

The response of the northern hemisphere storm tracks and jet streams to climate change in the CMIP3, CMIP5, and CMIP6 climate models

Article

Published Version

Creative Commons: Attribution 4.0 (CC-BY)

Open Access

Harvey, B. J. ORCID: <https://orcid.org/0000-0002-6510-8181>, Cook, P., Shaffrey, L. C. ORCID: <https://orcid.org/0000-0003-2696-752X> and Schiemann, R. ORCID: <https://orcid.org/0000-0003-3095-9856> (2020) The response of the northern hemisphere storm tracks and jet streams to climate change in the CMIP3, CMIP5, and CMIP6 climate models. *Journal of Geophysical Research: Atmospheres*, 125 (23). e2020JD032701. ISSN 2169-8996 doi: 10.1029/2020JD032701 Available at <https://centaur.reading.ac.uk/94525/>

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

To link to this article DOI: <http://dx.doi.org/10.1029/2020JD032701>

Publisher: American Geophysical Union

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the [End User Agreement](#).

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

JGR Atmospheres

RESEARCH ARTICLE

10.1029/2020JD032701

Key Points:

- The spatial patterns of the historical CMIP3, CMIP5, and CMIP6 biases are similar, but their magnitudes are substantially smaller in CMIP6
- The spatial patterns of the climate change responses in the NH storm tracks and jet streams are similar in the CMIP3, CMIP5, and CMIP6 models
- The SSP2-4.5 response in CMIP6 is larger than the RCP4.5 CMIP5 response, consistent with the larger climate sensitivities in CMIP6

Supporting Information:

- Supporting Information S1

Correspondence to:

B. J. Harvey,
b.j.harvey@reading.ac.uk

Citation:

Harvey, B. J., Cook, P., Shaffrey, L. C., & Schiemann, R. (2020). The response of the northern hemisphere storm tracks and jet streams to climate change in the CMIP3, CMIP5, and CMIP6 climate models. *Journal of Geophysical Research: Atmospheres*, 125, e2020JD032701. <https://doi.org/10.1029/2020JD032701>

Received 31 DEC 2019

Accepted 2 NOV 2020

Accepted article online 9 NOV 2020

The Response of the Northern Hemisphere Storm Tracks and Jet Streams to Climate Change in the CMIP3, CMIP5, and CMIP6 Climate Models

B. J. Harvey¹ , P. Cook¹, L. C. Shaffrey¹ , and R. Schiemann¹ 

¹National Centre for Atmospheric Science, University of Reading, Reading, UK

Abstract The representation of the Northern Hemisphere (NH) storm tracks and jet streams and their response to climate change have been evaluated in climate model simulations from Phases 3, 5, and 6 of the Coupled Model Intercomparison Project (CMIP3, CMIP5, and CMIP6, respectively). The spatial patterns of the multimodel biases in CMIP3, CMIP5, and CMIP6 are similar; however, the magnitudes of the biases in the CMIP6 models are substantially lower. For instance, the multimodel mean RMSE of the North Atlantic storm track for the CMIP6 models (as measured by time-filtered sea-level pressure variance) is over 50% smaller than that of the CMIP3 models in both winter and summer, and over 40% smaller for the North Pacific. The magnitude of the jet stream biases is also reduced in CMIP6, but by a lesser extent. Despite this improved representation of the current climate, the spatial patterns of the climate change response of the NH storm tracks and jet streams remain similar in the CMIP3, CMIP5, and CMIP6 models. The SSP2-4.5 scenario responses in the CMIP6 models are substantially larger than in the RCP4.5 CMIP5 models, which is consistent with the larger climate sensitivities of the CMIP6 models compared to CMIP5.

Plain Language Summary The ability of the climate models used in the CMIP (Coupled Model Intercomparison Project) models to represent the Northern Hemisphere storm tracks and jet streams is evaluated. The newest models (CMIP6) are found to have a much better representation, especially for the North Atlantic. The response of the storm tracks and jet streams to climate change is also assessed. The spatial pattern of the response in the CMIP3, CMIP5, and CMIP6 models is found to be similar. The response of the new CMIP6 models to the SSP2-4.5 scenario is found to be larger than that of the very similar RCP4.5 scenario in the CMIP5 models, suggesting the CMIP6 models are more sensitive to the forcing from climate change.

1. Introduction

Midlatitude storms are one of the major weather risks in the extratropics of the Northern Hemisphere (NH). Strong winds and heavy precipitation associated with extratropical cyclones can lead to wind damage, inland flooding, and coastal flooding. Understanding and predicting how midlatitude storms might respond to climate change is essential for assessing future weather risks and informing climate change adaptation strategies.

In the NH, midlatitude storms (also known as extratropical cyclones) are primarily found over the Northern Atlantic and Northern Pacific Oceans (e.g., Blackmon et al., 1977). Midlatitude storms typically develop in the strongly baroclinic regions over the western North Pacific and North Atlantic where strong meridional temperature gradients exist associated with the land/sea contrast and the Kuroshio and Gulf Stream ocean currents. These regions of midlatitude storm activity are known as the storm tracks. The NH jet streams are strongly associated with the storm tracks, with the North Atlantic jet stream and the North Pacific jet stream positioned just to the south of the storm tracks.

The ability of climate models to represent the NH storm tracks and jet streams has been previously assessed in the CMIP3 and CMIP5 (Coupled Model Intercomparison Project) models. Using 2–6 day band passed MSLP (mean sea level pressure) variance, Ulbrich et al. (2009) found that the wintertime North Atlantic storm track was too zonal in the CMIP3 models, that is, placed over central Europe rather than over Scandinavia. In addition, the North Pacific storm track was found to be located slightly further north than observed. A similar spatial wintertime bias was found in the CMIP5 climate models (Chang, 2013;

©2020. The Authors.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Mizuta, 2012; Zappa, Shaffrey, & Hodges, 2013), and Zappa et al. (2014) demonstrated how these storm track biases tend to be related to biases in atmospheric blocking, particularly in winter. However, Zappa, Shaffrey, and Hodges (2013) also found that the storm track biases were smaller on average in the CMIP5 models than in the CMIP3 models. A number of reasons have been put forward for why climate models have biases in the NH storm tracks and jet streams. These include their relatively low-grid resolution (Zappa, Shaffrey, & Hodges, 2013), the representation of orography (Berckmans et al., 2013), and biases in surface temperature (Keeley et al., 2012).

