Can low-resolution CMIP6 models provide insight into future European Post-Tropical Cyclone risk?

Supplementary material

Elliott M. Sainsbury[1], Reinhard K. H. Schiemann[2], Kevin I. Hodges[2], Alexander J. Baker[2], Len C. Shaffrey[2], Kieran T. Bhatia[3], Stella Bourdin[4]

[1]Department of Meteorology, University of Reading, Reading, Berkshire, UK

[2]National Centre for Atmospheric Science, University of Reading, Reading, Berkshire, UK

[3]Guy Carpenter, New York, United States

[4]Laboratoire des Sciences du Climat et de l'Environnement, LSCE/IPSL, CEA-CNRS-UVSQ-Université Paris-Saclay, Gif-sur-Yvette, France

Corresponding author: e.sainsbury@pgr.reading.ac.uk

1. CMIP6 model selection

The CMIP6 models used in the main manuscript were chosen from a large sample of CMIP6 models, based on their ability to simulate TCs in the North Atlantic. Figure S1 shows boxplots for North Atlantic TC frequency (hurricane season only), averaged over the entire historical run, for all the considered models. The green lines show the median, the box contains the interquartile range, and the maxima and minima are shown by the bars above and below the boxes. Between 1950 and 2014 (HURDAT2), there are approximately 12 TCs per year in the North Atlantic which have winds more than 17ms^{-1}, and 6.4 hurricanes per year (winds > 33ms^{-1}) These are labels 'Ob_ts' and 'Obs_hurr' in Figure S1 respectively.

Based on Figure S1, we select all CMIP6 models which have a median annual North Atlantic TC frequency greater than the average number of hurricanes per year (6.4) in observations. Boxplots are colored blue if SSP5-85 scenario data is also available, and white if not. Of the 8 models which have a median TC frequency greater than 6.4 TCs per year, 5 of these models also have the high-frequency relative vorticity fields needed for tracking for the SSP5-85 scenario. These models are therefore retained for further analysis.

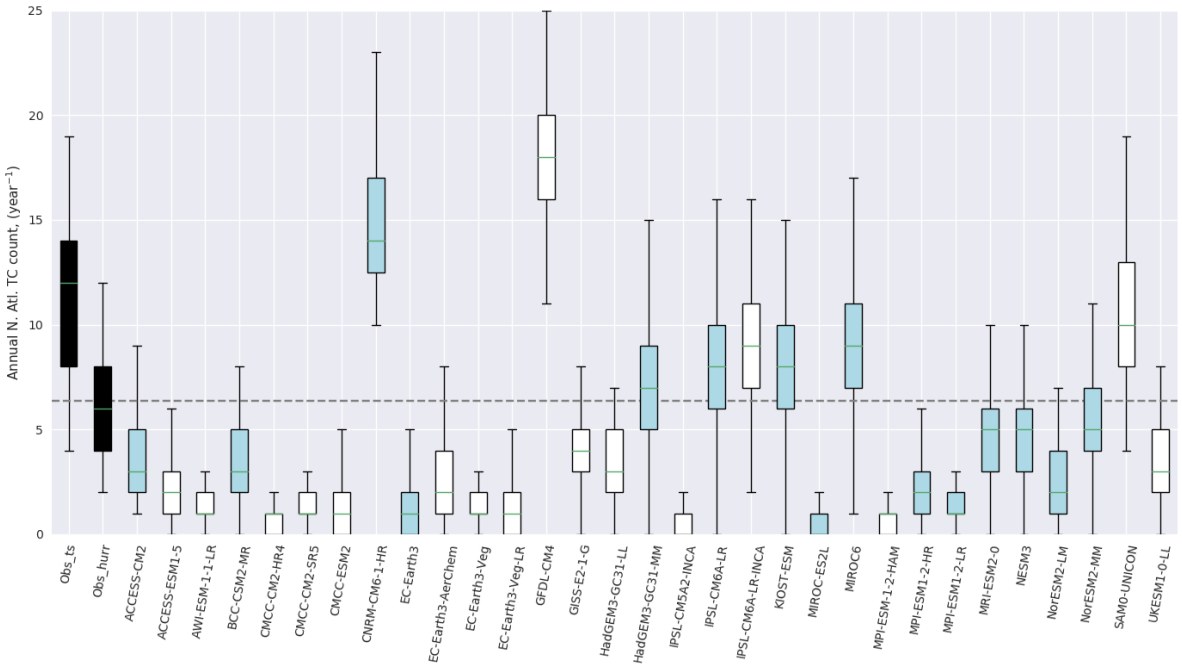


Figure S1: Boxplot showing the North Atlantic TC frequency (June 1st – Nov 30th), averaged over the entire historical run, for a selection of CMIP6 models and HURDAT2 (first two boxes, black). 'Obs_ts' represents all cyclones present in HURDAT2 with winds >= 17ms^{-1}. 'Obs_hurr' represents all hurricanes (winds >= 33ms^{-1}) present in HURDAT2. Grey horizontal line represents the threshold of 6.4 TCs per year used to select models for further analysis, which is the average number of hurricanes per year in HURDAT2 between 1950 and 2014. Blue boxes represent models which have sufficient data available for the SSP5-85 scenario to perform cyclone tracking and TC identification, and boxes which are not filled blue do not have sufficient SSP5-85 data available, and so cannot be considered in this study.

2. TC identification method

Objective TC identification methods such as used in Hodges et al. (2017) identify cyclone tracks as TCs if (among other things) they have a coherent vertical structure and a warm core, diagnosed using the vertical profile of relative vorticity. To achieve a balance between hit rate and false alarm rate, a latitude constraint is imposed such that the TC must have a genesis location equatorward of 30N. In the present climate, this is quite suitable as most TCs forming in the North Atlantic basin form equatorward of 30N (though there are notable exceptions, such as Ophelia in 2017). However, climate change may extend the genesis region of TCs further polewards and eastwards (e.g., Haarsma et al. 2013), and so a 30N threshold may lead to many future TCs being misidentified as non-tropical by virtue of where they form.

