections. In this way, the dominant drivers of the inter-model spread in the storm track response are determined. Whilst we restrict attention to the CMIP3 models in the present study, we note that both the magnitude of the mean storm track response in the CMIP5 models and the inter-model spread between the CMIP5 models have similar magnitudes and spatial distributions as the CMIP3 models (see Harvey et al. 2013, Figure 2**b**), and therefore the results of this study are most likely also relevant to the CMIP5 ensemble.

The experimental design is described in Section 2, followed by a description of the control experiments in Section 3. Two sets of perturbation experiments are presented: the first aims to test the hypothesis discussed in Harvey et al. (2013) that the large-scale equator-to-pole temperature difference at upper- and lower-tropospheric levels influences the storm track, and these are presented in Section 4. The second

set of perturbation experiments consider the North Atlantic in more detail and these are presented in Section 5. A summary and discussion are presented in Section 6.

2 Model and Experimental Design

2.1 HadGAM1

The model used in this study is HadGAM1, the atmospheric component of the Hadley Centre's HadGEM1 global climate model (Johns et al. 2006; Martin et al. 2006) which was used in the CMIP3 ensemble. It has a grid-point resolution of $1.875° \times 1.25°$, and 38 layers between the surface and the model top at 39 km altitude. Around 10 of the layers are globally above the tropopause.

The experiments differ only in the prescribed SST, SIC and well-mixed greenhouse gas (GHG) fields (the land surface is allowed to evolve). The GHGs are kept fixed during each experiment and are specified using either late 20th century or projected late 21st century values from the SRESA1B scenario as appropriate. Annually-repeating monthly SST and SIC fields are constructed for each experiment from CMIP3 output as described in Section 2.2. All other elements of the model setup are unchanged including the land-use specifications, and the aerosol and ozone concentrations.

Each experiment consists of two 21 year runs (single 21 year runs were found to be insufficient to provide robust statistics), all starting from the same pair of initial conditions and all having the first 12 months discarded as spin-up. Therefore there are 40 complete years of data for each experiment. Additional runs of 21 years are performed for the two control experiments, as described below, and in those cases 60 years of data are used.

2.2 Experimental design

Table 1 details the experiments presented in this paper, and the relationships between them. There are two control runs (CON-20C and CON-A1B) representing the 20C3M and SRESA1B mean climates respectively, and four pairs of perturbation runs (ARC±, UFM±, UFM3± and NATL±) designed to isolate different drivers of storm track change. Various methods have been used to construct the SST and SIC fields for the perturbation experiments, as described below. In each case the aim is to capture one standard deviation of the spread in the CMIP3 model responses of a particular driver of storm track change. In this way, the relative contributions of the different drivers to the overall spread of the CMIP3 ensemble are assessed.

Data from 21 CMIP3 models is used to construct the SST and SIC fields, as listed in Appendix A. One ensemble member per model is used in order to provide equal model weighting in the multi-model mean. Data from two periods is used: years 1961-2000 of the 20C3M scenario and years 2061-2100 of the SRESA1B scenario. Monthly SST and SIC climatologies are calculated for both 40 year periods of each model, and these are interpolated onto the HadGAM1 grid. Nearest grid point extrapolation is used to fill gaps resulting from land-sea mask mismatches between the models.

2.2.1 Control experiments (CON-20C and CON-A1B)

The CON-20C and CON-A1B experiments represent the CMIP3 20C3M and SRESA1B climates respectively. The SST fields used are the multi-model mean values from the 20C3M and SRESA1B experiments respectively. A similar method is desired for creating monthly SIC fields; however, due to the large range of 20C3M ice extents in the CMIP3 models (see, e.g., Stroeve et al. 2007), simply taking the multi-model mean of the SIC values produces an ice distribution with an unrealistically smooth ice edge region. To avoid this problem, artificial SIC distributions have been created to mimic the typical ice distribution

of the CMIP3 models in the 20C3M and SRESA1B scenarios whilst retaining a realistic ice-edge structure. The details of this procedure are presented in Appendix B. In brief, a multi-model mean ice edge position is constructed from the 20C3M ice distributions and then converted into an ice distribution which transitions smoothly from 100% within the ice edge to 0% outside of it. This distribution is used for the CON-20C experiment. The CON-20C ice edge position is then retreated by a distance which is representative of the ice retreats in the models, and then similar converted into a smooth ice distribution which is used for the CON-A1B experiment. This process is carried out separately in the Northern and Southern Hemispheres.

Figures 2**a** and **c** show the resulting DJF mean SST and SIC values for the CON-20C experiment as well as the CON-A1B minus CON-20C difference. The SST increases almost everywhere between CON-20C and CON-A1B, with a peak around 2.3 deg C in the equatorial regions. However, there are spatial variations and in particular there is a small region of cooling in the west side of the North-Atlantic sub-polar gyre, surrounded by a region of mediated warming covering much of the sub-polar gyre. This region of cooling has been linked to a slow down of the AMOC in the future, associated with changes in wind and precipitation in the regions of deep oceanic convection (Drijfhout et al. 2012). The SIC decreases between CON-20C and CON-A1B near the ice edge at all longitudes, with the largest changes in the Barents Sea and the Bering Strait.

2.2.2 Global perturbation experiments (ARC, UFM and UFM3)

Two pairs of perturbation experiments, ARC± and UFM±, are configured to reflect the spread in the lower- and upper-tropospheric equator-to-pole temperature differences respectively. In the ARC± experiments a perturbation is made to the Arctic sea-ice distributions, as described below. A sea ice reduction generally acts to warm the lower atmosphere in the Arctic, and therefore reduces the lower-tropospheric

equator-to-pole temperature gradient. The aim here is to modify the lower-tropospheric equator-to-pole temperature difference by approximately one standard deviation of the spread of the responses of the lower-tropospheric equator-to-pole temperature differences in the CMIP3 models. The Antarctic sea-ice extent is not changed as the focus here is on the Northern Hemisphere circulation. The method for perturbing the Arctic sea-ice extent is similar to that used to generate the CON-A1B ice distribution. The distance retreated for the ARC+ and ARC- experiments is equal to the distance retreated for the CON-A1B experiment plus and minus one standard deviation of the range of distances retreated by the ice edges in the CMIP3 models (see Appendix B for details). These experiments are therefore designed to capture the spread in the ice edge responses present in the CMIP3 models. The difference between the ARC+ and ARC- DJF mean SIC fields are shown in Figure 2**d**. In addition, the SST field is perturbed in the ice edge region as described below in section 2.2.3. The SST perturbations act to fill-in the region of exposed sea surface in these experiments, whilst also modifying the surrounding SST in expected sense, with ARC+ having a warmer ice edge region than A1B, and ARC- a cooler ice edge region. We do not attempt to separate the relative influence of the SST and ice perturbations as it is their combined effect that is of interest to this study.

Turning to the UFM± experiments,Harvey et al. (2013) show that the upper-tropospheric equator-to-pole temperature difference is closely related to the tropical SSTs in the CMIP5 models. Therefore, to modify the upper-tropospheric equator-to-pole temperature difference, a uniform SST anomaly ΔT_{UFM} is applied globally to the CON-A1B SST field. The magnitude of ΔT_{UFM} is chosen initially as the inter-model standard deviation of the tropical SST responses in the CMIP3 models, where the Tropics are defined here as the region with |latitude| < 30 degrees. The value of this is $\Delta T_{\mathrm{UFM}} = 0.29$K (see Figure 4**b**), and the UFM+ and UFM- experiment SST values are given by $SST_{\mathrm{CON-A1B}} \pm \Delta T_{\mathrm{UFM}}$ respectively. In addition to these, a further pair of experiments are performed, UFM3+ and UFM3-, with SST values

given by $SST_{\mathrm{CON-A1B}} \pm 3\Delta T_{\mathrm{UFM}}$ respectively. Other studies which consider the influence of a uniform SST perturbation on the atmosphere include Kodama & Iwasaki (2009) and Chen et al. (2013) (both using aquaplanet models), and Graff & LaCasce (2012) (using a full AGCM). It should be noted that as well as perturbing the upper-level equator-to-pole temperature difference this experimental design may introduce influences from other mechanisms, most notably changes in diabatic heating within the storm tracks due to changes in precipitable water (as discussed in Schneider et al. 2010) and changes in the lower-level equator-to-pole temperature difference due to remotely-forced polar amplification (see, e.g., Screen et al. 2012).