The response of the NH storm tracks and jet streams to climate change have also been assessed in the CMIP3 and CMIP5 models (Chang, 2013; Mizuta, 2012; Ulbrich et al., 2009; Zappa, Shaffrey, Hodges, Sansom, et al., 2013). In general, the response can be described as a poleward shift in the upper level jet stream associated with changes in the zonal mean circulation (Shaw, 2019). However, in the winter-time North Atlantic, the storm track extends further into Europe, and there is a high-latitude reduction in storminess associated with the reduction in meridional temperature gradient and amplified warming over the Arctic (Harvey et al., 2014).

The aim of this paper is to further our understanding of the response of the NH storm tracks and jet streams to climate change. This will be achieved by addressing the following questions:

1. How well are the NH storm tracks and jet streams represented in the CMIP3, CMIP5, and CMIP6 models?
2. How are NH storm tracks and jet streams projected to respond in the CMIP3, CMIP5, and CMIP6 models?

In section 2, the methods and data sets used are described. Model biases and projections are discussed in sections 3 and 4, respectively. In section 5, conclusions are discussed.

2. Methods and Data Sets

2.1. The CMIP3, CMIP5, and CMIP6 Ensembles

This study uses the 20C3M CMIP3 simulations, the Historical CMIP5 and CMIP6 simulations, the SRESA1B simulations from CMIP3, the RCP4.5 simulations from CMIP5, and the SSP2-4.5 simulations from CMIP6. The number of climate models used is limited by the availability of daily MSLP output (19 from CMIP3, 38 from CMIP5, and currently 14 from CMIP6 for the present-day simulations, and slightly fewer for the future simulations). Where data are available for the present-day simulations but not for the future simulations, the full complement of models is used to assess the present-day biases, but only the subset with data available for both present-day and future simulations is used to assess the future changes. The precise models used in each case are described in Tables S1 and S2 in the supporting information. The choice of future scenarios represents similar middle-of-the-road projections from the different CMIPs, but they are not identical (see Collins et al., 2013, [section 12.4.9] for a description of the differences between SRESA1B and RCP4.5 and O'Neill et al., 2016, for a description of the relationship between RCP4.5 and SSP2-4.5). For the purposes of the present study, it should be noted that the ensemble mean changes in global mean surface temperature for the simulations employed here are 2.9 K (CMIP3), 1.9 K (CMIP5), and 2.6 K (CMIP6). The weaker warming in CMIP5 than CMIP3 is likely due to differences between the SRESA1B and RCP4.5 forcing scenarios, whereas the larger warming in CMIP6 than CMIP5 may be associated with the higher climate sensitivities exhibited by the CMIP6 models (Zelinka et al., 2020).

Since many of the CMIP3 models only have single realizations, only a single ensemble member is included from each available model. The analysis has been repeated using the ensemble mean wherever multiple realizations are available, and the results are insensitive to this choice. The dates used are the most recent 30-year period from the present-day simulations (December 1969 to November 1999 for CMIP3, December 1975 to November 2005 for CMIP5, and December 1979 to November 2009 for CMIP6), and the last 30 years of the 21st century from the future simulations (December 2070 to November 2100) except for CMIP3 where daily MSLP output is only available for the 19-year period December 2081 to November 2100. The analysis of the “present-day” simulations has been repeated using a common period for each CMIP (December 1969 to November 1999), and the results are also insensitive to this choice.

2.2. Observational Data Sets

The climate model results in this study are compared to those from the ERA5 reanalysis during 1979–2010. ERA5 has a resolution of 30 km and uses 137 levels up to a height of 80 km. Additional details of ERA5 can be found at <https://cds.climate.copernicus.eu/>. Previous work (e.g., Zappa, Shaffrey, & Hodges, 2013) has shown that the representation of the storm tracks in several commonly used reanalysis products are all much more similar than current climate model biases, and as such we would expect similar results if we used a different reanalysis data set.

2.3. Assessing Storm Tracks and Jet Streams

The main diagnostic used to evaluate the storm tracks is the root mean square (RMS) 2–6 day bandpass-filtered MSLP. Daily mean MSLP data are used, filtered with a 61-day Lanczos filter (Duchon, 1979) and then RMS values computed over each available season. The computation is performed on the native grid of each model, and the resulting storm track field is interpolated to a common $1 \times 1^\circ$ grid for plotting. The reason for using this variable is the limited high-frequency output from CMIP3, which precludes the use of cyclone tracking schemes such as that used in Zappa, Shaffrey, and Hodges (2013). The jet streams are evaluated using monthly mean zonal wind at 250 hPa. Only boreal winter (DJF) and boreal summer (JJA) results are discussed for brevity, as the results in the equinoctial seasons tend to lie between those for winter and summer.

3. The Representation of the NH Storm Tracks and Jet Streams in CMIP3, CMIP5, and CMIP6

3.1. Winter Storm Tracks and Jet Streams

Figure 1 shows the NH DJF storm tracks and jet streams from the ERA5 reanalysis, and the DJF biases in the CMIP3, CMIP5, and CMIP6 multimodel means. The two maxima in the MSLP variance are the Northern Atlantic and North Pacific storm tracks (Figure 1a) and the two maxima in the zonal wind field are the North Atlantic and North Pacific jet streams which are located to the south of the storm tracks (Figure 1b).

In the CMIP3 models (Figures 1c and 1d), the North Atlantic storm track is too zonal. That is, the storm track is too strong over Europe and too weak over the Nordic Seas region compared to the ERA5 reanalysis. An associated southward bias is also present in the North Atlantic jet stream (Figure 1d). In the North Pacific, the CMIP3 storm track is too strong on both its northeastern and northwestern flanks. This contrasts with the jet stream which has an equatorward bias (Ulbrich et al., 2009). The reason for the discrepancy between the spatial structures of the North Pacific storm track and jet stream biases is not clear. However, the relationship between baroclinicity and the storm track is known to be complex in this region during DJF as evidenced by the numerous proposed mechanisms for the so-called “midwinter minimum” (see Park & Lee, 2020, and references therein).

