Serial clustering of extratropical cyclones in historical and future CMIP5 model simulations

T. Economou[a]*,D.B. Stephenson[a],J. Pinto[b],Len.C. Shaffrey[b],G. Zappa[b]

[a]*College of Engineering, Mathematics and Physical Sciences, University of Exeter*

[b]*Department of Meteorology, University of Reading, Reading, UK*

*Correspondence to: t.economou@exeter.ac.uk

Temporal clustering of extratropical cyclones over Europe poses a major societal risk. Single runs from 30 historical and future winters, from 17 climate models are used to quantify clustering. Results show that models broadly capture patterns seen in reanalyses and that the North Atlantic Oscillation (NAO) explains clustering over the North Atlantic and parts of Europe, but has negligible effect on future changes. Sampling uncertainty in measures of clustering over 30 years is larger than the change due to climate change. Near London, the standard deviation of the clustering measure from the Hadley Centre Global Environment Model version 2 (HadGEM2), is 0.37 whereas the change is an increase of 0.1.

Key Words: Storm clustering; CMIP5; Poisson process; extratropical cyclones.

Received . . .

1. Introduction

Extratropical cyclones (ETCs) pose a major societal risk in Europe, especially when they occur successively in clusters, which then leads to large aggregate losses. For example, in the recent winter of 2013/14, numerous ETCs caused destruction of infrastructure and disruption to transport and business in Europe: windstorms Christian, Xaver, Dirk and Tini caused insured losses of $1.382, 0.961, 0.468 and 0.356 billion resulting in a total insured loss of more than $3 billion (source: http://www.perils.rog).

There are three main reasons why storms might cluster in time:

1. By chance – even if storm occurrences happen completely at random, some of them will occur in clusters;
2. Due to modulation by large-scale climate modes. Time varying background climatic conditions can cause the rate of occurrence to be non-stationary (Mailier *et al.* 2006);
3. Dependence between successive storms – e.g. a "parent" storm might generate one or more "offspring" (secondary cyclogenesis; Parker (1998)).

Previous studies have used Poisson process concepts to investigate temporal clustering of storms using observations and reanalyses data. For historical windstorms affecting the North Atlantic and Europe, Mailier *et al.* (2006) has found that there is more clustering than can be expected by chance at both flanks and downstream areas of the main North Atlantic storm track. Moreover, evidence was provided that more extreme storms cluster to a greater extent (Vitolo *et al.* 2009). Using a different tracking methodology, Pinto *et al.* (2013) confirmed that the identified spatial pattern of clustering and its intensification for more severe storms is a robust feature in reanalysis data.

Mailier *et al.* (2006) also showed that clustering can be largely explained by modulation of storm counts by large-scale climate modes. A large part of the contribution from the large-scale modes to the storm clustering derives from the modulation of the eddy-driven jet over the North Atlantic. Indeed, the occurrence of historical storm series affecting Western Europe is related with a recurrent extension of an intensified eddy driven jet towards Western Europe for periods of one or more weeks (Pinto *et al.* 2014). Moreover, upstream cyclone development (secondary cyclogenesis (Parker 1998) is strongly related to cyclone clustering, leading to the development of multiple cyclones on a single jet streak (Pinto *et al.* 2014).

Kvamstø *et al.* (2008) showed that a simulation from a general circulation model (GCM) underestimated clustering compared to reanalysis and failed to capture the relationship between clustering and modes of climate variability. Large differences between three GCM simulations and reanalysis were also noted in Mailier (2007, ch. 6) who also found that differences in models responses prohibited conclusive statements about the future. Pinto *et al.* (2013) investigated future changes in clustering using an ensemble of 20 runs from a single GCM and found evidence, based on the ensemble mean, pointing to a possible decrease in cyclone clustering over parts of the Northern Atlantic and Western Europe, particularly North of 50°N. Note however that Pinto *et al.* (2013) also found evidence of considerable sampling variability, between ensemble members, in estimates of clustering statistics (their Fig. 6). This implies that such estimates based on relatively short time periods (e.g. 30 years) might be quite uncertain which in turn has an effect on the uncertainty regarding future changes of such statistics. This is investigated in detail later, in section 3.3.

This study provides a first investigation of clustering of North Atlantic extratropical cyclones and its future changes using a large ensemble of climate models. Simulations from 17 different models computed for the recent 5th phase of the Coupled Model Inter-comparison Project (CMIP5) multi-model experiment are considered (Taylor *et al.* 2012). The following questions are addressed:

1. How well do the CMIP5 climate models capture the clustering seen in historical reanalyses?
2. How does clustering change in future climate model projections and how model-dependent is the response?
3. Can the changes be understood in terms of changes in physical drivers such as the North Atlantic Oscillation (NAO) or are they simply due to natural variability in storm counts?

Section 2 provides a brief description of the tracking algorithm used to produce cyclone tracks from CMIP5 and reanalysis data, and how the NAO was defined and calculated. Section 3 shows results from quantifying clustering in historical and future model simulations but also reanalysis. In addition, a simple statistical model is implemented to investigate variation and future changes in clustering in terms of the NAO. Section 4 gives the conclusions and a discussion of the results.

2. Data

2.1. *Storm tracks from reanalyses*

Objective feature-identification software (Hodges 1994) was used to extract ETC tracks from 6-hourly ECMWF reanalysis (ERA40) data, for the period December 1958 until November 2001, i.e. 43 historical winters. Individual cyclone tracks are identified by tracking local maxima in relative vorticity at 850hPa just above the boundary layer (about 1.5km above sea-level).

Cyclone tracks are defined as 6-hourly measurements of cyclone intensity and location in winters December–February (DJF). Minimum sea-level pressure (MSLP) is taken here as the intensity measure used to define extreme storms and note that for extremes, MSLP is strongly related to other variables such

as wind speed (see Fig 2 in Economou *et al.* (2014)). A region defined by [90°W, 25°E] and [30°N, 70°N] is discretised using a $2.5^\circ \times 2.5^\circ$ grid, and at each grid point a 6.3° (700km) radius circle is considered following the approach of Pinto *et al.* (2013). This choice of radius is within the range of effective radius for extra-tropical cyclones (600 - 1000 km; Rudeva and Gulev (2007)) and corresponds to a plateau of quasi-constant values of a particular dispersion statistic (see section 2.4) over most of the study area (Pinto *et al.* (2013), their Fig. S1). To avoid problems with multiple counting, only the highest intensity event is counted for each track passing through the circle.

2.2. *Model simulated storm tracks*

CMIP5 is a comprehensive collection of experiments in which many climate modelling groups around the globe produced a suite of past and future climate simulations (Taylor *et al.* 2012). The tracking algorithm described above was then used to extract storm tracks for 17 CMIP5 models (Zappa *et al.* 2013a) based on simulations of 30 historical winters (December 1975 to March 2005) and 30 future winters (December 2069 to March 2099). The future simulations used the RCP4.5 forcing scenario which gives around 2°C global mean warming by 2100 (IPCC 2013). See Zappa *et al.* (2013a), Zappa *et al.* (2013b) and Sansom *et al.* (2013) for details on how well CMIP5 models capture North Atlantic ETCs and simulate future projections. For some models there are multiple simulations for each period, based on different model initialisations, however here only a single simulation is considered per model – thus avoiding the problem of having to compare results from models with different number of runs. This is also more comparable to the real world where there is only one realisation in the past and future periods. For illustration, Figure 1 shows time series of storm counts from ERA40 and HadGEM2 (Hadley Centre Global Environment Model version 2) for a grid point near London (1.25°W, 51.25°N). The mean storm count for ERA40 is slightly lower than the one from the historical period in HadGEM2 but overall the model counts show similar behaviour to the reanalysis counts. Both the mean count and the number of storms with very low MSLP increase in the future over this location. This is consistent with Zappa *et al.* (2013b) and Sansom *et al.* (2013) who found a slight tendency of an increase in the number of extratropical cyclones in the multi model mean of CMIP5 future projections.