We note that additional factors may also influence the climate change response of the storm tracks. For instance Scaife et al (2012) showed that the representation of the stratosphere can have a large impact on the response of the North Atlantic storm track in particular. We do not attempt to capture this effect directly, since none of the CMIP3 models have well-resolved stratospheres, however any impact of this process may be at least partially captured via its impact of the upper-tropospheric equator-to-pole temperature difference.

2.2.3 North Atlantic perturbation experiments (NATL)

There is a particularly large spread of North Atlantic SST responses in the CMIP3 models. This spread may be interpreted as a combination of differences in the magnitude and location of the region of mediated warming in the sub-polar gyre (see Figure 2**a**), along with differences associated with sea-ice retreat (Woollings et al. (2012)). As a result there is spatial structure present in the inter-model spread of the SST responses. To capture this spatial structure, the SST pattern which explains the maximum of the variance between the SST response fields of the 21 CMIP3 models is calculated (this is effectively the

first "inter-model EOF" of the responses). This is denoted $\Delta T_{\mathrm{EOF}}(lon, lat)$ and is illustrated in Figure **2b**.

The pattern of ΔT_{EOF} is close to zero along the southern boundary of the domain, which justifies the choice of domain used, and has two distinct maxima: one in the sub-polar gyre region and the other along the ice edge. Both of these maxima are present in composites representing the impact of AMOC variability on the North Atlantic SSTs (see, e.g., Drijfhout et al. 2012, Figure 2), so their presence is consistent with the inter-model SST differences being associated with variations in the amount of AMOC reduction between the models. Since the maxima along the ice edge is most likely associated with ice retreat, the ΔT_{EOF} pattern has been split into two regions (indicated A and B in the figure) capturing the SST changes in the sub polar gyre and along the ice edge respectively. Perturbation experiments are performed with SST values using that part of ΔT_{EOF} in region A, denoted ΔT_{NATL}. These are denoted as NATL+ and NATL- and have the SST distributions $SST_{\mathrm{CON-A1B}} \pm \Delta T_{\mathrm{NATL}}$ respectively. The part of ΔT_{EOF} in region B is used as the SST perturbation around the ice edge region in the ARC+ and ARC- experiments, thus providing an SST value in regions of ice retreat.

3 Results I: Control experiments (CON-20C and CON-A1B)

Before considering the results of the perturbation experiments, we first consider the extent to which the CON-20C and CON-A1B experiments reproduce the corresponding CMIP3 multi-model mean 20C3M and SRESA1B climates.

3.1 The CON-20C climate

The CON-20C DJF climate is shown by the *grey contours* in Figure 3, which shows the zonal mean temperature (panel **a**), the zonal wind at 850 hPa (panel **b**) and the storm track (panel **c**). These three

variables are used throughout as they provide a simple overview of the thermodynamic and circulation response of the atmosphere. The CON-20C DJF zonal mean temperature is very similar to the CMIP3 multi-model mean zonal mean temperature (*contours* in Figure 1**c**). The CON-20C DJF U850 and storm track (panels **b** and **c**) are also both broadly similar to the corresponding CMIP3 multi-model mean values (*contours* in Figures 1**d** and **a** respectively). However in both ocean basins the CON-20C DJF U850 and storm track are slightly stronger than the corresponding CMIP3 multi-model mean values, and in addition the North Atlantic and North Pacific U850, and the North Atlantic storm track, have stronger SW-NE tilts in the CON-20C experiment than in the CMIP3 multi-model mean. It is a well-known problem that many climate models have jet streams and storm tracks that are oriented too zonally, particularly in the North Atlantic region (Zappa et al. 2013a). HadGAM1 appears to here have a more strongly-tilted jet and storm track than the CMIP3 multi-model mean, despite being forced with CMIP3 mean SST and SIC fields. Other features of the CON-20C climate are broadly similar to the CMIP3 multi-model mean, justifying its use here as a control with which to compare the perturbation experiments.

3.2 The difference between the CON-A1B and CON-20C climates

The difference between the CON-A1B and CON-20C DJF climates is shown by the *shading* in Figure 3. There is warming throughout the troposphere, and cooling in the stratosphere, accompanied by regions of enhanced tropospheric warming in the tropical upper-troposphere and in the Arctic lower-troposphere (panel **a**). The spatial structure and magnitudes of these features closely match those of the CMIP3 multi-model mean response (Figure 1**c**), giving confidence that the thermal structure of the atmosphere in the HadGAM1 model is responding in a similar manner to that seen in the CMIP3 models.

The difference between the CON-A1B and CON-20C DJF U850 fields (Figure 3**b**) also shares similarity with the CMIP3 multi-model mean U850 response (Figure 1**c**). The spatial pattern and magnitude of

the response is reproduced. In addition, the difference between the CON-A1B and CON-20C storm tracks (Figure 3**c**) is also remarkably similar to the CMIP3 multi-model mean storm track response. The North Pacific storm track intensifies and shifts poleward, whereas the North Atlantic storm track has an intensification on its southern flank. However, the magnitudes of these features in the CON-A1B minus CON-20C difference is larger than in the CMIP3 multi-model mean response, and the North Atlantic intensification over Europe is more localised spatially and located further north. Some of these differences may be expected since (i) a single model experiment (HadGAM1) is being compared with the smoother CMIP3 multi-model mean, and (ii) the northerly position of the intensification over western Europe in Figure 1**a** is consistent with the stronger SW-NE tilt of the North Atlantic storm track in the CON-20C experiment. Therefore we conclude that the CON-A1B minus CON-20C difference is qualitatively similar to the CMIP3 multi-model mean response, and whilst these are not identical they are similar enough to provide confidence in the results of the perturbation experiments.

4 Results II: Global perturbation experiments (UFM, ARC and UFM3)

In this section results from the three global perturbation experiments (ARC, UFM and UFM3) are presented.

4.1 ARC experiments

The DJF temperature, circulation and storm track for the ARC experiments are shown in Figures 5**a-c**. These plots show the difference between the ARC+ and ARC- experiments (which will be referred to as the 'ARC response') and therefore represent a two standard deviation perturbation to the sea ice edge position, centred on the CON-A1B climate. Before discussing these plots we first consider Figure 4 which shows the DJF climate change responses of several key variables in both the CMIP3 ensemble and in the

experiments presented in this paper. In particular, panel **a** shows the SRESA1B minus 20C3M difference for each CMIP3 model (*diamonds*) and the EXPT minus CON-20C difference for the ARC+, CON-A1B and ARC- experiments (*squares with labels*). The variables shown are the change in Arctic sea ice extent (defined as the area with SIC greater than 15%) and the following measure of the lower-tropospheric equator-to-pole temperature difference (as used in Harvey et al. (2013)):

$$\Delta T850 = T850_{(30\mathrm{S}-30\mathrm{N})} - T850_{(60\mathrm{N}-90\mathrm{N}),} \quad (1)$$

where $T850_{(30S-30N)}$ and $T850_{(60N-90N)}$ are the area average zonal mean temperature at the 850 hPa level in the latitude ranges indicated.

There is a wide range of sea-ice extent responses in the CMIP3 ensemble, from around $-1 \times 10^6 \mathrm{km}^2$ to $-7 \times 10^6 \mathrm{km}^2$. By construction, the sea-ice extents of the two ARC experiments span a large part of this range: the difference between ARC+ and ARC- is a little over two standard deviations of the responses in the CMIP3 ensemble. There is also a wide range of responses of the lower-tropospheric equator-to-pole temperature difference in the CMIP3 models. The plot shows that the ARC + and ARC- experiments capture a similar fraction the lower-tropospheric equator-to-pole temperature difference (as measured by $\Delta T850$) with a span of slightly over two standard deviations of the CMIP3 ensemble. Therefore the ARC perturbation experiments are capturing the spread of the desired driver of storm track change.