Figure S2 shows the genesis density for the 5 CMIP6 models in the historical (1984-2014) and future (2069-99) periods, along with the differences (future-historical). Here we identify the TCs using criteria in section 2.3, however criteria 1) has been relaxed to 45N, to investigate whether there is any model consensus for increased genesis of TC-like features between 30 and 45N. Figure S2 shows no robust increase in TC genesis between 30 and 45N, though there is some increase in the MIROC along the US East Coast. We therefore use a 30N constraint throughout the main manuscript with confidence that we are not missing a large trend in subtropical TC genesis poleward of 30N.

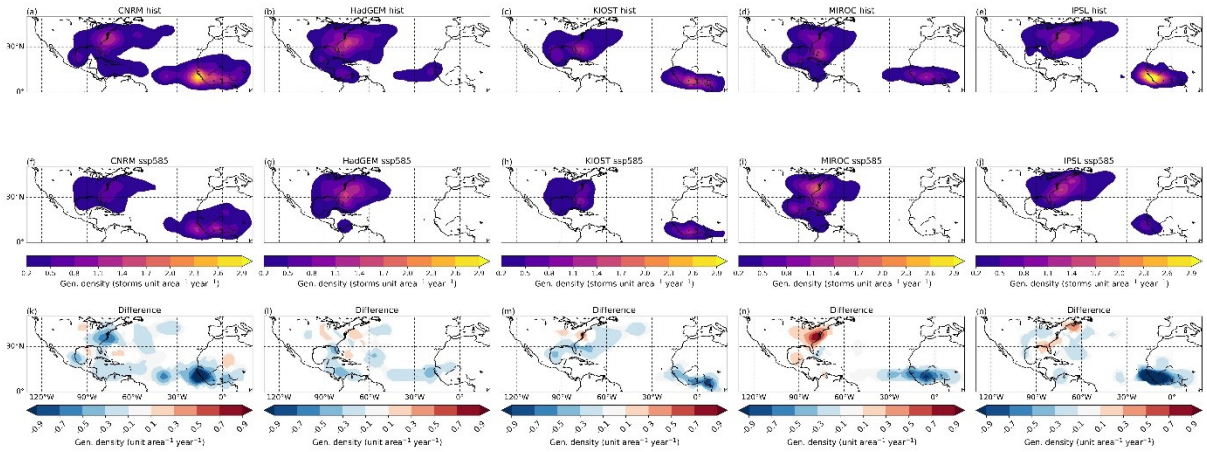


Figure S2: Genesis density for the historical (1984-2014, a-e) and future (2069-99, f-j) periods for the five selected CMIP6 models along with the differences (future-historical, k-o). TCs are identified using the criteria described in section 2.3 of the main manuscript, but the latitude constraint has been relaxed to 45N.

3. Historical Lysis Density

Figures 2 and 4 in the main manuscript indicate that the TCs forming in the main development region in the IPSL model are weak, short-lived TCs. To investigate this further, a spatial map of the lysis density is created for the 5 models and ERA5, shown in Figure S3. There is a large region of non-zero lysis densities in the tropics, subtropics and midlatitudes in all the models and ERA5, highlighting that vast region in which TCs travel and dissipate. Many models (and ERA5) have their highest (or joint highest) lysis density values in the subtropics and midlatitudes. However, the IPSL is very different from the other models. In the IPSL, the highest lysis density values occur in the MDR, very close to where the storms form. These TCs form on the west coast

of Africa (Fig. 1), indicating that the storms are dissipating with little opportunity to move from their place of origin. This further indicates the severe bias in the MDR of the IPSL.

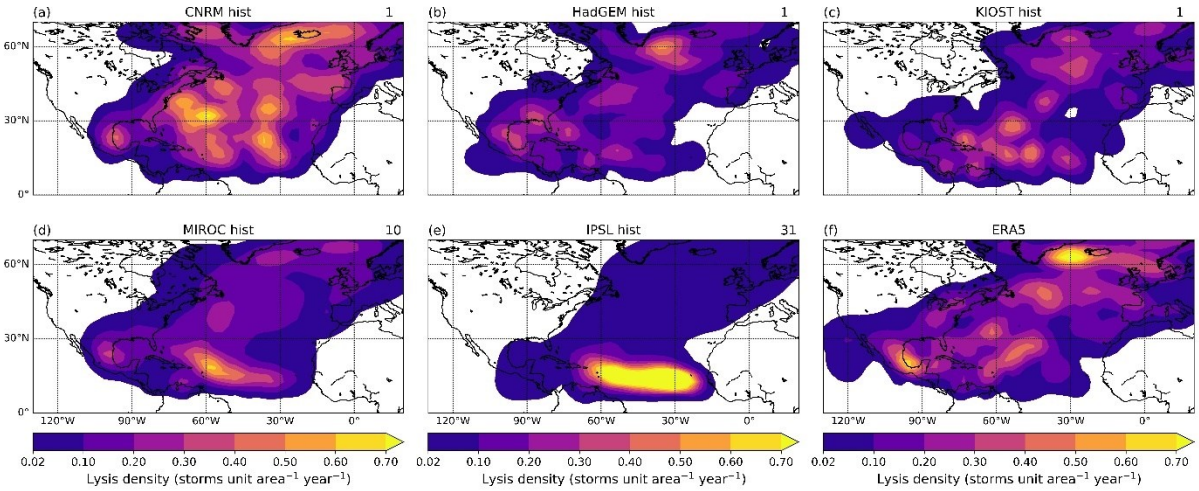


Figure S3: Lysis density (storms per unit area per year, where the unit area is equal to a spherical cap with a 5-degree radius) for the 1984-2015 period from the historical runs of 5 CMIP6 models (a-e) and ERA5 (f). All ensemble members are used where available. Densities less than 0.02 have been masked for clarity.

4. Historical GPI bias

Figure S4 shows the bias in GPI (and individual terms) for the fully historical coupled runs from 1984-2014. Biases are calculated relative to ERA5. All models except CNRM have a negative GPI bias or no GPI bias (top row) in the main development region. All models except the CNRM have too much wind shear in the western MDR, increasing hostility (row 5).