2.3. *NAO indices*

The North Atlantic Oscillation (NAO) index has been shown to be a major driver in ETC clustering (Mailier *et al.* 2006; Seierstad *et al.* 2007). The 5th IPCC assessment of ETCs (see 14.5.1; Christensen *et al.* (2014)) reports that the mean NAO is likely to increase slightly in the future whereas the variance shows little change (see IPCC AR5, Ch. 14, Figure 14.16a). It is therefore of interest to investigate whether the projected changes in NAO can explain changes in clustering and this is investigated here using a simple statistical model (see section 3).

Observed values of the NAO index matching the ERA40 reanalysis tracks were taken from the National Oceanic and Atmospheric Administration (NOAA) webpage http://www.noaa.gov. The index is defined as the difference between sea level pressure (SLP) in Iceland and the Azores, and is standardised to have zero mean and unit variance. For CMIP5 runs, an index was calculated by subtracting the mean SLP based on an 1861–1900 climatology, over a region north of 55°N, and between 90°W and 60°E from the mean SLP over a region between 20°N and 55°N and between 90°W and 60°E (Gillett and Fyfe 2013). The NAO index was then calculated by standardising using the sample variance in the historical period 1975–2005.

2.4. *Dispersion of counts as a measure of clustering*

If storms were to occur completely at random, then the simplest model to describe this is the homogeneous Poisson process with constant intensity λ (Cox and Isham 1980), which implies that the number of events in any time interval of length T is Poisson distributed with mean $\mu = \lambda T$ and variance $\sigma^2 = \lambda T$. Deviations from this process can give rise to events that either appear more clustered or more regular. The number of events per unit time is then either overdispersed ($\sigma^2 > \mu$) or underdispersed ($\sigma^2 < \mu$) (Mailier *et al.* 2006).

Let Y be the number of storms passing through a region in a winter. A dispersion measure $\phi = Var(Y)/\mathbb{E}[Y] - 1$ can be estimated by $s_y^2/\bar{y} - 1$, where s_y^2 is the sample variance of Y.

Positive (negative) values of ϕ suggest clustering (regularity) whereas $\phi = 0$ suggests complete serial randomness.

Mailier *et al.* (2006) showed that a large fraction of overdispersion over Europe (in reanalysis data) can be captured by characterising the mean of Y as a function of climate indices such as the NAO. More recently, Blender *et al.* (2014) have used a Weibull renewal process to also quantify clustering in ERA interim reanalysis, and have confirmed high clustering (overdispersion) over Europe and Scandinavia.

Extreme storms at a grid point are defined as ones whose MSLP value went below a threshold, specifically the 0.2 empirical quantile based on all MSLP minima associated with storms at that grid point. The thresholds are different for historical and future projections to allow for the fact that the distribution of storm intensity or the background MSLP might change (Chang 2014).

3. Results

In what follows, letters "H" and "F" are used to refer to statistics from either the 30 historical or the 30 future winters from each climate simulation.

3.1. *Historical clustering*

Figure 2 shows dispersion in historical runs of each CMIP5 model and the ERA40 reanalysis, as well as the multi-model mean. The models qualitatively capture the gross features seen in ERA40 reanalysis: regularity along the storm track axis (the upper tropospheric jet axis are indicated on all plots for reference) and clustering around the edges of the track. These features are also consistent with previous clustering studies (Mailier *et al.* 2006; Vitolo *et al.* 2009; Pinto *et al.* 2013). Figure 3 shows estimates of ϕ_H but for extreme storms. Again, models largely capture the overall features seen in ERA40 reanalysis: ϕ increases everywhere, and the overdispersion becomes particularly large over Northern Europe, Scandinavia and the Azores. This increase in dispersion for more extreme storms was previously noted for reanalysis in Vitolo *et al.* (2009) and Pinto *et al.* (2013). Notice however that the models tend to underestimate the dispersion statistic over the UK/Central Europe, something that is consistent with the a zonal North Atlantic jet simulated by the CMIP5 models, which is associated with a too strong and southward displaced jet over Central Europe compared to reanalysis data (fig 9a in Zappa *et al.* (2013a)).

3.2. *Future projections*

Figure 4 shows estimates of changes in dispersion $\Delta\phi = \phi_F - \phi_H$. There is considerable variation with little agreement between models. However, the multi-model mean (upper left panel but also more clearly in Figure 5a) shows a similar response to the ensemble mean of ECHAM5 GCM simulations in Figure 5e of Pinto *et al.* (2013): i.e. increase in underdispersion over the region of cyclone genesis and an increase of clustering south of the storm track.

Figure 5b shows multi-model mean $\Delta\phi$ estimates for extreme storms (again there is substantial variability between and the multi-panel is given in the supplementary material in Figure S1). The multi-model mean plot shows increase in dispersion near Greenland and over Northern Europe which is different from the ensemble mean of ECHAM5 in Figure 5f of Pinto *et al.* (2013), who detect a decrease in dispersion over the same areas. This may be attributed to the following reasons: a) different GCM, in this case ECHAM5 GCM, b) different cyclone tracking mothodology and c) the definition of extreme storms: Pinto *et al.* (2013) use background pressure to calculate local thresholds whereas we use a threshold based on the lowest MSLP measurement of each storm that went through a grid point. Therefore the extremes here are relative to the maximum intensity of all other storms within the neighbourhood, as calculated by the tracking algorithm – so results regarding the extremes may not be comparable to ones created by other tracking algorithms. Multi-model means are not very informative with respect to an actual prediction of the future, especially if the variability about these means is quite large, as is apparent from Figure 4.

Note that all of the above analyses were repeated using the more severe RCP8.5 scenario but the results are not shown for conciseness. The main findings regarding future changes were that: 1) as with RCP4.5 there was considerable variability between the models, 2) very little qualitative agreement when comparing future changes from RCP4.5 and RCP8.5 maps of individual models and 3) the multi-model mean plot of $\Delta\phi$ for the two scenarios was qualitatively similar but with RCP8.5 changes being

more pronounced. Point 2 indicates that much of the individual model responses is dominated by internal climate variability. This will be investigated in the next section.

3.3. *Understanding the changes*

3.3.1. *Are changes in dispersion due to changes in the mean counts?*

Changes in $\phi = s^2/\bar{y}$ are either due to changes in the sample mean $\bar{y}$ or the sample variance s^2 of the counts, or both. To investigate this, we recalculated $\Delta\phi$ in two different ways: 1) Assuming $\bar{y}$ remains the same in the future so that $\Delta\phi = \left(s_F^2 - s_H^2\right)/\bar{y}_H$ shown in Figure S2 and 2) assuming the variance remains the same so that $\Delta\phi = s_H^2\left[1/\bar{y}_F - 1/\bar{y}_H\right]$ shown in Figure S3. Comparing Figures S2 and S3 with Figure 4, it is evident that changes in dispersion are mostly driven by changes in the variance of the counts. The same holds for extreme storms (not shown for brevity).

3.3.2. *Why did the variance of the counts change?*

For count data, it can be shown by Taylor expansion that $4Var(\sqrt{Y}) \approx Var(Y)/\mathbb{E}[Y]$ (Anscombe 1948; Yu 2009). The square root is a variance stabilising transformation (i.e. such that variance is unaffected by changes in mean) and if the data are over- or underdispersed, the variance of $\sqrt{Y}$ will be larger/smaller than $1/4$ implying that the dispersion measure can be approximated by $\phi \approx 4Var(\sqrt{Y}) - 1$ and estimated by $4s^2_{\sqrt{Y}} - 1$ where $s^2_{\sqrt{Y}}$ is the sample variance of $\sqrt{Y}$. Figure 6 shows a scatter plot of $s^2_y/\bar{y} - 1$ versus and $4s^2_{\sqrt{Y}} - 1$ based on ERA40 counts at each grid point. The plot indicates that the approximation is good, except for a handful points that are located near the edge of the map where storm counts are zero for most winters. This approximation to ϕ is now used to build a simple statistical model to characterise $\sqrt{Y}$ and thus clustering.