Turning to the spatial patterns of the ARC response, shown in Figures 5**a-c**, the ARC+ minus ARC- zonal mean temperature difference has a localised warm anomaly in the lower-troposphere north of around 60N. The anomaly is strongest below 850hPa, although there is some penetration throughout the depth of the Arctic troposphere. There are differences in both U850 and the storm track in the ARC response (panels **b** and **c** respectively), particularly in the North Atlantic. There is a southward shift of the North Atlantic jet accompanied by a reduction in storm activity over the northern North Atlantic, similar to the negative phase of the NAO. In this case, however, the storm track response is widespread across all

of the sub-polar regions, and not confined solely to the North Atlantic sector. The sign of the storm track response is consistent with the storm track responding directly to the reduced baroclinicity in the lower-troposphere. The magnitude of the storm track response in the northern North Atlantic is of the order 0.6 hPa, which is larger than the inter-model standard deviation of the responses in the CMIP3 models, whereas it is weak across the entire North Pacific. This suggests that the range of sea ice responses in the CMIP3 models contributes towards the particularly large spread in the CMIP3 storm track responses in the northern North Atlantic region, but not in the North Pacific.

Given that the response of the storm track is strong in the ARC experiments, we briefly consider whether there is a relationship present between the responses of the lower-tropospheric equator-to-pole temperature difference and the responses of the storm track in the CMIP3 models. To this end, Figure 6**a** shows the slope of the linear regression between the local storm track responses in each model and the corresponding ΔT indices in each model. Similar plots are presented for the CMIP5 models and discussed in detail in Harvey et al. (2013). The values in Figure 6 are scaled by the magnitude of two standard deviations of the corresponding ΔT responses (as illustrated by the vertical bars in Figure 4) and the sign of the mean ΔT response. Therefore the quantity shown in Figure 6**a** may be expected to correspond in both magnitude and sign with the ARC storm track response of Figure 5**c**. There is a close resemblance between this regression plot and the experiment results of Figure 5**c**, and this provides further evidence that the lower-tropospheric equator-to-pole temperature difference has an influence on the storm track, with a reduction in the temperature difference resulting in a wide spread storm track reduction over much of the hemisphere. Interestingly, there is only a weak signal in the North Atlantic region in Figure 6**a**. However, Harvey et al. (2013) show that using instead the equator-to-pole temperature difference from the North Atlantic region only (rather than the full zonal mean) results in much stronger regression values in the North Atlantic region, as this is also the case for CMIP3 models (not shown).

312 4.2 UFM and UFM3 experiments

Panel **b** of Figure 4 shows SRESA1B minus 20C3M differences for each CMIP3 model (*diamonds*) and the EXPT minus CON-20C differences for the CON-A1B, UFM± and UFM3± experiments (*squares with labels*). The variables shown are the change in tropical SST and the following measure of the upper-tropospheric equator-to-pole temperature difference:

$$\Delta T250 = T250_{(30\mathrm{S}-30\mathrm{N})} - T250_{(60\mathrm{N}-90\mathrm{N}).} \quad (2)$$

313 By construction, the UFM+ and UFM- experiments span exactly two standard deviations of the CMIP3
314 tropical SST responses. However, the corresponding span of $\Delta T250$ is less than one standard deviation
315 of the CMIP3 models. Therefore the UFM+ and UFM- experiments fail to capture the spread in the
316 desired driver of storm track change. This suggests that, whilst the tropical SST does play a role in the
317 climate change response of the equator-to-pole temperature difference, other factors also contribute to
318 the differences between the $\Delta T250$ responses in the CMIP3 models. For instance, the diverse range of
319 subgrid parametrizations present amongst the CMIP3 models may amplify the spread from the tropical
320 SST response to the upper-tropospheric temperature response. Alternatively, it may be the case that
321 HadGAM1 only weakly responds to the tropical SST values compared to the other CMIP3 models. In
322 an attempt to capture the spread of $\Delta T250$, the pair of more-strongly-forced experiments UFM3± are
323 performed. These are identical to the UFM± experiments, except the uniform SST perturbation is three
324 times stronger, and therefore the UFM3+ minus UFM3- difference spans six standard deviations of the
325 CMIP3 tropical SST responses. This magnitude was chosen with the aim of spanning a similar fraction
326 of the CMIP3 model spread in $\Delta T250$ as the ARC experiments do in $\Delta T850$, that is, a little over two
327 standard deviations. Figure 4**b** shows that this is achieved.

The spatial patterns of the UFM+ minus UFM- difference (the 'UFM response') are shown in Figures 5**d-f**, and the UFM3+ minus UFM3- difference (the 'UFM3 response') in Figures 5**g-i**. The pattern of tropospheric warming in each case (panels **d** and **g**) have the expected zonal mean structure: there is warming throughout the troposphere with a maximum in the tropical upper-troposphere. The equator-to-pole temperature difference therefore increases at upper-tropospheric levels in each case. There is also a small amount of northern-hemisphere polar amplification present in the UFM3 response, resulting in a slight reduction in the lower-tropospheric equator-to-pole temperature difference. Such remotely-forced polar amplification is a well-known feature of the climate system and possible causes include changes in the Northern Hemisphere poleward heat transport and changes in Arctic cloudiness (Screen et al. (2012)). Given the results of the ARC experiments, this may be expected to have some influence on the circulation and storm track.

The thermal changes are statistically significant for both the UFM and UFM3 responses. However, the corresponding U850 and storm track responses in UFM (panels **e** and **f**) are too weak to be distinguished from internal variability in these runs (as indicated by the limited areas of stippling in these panels), and the pattern of the response of neither variable is robust to sub-sampling of the yearly data (not shown). Therefore the UFM SST perturbations appear to be too small to significantly influence the storm activity in this experimental setup.

In contrast, the UFM3 response does contain regions of distinguishable circulation response (Figures 5**h** and **i**). Both the U850 and the storm track exhibit a tri-polar pattern in the North Atlantic region, with a decrease in both variables in the subtropics and the Arctic, and an increase over north-western Europe. In the North Pacific there is an intensification of both U850 and the storm track on the southern downstream flank. These patterns are qualitatively very similar to the uniform SST experiment of Graff & LaCasce (2012) (e.g. their Figure 11**b**). It is of note that the spatial patterns of the UFM3 U850 and

storm track responses are very similar. This is in contrast to the ARC experiments where there was a meridional shift in the U850 pattern and a general decrease in the storm track.

The UFM3 storm track differences are in some places of similar magnitude to the CMIP3 inter-model standard deviation of Figure 1**b**. However, the magnitude is smaller than that of the ARC experiments. In addition, the spatial pattern of the UFM3 storm track response does not match the pattern of the regions of large inter-model spread, particularly in the northern North Atlantic. Both of these features suggest that the upper-level equator-to-pole temperature difference is not the dominant driver of the inter-model spread in the CMIP3 ensemble. Similar to the ARC response and Figure 6**a**, Figure 6**b** shows the influence of $\Delta T250$ on the storm track responses as diagnosed from a linear inter-model regression. In this case, there is very little area exhibiting a significant correlation. This further suggests that the spread in this variable is not a dominant driver of the spread between the CMIP3 model storm track responses. Interestingly, Harvey et al. (2013) find some regions of significant correlation in the CMIP5 version of Figure 6**b**; the reason for this difference between CMIP3 and CMIP5 is not clear.