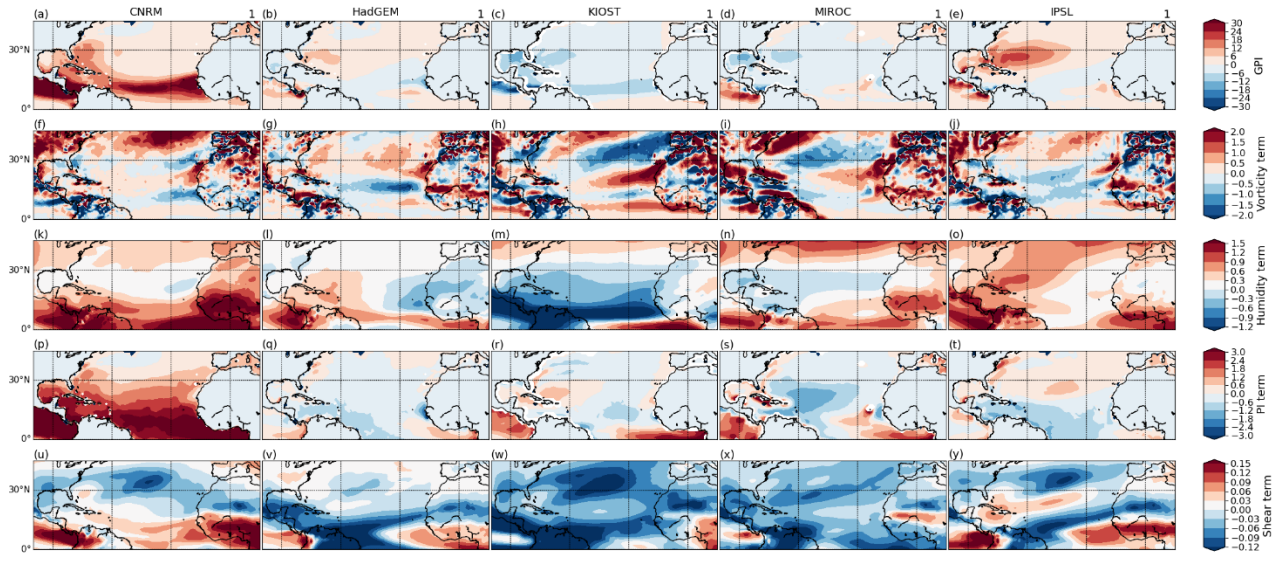


Figure S4: Bias (model minus ERA5) in GPI (top row) and the terms comprising GPI (rows 2-5). Data from 1984-2014 is used for ERA5 and the fully-coupled historical runs. Note that a positive bias in shear means a low-bias in wind shear.

5. Wind-pressure relationship

The wind-pressure relationships for the warm-core part of the tracks (i.e., the part of the track at which the criteria in section 2.3 are satisfied) over the historical period is compared with ERA5 and HURDAT2 in Figure S5. The relationship between wind and pressure in TCs is well-established, and using one point for each TC, or all points for each TC, should not change the results. ERA5 and all of the CMIP6 models have a slope which is too steep, where cyclones do not reach the expected (from HURDAT2 observations) wind speeds for a given sea level pressure. The apparent intensity of the TCs is greater in terms of their sea level pressure than in terms of their maximum winds. This result has also been found in other reanalyses in previous studies (Hodges et al., 2017). This difference is smallest for CNRM, which simulates stronger TCs than the other 4 models. Despite this, in all 5 models (and ERA5), there is a clear relationship between wind and pressure in the tropical stage of the identified TCs, a feature present in observed TCs.

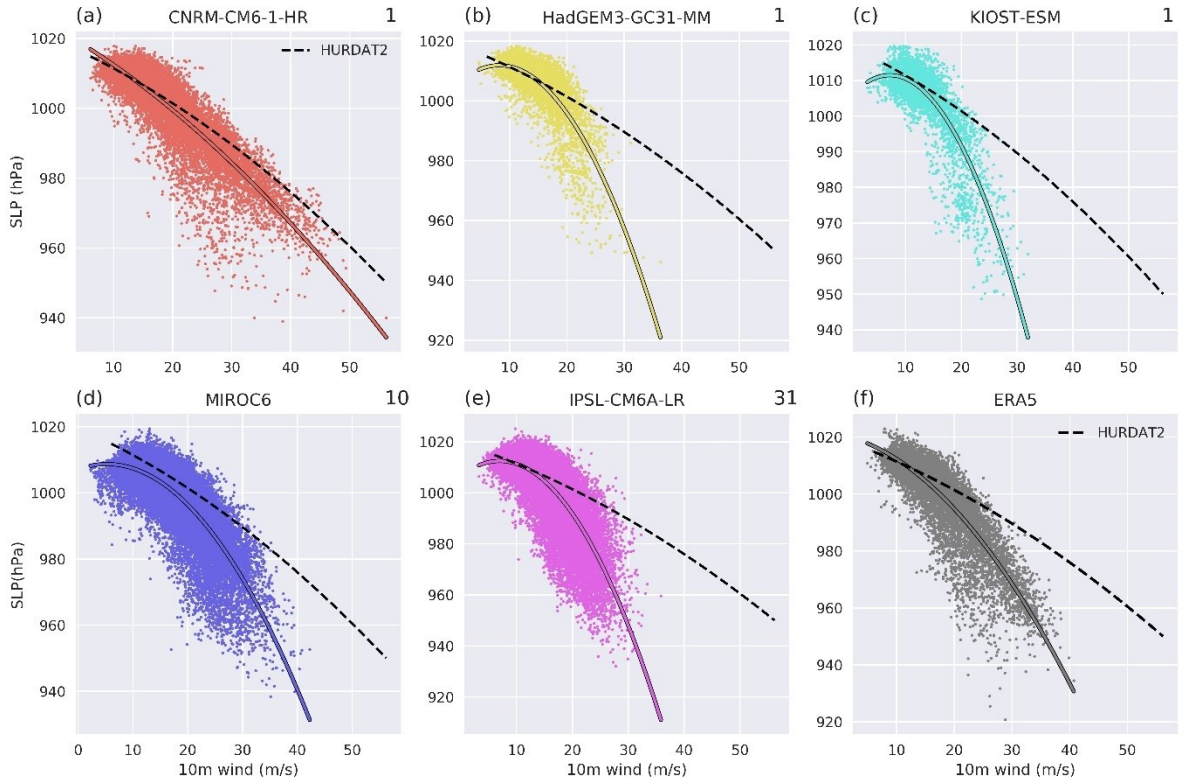


Figure S5: Wind-pressure relationship for the TCs (during their TC phase) for selected CMIP6 models (a-e) and ERA5 (grey, f). Coloured lines on each panel show a quadratic polynomial fit for each CMIP6 model and ERA5, dashed black line shows the quadratic polynomial fit for HURDAT2. The number of ensemble members used for each model is shown to the upper-right of each panel.

