

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21

Do CMIP5 models reproduce observed low-frequency North Atlantic jet variability?

Thomas J. Bracegirdle¹, Hua Lu¹, Rosie Eade², Tim Woollings³

¹British Antarctic Survey, Cambridge, UK.

²Met Office Hadley Centre, Exeter, UK

³Department of Physics, Atmospheric, Oceanic and Planetary Physics, University of Oxford, Oxford, UK

Corresponding author: Thomas J. Bracegirdle (tjbra@bas.ac.uk)

Key Points:

- Reproducing post-1862 N Atlantic (NA) multi-decadal jet strength variability is a tougher test for models than reproducing NAO variability
- An observed strong multi-decadal correlation between NA jet strength and NAO may be specific to the 1862-2005 period.
- Efforts to improve models' representation of NA low-frequency variability should involve understanding drivers of jet strength variability.

22 **Abstract**

23 The magnitude of observed multi-decadal variations in the North Atlantic Oscillation (NAO) is
24 at the upper end of the range simulated by climate models and a clear explanation for this
25 remains elusive. Recent research shows that observed multi-decadal NAO variability is more
26 strongly associated with North Atlantic (NA) jet strength than latitude, thus motivating a
27 comprehensive comparison of NA jet and NAO variability across the CMIP5 models. Our results
28 show that the observed peak in multi-decadal jet strength variability is even more unusual than
29 NAO variability when compared to the model-simulated range across 133 historical CMIP5
30 simulations. Some CMIP5 models appear capable of reproducing the observed low-frequency
31 peak in jet strength, but there are too few simulations of each model to clearly identify which. It
32 is also found that an observed strong multi-decadal correlation between jet strength and NAO
33 since the mid-19th century may be specific to this period.

34

35 **Plain Language Summary**

36 The dominant pattern in sea-level pressure variability over the North Atlantic is the North
37 Atlantic Oscillation (NAO), which is strongly associated with climate variations over western
38 Europe. However, major focus of current research is that climate models broadly do not
39 reproduce large enough low-frequency (longer than 30 years) NAO variability, with the
40 implication that significant improvements in the skill of decadal weather prediction are possible.

41 Motivated by this, we analyzed data from 44 of the world's leading models and delved more
42 deeply into major underlying characteristics of low-frequency winter NAO variability,
43 specifically: (1) strength and (2) latitude of the belt of prevailing westerly winds that blow across
44 the North Atlantic from the Eastern USA to Western Europe. Our study revealed that it is the
45 strength of the westerlies, rather than latitude, which underlies models' difficulty in reproducing
46 low-frequency NAO variability.

47 This more detailed picture also really matters in terms of impacts on European climate, since a
48 more positive NAO associated mainly with stronger westerlies has different climate impacts
49 compared to one associated mainly with more poleward westerlies. This is emphasized by a
50 further finding that either component of the westerlies can dominate NAO variability across
51 different decades and even centuries.

52

53 **1 Introduction**

54 It is important to understand low-frequency variability of the North Atlantic (NA) climate
55 system since it exerts a strong influence on European climate (Sutton et al., 2018). This is highly
56 relevant to both decadal predictability of future change as well as efforts to improve detection
57 and attribution of climate trends that have occurred in the past. Recent studies have highlighted
58 that a large proportion of contemporary climate models exhibit weak multi-decadal 20th century
59 NA climate variability when compared to observations (Cheung et al., 2017; Han et al., 2016;
60 Kravtsov, 2017; Wang et al., 2017; Zhang & Wang, 2013). Further, many models exhibit
61 weaker-than-observed links between the NAO and low-frequency phenomena such as Atlantic
62 multi-decadal variability (AMV) (Ba et al., 2014; Keeley et al., 2012; Kim et al., 2018; Omrani
63 et al., 2014; Peings et al., 2016). These studies suggest that multi-decadal internal variability of

64 the NAO is in fact too weak in most contemporary climate models (Kim et al., 2018; Kravtsov,
65 2017). However, the roles of different drivers associated with both internal variability and
66 external forcing remain unclear. For example there are still questions over possible drivers of the
67 dramatic increase in the winter NAO during the 30-year period between the 1960s and 1990s.

68 Recent research shows that diagnostics of the regional NA eddy-driven tropospheric
69 westerly jet can help to improve understanding of key drivers and impacts of NAO variability.
70 Woollings et al. (2015) found that 20th century inter-annual winter NAO variability is largely
71 associated with variability in jet latitude whereas multi-decadal NAO variability (> 30 years) is
72 more strongly related to variability in jet strength. To illustrate the importance of this, Figure 1
73 shows the climatological winter-mean lower-tropospheric westerly wind structure of the NA
74 region, for which a basin-wide (~60°W to ~0°E) strengthening and/or poleward shifting of the
75 mid-latitude wind maximum (the jet) are both associated with a more positive NAO (note that
76 only the shorter-term linkage is shown here and the interested reader is referred to Woollings et
77 al. (2015) for an in-depth evaluation of such relationships at different timescales). A key point is
78 that for a given change in the NAO the impacts or drivers of that change depend strongly on the
79 ratio of associated jet shifting and/or strengthening, which exhibit contrasting spatial correlation
80 patterns with local westerly wind across the NA and over Europe (compare Figures 1b and 1c).
81 The above-mentioned multi-decadal winter jet strength variability exhibits a distinct maximum
82 in power at low frequencies that is strongly correlated with AMV (Woollings et al., 2014).
83 However, for jet latitude there is no such low-frequency maximum in winter variability.
84 Furthermore, idealized model experiments produce different sensitivities of jet latitude and
85 strength to different external drivers (Baker et al., 2017; McGraw & Barnes, 2016).