Finally, it is of note that the magnitude of the difference between the UFM3 SST fields is similar to the CON-A1B minus CON-20C tropical SST difference (Figure 4**a**). Whilst the basic states of these experiments are not the same, we briefly compare the CON-A1B minus CON-20C response of Figure 3 to the UFM3 response plots of Figure 5**g-i**. Which features, of any, of the mean climate change response are produced in the uniform SST perturbation experiments? Interestingly, the UFM3 experiments do not produce a poleward shift in the upstream end of the North Pacific storm track, a feature which is present in both the CMIP3 multi-model mean and the CON-A1B minus CON-20C difference. This is unexpected as (i) many idealised modelling studies have suggested a link between a poleward shifted storm track and an increase in the upper-level equator-to-pole temperature difference, and (ii) the regression analysis of Harvey et al. (2013) shows that those CMIP5 models which have larger upper-level $\Delta T250$ increases tend

to have more poleward-shifted storm tracks. On inspection of the U850 responses in each case (Figures 3**b** and 5**h**), however, there does appear to be similarities between the responses. Both have U850 increases in two regions: one on the poleward side of the upstream end of the North Pacific jet, and the other on the equatorward side at the downstream end. The weighting between these two regions is different in the two cases. The reason for this is unclear, but may be related to differences between the SSTs in the tropical Pacific driving differences in the tropical convection. In the North Atlantic, the storm track increase over north-western Europe is remarkably similar to the CON-A1B minus CON-20C mean response. Several studies have suggested that this pattern of response, an intensification on the southern flank of the North Atlantic storm track, may be due to changes in the local SST gradient influencing the baroclinicity. These experiments (and those of Graff & LaCasce (2012)) suggest however, that such a response can also be achieved by a globally uniform SST perturbation. This point is discussed further below.

5 Results III: North Atlantic perturbation experiments (NATL)

The global perturbation experiments presented above show that the particularly large spread in the CMIP3 storm track responses in the North Atlantic region is associated with differences in the responses of the equator-to-pole temperature difference, particularly at low levels. The study of Woollings et al. (2012) showed that the large spread is also correlated with changes in the AMOC in the models. However, it is not clear from that study whether the AMOC influences the storm track via the large scale temperature structure or via more local SST changes. The NATL experiments test this by perturbing the North Atlantic SSTs, capturing the spread in the changes in the models without changing the equator-to-pole temperature difference. The SST perturbation used is described above in Section 2.2.3, and plotted in Figure 2**b**.

Figure 7 shows the difference between the pair of North Atlantic perturbation experiments (NATL+ and NATL-). Panel **a** shows the spatial distribution of the lower-tropospheric temperature difference (T850), and panels **b** and **c** show the U850 and storm track responses. To compare the magnitude of the NATL SST perturbations with the CMIP3 inter-model spread, Figure 4**c** shows the responses of the area-average SST in the region indicated by the box in Figure 7**a**. The pair of NATL experiments span almost two standard deviations of this measure of the CMIP3 inter-model spread.

The NATL T850 response consists of a warming above the North Atlantic SST anomaly, and little change elsewhere (Figure 7**a**). However, the magnitude of the warming at 850 hPa is much smaller than the polar warming in the ARC experiments. There is very little significant response in the U850 or storm track in the NATL experiments (panels **b** and **c**): there is a weak equatorward shift of the jet, and an increase in storm activity coinciding with, and extending downstream of, the maximum in warming.

Woollings et al. (2012) showed that the responses of the North Atlantic wintertime storm track in the CMIP3 models depend strongly on changes in the AMOC, and suggested that the North Atlantic SSTs may play a role in this relationship. The small response of the circulation in the NATL experiments compared with the ARC experiments suggests that this is not the case in our experimental setup, and instead the mechanism behind the AMOC link is via changes in Arctic sea ice and/or the SST anomalies in the ice edge region.

A caveat is that there may be important features of the North Atlantic SST which are not captured by our experimental design. For instance, the regions of strongest SST gradient are smoothed out in the multi-model mean SST fields used here and this may reduce the sensitivity of the storm track to SST perturbations. Alternatively, the atmospheric response to a North Atlantic SST perturbation may depend on the position and/or tilt of the North Atlantic storm track. As noted above, the CMIP3 multi-model

mean has a more zonal North Atlantic storm track than is present in our experimental setup, and this may align more closely with the SST gradients in the NATL experiments.

6 Summary and Discussion

The experiments presented here aim to elucidate the causes of the inter-model differences in the responses of the Northern Hemisphere storm tracks in the CMIP3 multi-model ensemble, with a particular focus on the North Atlantic region where the inter-model spread is largest. A number of possible 'drivers' of the storm track response are considered in turn, and their role in causing the inter-model differences assessed:

- The lower-tropospheric equator-to-pole temperature difference has a strong impact on the storm track, as inferred from an experiment forced by a perturbation to the Arctic sea ice. The magnitude and spatial pattern of the response to this forcing suggest that there is a substantial contribution from this driver to the inter-model differences in CMIP3, particularly in the northern North Atlantic region.
- The upper-tropospheric equator-to-pole temperature difference is found to have some impact on the storm track, as inferred from an experiment forced by a globally-uniform SST perturbation. However, the magnitude and spatial pattern of the response to this forcing are not consistent with this being the dominant driver of the inter-model differences in CMIP3.
- A local SST anomaly in the North Atlantic ocean, reminiscent of the sub-polar gyre part of the pattern of mediated warming in the North Atlantic associated with an AMOC reduction, has little impact on the North Atlantic storm track.

In each case, the magnitude of the perturbations were chosen with the aim of representing one standard deviation of the spread in the CMIP3 models of the corresponding driver of change. However, as different mechanisms cannot be completely isolated in this experimental setup, no attempt is made to precisely

quantify the impacts of each driver, and instead we restrict attention to a more qualitative assessment of the responses in each case.

The spatial pattern of the response of the circulation to the globally-uniform SST perturbation captures features of the mean climate change response, particularly in the North Atlantic region where there is a tri-polar pattern in U850 and storm track responses. The response to the Arctic sea-ice perturbations, in contrast, has an NAO-like pattern which is qualitatively different to the mean response in the CMIP3 models. Therefore whilst the experiments show that the loss of Arctic sea ice contributes to the large inter-model differences in the CMIP3 storm track responses, it is not the dominant cause of the mean climate change response. In addition there are some regions, particularly in the latitude band 30N-40N, where neither perturbation captures the magnitude of the inter-model spread, indicating that additional factors influence the storm track responses there. It is interesting that the eddy-driven jet and the storm track responses have very similar spatial patterns in both the control climate change experiment and the globally-uniform SST experiments, whereas there is little relationship between them in the ARC experiments.

The lack of impact of the local North Atlantic SST anomaly suggests that the relationship between the responses of the AMOC and the responses of the North Atlantic storm track (Woollings et al. (2012)) is via changes in Arctic sea ice and related SSTs, rather than sub-polar gyre SST anomalies. However, there are two important caveats to this conclusion: model biases in the position and tilt of the North Atlantic storm track may cause the atmospheric response to a given SST anomaly to be model dependent, and the experimental design used here may not capture the spread in the most relevant drivers of storm track change.

This study has focused on the CMIP3 models, however both the magnitude of the mean storm track response in the CMIP5 models and the inter-model spread between the CMIP5 models have similar

magnitudes and spatial distributions as the CMIP3 models (see Harvey et al. 2013, Figure 2**b**). Therefore the results of this study are most likely also relevant to the CMIP5 ensemble, although this remains to be checked.

A List of CMIP3 models used in this study

Data from the following 21 models are used in this study. One run per model is used, and this is the run denoted 'run1' in almost all cases. It is noted in the main text if particular fields are unavailable for any of the models.

BCCR-BCM2.0 (Bjerknes Centre for Climate Research), CGCM3.1(T47) and CGCM3.1(T63) (Canadian Centre for Climate Modelling & Analysis), CNRM-CM3 (Mètèo-France / Centre National de Recherches Mètèorologiques), CSIRO-Mk3.0 and CSIRO-Mk3.5 (CSIRO Atmospheric Research), ECHAM5/MPI-OM (Max Planck Institute for Meteorology), ECHO-G (Meteorological Institute of the University of Bonn, Meteorological Research Institute of KMA, and Model and Data group), GFDL-CM2.0 and GFDL-CM2.1 (US Dept. of Commerce / NOAA / Geophysical Fluid Dynamics Laboratory), GISS-AOM and GISS-ER (NASA / Goddard Institute for Space Studies), INGV-SXG (Instituto Nazionale di Geofisica e Vulcanologia), INM-CM3.0 (Institute for Numerical Mathematics), IPSL-CM4 (Institut Pierre Simon Laplace), MIROC3.2(hires) and MIROC3.2(medres) (Center for Climate System Research (The University of Tokyo), National Institute for Environmental Studies, and Frontier Research Center for Global Change (JAMSTEC)), MRI-CGCM2.3.2 (Meteorological Research Institute), NCAR CCSM3 (National Center for Atmospheric Research), UKMO-HadCM3 and UKMO-HadGEM1 (Hadley Centre for Climate Prediction and Research / Met Office).