6. Relationship between TC frequency and recurving TC frequency

To reach Europe as a PTC, a TC must first interact with the midlatitudes (i.e., enter the 'recurvature' domain shown in Figure 1 of the main manuscript). Previous work has shown that there is a strong relationship between TC activity and recurving TC activity (Sainsbury et al., 2022). It is therefore important that CMIP6 models can capture this relationship, as it may have implications for the frequency of Europe-impacting PTCs. The ability of the selected CMIP6 models to capture this relationship is explored in Figure S6. In ERA5 (TCs identified using track matching with HURDAT2), there is a significant relationship between TC frequency and recurving TC frequency (top left panel, r=0.85). A strong relationship is also found in the CMIP6 models (correlations 0.61-0.76), which are statistically significant in all cases to the 95% level. CNRM, HadGEM and KIOST all have a slope like ERA5, whereas the IPSL and MIROC have a slope that is too shallow, indicating that too few TCs recurve in these two models.

Overall, all five CMIP6 models capture the relationship between TC frequency and recurving TC frequency despite TC intensity biases, indicating that the models can capture the main driver of the interannual variability of recurving TC frequency, which may have important implications for European PTC risk.

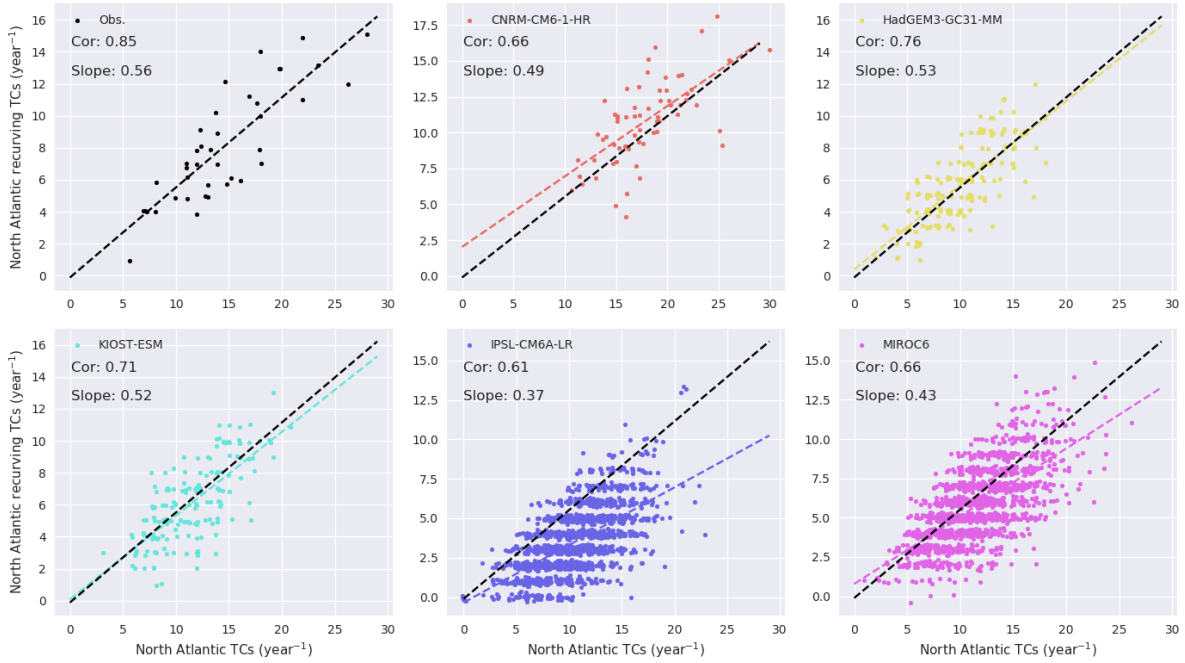


Figure S6: Relationship between TC frequency and recurving TC frequency in ERA5 (top left) and the five selected CMIP6 models. For ERA5, data from 1979-2018 is used. For the other models, all data from the fully-coupled historical runs are used (1950-2014 for CNRM, IPSL, 1850-2014 for HadGEM, KIOST, and MIROC6). Random noise has been added to the data points so that all points can be seen, but the correlations and slopes are identified using the original data.

7. European-impacting PTC statistics - only considering ensemble members present in both the historical and future periods

Finite computational resources force a compromise to be made between model resolution and length of simulation/number of ensemble members. In this work, we are using relatively coarse resolution models. This limitation is compensated by the large number of ensemble members available for some of the models, particularly during the historical period for the MIROC6 (10) and IPSL-CM6A-lR (31). However, data is not available for this many ensemble members in the future runs under the SSP585 scenario.

To ensure that the projected changes shown in the paper are not the result of different ensemble members being considered in the historical and future periods, we repeat the analysis, this time just considering the common ensemble members over the historical and future period. This is shown in Figure S7. The CNRM-CM6-1-HR, HadGEM3-GC31-MM and KIOST-ESM only have one ensemble member for the historical and future period and so see no change. However, for the MIROC6 we only consider the first three ensemble members for the historical period (ensuring that a direct comparison can be drawn with the three ensemble members available for the future period). For the IPSL-CM6A-LR, we only use the first ensemble member for the historical period, as this is the only ensemble member which is available for the future period under the SSP585 scenario.