As a simple model for the relationship between $\sqrt{Y}$ and the NAO, we consider

$$\sqrt{Y} = \alpha + \beta X + \epsilon \qquad (1)$$

where X is the seasonal mean NAO index and ϵ is independent normally distributed noise. To test the assumption of a linear relationship between $\sqrt{Y}$ and X we actually considered a more general model where $\sqrt{Y} = \alpha + f(X)$ and $f(\cdot)$ is a smooth (possibly non-linear) function and $\epsilon \sim N(0, \tau^2)$. This a Generalised Additive Model (GAM) which was implemented in R (R Core Team 2014) using the 'mgcv' package (see Wood (2006) for details). The model was fit at each grid point and CMIP5 model, and Fig. 7 shows the fit for ERA40 and HADGem2 at the grid cell near London and another cell near Scotland (8.75°W, 58.75°N). There is little evidence supporting a non-linear relationship and this is true for most model/grid cell combinations investigated (not shown for brevity). The relationship between NAO and $\sqrt{Y}$ is stronger in Scotland implying that NAO explains more of the clustering than it does near London (see also map in Figure 8). Furthermore, the HADGem2 plots indicate a change in the relationship between $\sqrt{Y}$ and NAO, however it is clear from the 95% confidence intervals that there is no significant difference in the historical and future curves.

Eqn (1) can be shown to imply

$$\phi \approx 4Var(\sqrt{Y}) - 1 = 4\beta^2 Var(X) + 4Var(\epsilon) - 1 \qquad (2)$$

The dispersion is simply the sum of two parts due to the variability of NAO and the variance of the error term ϵ. The error term accounts for influential factors other than the NAO as well as the random variations in $\sqrt{Y}$. To investigate the effect of NAO variability on clustering, Figure 8 shows maps of $4\hat{\beta}_H^2 s^2_{X_H}$, i.e. the estimate of the first part of equation (2), for the historical period (based on historical estimates of β and $Var(X)$). Comparing this with the original ϕ estimates in Figure 2, it is evident that a large proportion of the historical (and also future but not shown) patterns can be attributed to NAO variability – in particular the overdispersion north and south of the storm track. So now the question is whether future changes in NAO are the major driver behind the apparent changes in storm clustering.

Assuming the relationship between storm counts and NAO remains the same in the future (i.e. α and β unchanged), Eqn (2)

implies that the changes in clustering are approximately given by

$$\Delta\phi = 4\beta^2\big(Var(X)_F - Var(X)_H\big) + 4\big(Var(\epsilon)_F - Var(\epsilon)_H\big). \quad (3)$$

In other words, changes in clustering are driven by changes in the variability of NAO and also by changes in the variance of ϵ. The estimate of the first term in (3), i.e. the change in NAO variance, is plotted in Fig. 9 (based on estimates of β from historical data). This is rather different to Fig. 4, which depicts $\Delta\phi$ estimates, implying that changes in NAO explain little of the changes in clustering. The term $Var(X)_F - Var(X)_H$ is very small (which is in agreement with the IPCC findings) and the changes are dominated by the $\big(Var(\epsilon)_F - Var(\epsilon)_H\big)$ term. So although the NAO explains much of the clustering, future changes are dominated by changes in the variance of ϵ – i.e. random variations not related to NAO.

Future changes in ϕ are explained by non-NAO related variations in $\sqrt{Y}$. Simulation experiments are now used to show that these changes can be explained by natural variation in the year to year counts. Considering the London grid point again, the clustering measure ϕ increased from 0.4 to 0.5 in the future for HadGEM2. Model (1) was fitted to the historical data and then 1000 stochastic simulations of the square root of storm counts for the 30-year historical period were produced, calculating the estimate of ϕ each time. These 1000 samples were used to produce a histogram of the sampling distribution of estimated ϕ, shown in Fig. 10a. The apparent increase in dispersion would be hard to detect because of the large variance of the distribution, which is essentially due to the large natural inter-annual variability in the counts. The same simulation experiment was performed, this time forcing NAO to be fixed in each 30-year period (i.e. $Var(X) = 0$), and the resulting histogram of ϕ is shown in Figure 10b. The distribution shifted to the left but is not quite centred at zero as some dispersion is being captured by $Var(\epsilon)$ (estimated to be more than $1/4$). The plot shows that even with NAO fixed, there is a large amount of uncertainty in estimating ϕ from short records of 30 winters. The standard deviation in the ϕ estimates for London are 0.37 and 0.34 for NAO random and NAO fixed respectively – much larger than the change of 0.1 in ϕ.

4. Conclusions

Clustering has been investigated in historical and future runs from 17 CMIP5 models. This study has found that:

- models qualitatively capture clustering seen in ERA40 reanalysis – also for extreme storms;
- there is not much agreement between CMIP5 models with regards to future changes in clustering – any signal appears to be overwhelmed by the variation between model responses. The multi-model mean map resembles the results from Pinto *et al.* (2013) ensemble mean response, although the variability between models is quite large. The equivalent multi-model mean map for the extremes indicates that clustering in extreme storms will increase in the future (RCP4.5) over Northern Europe and Scandinavia;
- NAO accounts for much of the historical clustering but because the variance of NAO changes little in future projections, it has negligible effect on the changes in clustering. So the changes in individual runs (and future observations) are dominated by natural variability in counts.

Sampling uncertainty in the estimate of dispersion over 30 years is large, which prohibits detection of any changes. Multiple realisations are required to reasonably detect a signal, which is what Pinto *et al.* (2013) found using an ensemble of 20 ECHAM5 model runs. Here a collection models was used instead, however the variability of ϕ estimates characterised by differences in CMIP5 models, masks any possible future change. This "disagreement" between models is partly due to actual differences between them, but also due to the large sampling variability of ϕ. The fact that multi-model mean maps and ensemble mean maps from Pinto *et al.* (2013) are comparable is encouraging, unfortunately however future estimates of dispersion are unlikely to match the multi-model mean response.

It is worth mentioning that while the NAO variability can explain the occurrence of cyclone clustering over the North Atlantic basin, it does not explain all of the variability over Europe (see Figure 8). European windstorm occurrence is related to the occurrence of an extended and intensified jet towards Europe (e.g.

Hanley and Caballero (2012); Gómara *et al.* (2014)) and hence may have an effect upon clustering – see Figure 2 where jet stream core area corresponds to the areas with underdispersion in the Western North Atlantic and also the ERA40 map in Figure 2 where overdispersion is evident on both flanks and downstream of the North Atlantic jet, where the variance of the jet is enhanced. As such, jet stream biases in CMIP5 models Zappa *et al.* (2013a) may explain the biases in clustering noted in Figure 2. Moreover, Pinto *et al.* (2014) have recently provided evidence that a recurrent extension of an intensified eddy driven jet towards Western Europe lasting at least one week is one of the main factors leading to the occurrence of historical storm series.

The jet stream is projected to intensify and extend towards Europe under future climate conditions (e.g. Figure 5a). These changes may have implications for storm clustering – for example the increase in clustering over Central Europe for extreme cyclones in Figure 5b. More work would be necessary to actually estimate the relative role of jet stream changes to the possible changes in clustering in CMIP5 models, and in how far other mechanisms (like secondary cyclogenesis) play a more important role.

Acknowledgements

This research was supported by NERC project CREDIBLE. We thank Nathan Gillett for kindly providing the NAO indices for the CMIP5 models. We also thank Kevin Hodges for providing ERA40 cyclone tracks and Phil Sansom for valuable input.

Supporting information

Figure S1: Multi-model mean and individual model changes in dispersion statistic for extreme storms.

Figure S2: Projected dispersion changes assuming no change in mean storm counts.

Figure S3: Projected dispersion changes assuming no change in variance of storm counts.

References

Anscombe FJ. 1948. The transformation of Poisson, binomial and negative binomial data. *Biometrika* **35**(3-4): 246–254, doi:10.1093/biomet/35.3-4.246, URL http://biomet.oxfordjournals.org/content/35/3-4/246.short.

Blender R, Raible CC, Lunkeit F. 2014. Non-exponential return time distributions for vorticity extremes explained by fractional poisson processes. *Quarterly Journal of the Royal Meteorological Society* doi:10.1002/qj.2354, URL http://dx.doi.org/10.1002/qj.2354.

Chang EKM. 2014. Impacts of background field removal on cmip5 projected changes in pacific winter cyclone activity. *Journal of Geophysical Research: Atmospheres* **118**: 4626–4639.