B Design of the sea-ice concentration fields

As noted in Section 2.2, due to the large range of 20C3M ice extents in the CMIP3 models (Stroeve et al. 2007), simply taking the multi-model mean of the SIC values (as was done for the SSTs) produces an ice distribution with an unrealistically smooth ice edge region. To avoid this problem artificial SIC distributions have been created to mimic the typical ice distribution of the CMIP3 models, whilst retaining a realistic ice-edge structure. Here we describe the algorithm for constructing these SIC fields.

Four ice distributions are used in this study, one representing the 20C3M ice distribution (used in CON-20C), one representing the mean SRESA1B response (used in CON-A1B), and two capturing the spread of the responses between the

models (used in ARC±). For each month of the CON-20C experiment, a multi-model mean ice edge position is defined as the 50% contour of the multi-model mean SIC distribution. The artificial SIC distribution is then constructed to equal 100% poleward of this ice edge, linearly reducing over a distance of 5 degrees to 0% equatorward of the ice edge. The width of this ice edge region was chosen by a consideration of typical ice edge regions in the CMIP3 20C3M wintertime ice distributions.

The remaining three ice distributions are then constructed by retreating the ice edge in CON-20C towards the pole. The distance retreated, a function of longitudinal grid point, is first calculated separately for each CMIP3 model. The distances retreated for CON-A1B, ARC+ and ARC- are then taken as the median and standard deviation of the retreats in the models. This process is illustrated in Figure 8 which shows, for the January values, the number of degrees latitude retreated by each the ice edge in each model. Also shown, in the *lower panel*, is the resulting ice edge positions of the four ice distributions. Once the position of the ice edge is found, artificial SIC distributions are then constructed, as above, to equal 100% poleward of this ice edge, linearly reducing over a distance of 5 degrees to 0% equatorward of the ice edge.

Acknowledgements BJH was supported by the Natural Environment Research Councils project Testing and Evaluating Model Predictions of European Storms (TEMPEST) during the course of this work. The authors acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate modeling groups for producing and making available their model output. For CMIP the U.S. Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals.

References

Barnes, E. A. & Polvani, L. 2013 Response of the Midlatitude Jets, and of Their Variability, to Increased Greenhouse Gases in the CMIP5 Models. *J. Clim.* **26**, pp. 7117–7135.

Brayshaw, D. J., Hoskins, B. & Blackburn, M. 2011 The basic ingredients of the North Atlantic storm track. Part II: sea surface temperatures. *J. Atmos. Sci.* **68**, pp. 1784–1805.

Brayshaw, D. J., Woollings, T. & Vellinga, M. 2009 Tropical and extratropical responses of the North Atlantic atmospheric circulation to a sustained weakening of the MOC. *J. Clim.* **22**, pp. 3146–3155.

Butler, A. H., Thompson, D. W. J. & Heikes, R. 2010 The steady-state atmospheric circulation response to climate change-like thermal forcings in a simple general circulation model. *J. Clim.* **23**, pp. 3474–3496.

512 Chang, E. K. M. 2009 Are band-pass variance statistics useful measures of storm track activity? Re-examining storm track
513 variability associated with the NAO using multiple storm track measures. *Clim. Dyn.* **33**, pp. 277–296.

514 Chen, G., Lu, J. & Frierson, D. M. W. 2008 Phase speed spectra and the latitude of surface westerlies: Interannual variability
515 and global warming trend. *J. Clim.* **21**, pp. 5942–5959.

516 Chen, G., Lu, J. & Sun, L. 2013 Delineating the Eddy-Zonal Flow Interaction in the Atmospheric Circulation Response to
517 Climate Forcing: Uniform SST Warming in an Idealized Aquaplanet Model. *J. Atmos. Sci.* **70**.

518 Deser, C., Walsh, J. E. & Timlin, M. S. 2000 Arctic Sea Ice Variability in the Context of Recent Atmospheric Circulation
519 Trends. *J. Clim.* **13**, pp. 617–633.

520 Drijfhout, S., van Oldenburgh, G. V. & Cimatoribus, A. 2012 Is a Decline of AMOC Causing the Warming Hole above the
521 North Atlantic in Observed and Modeled Warming Patterns? *J. Clim.* **25**, pp. 8373–8379.

522 Frierson, D. M. W. 2006 Robust increases in midlatitude static stability in simulations of global warming. *Geophys. Res.*
523 *Lett.* **33**, p. L24,816.

524 Graff, L. S. & LaCasce, J. H. 2012 Changes in the Extratropical Storm Tracks in Response to Changes in SST in an AGCM.
525 *J. Clim.* **25**, pp. 1854–1870.

526 Hartmann, D. L., Wallace, J. M., Limpasuvan, V., Thompson, D. W. J. & Holton, J. R. 2000 Can ozone depletion and
527 global warming interact to produce rapid climate change? *Proc. Nat. Acad. Sci.* **97**, pp. 1412–1417.

528 Harvey, B. J., Shaffrey, L. C. & Woollings, T. J. 2013 Equator-to-pole temperature differences and the extra-tropical storm
529 track responses of the CMIP5 climate models. *Clim. Dyn.* , pp. 1–12.

530 Harvey, B. J., Shaffrey, L. C., Woollings, T. J., Zappa, G. & Hodges, K. I. 2012 How large are projected 21st century storm
531 track changes? *Geophys. Res. Lett.* **39**, p. L052,873.

532 Hoskins, B. J. & Hodges, K. I. 2002 New perspectives on the Northern Hemisphere winter storm tracks. *J. Atmos. Sci.* **59**,
533 pp. 1041–1061.

534 Johns, T. C., Durman, C. F., Banks, H. T., Roberts, M. J., McLaren, A. J., Ridley, J. K., Senior, C. A., Williams,
535 K. D., Jones, A., Rickard, G. J. et al. 2006 The new Hadley Centre climate model (HadGEM1): Evaluation of coupled
536 simulations. *J. Clim.* **19**, pp. 1327–1353.

537 Kidston, J., Vallis, G. K., Dean, S. M. & Renwick, J. A. 2011 Can the increase in the eddy length scale under global warming
538 cause the poleward shift of the jet streams? *J. Clim.* **24**, pp. 3764–3780.

539 Kodama, C. & Iwasaki, T. 2009 Influence of the SST rise on baroclinic instability wave activity under an aquaplanet
540 condition. *J. Atmos. Sci.* **66**, pp. 2272–2287.

Laîné, A., Kageyama, M., Salas-Mélia, D., Ramstein, G., Planton, S., Denvil, S. & Tyteca, S. 2009 An energetics study of wintertime Northern Hemisphere storm tracks under 4× CO2 conditions in two ocean-atmosphere coupled models. *J. Clim.* **22**, pp. 819–839.

Lorenz, D. J. & DeWeaver, E. T. 2007 Tropopause height and zonal wind response to global warming in the IPCC scenario integrations. *J. Geophys. Res.: Atmos.* **112**.

Magnusdottir, G., Deser, C. & Saravanan, R. 2004 The effects of North Atlantic SST and sea ice anomalies on the winter circulation in CCM3. Part I: Main features and storm track characteristics of the response. *J. Clim.* **17**, pp. 857–876.

Martin, G. M., Ringer, M. A., Pope, V. D., Jones, A., Dearden, C. & Hinton, T. J. 2006 The physical properties of the atmosphere in the new Hadley Centre Global Environmental Model (HadGEM1). Part I: Model description and global climatology. *J. Clim.* **19**, pp. 1274–1301.