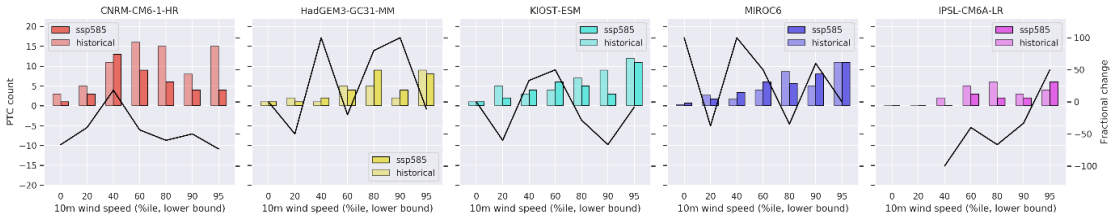


Figure S7: As in Figure 9, but only considering common ensemble members over the historical and future periods.

There are no changes to the data (and hence results) for the CNRM, HadGEM and KIOST. When just using common ensemble members, the projected change in the number of European-impacting PTCs in the highest intensity bin is smaller (fractional change down from +20% to +5%) in the MIROC and has increased for the IPSL (fractional change up to +50% from +40%). These changes are relatively small and do not alter the key result, which is that there is no robust model response in terms of European-impacting PTC intensity changes.

Table S1 shows the repeated analysis for Table 2, again only considering the common ensemble members. The results are largely unchanged. All fractional changes which were shown to be significant in Table 2 (except the increase in European-impacting PTC counts in the MIROC6) are still significant when only considering common ensemble members, and therefore the robust responses seen in the models (reduction in North Atlantic TC count, increase in likelihood of recurvature) are not the result of differing ensemble sizes in the historical and future periods.

	N_{Europe}			N_{TC}			F_{rec}			$F_{Europe\|rec}$		
	Hist	SSP	Diff	Hist	SSP	Diff	Hist	SSP	Diff	Hist	SSP	Diff
CNRM	2.35	1.29	**-45%**	14.68	9.00	**-39%**	0.55	0.59	7%	0.29	0.24	-17%
HadGEM	0.81	0.94	16%	7.29	4.35	**-40%**	0.46	0.61	**31%**	0.24	0.35	**49%**

KIOST	1.32	1.03	-22%	7.97	5.19	**-35%**	0.41	0.40	-4%	0.40	0.50	24%
MIROC	1.08	1.17	8%	8.37	5.53	**-34%**	0.43	0.68	**58%**	0.30	0.31	3%
IPSL	0.65	0.42	-35%	8.84	3.16	**-64%**	0.20	0.42	**110 %**	0.36	0.32	-11%

Table S1: As in Table 2, but only considering common ensemble members between the historical and future periods.

While there are some differences between results produced using all ensemble members and common ensemble members, they do not differ significantly, and they do not change the results of the paper.

8. Normalized genesis density changes

Figure S8 shows the change in the normalized genesis density in the five CMIP6 models. There is a robust model response for a shift in genesis (proportionally) away from the MDR towards the WEST and SUB regions. This difference is largest in the HadGEM, MIROC and IPSL - the three models in which there is the largest increase in the likelihood of recurvature - and is consistent with the results of Table 3.

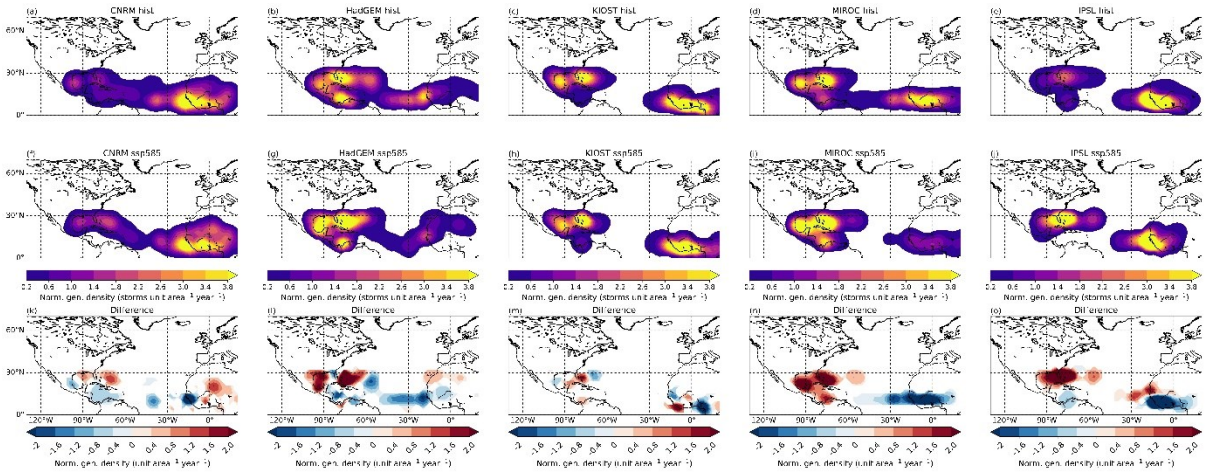


Figure S8: Normalized genesis density for the 5 selected CMIP6 models during the historical (top) period, towards the end of the century under SSP585 (middle), and the difference (future minus historical, bottom). Densities less than 0.2 have been masked for clarity.

9. TC LMI changes

While shifts in genesis explain most of the projected change in F_{Rec} in HadGEM, MIROC and IPSL, they do not explain all of the projected increase. To further investigate the projected changes in F_{Rec}, the relationship between projected TC LMI changes and projected F_{Rec} changes is explored. Figure S9 shows the change in TC

LMI for the 5 selected models, for each of the three regions in equation (3) and Table 3 (main manuscript), along with a scatterplot of TC LMI changes against recurvature likelihood changes within the regions.

In the MDR, there is a large increase in the proportion of North Atlantic TCs which are strong (TC LMI > mean of the combined (historical + future) North Atlantic TC distribution) in MIROC. In the WEST region, this is also true for HadGEM, MIROC and IPSL. These are also the only four models/regions in our selection of CMIP6 models where there is a statistically significant increase in the likelihood of recurvature (Table 3, main manuscript). Figure S9d shows a strong relationship between TC LMI changes and changes in recurvature likelihood within regions. It is therefore possible that changes in TC LMI – particularly in the WEST regions of MIROC, HadGEM and IPSL, and MDR of MIROC – are associated with the projected increase in F_{Rec}. For Fig. S9d, the Pearson's correlation coefficient is 0.52 and is significant at the 95% level. However, it should be noted that this is only a correlation over 15 data points. It is not completely clear from the GPI (Fig. 6) why the TC LMI is seen to increase in the WEST regions in HadGEM and IPSL, and the MDR and WEST regions in MIROC.