Christensen JH, Krishna Kumar K, Aldrian E, An SI, Cavalcanti IFA, de Castro M, Dong W, Goswami P, Hall A, Kanyanga JK, Kitoh A, Kossin J, Lau NC, Renwick J, Stephenson DB, Xie SP, Zhou T. 2014. Climate phenomena and their relevance for future regional climate change. 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 : 1217–1308.

Cox DR, Isham V. 1980. *Stochastic processes*. Chapman & Hall/CRC Monographs on Statistics & Applied Probability.

Economou T, Stephenson DB, Ferro CAT. 2014. Spatio-temporal modelling of extreme storms. *Annals of Applied Statistics* (To appear).

Gillett NP, Fyfe JC. 2013. Annular mode changes in the CMIP5 simulations. *Geophysical Research Letters* **40**(6): 1189–1193, doi:10.1002/grl.50249, URL http://dx.doi.org/10.1002/grl.50249.

Gómara I, Pinto JG, Woollings T, Masato G, Zurita-Gotor P, Rodríguez-Fonseca B. 2014. Rossby wave-breaking analysis of explosive cyclones in the Euro-Atlantic sector. *Quarterly Journal of the Royal Meteorological Society* **140**(680): 738–753, doi:10.1002/qj.2190, URL http://dx.doi.org/10.1002/qj.2190.

Hanley J, Caballero R. 2012. The role of large-scale atmospheric flow and Rossby wave breaking in the evolution of extreme windstorms over Europe. *Geophysical Research Letters* **39**(21), doi:10.1029/2012GL053408, URL http://dx.doi.org/10.1029/2012GL053408.

Hodges K. 1994. A general method for tracking analysis and its application to meteorological data. *Monthly Weather Review* **122**(11): 2573–2585.

IPCC. 2013. 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*, Stocker TF, Qin D, Plattner GK, Tignor M, K AS, Boschung J, Nauels A, Xia Y, Bex V, M MP (eds), Cambridge University Press: Cambridge, UK and New York, NY, USA, p. 1535 pp, doi:10.1017/CBO9781107415324.

Kvamstø NG, Song Y, Seierstad IA, Sorteberg A, Stephenson DB. 2008. Clustering of cyclones in the ARPEGE general circulation model. *Tellus A* **60**(3): 547–556, doi:10.1073/pnas.1100436108, URL http://www.pnas.org/content/108/43/17626.abstract.

Mailier PJ. 2007. Serial clustering of extratropical cyclones. PhD thesis, University of Reading.

Mailier PJ, Stephenson DB, Ferro CAT, Hodges KI. 2006. Serial clustering of extratropical cyclones. *Monthly Weather Review* **134**(8): 2224–2240, URL http://centaur.reading.ac.uk/119/.

Parker DJ. 1998. Secondary frontal waves in the north atlantic region: A dynamical perspective of current ideas. *Quarterly Journal of the Royal Meteorological Society* **124**(547): 829–856.

Pinto JG, Bellenbaum N, Karremann MK, Della-Marta PM. 2013. Serial clustering of extratropical cyclones over the North Atlantic and Europe under recent and future climate conditions. *Journal of Geophysical Research: Atmospheres* **118**(22): 12,476–12,485, doi:10.1002/2013JD020564, URL http://dx.doi.org/10.1002/2013JD020564.

Pinto JG, Gómara I, Masato G, Dacre HF, Woollings T, Caballero R. 2014. Large-scale dynamics associated with clustering of extra-tropical cyclones affecting Western Europe. *Journal of Geophysical Research: Atmospheres* (To appear), doi:10.1002/2014JD022305, URL http://dx.doi.org/10.1002/2014JD022305.

R Core Team. 2014. R: A language and environment for statistical computing. URL http://www.R-project.org/.

Rudeva I, Gulev SK. 2007. Climatology of cyclone size characteristics and their changes during the cyclone life cycle. *Monthly Weather Review* **135**: 2568–2587.

Sansom PG, Stephenson DB, Ferro CAT, Zappa G, Shaffrey L. 2013. Simple uncertainty frameworks for selecting weighting schemes and interpreting multimodel ensemble climate change experiments. *Journal of Climate* **26**(12): 4017–4037, doi:http://dx.doi.org/10.1175/JCLI-D-12-00462.1.

Seierstad IA, Stephenson DB, Kvamstø NG. 2007. How useful are teleconnection patterns for explaining variability in extratropical storminess? *Tellus A* **59**(2): 170–181, doi:10.1111/j.1600-0870.2007.00226.x, URL http://dx.doi.org/10.1111/j.1600-0870.2007.00226.x.

Taylor KE, Stouffer RJ, Meehl GA. 2012. An overview of CMIP5 and the experimental design. *Bulletin of the American Meteorological Society* **93**: 485–498.

Vitolo R, Stephenson DB, Cook IM, Mitchell-Wallace K. 2009. Serial clustering of intense European storms. *Meteorologische Zeitschrift* **18**(4): 411–424, URL http://dx.doi.org/10.1127/0941-2948/2009/0393.

Wood SN. 2006. *Generalized additive models, an introduction with r*. Chapman & Hall/CRC.

Yu G. 2009. Variance stabilizing transformations of Poisson, binomial and negative binomial distributions. *Statistics & Probability Letters* **79**(14): 1621–1629, doi:10.1016/j.spl.2009.04.010, URL http://www.sciencedirect.com/science/article/pii/S0167715209001473.

Zappa G, Shaffrey LC, Hodges KI. 2013a. The ability of CMIP5 models to simulate north atlantic extratropical cyclones. *Journal of Climate* **26**(15): 5379–5396.

Zappa G, Shaffrey LC, Hodges KI, Sansom PG, Stephenson DB. 2013b. A multimodel assessment of future projections of North Atlantic and European extratropical cyclones in the CMIP5 climate models. *Journal of Climate* **26**(16): 5846–5862.