The spatial distribution of biases in the CMIP5 models (Figures 1e and 1f) is similar to that in the CMIP3 models. As noted by Zappa, Shaffrey, and Hodges (2013), there is some improvement in the magnitude of the North Atlantic biases in CMIP5 compared to CMIP3. There is a less strong equatorward bias in the jet stream, a weaker negative bias in the storm track in the Nordic Seas region, and a slightly weaker positive storm track bias over Europe. The change in the North Pacific biases between the CMIP3 and CMIP5 models is mixed; while there is an improvement in the storm track biases on the northeastern and northwestern flanks, the equatorward bias of the jet is stronger in the CMIP5 models, and the core of the storm track has become too weak.

In the CMIP6 models, there has been a further improvement in the representation of the North Atlantic storm track and jet stream (Figures 1g and 1h) compared to the CMIP5 and CMIP3 models, although the same spatial pattern remains with the storm track placed slightly too far south over Europe. This improvement is examined further in section 3.3 in terms of the latitude of the storm track maximum. In the North Pacific, the biases in the CMIP6 models do not appear to be improved over those in the CMIP3 and CMIP5 models.

To summarize this evolution of biases across the CMIP generations, Table 1 shows the spatial RMSE of the storm track and jet stream biases computed over regions loosely representing the North Atlantic and North

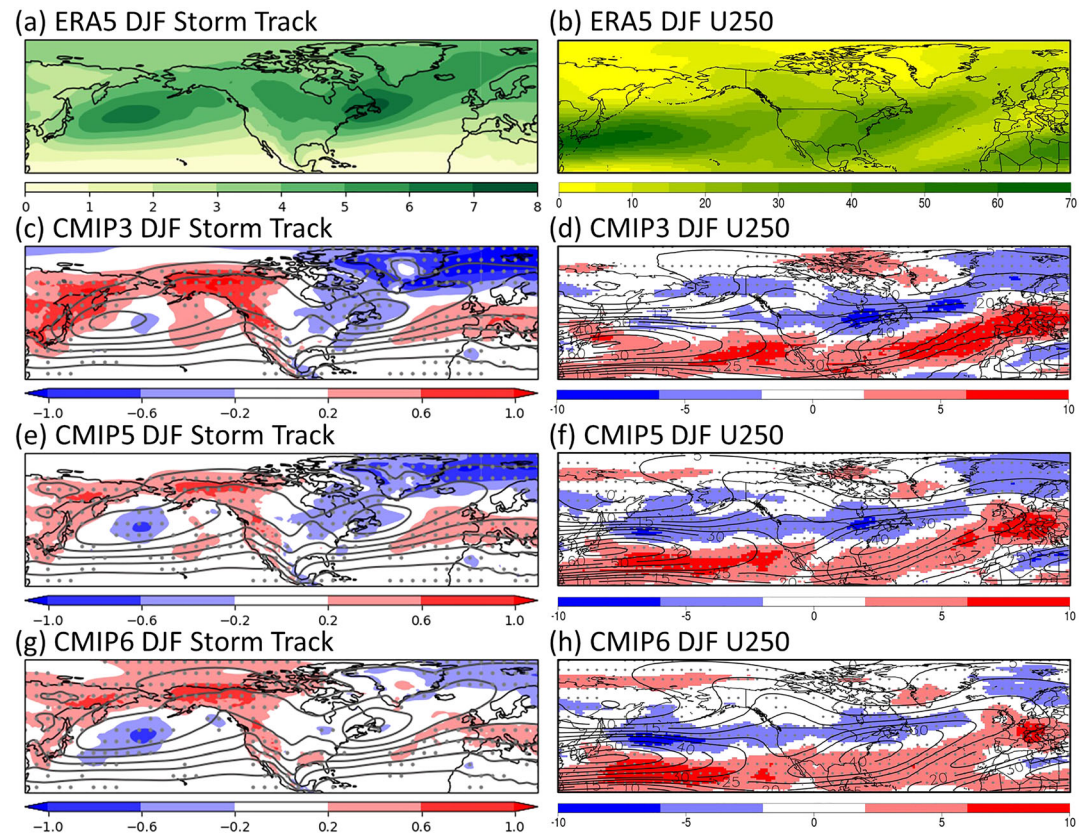


Figure 1. DJF present-day storm tracks (left panels; units: hPa) and 250 hPa zonal winds (right panels; units: ms^{-1}) from (a, b) ERA5, and multimodel mean CMIP minus ERA5 differences for (c, d) CMIP3, (e, f) CMIP5, and (g, h) CMIP6. The contours in (c–h) show the corresponding present-day multimodel mean climatology for each CMIP (contour intervals: 1 hPa and 5 ms^{-1}). The details of the time periods used are in the text. Stippling denotes statistical significance at the 5% level using a t test.

Pacific storm tracks, defined here as (0–60°W, 30–90°N) and (140–220°E, 30–90°N), respectively. These diagnostics represent the typical magnitudes of the biases in each region. They are computed separately for each model, and then the multimodel mean is found. This multimodel mean RMSE value for the North Atlantic storm track in the CMIP6 models is less than half that of the CMIP3 models. In fact, in both regions, this measure of the storm track bias shows a significant improvement with each new CMIP (see Table 1 caption for details). The multimodel mean jet stream RMSEs likewise show an improvement with each new CMIP; however, the improvement is smaller in percentage terms and only the CMIP5 to CMIP6 North Atlantic improvement passes the significance test.

Table 1

Multimodel Mean of the Spatial RMSE of the DJF Storm Track and Jet Stream Biases Computed Over the North Atlantic (0–60°W, 30–90°N) and North Pacific (140–220°E, 30–90°N) Regions

DJF	North Atlantic		North Pacific	
	Storm track (hPa)	Jet stream (ms^{-1})	Storm track (hPa)	Jet stream (ms^{-1})
CMIP3	0.835	5.57	0.755	4.42
CMIP5	0.653*	4.69	0.588*	4.37
CMIP6	0.389*	3.32*	0.434*	3.80

Note. Asterisks indicate a significant improvement over the previous CMIP (i.e., CMIP3 to CMIP5, and CMIP5 to CMIP6) according to a two-sample t test at the 0.05 level.