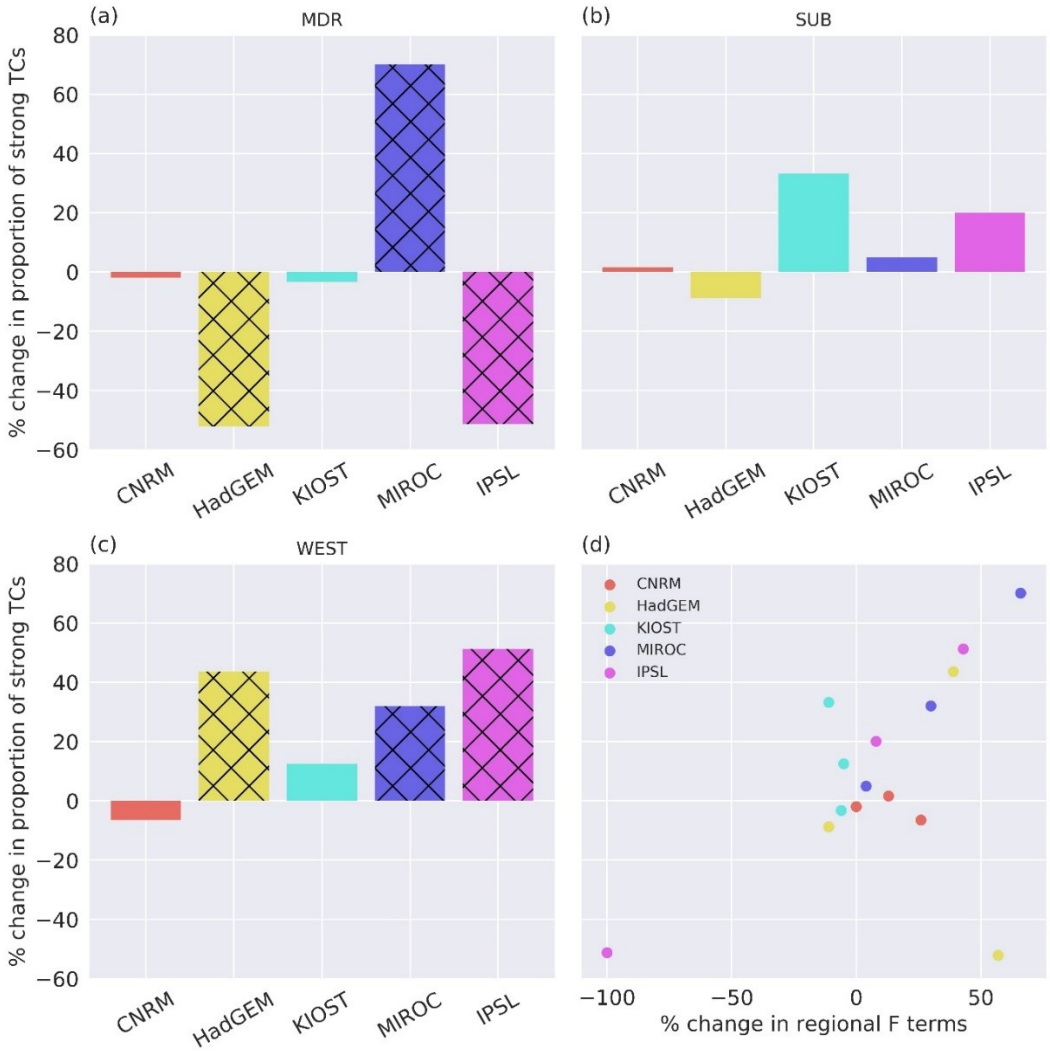


Figure S9: Bar charts showing the percentage change in the proportion of strong North Atlantic TCs (TCs which have a TC LMI greater than the mean of the combined (historical + future) distribution) in the MDR (a), SUB (b) and WEST (c) regions. (d) shows a scatterplot of the percentage change in the proportion of strong North Atlantic TCs against the percentage change in the F terms within the individual regions (columns 3, 5, and 7 of the bottom of Table 3, main manuscript). Hatching represents the models and regions in which the % change is significant to 95% using a bootstrapping method.

10. Direct comparison with Haarsma et al. (2013)

The only previous studies which investigate projected changes in European-impacting PTCs find a large increase in the number of high-intensity PTCs in the future (Haarsma et al. 2013, hereafter H13; Baatsen et al. 2015). None of the CMIP6 models investigated in this study show an increase as large as found in these previous studies (driven by a very large decrease in TC counts basin-wide). However, the results presented in section 3.2. do not use the same methodology as used in H13.

To make our work as comparable as possible to H13, the analysis of Figures 9 and 10 are repeated with some methodological changes. As in H13, we only consider cyclones which impact Europe (Fig. S10) or form (Figs S11, S12) during August, September, and October. We also change the domain of interest to look at the same four regions which were used in H13: Norway (60-70N, 0-15E), North Sea (50-60N, 3W-8E), Western UK and Ireland (50-60N, 3-15W) and the Gulf of Biscay (43-50N, 0-15W).

Due to the lower resolution of the CMIP6 models than used in H13, we cannot use the same intensity threshold of hurricane force (>33ms^{-1}) winds. We instead just consider cyclones impacting each of the four regions in ASO which have maximum winds in the region greater than the 90th percentile of the combined (PTC+MLC, historical + future) distribution of cyclone maximum winds in the region. This gives us a similar PTC frequency as in H13 (Fig 2f therein).

Figure S10 shows mixed changes in the number of strong PTCs impacting each region in the future, though more regions and models project decreases than increases. Sample sizes are extremely small, as the Europe sub-regions are considerably smaller than the European domain used in the main manuscript. The result of Figure S10 is different from the result of H13, in which there was a substantial increase in strong PTCs in the future across three of these four regions. However, when considering the fraction of North Atlantic TCs forming in ASO which impact each of these regions (Figure S11 for all PTCs, Figure S12 for strong PTCs (> 90th percentile)), there's a general increase, in line with Figure 10 from the main manuscript. This suggests that in the future, North Atlantic TCs are more likely to reach these four regions of Europe, and future risk to these regions depends strongly on the projected change in basin-wide TC counts.