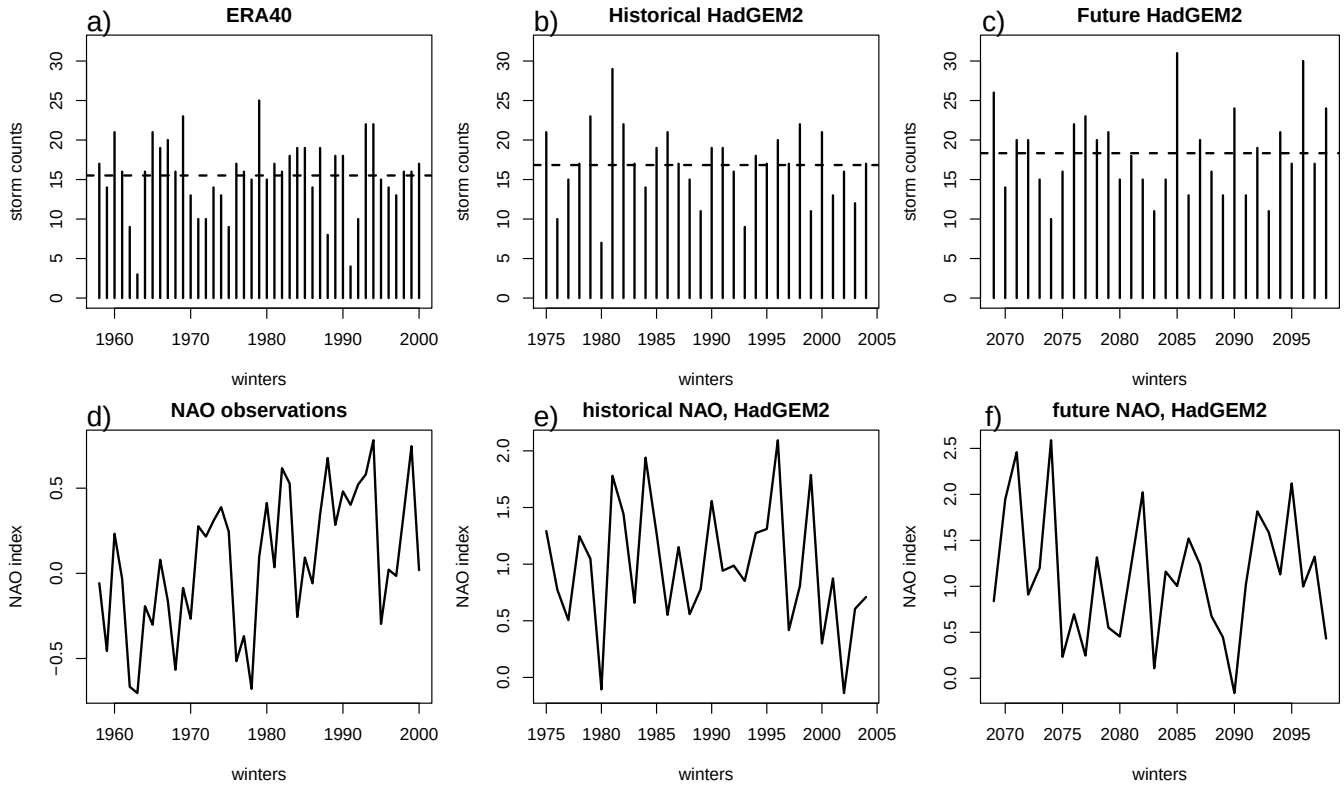

Figure 1. Storm counts and NAO indices for reanalysis (a and d), historical simulation from HadGEM2 (b and e), and future simulation from HadGEM2 (c and f).

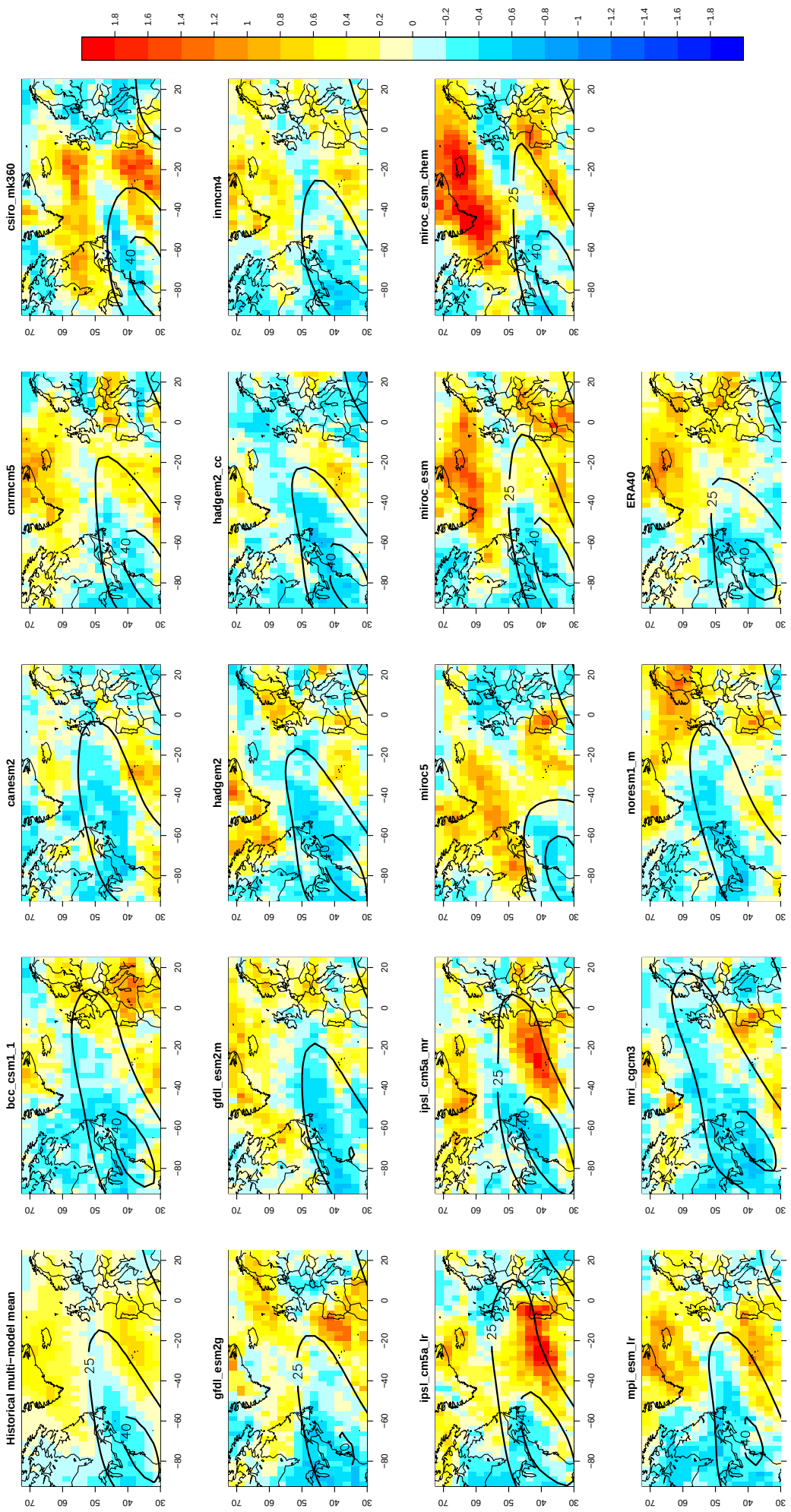

Figure 2. Maps of dispersion statistic ϕ_H. Upper left panel is the multi-model mean, followed by each model in alphabetical order from left to right. The bottom right panel relates to ERA40. The upper tropospheric jet axis is indicated by contours (25 and 40ms^{-1}) of the mean zonal wind at 250hPa for each model.