3.2. Summer Storm Tracks and Jet Streams

Figure 2 shows the NH JJA storm tracks and jet streams from the ERA5 reanalysis and the corresponding biases in the CMIP3, CMIP5, and CMIP6 multimodel means. In summer, the storm tracks and jet streams shift northward of their wintertime positions, consistent with the seasonality of the Hadley circulation. In addition, the summertime upper level jet streams are orientated more zonally than they are in winter.

Both the North Pacific and North Atlantic summertime storm tracks are weaker than observed in the CMIP3 models, with this negative bias as large as 20% in places, while there is more storm activity over North America and eastern Asia than in ERA5 (Figure 2c). The North Pacific summertime jet stream bias has a banded structure with values up to 6 ms^{-1} too weak in the jet core and up to 6 ms^{-1} too strong on both the

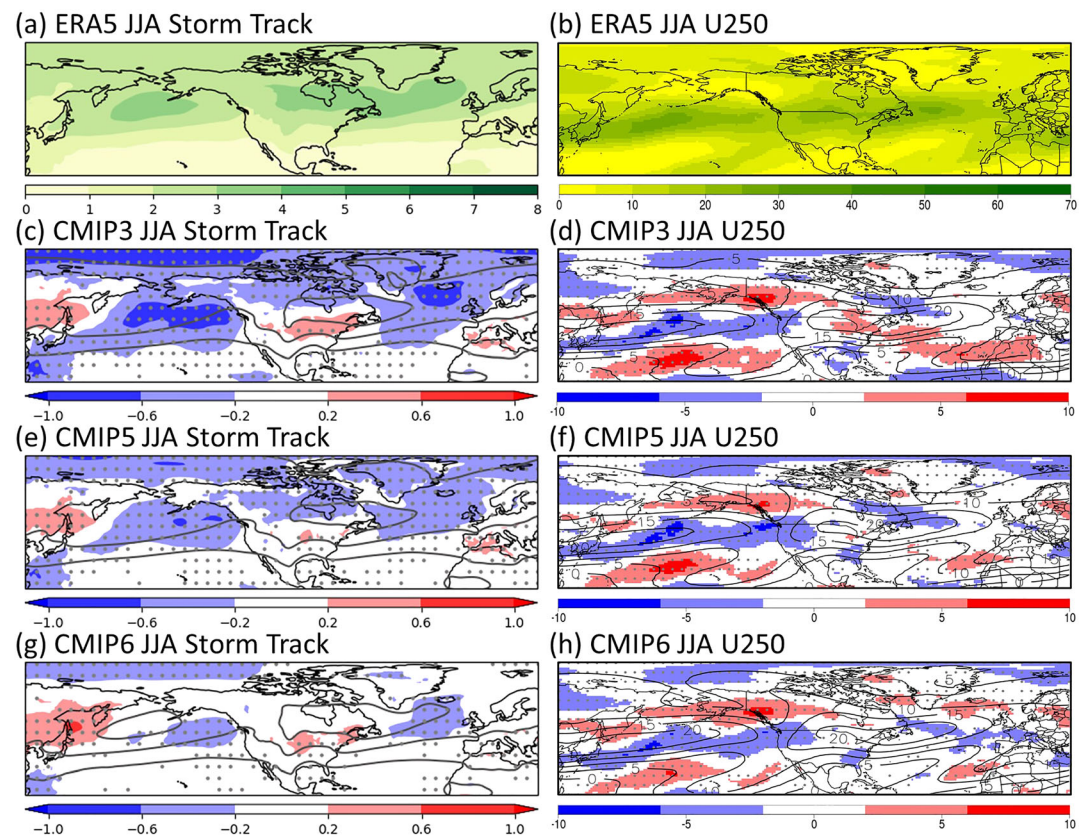


Figure 2. As Figure 1, but for JJA.

northern and southern flanks. Over the North Atlantic, the summertime jet stream, in contrast, is shifted slightly to the south (Figure 2d).

In the CMIP5 models, the summertime biases are somewhat improved over CMIP3 (Figures 2e and 2f). The storm tracks in the CMIP5 models are generally stronger than in the CMIP3 models, although they are still weaker than those in the ERA5 reanalysis by around 15%. The jet stream biases in the North Pacific are similar to those in the CMIP3 models, but the biases in the North Atlantic are much improved and in close agreement with the jet streams in the ERA5 reanalysis.

The biases in CMIP6 in the North Pacific and North Atlantic storm tracks are even more improved compared to those in the CMIP5 models, except over eastern Asia where the positive bias has increased. The storm tracks in the North Pacific and the North Atlantic are only slightly weaker than observed in the ERA5 reanalysis (Figure 2g). However, the biases in the jet streams are similar to those found in the CMIP5 models (Figure 2h).

Table 2
As Table 1, But for JJA

JJA	North Atlantic		North Pacific	
	Storm track (hPa)	Jet stream (ms^{-1})	Storm track (hPa)	Jet stream (ms^{-1})
CMIP3	0.571	3.21	0.680	4.21
CMIP5	0.411*	2.86	0.512*	4.31
CMIP6	0.271*	2.48 [#]	0.348*	3.51*

Note. The additional # symbol here indicates that the CMIP6 North Atlantic jet stream improvement is significant relative to CMIP3, but not CMIP5.

As before, these changes are summarized in Table 2 in terms of the spatial RMSE of the biases in each model in both the North Atlantic and North Pacific regions. Like DJF, the North Atlantic summertime storm track shows the largest percentage improvement according to this metric, with the CMIP6 value less than half of that in CMIP3. Again, this measure of the storm track bias shows a significant improvement in both regions with each new CMIP. The corresponding jet stream improvement is smaller in percentage terms, and this time, only the CMIP5 to CMIP6 North Pacific improvement passes the significance test.

3.3. Intermodel Spread in North Atlantic Storm Track Biases

In this subsection, the improvement in the zonal bias of the North Atlantic storm tracks is briefly assessed in terms of intermodel spread.