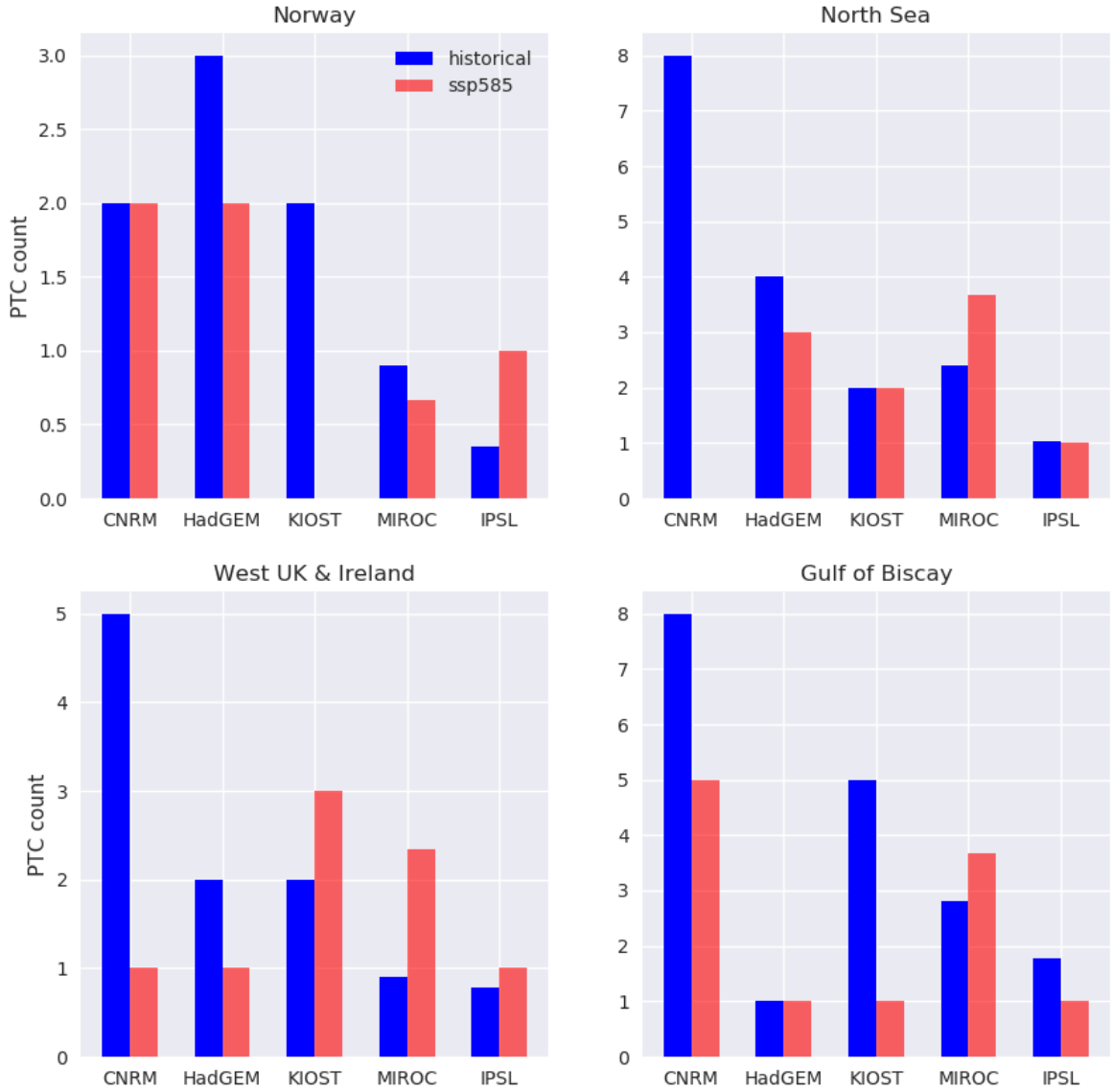


Figure S10: Number of high intensity PTCs (max wind in region > 90th percentile of combined distribution) impacting each of the four regions during Aug-Oct in the historical period (blue) and future period under the SSP585 scenario (red).