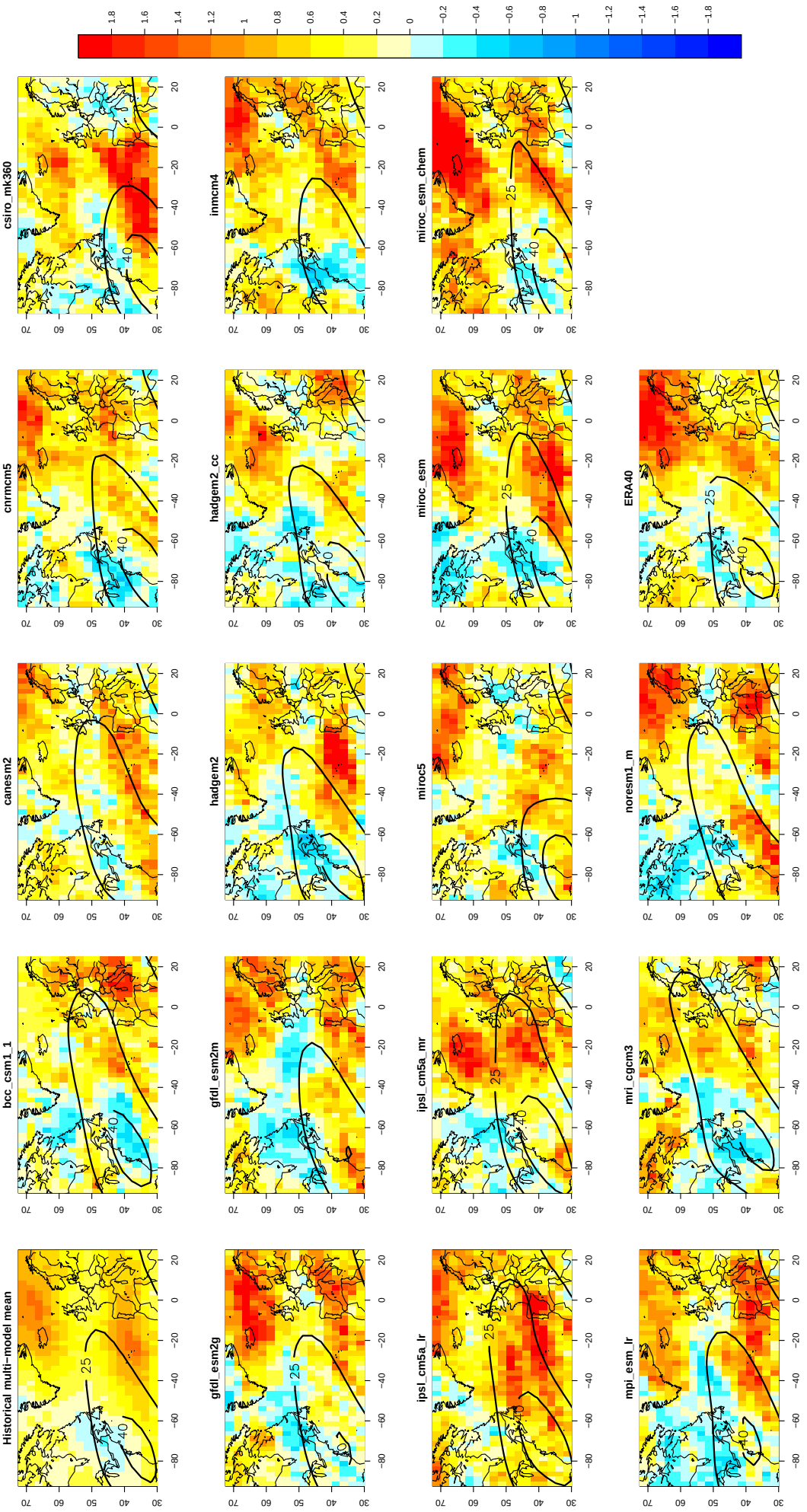

Figure 3. Maps of dispersion statistic ϕ_H for extreme storms. Upper left panel is the multi-model mean, followed by each model in alphabetical order from left to right. The bottom right panel relates to ERA40.

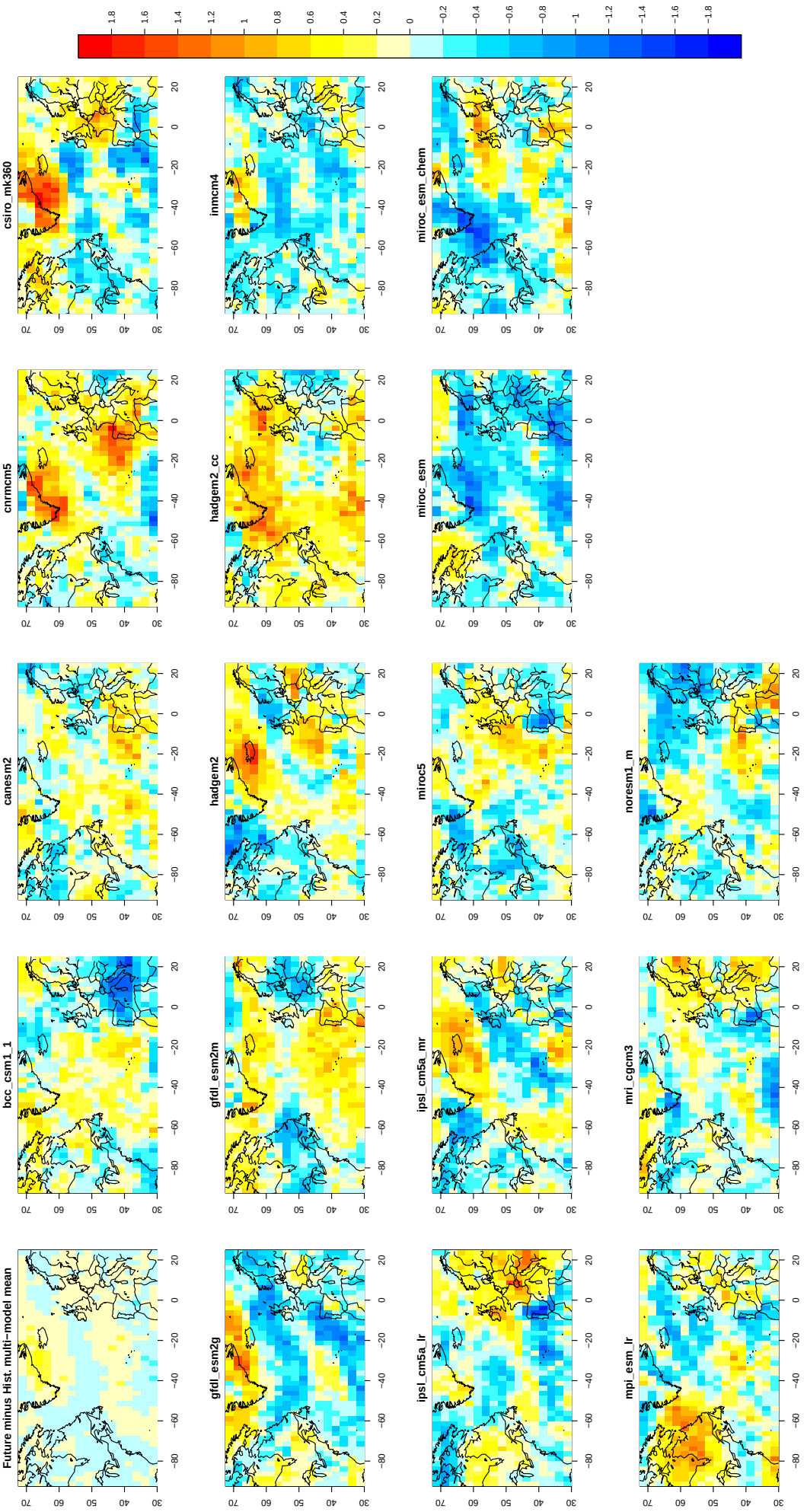

Figure 4. Maps of projected changes in dispersion $\Delta\phi = \phi_F - \phi_H$. Upper left panel is the multi-model mean, followed by each model in alphabetical order from left to right. Lower right panel reflects the percentage change in multi-model ϕ.

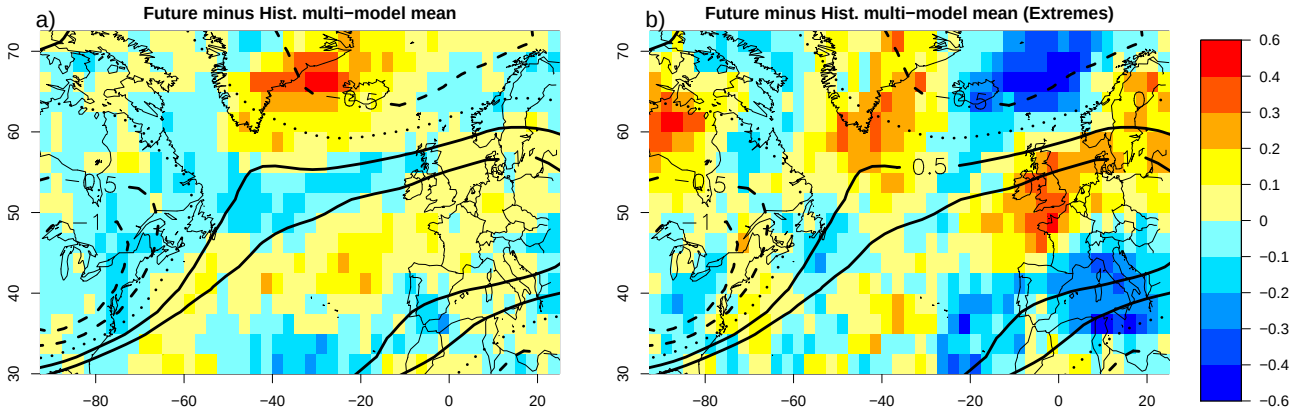

Figure 5. Maps of multi-model means of projected changes in dispersion $\Delta\phi = \phi_F - \phi_H$ for all storms (left panel) and extreme storms (right panel). Contours of multi-model mean jet changes are indicated at -1, -0.5, 0, 0.5 and 1ms^{-1}, using dashed, dotted and solid lines for negative, zero and positive values respectively.