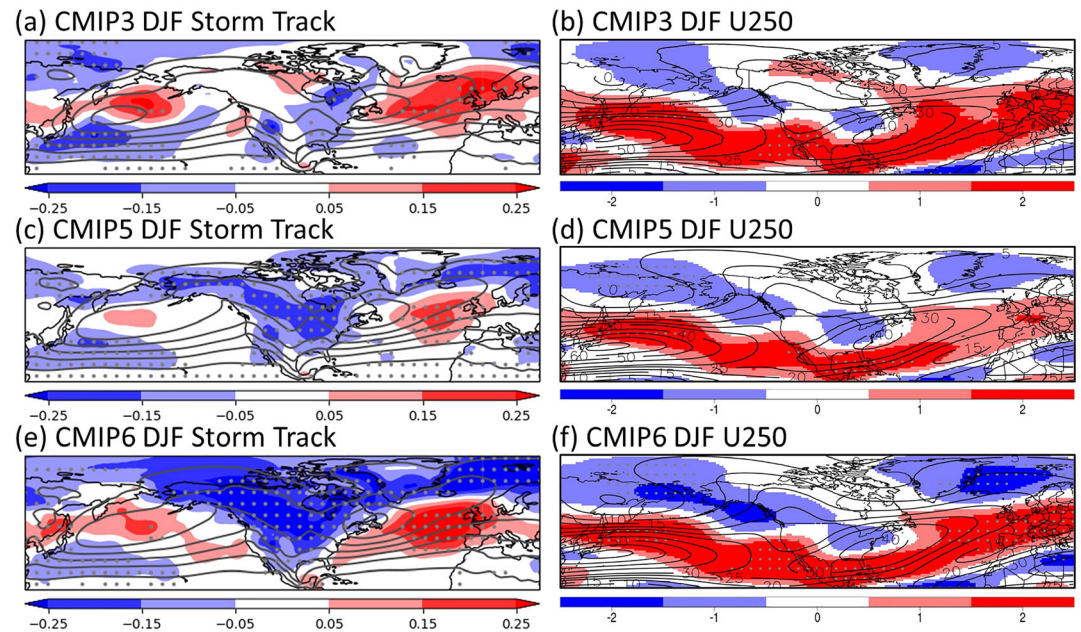


Figure 3. DJF multimodel mean future minus present-day storm tracks (left panels; units: hPa) and 250 hPa zonal winds (right panels; units: ms^{-1}) from (a, b) CMIP3, (c, d) CMIP5, and (e, f) CMIP6. The contours show the corresponding present-day multimodel mean climatology for each CMIP (contour intervals: 1 hPa and 5 ms^{-1}). The details of the time periods are in the text. Stippling denotes statistical significance at the 5% level using a *t* test.

Figure S1d shows the distribution of DJF values of the latitude of the North Atlantic storm track at the Greenwich Meridian (that is, at 0° longitude) in the CMIP3, CMIP5, and CMIP6 models. It can be seen from Figure S1d that most models place the storm track too far south in this location, but that the mean bias reduces significantly between consecutive CMIPs (bold arrows) and that the intermodel standard deviations for CMIP3 (5.6°) and CMIP5 (4.1°) are substantially larger than for CMIP6 (2.7°). To test the robustness of the reduction in standard deviation, a 10,000 member random resampling of 14 (i.e., the number of CMIP6 models) of the CMIP3 and CMIP5 models has been performed. None of the CMIP3 samples exhibited a standard deviation as small as the CMIP6 value, and only 8.6% of the CMIP5 samples exhibited a standard deviation as small as the CMIP6 value. This reduction in the intermodel spread is shown for individual models in Figures S1a–S1c. The improvement in the ability of individual CMIP6 models to capture the observed structure of the North Atlantic storm track can be clearly seen.

The distribution of JJA values of the latitude of the North Atlantic storm track is shown in Figure S2. Figure S2 also shows that there is a substantial reduction in both the mean bias and intermodel standard deviation in the position of the JJA North Atlantic storm track in CMIP6 compared to the CMIP5 and CMIP3 models. While the reduction in bias between CMIPs does not pass the significance test, the reduction in intermodel standard deviation is robust in the sense that, under the sampling strategy described above, 0% of the CMIP3 or CMIP5 samples exhibit a standard deviation as small as the CMIP6 value. In summary, the analysis of intermodel spread suggests that the reduction in the multimodel mean biases in the CMIP6 models is occurring from a systematic improvement in the representation of the North Atlantic storm track in the CMIP6 ensemble.

4. Projected Response of the NH Storm Tracks and Jet Streams to Climate Change

In this section, the end of the 21st century response of the NH storm tracks and jet streams to climate change is assessed in the CMIP3, CMIP5, and CMIP6 models. Figure 3 shows the DJF responses to the SRESA1B scenario in the CMIP3 models and to the RCP4.5 (SSP2-4.5) scenario in the CMIP5 and CMIP6 models. In the CMIP3 models, the DJF North Pacific storm track weakens on its southern flank and strengthens on its northern flank (Figure 3a). This results in a northward shift in its position and is associated with a