Rivière, G. 2011 A dynamical interpretation of the poleward shift of the jet streams in global warming scenarios. *J. Atmos. Sci.* **68**, pp. 1253–1272.

Scaife A, T S, Fereday DR, Cubasch U, Langematz U, Akiyoshi H, Bekki S, Braesicke P, Butchart N, Chipperfield MP, et al (2012) Climate change projections and stratosphere–troposphere interactions. Climate Dynamics 38:2089–2097

Schneider, T., O'Gorman, P. A. & Levine, X. J. 2010 Water vapor and the dynamics of climate changes. *Rev. Geophys.* **48**, p. RG3001.

Screen, J. A., Deser, C. & Simmonds, I. 2012 Local and remote controls on observed Arctic warming. *Geophys. Res. Lett.* **39**, p. L10,709.

Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K., M.Tignor & Miller, H., editors 2007 *IPCC, 2007: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Stroeve, J., Holland, M. M., Meier, W., Scambos, T. & Serreze, M. 2007 Arctic sea ice decline: Faster than forecast. *Geophys. Res. Lett.* **34**.

Ulbrich, U., Pinto, J., Kupfer, H., Leckebusch, G., Spangehl, T. & Reyers, M. 2008 Changing northern hemisphere storm tracks in an ensemble of IPCC climate change simulations. *J. Clim.* **21**, pp. 1669–1679.

Williams, G. P. 2006 Circulation sensitivity to tropopause height. *J. Atmos. Sci.* **63**.

Wilson, C., Sinha, B. & Williams, R. G. 2009 The effect of ocean dynamics and orography on atmospheric storm tracks. *J. Clim.* **22**, pp. 3689–3702.

570 Woollings, T. & Blackburn, M. 2012 The North Atlantic jet stream under climate change and its relation to the NAO and
571 EA patterns. *J. Clim.* **25**, pp. 886–902.
572 Woollings, T., Gregory, J. M., Pinto, J. G., Reyers, M. & Brayshaw, D. J. 2012 Response of the North Atlantic storm track
573 to climate change shaped by ocean-atmosphere coupling. *Nat. Geosci.* **5**, pp. 313–317.
574 Yin, J. H. 2005 A consistent poleward shift of the storm tracks in simulations of 21st century climate. *Geophys. Res. Lett.*
575 **32**, p. L18,701.
576 Zappa, G., Shaffrey, L. C. & Hodges, K. I. 2013a The Ability of CMIP5 Models to Simulate North Atlantic Extratropical
577 Cyclones. *J. Clim.* **26**, pp. 5379–5396.
578 Zappa, G., Shaffrey, L. C., Hodges, K. I., Sansom, P. G. & Stephenson, D. B. 2013b A Multimodel Assessment of Future
579 Projections of North Atlantic and European Extratropical Cyclones in the CMIP5 Climate Models. *J. Clim.* **26**, pp.
580 5846–5862.

List of Figures

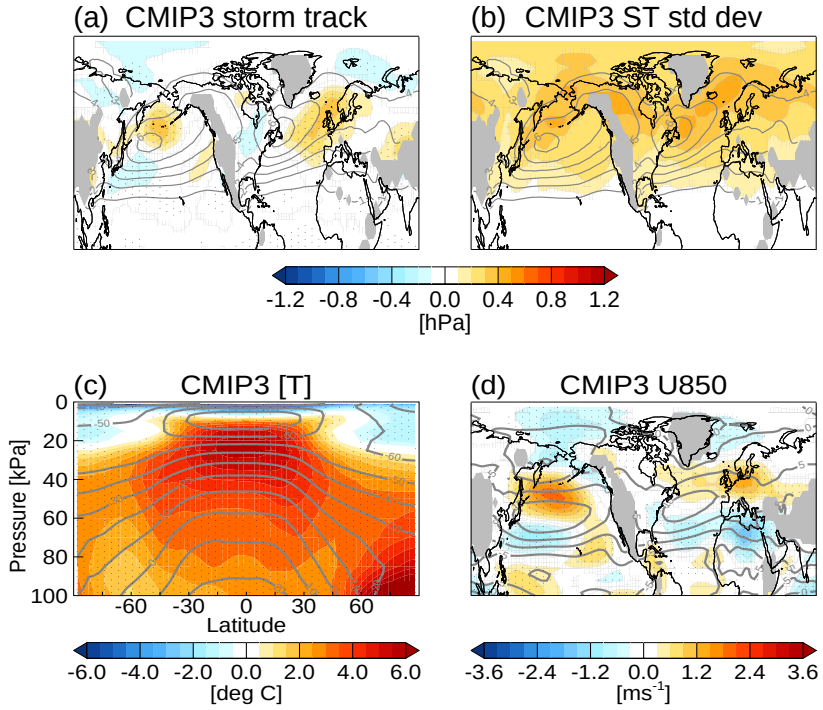

Fig. 1 Multi-model mean CMIP3 climate change response. *Shading* shows CMIP3 DJF values of **a** the multi-model mean response of the storm track, **b** the inter-model standard deviation of the storm track responses, **c** the multi-model mean response of the zonal mean temperature, and **d** the multi-model mean response of the zonal wind at 850 hPa. *Contours* in **a** and **b** show the multi-model mean 20C3M storm track (units: hPa), in **c** the multi-model mean 20C3M zonal mean temperature (units: 10 deg C), and in **d** the multi-model mean 20C3M zonal wind at 850 hPa (units: 5 ms^{-1}). The stippling in **a**, **c** and **d** indicates regions where the multi-model mean responses are non-zero at the 95% confidence level, relative to the inter-model spread, according to a Student's t-test. One run per model is used, as listed in Appendix A, with the exception of the two UKMO models which are omitted here due to the lack of available storm track data.

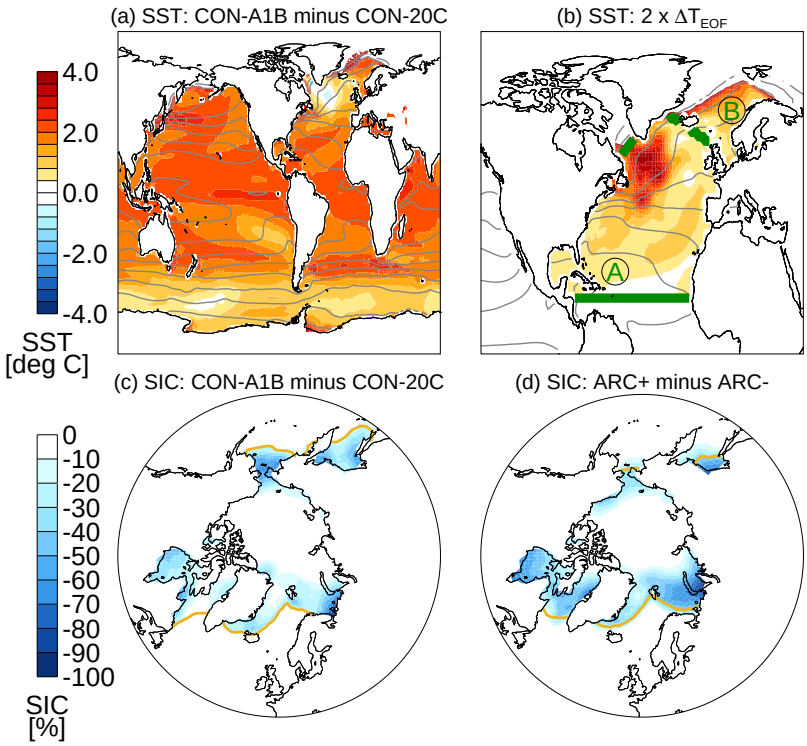

Fig. 2 DJF mean values of the SST and SIC forcing fields. **a** shows the CON-20C SST (*contours*; units: 4C) and the CON-A1B minus CON-20C SST difference (*shading*), **b** shows the CON-A1B SST (*contours*; units: 4C) and 2 × ΔT_{EOF} (*shading*) which equal to 2 x the first inter-model EOF of the CMIP3 SST responses (the EOF is normalised by the inter-model standard deviation - see text for details), **c** shows the CON-20C ice edge position (*contour*) and the CON-A1B minus CON-20C SIC difference (*shading*), and **d** shows the CON-A1B ice edge position (*contour*) and the ARC+ minus ARC- SIC difference (*shading*).