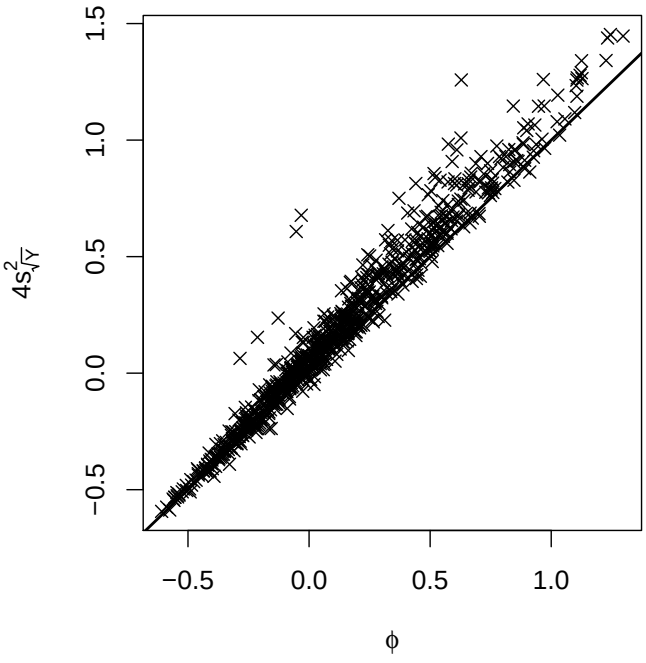

Figure 6. Plot of dispersion statistic ϕ and its approximation $4s^2_{\sqrt{Y}} - 1$ for ERA40 reanalysis. Each cross corresponds to a grid point.

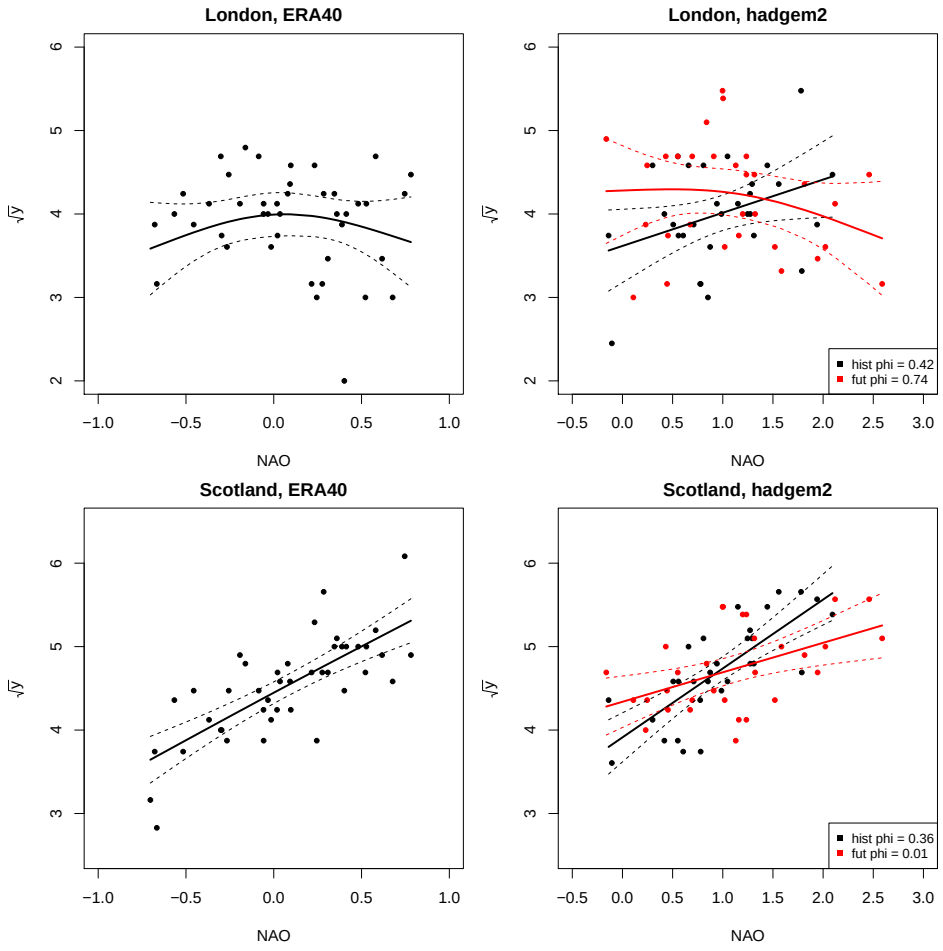

Figure 7. Estimated relationship between square root of counts and NAO with associated 95% confidence intervals, for counts near London in top panels and Scotland in bottom panels. Left panels relates to ERA40 reanalysis and right panels to HadGEM2.

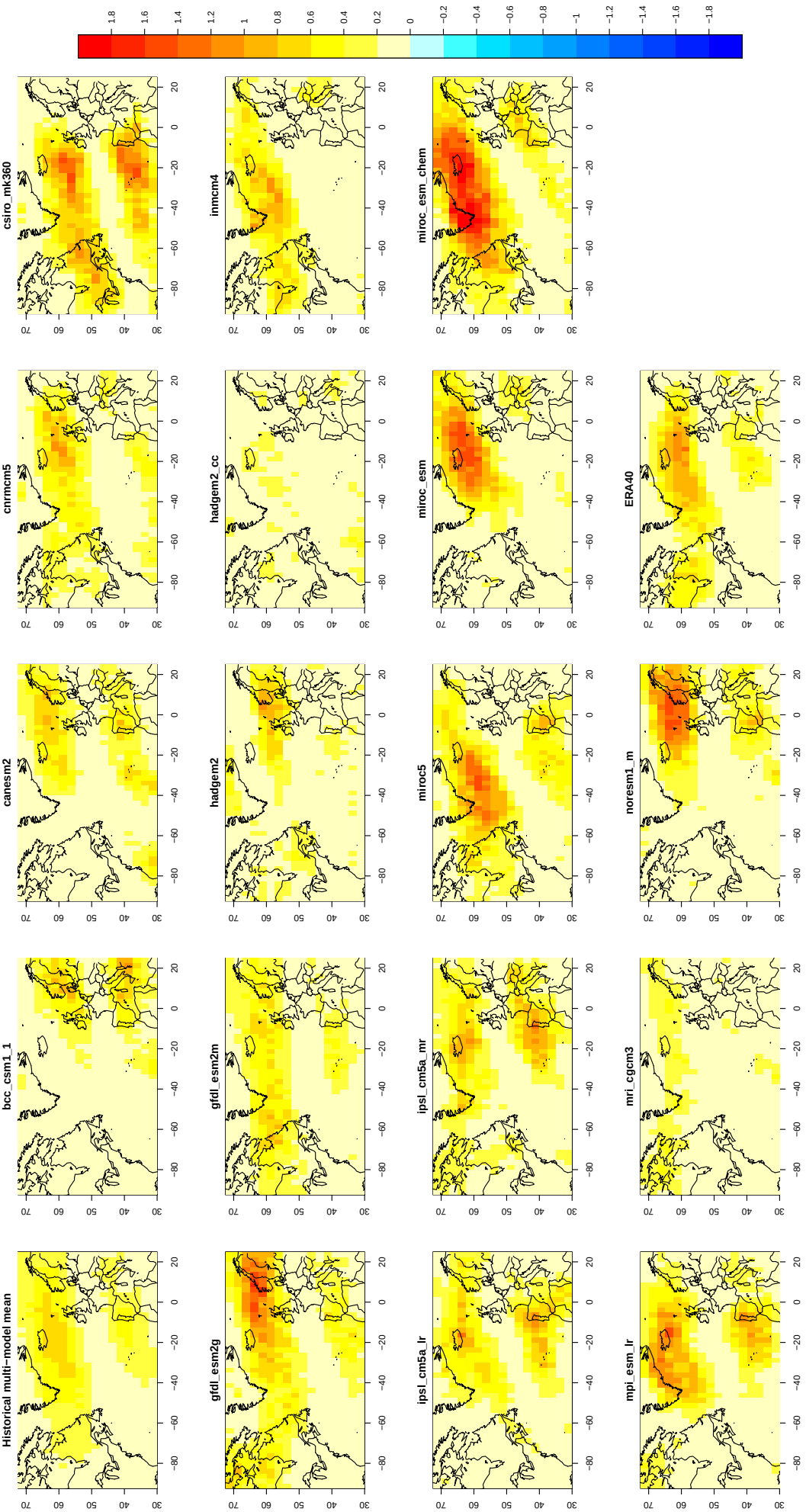

Figure 8. Estimates of $4\beta^2 Var(X)$ for the historical data. Upper left panel is the multi-model mean, followed by each model in alphabetical order from left to right. The bottom right panel relates to ERA40.