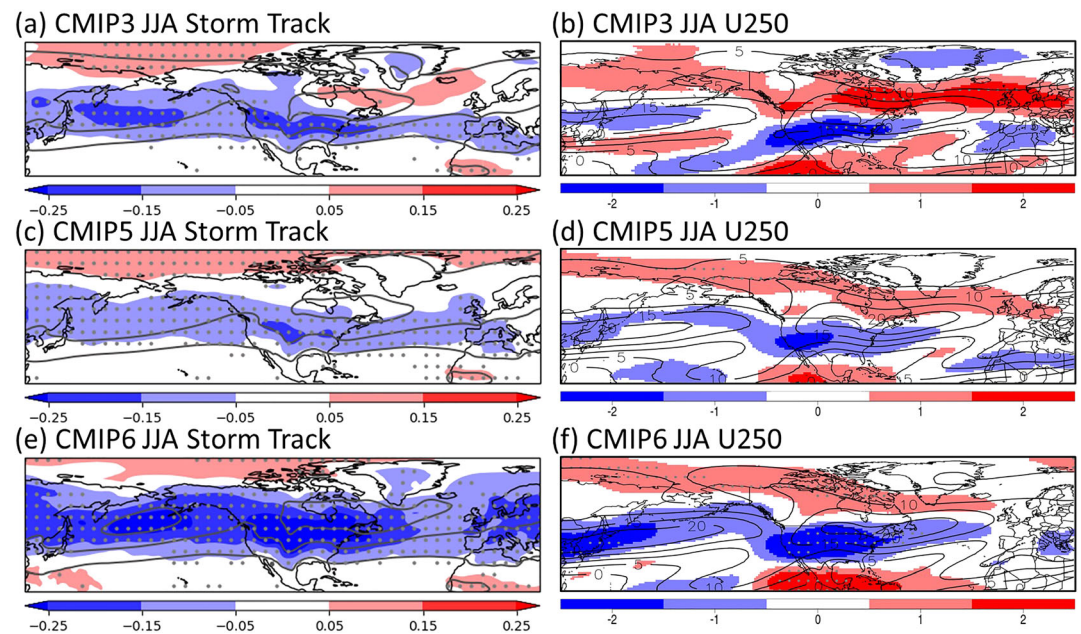


Figure 4. From (a, b) CMIP3, (c, d) CMIP5, and (e, f) CMIP6. As Figure 3 but for JJA.

northward shift in the North Pacific jet stream (Figure 3b) (Shaw, 2019). The magnitudes of the changes are typically 5%–10% of the mean. In the North Atlantic, the main response of the CMIP3 models is a strengthening and extension of the DJF storm track. In addition, the North Atlantic storm track weakens slightly on its northern and southern flanks. Again, the changes are typically 5%–10% of the mean fields. The response of the North Atlantic storm track is reflected in the change in the North Atlantic jet stream, which also strengthens and extends further into Europe. Ulbrich et al. (2009) suggested the extension of the North Atlantic storm track was a result of the strengthening of baroclinicity associated with changes in North Atlantic Ocean currents.

In the CMIP5 models, the spatial response of the NH storm tracks and jet streams is generally similar to that in the CMIP3 models (Figures 3c and 3d), even though the SRESA1B and RCP4.5 scenarios are not the same. In the DJF North Atlantic, however, the weakening on the northern flank of the storm track is stronger and more widespread, while the extension of the storm track toward Europe is weaker (Harvey et al., 2012; Zappa, Shaffrey, Hodges, Sansom, & et al., 2013). The response of the CMIP6 models is shown in Figures 3e and 3f. Again, the spatial response of the NH storm tracks and jet streams is very similar to that found in the CMIP5 models. Given the forcing in RCP4.5 and SSP2-4.5 scenario is the same, it is interesting to note that the CMIP6 response is substantially larger in magnitude than in the CMIP5 response (typically 50% larger in the North Atlantic storm track). This is likely associated, at least partly, with the larger climate sensitivities of the CMIP6 models noted above. It is indeed found that simply scaling by the storm track responses with the global mean temperature responses does bring the CMIP5 and CMIP6 multimodel mean responses closer together, however, the scaled CMIP6 response remains stronger than CMIP5, particularly over the North Atlantic (not shown). The reason for this difference is unclear, and there are several candidates to be explored, most notably changes in Arctic sea ice, North Atlantic SSTs, and the land-sea contrast in surface air temperature.

Figure 4 shows the JJA storm track and jet stream response to the SRESA1B scenario in the CMIP3 models and to the RCP4.5 (SSP2-4.5) scenario in the CMIP5 and CMIP6 models. In the North Atlantic, the spatial pattern of the response is primarily a northward shift of the jet stream in all CMIPs. There is a notable weakening of the storm track in CMIP5 and CMIP6 and a slight northward shift in CMIP3. It should be noted, however, that because the strongest weakening in CMIP5 and CMIP6 is located to the south of the storm track maximum, there is a gradient in the response pattern across the storm track and the net effect is a combination of a weakening and a poleward shift. In the North Pacific, the response is dominated by a

weakening of the storm track with less evidence of a shift. As for DJF, it is interesting to note that for JJA the CMIP6 response is substantially larger in magnitude than in the CMIP5 response in both basins. Scaling the response maps by the global mean temperature response again accounts for much, but not all, of the difference between the CMIP5 and CMIP6 responses, with the CMIP6 responses remaining slightly stronger even when scaled.

5. Conclusions

In this study, the representation of the NH storm tracks and jet streams and their response to climate change have been evaluated in the CMIP3, CMIP5, and CMIP6 climate models. The main conclusions are:

1. The spatial pattern of the biases in the multimodel mean of the CMIP3, CMIP5, and CMIP6 models are similar. In winter, the DJF North Atlantic storm track and jet stream is too zonal and placed further south over Europe, while the DJF North Pacific storm track is placed further north of its observed position. In summer, the JJA North Atlantic and JJA North Pacific storm tracks are weaker than observed.
2. The multimodel mean bias in the CMIP6 models is substantially lower in the DJF North Atlantic storm track and jet stream than in the CMIP3 and CMIP5 models. In summer, the biases in the JJA North Atlantic and North Pacific storm tracks are also much reduced in the CMIP6 models compared to the CMIP3 and CMIP5 models. However, the biases in the DJF North Pacific storm track and jet stream in the CMIP3, CMIP5, and CMIP6 models are similar. An analysis of the intermodel spread of biases suggests the improvement in the North Atlantic arises from a systematic improvement in the models in the CMIP6 ensemble.
3. The spatial pattern of the climate change response in the NH storm tracks and jet streams are similar in the CMIP3, CMIP5, and CMIP6 models. The SSP2-4.5 scenario responses in the CMIP6 models are substantially larger than in the RCP4.5 CMIP5 models, which is consistent with the larger climate sensitivities of the CMIP6 models compared to CMIP5.

The choice of future emissions scenarios used here are the middle-of-the-road experiments from each CMIP, and the focus has been on the late-21st century minus late-20th century change. No attempt has been made to compare different emissions scenarios within each CMIP, nor the time evolution of the response. While at leading order the storm track responses might be expected to scale with the global mean temperature response, for instance, Zappa, Shaffrey, Hodges, Sansom, et al., (2013) showed that the CMIP5 RCP8.5 storm track response is approximately double the RCP4.5 response, some studies also present evidence for a non-linear response of the storm tracks to forcing (e.g., Catto et al., 2011; Li et al., 2018). The extent to which the results presented here apply more generally is left to future work.