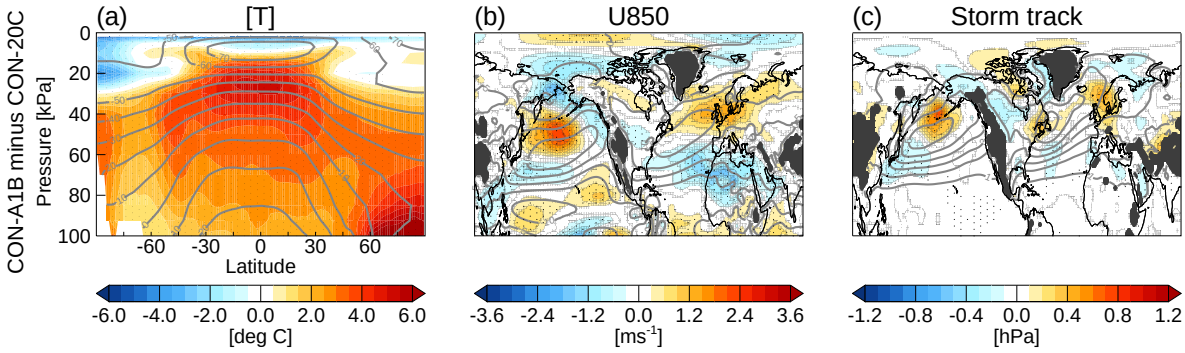

Fig. 3 DJF mean values of the CON-A1B minus CON-20C difference (*shading*). The *panels* show **a** the zonal mean temperature response, **b** the U850 response, and **c** the storm track response. The *grey contours* show the corresponding climatology fields from CON-20C for reference (*contour intervals:* 10 deg C, 5 ms^{-}1, and 1 hPa respectively), and the stippling indicates regions where the two experiments differ at the 95% confidence level, relative to the inter-annual variability, according to a Student's t-test.

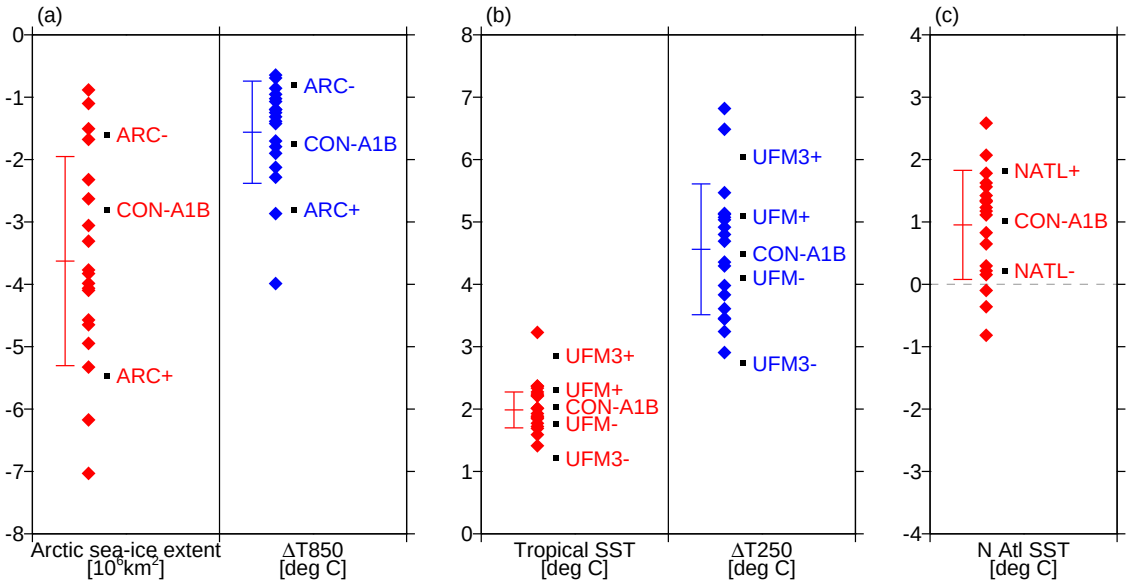

Fig. 4 The DJF climate change responses of several key variables in the CMIP3 models and in the experiments presented in this paper. Diamonds indicate the responses of the individual CMIP3 models (SRESA1B minus 20C3M), vertical bars indicate the mean and inter-model standard deviation of the CMIP3 model responses, and the squares with labels indicate the EXPT minus CON-20C difference for selected experiments. The variables shown are defined in the main text.

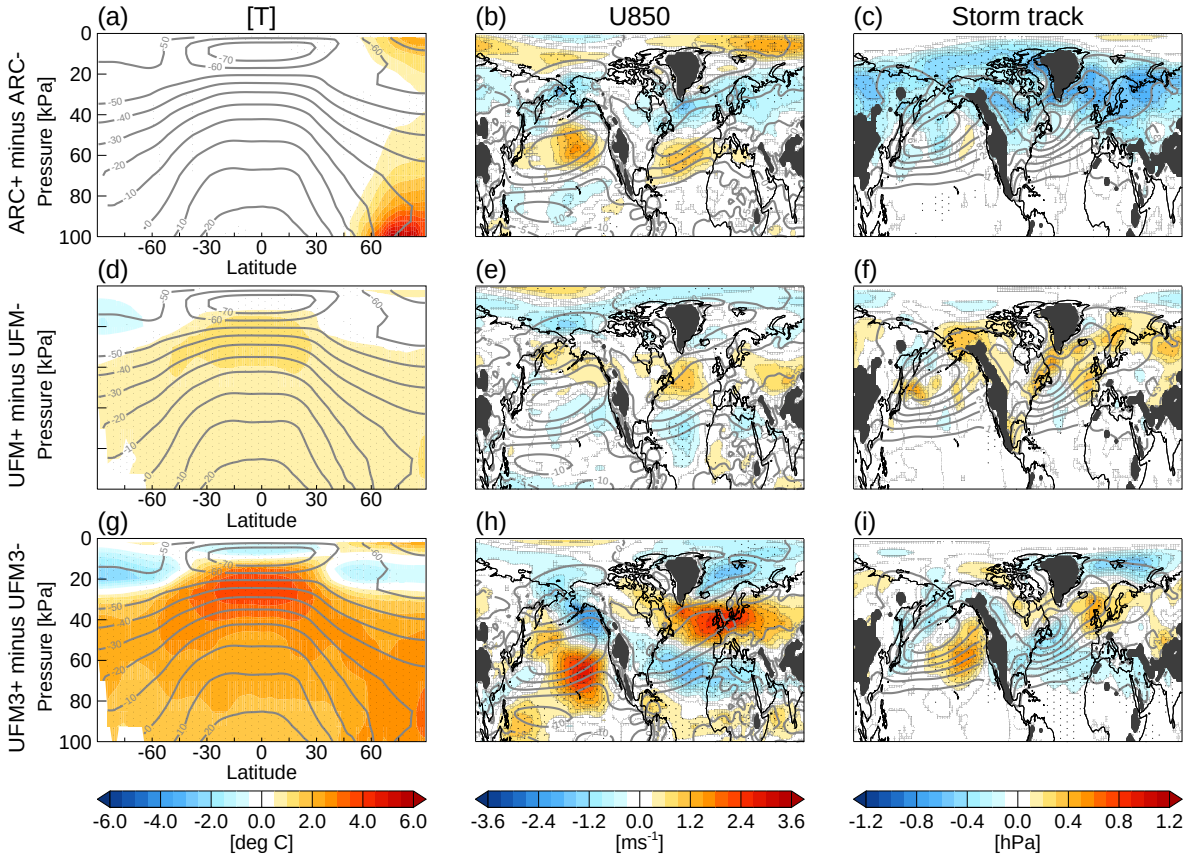

Fig. 5 DJF mean responses in the global perturbation experiments. The *columns* show, from left to right, the zonal mean temperature response, the U850 response, and the 6-2 day MSLP storm track response. The *rows* show, from top to bottom, the ARC+ minus ARC- difference, the UFM+ minus UFM- difference, and the UFM3+ minus UFM3- difference. The *grey contours* show the corresponding climatology fields from CON-A1B for reference, with *contour intervals* as in Figure 3, and the stippling indicates regions where the two experiments shown differ at the 95% confidence level, relative to the inter-annual variability, according to a Student's t-test.