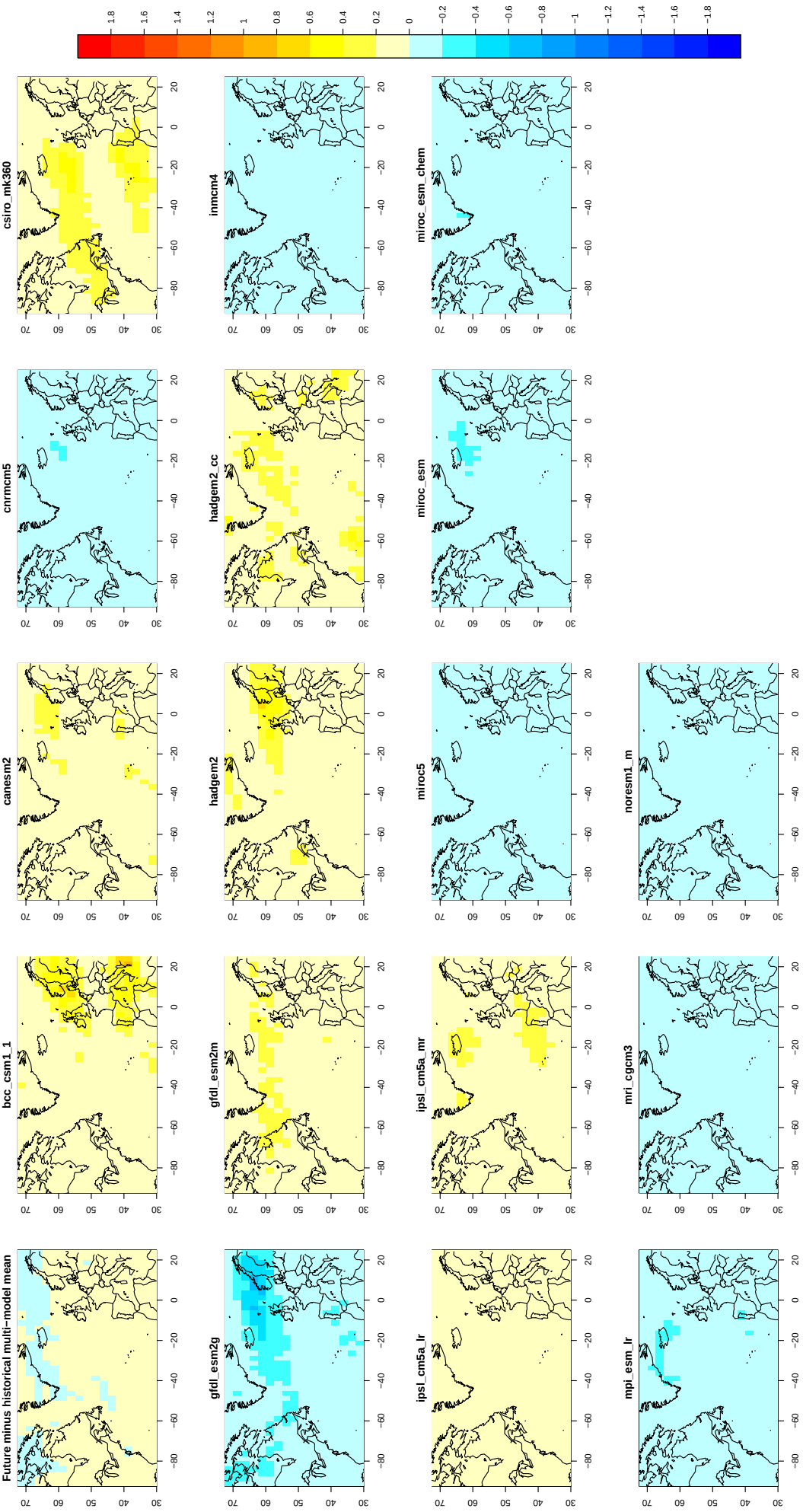

Figure 9. Estimates of $4\beta^2(Var(X)_F - Var(X)_H)$. Upper left panel is the multi-model mean, followed by each model in alphabetical order from left to right.

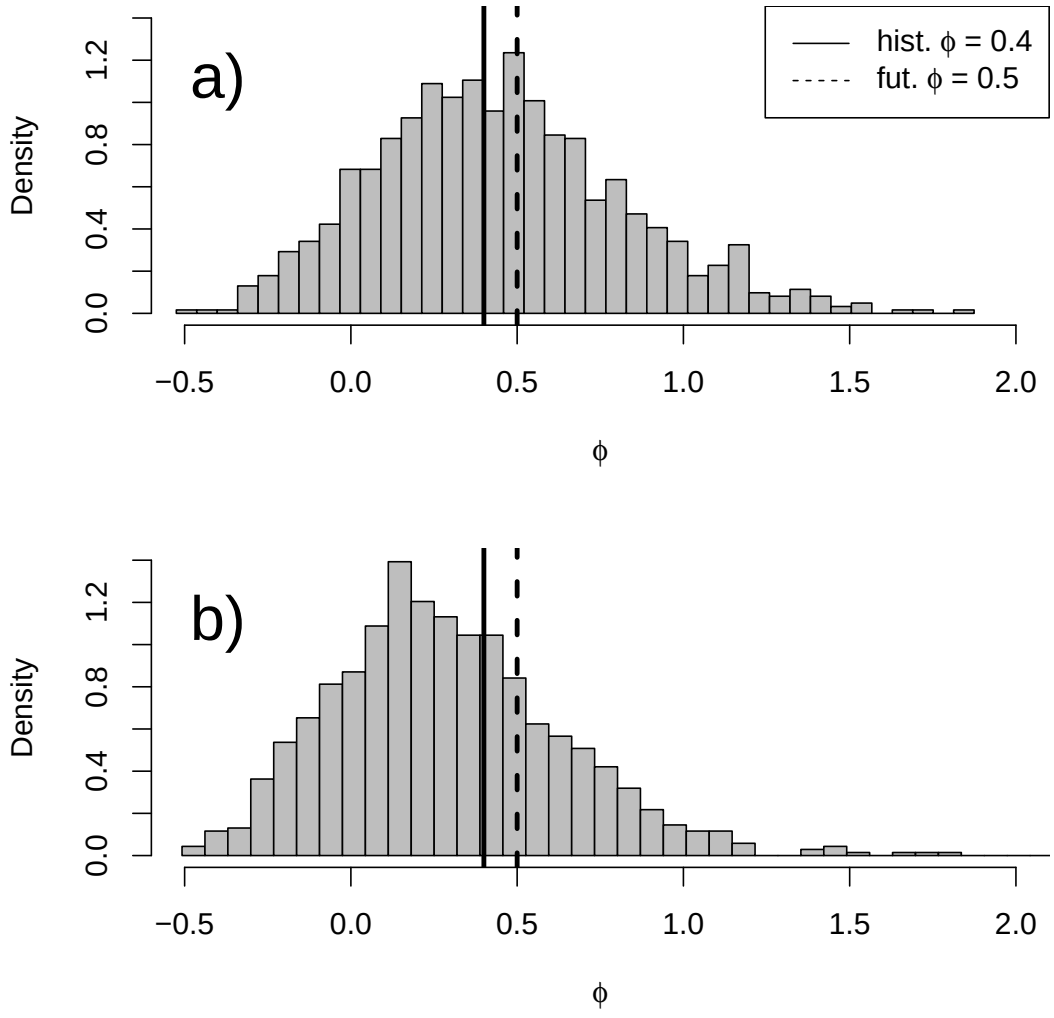

Figure 10. a) Histogram showing the sampling distribution of the dispersion statistic ϕ near London for HadGEM2. b) Histogram The vertical lines show the historical (solid) and future (dashed) ϕ estimates.