In addition, the choice of storm track diagnostic used here is constrained by the availability of data stored from CMIP3. Priestley et al. (2020) recently analyzed historical biases in just the CMIP5 and CMIP6 models using a feature-tracking algorithm and found similar results to here, with a reduction of the wintertime zonal bias in the North Atlantic in CMIP6. Another avenue for future work is to analyze the future projections of the storm tracks in more detail with additional storm track diagnostics. Numerous studies have compared Eulerian measures of storm activity, such as that used here, with more sophisticated feature-tracking algorithms and found that biases and/or changes in the large-scale features tend to be closely mirrored between the two (e.g., Hoskins & Hodges, 2002; Zappa, Shaffrey, & Hodges, 2013; Zappa, Shaffrey, Hodges, Sansom, et al., 2013). However, by their nature feature-tracking algorithms can provide more small-scale detail (e.g., the Mediterranean storm track) and more storm-specific diagnostics than Eulerian variance methods (e.g., cyclone frequency, intensity, and genesis locations), aspects which are invaluable for detailed impact studies as well as understanding changes in the dynamical properties of cyclones.

The main unresolved question is to understand the causes of the improvements between the CMIP generations. One likely candidate is improvements in model parametrization schemes such as convection (Jung et al., 2010) and drag (Pithan et al., 2016; Sandu et al., 2019) which can influence planetary-scale stationary wave patterns. Another candidate is the general increase in model resolution, for example, Zappa, Shaffrey, and Hodges (2013) and Anstey et al. (2013) find evidence that higher resolution climate models have an improved representation of the North Atlantic storm track. This may be a result of an improved representation of orography (e.g., Berckmans et al., 2013), a better representation of tropical rainfall and corresponding stationary waves (e.g., Shaffrey et al., 2009), or improved weather regime dynamics (Dawson & Palmer,

2015). A brief analysis of the relationship between the wintertime storm track biases (defined in section 3.1) and CMIP3, CMIP5, and CMIP6 model resolution has been performed. When the models from all three CMIPs are taken together, a linear least-squares fit produces a robust relationship ($p < 0.05$) but accounts for less than 20% of the intermodel variance in each basin. In contrast, the summertime storm track biases exhibit more robust correlations, with resolution explaining 26% (North Atlantic) and 37% (North Pacific) of the total intermodel variance. This suggests there may be some relationship between model resolution and biases in some aspects of mid-latitude circulation. However, it is difficult to disentangle cause and effect in the CMIP models, which are primarily “ensembles of opportunity.”

An additional aspect is that the North Atlantic improvements in the wintertime storm track and the jet stream occur concurrently and appear to be related to each other. In summer, however, the jet bias does not improve, whereas the storm tracks are improved in both basins. The storm tracks are known to respond strongly to low-level baroclinicity (e.g., Harvey et al., 2014), and it is possible this inconsistency can be explained by changes in the low-level baroclinicity across the CMIP ensembles. However, an analysis of the relationship between the RMSE storm track diagnostic and the low-level baroclinicity, defined (following Harvey et al., 2014) as the difference in temperature at 850 hPa in the tropical (30°S–30°N) and polar (60°N–90°N) regions, found no evidence for a relationship in either basin or season.

Finally, the representation of the DJF North Pacific storm track and jet stream in particular are likely to be associated with biases in the distribution of the precipitation in the Tropical Pacific Ocean, which are associated with long-standing issues such as the double Intertropical Convergence Zone (Li & Xie, 2014). Addressing these long-standing errors is essential to reducing systematic biases in climate models and increasing confidence in climate change projections.

Data Availability Statement

The CMIP3, CMIP5, and CMIP6 data sets can be obtained from the Earth Federation System Grid at <https://esgf-index1.ceda.ac.uk/projects/esgf-ceda/>. The CMIP6 data are publicly available, and the CMIP5 data are available under the following license <https://esgf-node.llnl.gov/ac/subscribe/CMIP5%20Research>. ERA5 is the Fifth ECMWF atmospheric reanalysis and is available on the Copernicus Climate Change Service Climate Data Store at <https://cds.climate.copernicus.eu/cdsapp#!/home>

Acknowledgments

This paper is supported by European Union's Horizon 2020 research and innovation program under grant agreement 727862 (project APPLICATE) and grant agreement 690462 (project WINDSURFER) under the ERA4CS ERA-NET. The authors were also supported by the National Centre for Atmospheric Science which is funded by the UK Natural Environment Research Council under contract R8/H12/83.