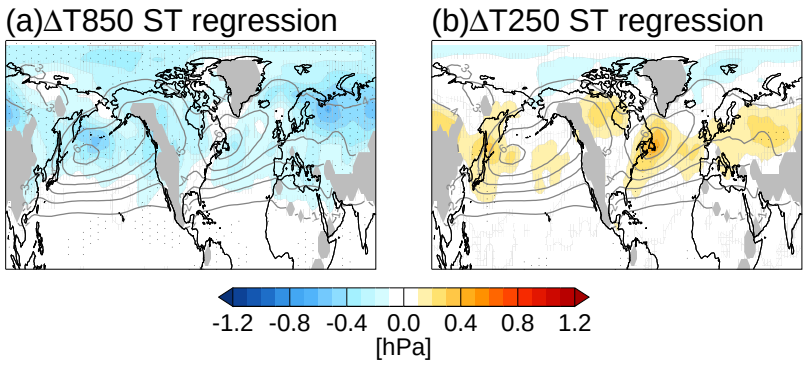

Fig. 6 Inter-model regression between the local CMIP3 DJF storm track responses and the DJF responses of **a** $\Delta T850$ and **b** $\Delta T250$. The quantity shown is the regression slope multiplied by two standard deviations of the corresponding ΔT responses and the sign of the mean ΔT response. The stippling indicates a correlation greater than 0.47, which is the 95% confidence level for a two-tailed test.

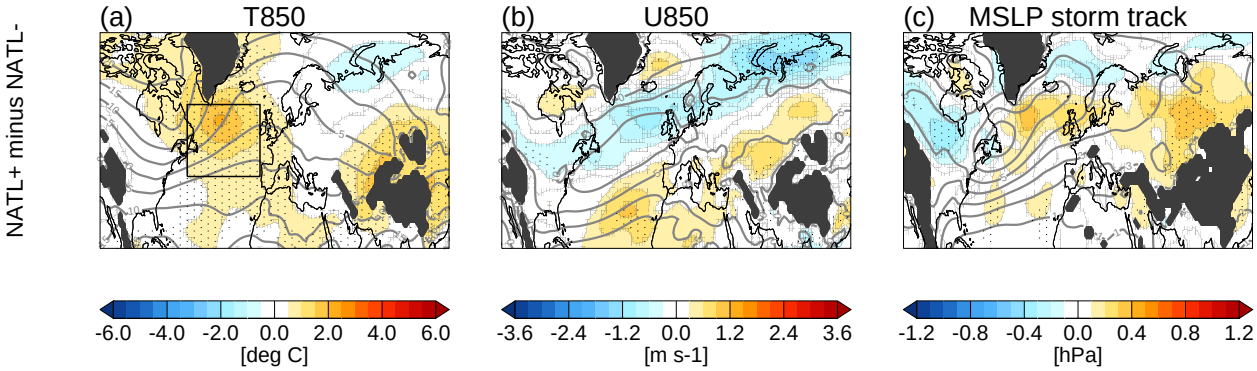

Fig. 7 DJF mean responses in the North Atlantic perturbation experiments. The *panels* show **a** the T850 response **b** the U850 response **c** and the 6-2 day MSLP storm track response. The *grey contours* show the corresponding climatology fields from CON-A1B for reference, with *contour intervals* as in Figure 3, and the stippling indicates regions where the two experiments shown differ at the 95% confidence level, relative to the inter-annual variability, according to a Student's t-test. The box in **a** indicates the region used to define the North Atlantic SST index in Figure 4 (10-60W, 40-60N).

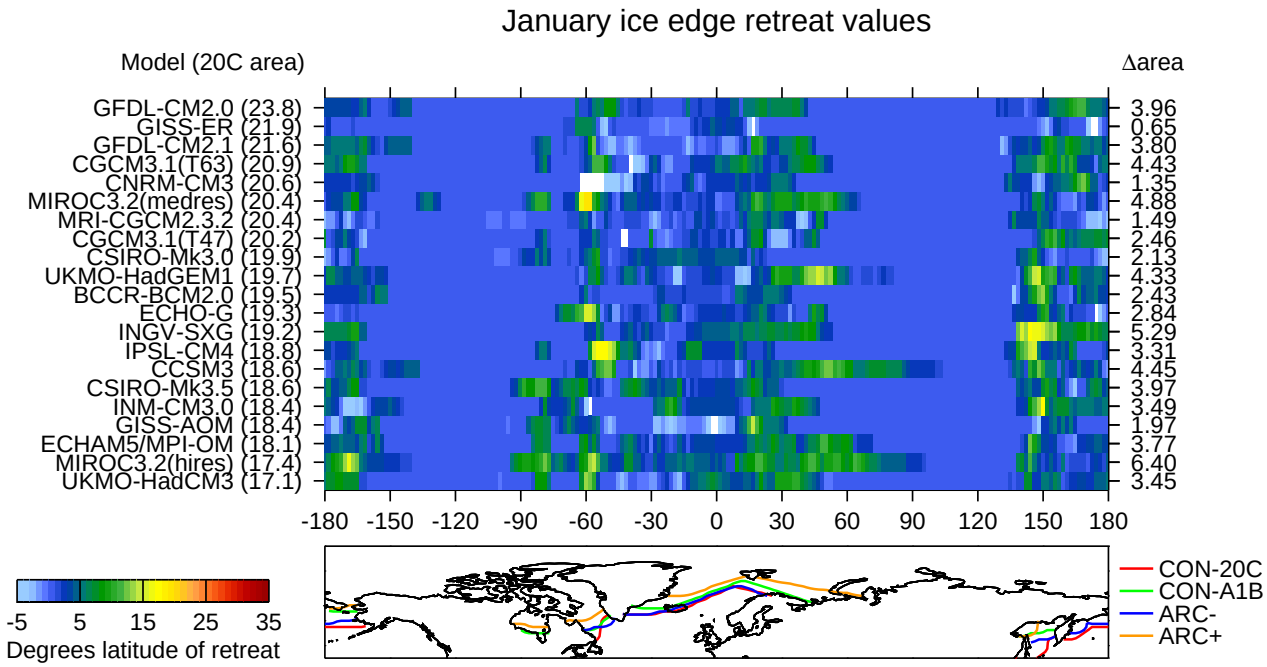

Fig. 8 The *large panel* shows the distance of ice edge retreat, as a function of longitude, for the monthly-mean January ice distributions of each CMIP3 model. The retreat is defined as the difference in ice edge position between years 1961-2000 from 20C3M and years 2061-2100 from SRESA1B and is measured in units of degrees of latitude. The models are ordered by the total area of monthly mean January ice in the 20C3M period, as indicated on the left-side axis; also indicated (right-side axis) is the change in January ice area between the 20C3M and SRESA1B periods. The *lower panel* indicates the constructed January ice edge positions for the CON-20C, CON-A1B, ARC- and ARC+ experiments, as indicated.

List of Tables

Experiment name	SST forcing field	SIC forcing field
CON-20C	Multi-model mean 20C3M	Reconstructed 20C3M
CON-A1B	Multi-model mean SRESA1B	Reconstructed SRESA1B
ARC±	CON-A1B	CON-A1B ± perturbation to Arctic sea-ice edge
UFM±	CON-A1B ± global uniform perturbation	CON-A1B
UFM3±	CON-A1B ± 3 × global uniform perturbation	CON-A1B
NATL±	CON-A1B ± North Atlantic perturbation	CON-A1B

Table 1 The experiments presented in this paper.