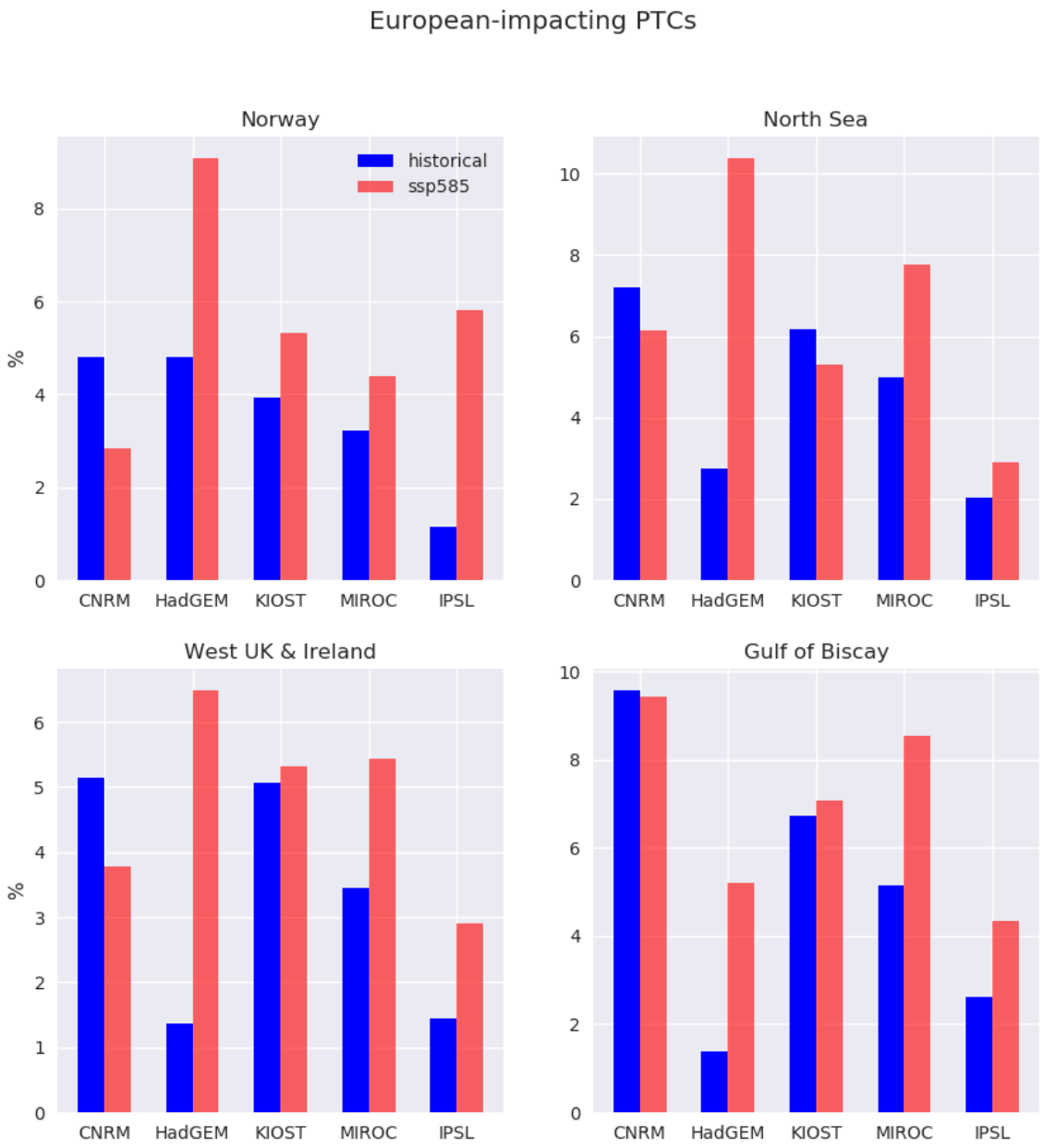
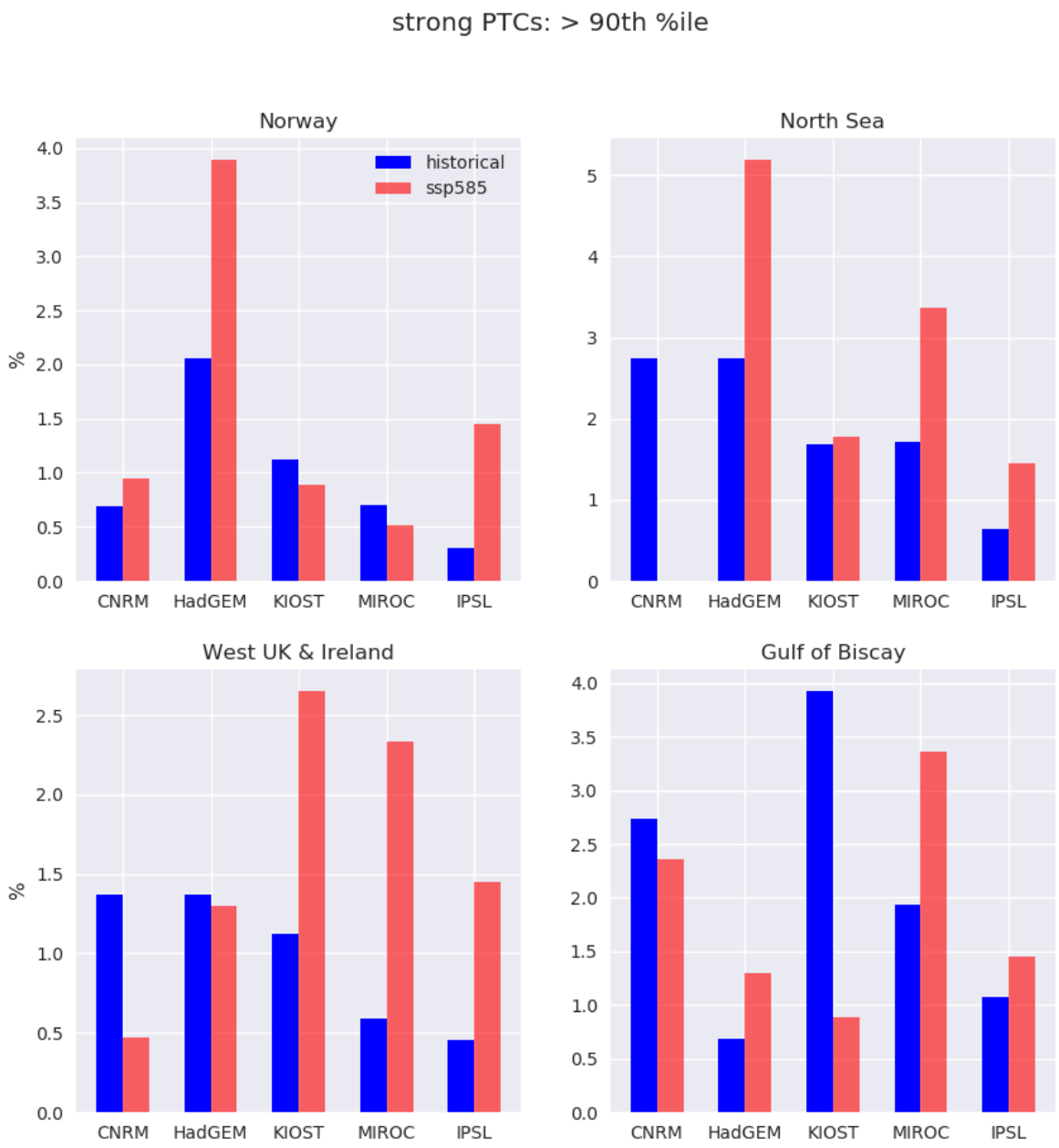


Figure S11: Fraction of North Atlantic TCs which impact the given four regions as PTCs, considering all North Atlantic TCs forming during August-October in the historical period (blue) and future period under the SSP585 scenario (red).



255

256 **Figure S12: As in Figure S11 but considering the fraction of North Atlantic TCs which reach the four**
257 **regions as strong PTCs (winds > 90th percentile of the max wind distribution in the region).**

258