References

- Anstey, J. A., Davini, P., Gray, L. J., Woollings, T. J., Butchart, N., Cagnazzo, C., et al. (2013). Multi-model analysis of northern hemisphere winter blocking: Model biases and the role of resolution. *Journal of Geophysical Research: Atmospheres*, 118, 3956–3971. <https://doi.org/10.1002/jgrd.50231>
- Berckmans, J., Woollings, T., Demory, M.-E., Vidale, P.-L., & Roberts, M. (2013). Atmospheric blocking in a high resolution climate model: Influences of mean state, orography and eddy forcing. *Atmospheric Science Letters*, 14, 34–40. <https://doi.org/10.1002/asl2.412>
- Blackmon, M. L., Wallace, J. M., Lau, N., & Mullen, S. L. (1977). An observational study of the northern hemisphere wintertime circulation. *Journal of the Atmospheric Sciences*, 34, 1040–1053. <https://doi.org/10.1175/1520-0469>
- Catto, J. L., Shaffrey, L. C., & Hodges, K. I. (2011). Northern hemisphere extratropical cyclones in a warming climate in the HiGEM high-resolution climate model. *Journal of Climate*, 24, 5336–5352. <https://doi.org/10.1175/2011JCLI4181.1>
- Chang, E. (2013). CMIP5 projection of significant reduction in extratropical cyclone activity over North America. *Journal of Climate*, 26, 9903–9922.
- Collins, M., Knutti, R., Arblaster, J., Dufresne, J.-L., Fichet, T., Friedlingstein, P., et al. (2013). Long-term climate change: Projections, commitments and irreversibility. In *Climate change 2013: The physical science basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (Chap. 12, pp. 1029–1136). Cambridge: Cambridge University Press.
- Duchon, C. E. (1979). Lanczos filtering in one and two dimensions. *Journal of Applied Meteorology*, 18(8), 1016–1022.
- Harvey, B., Shaffrey, L., & Woollings, T. (2014). Equator-to-pole temperature differences and the extra-tropical storm track responses of the CMIP5 climate models. *Climate Dynamics*, 43, 1171–1182. <https://doi.org/10.1007/s00382-013-1883-9>
- Harvey, B. J., Shaffrey, L. C., Woollings, T. J., Zappa, G., & Hodges, K. I. (2012). How large are projected 21st century storm track changes? *Geophysical Research Letters*, 39, L18707. <https://doi.org/10.1029/2012GL052873>
- Hoskins, B. J., & Hodges, K. I. (2002). New perspectives on the northern hemisphere winter storm tracks. *Journal of the Atmospheric Sciences*, 59(6), 1041–1061.
- Jung, T., Balsamo, G., Bechtold, P., Beljaars, A. C., Köhler, M., Miller, M. J., et al. (2010). The ECMWF model climate: Recent progress through improved physical parametrizations. *Quarterly Journal of the Royal Meteorological Society*, 136(650), 1145–1160. <https://doi.org/10.1002/qj.634>
- 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. *Quarterly Journal of the Royal Meteorological Society*, 138, 1774–1783. <https://doi.org/10.1002/qj.1912>
- Li, C., Michel, C., Graff, L. S., Bethke, I., Zappa, G., Bracegirdle, T. J., et al. (2018). Midlatitude atmospheric circulation responses under 1.5°C and 2.0°C warming and implications for regional impacts. *Earth System Dynamics*, 9(2), 359–382. <https://doi.org/10.5194/esd-9-359-2018>

- Li, G., & Xie, S. (2014). Tropical biases in CMIP5 multimodel ensemble: The excessive equatorial Pacific cold tongue and double ITCZ problems. *Journal of Climate*, 27, 1765–1780. <https://doi.org/10.1175/JCLI-D-13-00337.1>
- Mizuta, R. (2012). Intensification of extratropical cyclones associated with the polar jet change in the CMIP5 global warming projections. *Geophysical Research Letters*, 39, L19707. <https://doi.org/10.1029/2012GL053032>
- O'Neill, B. C., Tebaldi, C., van Vuuren, D. P., Eyring, V., Friedlingstein, P., Hurtt, G., et al. (2016). The scenario model intercomparison project (ScenarioMIP) for CMIP6. *Geoscientific Model Development*, 9(9), 3461–3482. <https://doi.org/10.5194/gmd-9-3461-2016>
- Dawson, A., & Palmer, T. N. (2015). Simulating weather regimes: Impact of model resolution and stochastic parameterization. *Climate Dynamics*, 44(7–8), 2177–2193. <https://doi.org/10.1007/s00382-014-2238-x>
- Park, M., & Lee, S. (2020). A mechanism for the midwinter minimum in North Pacific storm-track intensity from a global perspective. *Geophysical Research Letters*, 47, e2019GL086052. <https://doi.org/10.1029/2019GL086052>
- Pithan, F., Shepherd, T. G., Zappa, G., & Sandu, I. (2016). Climate model biases in jet streams, blocking and storm tracks resulting from missing orographic drag. *Geophysical Research Letters*, 43, 7231–7240. <https://doi.org/10.1002/2016GL069551>
- Priestley, M. D., Ackerley, D., Catto, J. L., Hodges, K. I., McDonald, R. E., & Lee, R. W. (2020). An overview of the extratropical storm tracks in CMIP6 historical simulations. *Journal of Climate*, 33(15), 6315–6343. <https://doi.org/10.1175/JCLI-D-19-0928.1>
- Sandu, I., van Niekerk, A., Shepherd, T. G., Vosper, S. B., Zadra, A., Bacmeister, J., et al. (2019). Impacts of orography on large-scale atmospheric circulation. *npj Climate and Atmospheric Science*, 2(1), 1, 10–8. <https://doi.org/10.1038/s41612-019-0065-9>
- Shaffrey, L. C., Stevens, I., Norton, W. A., Roberts, M. J., Vidale, P. L., Harle, J. D., et al. (2009). U.K. HiGEM: The New U.K. high-resolution global environment model—Model description and basic evaluation. *Journal of Climate*, 22(8), 1861–1896. <https://doi.org/10.1175/2008JCLI2508.1>
- Shaw, T. A. (2019). Mechanisms of future predicted changes in the zonal mean mid-latitude circulation. *Current Climate Change Reports*, 5(4), 345–357. <https://doi.org/10.1007/s40641-019-00145-8>
- Ulbrich, U., Leckebusch, G. C., & Pinto, J. G. (2009). Extra-tropical cyclones in the present and future climate: A review. *Theoretical and Applied Climatology*, 96, 117. <https://doi.org/10.1007/s00704-008-0083-8>
- Zappa, G., Masato, G., Shaffrey, L., Woollings, T., & Hodges, K. (2014). Linking northern hemisphere blocking and storm track biases in the CMIP5 climate models. *Geophysical Research Letters*, 41, 135–139. <https://doi.org/10.1002/2013GL058480>
- Zappa, G., Shaffrey, L. C., & Hodges, K. I. (2013). The ability of CMIP5 models to simulate North Atlantic extratropical cyclones. *Journal of Climate*, 26, 5379–5396. <https://doi.org/10.1175/JCLI-D-12-00501.1>
- Zappa, G., Shaffrey, L. C., Hodges, K. I., Sansom, P. G., & Stephenson, D. B. (2013). A multimodel assessment of future projections of North Atlantic and European extratropical cyclones in the CMIP5 climate models. *Journal of Climate*, 26, 5846–5862. <https://doi.org/10.1175/JCLI-D-12-00573.1>
- Zelinka, M. D., Myers, T. A., McCoy, D. T., Po-Chedley, S., Caldwell, P. M., Ceppi, P., et al. (2020). Causes of higher climate sensitivity in CMIP6 models. *Geophysical Research Letters*, 47, e2019GL085782. <https://doi.org/10.1029/2019GL085782>