86 These studies motivate us to carry out an analysis of NA jet variability across the CMIP5
87 models to help to identify reasons for the relatively weak simulated low-frequency NAO
88 variability. We additionally utilize the large ensemble size of the CMIP5 simulations to evaluate
89 the longer-term stationarity of the relationships between low-frequency jet and NAO variability,
90 which is important for understanding spatial patterns of impacts on the NA climate system. Our
91 analysis aims to address two research questions:

92 1. To what extent is the observed low-frequency variability in jet strength reproduced
93 across the CMIP5 historical simulations?

94 2. Is the observed link between jet strength and the NAO at multi-decadal time-scales
95 reproduced in the CMIP5 simulations?

96

97 **2 Data and methods**

98 2.1 CMIP5 data

99 Output from all-forcing historical coupled simulations of the World Climate Research
100 Programme's (WCRP's) Coupled Model Intercomparison Project Phase 5 (CMIP5) was used for
101 this analysis (Taylor et al., 2012). The fields analyzed were mean sea level pressure (MSLP) and
102 zonal wind at the 850 hPa level (U850 hereafter). Ideally daily fields would have been used for
103 the main analysis of U850, but these were generally only archived for the later part of the 20th
104 century in the historical simulations. The sensitivity to this limitation is evaluated using available
105 daily mean fields of U850 and the results are provided later in this section (see Sections 2.4 and

106 2.6). We analyzed all available realizations (i.e. individual historical simulations) with the above
107 variables and output available back to at least 1861. This comprised 133 members from 46
108 different models (Supporting Table S1).

109 2.2 Twentieth Century Reanalysis Version 2c

110 The NOAA/CIRES Twentieth Century Global Reanalysis Version 2c (20CRv2c) was
111 used to derive estimates of real-world jet variability (Compo et al., 2015). This version of the
112 20CR range of reanalysis datasets extends back to 1851, but the post-1861 period has been
113 described as more suitable for climate studies and is therefore used here. To provide a measure
114 of uncertainty in the re-analysis-derived results, all 56 ensemble members were evaluated.
115 Decadal variability of the NA jet in 20CR has been found to be similar to that in the century-long
116 ERA20C reanalysis (Woollings et al., 2018).

117 2.3 NAO indices

118 Two different NAO indices were used to help ensure robustness of the results and to aid
119 comparison with earlier studies (e.g. Kelley et al., 2012; Osborn, 2004). First, following Scaife et
120 al. (2009), a point index was defined as the non-normalized difference between winter season
121 (December to February; DJF hereafter with the month of January defining the year of each
122 winter) MSLP over the Azores (Ponta Delgada) and Iceland (Akureyri). For the gridded datasets
123 (20CRv2c and CMIP5 output) bilinear interpolation was used to extract MSLP values at these
124 locations. Second, an Empirical Orthogonal Function (EOF)-based approach was also used,
125 defined here as the principal component time series of the leading EOF of winter-mean MSLP
126 anomalies over each reanalysis or CMIP5 time series over the NA region (20-80N, 90W-40E)
127 (following Hurrell, 1995).

128 2.4 Jet strength and latitude diagnostics

129 The diagnostics of NA jet strength and latitude in the lower troposphere draw from
130 Woollings et al. (2015), who use U850. For each U850 field (monthly or daily), the longitudinal
131 mean between 60°W and 0° was first calculated at each latitude. This was then used to identify
132 the maximum (jet strength) and its location (jet latitude). Winter averages were then estimated as
133 a mean of the monthly or daily jet diagnostics through DJF. The same diagnostics were applied
134 to both the 20CRv2c and CMIP5 data. Due to a lack of daily-mean data covering the full length
135 of historical CMIP5 simulations, our analysis was conducted based on monthly mean U850
136 fields. A comparison between available overlapping monthly and daily CMIP5 data suggests that
137 diagnostics based on monthly mean fields are a close analogue for those derived from daily data
138 (see Supporting Information S2 for details).

139 2.5 Frequency power spectra

140 Power spectra were generated from unsmoothed time series of winter-mean NAO and jet
141 diagnostics. A common year range of 1862-2005 was used across the CMIP5 model realizations
142 and 20CRv2c for this part of the analysis to ensure comparable frequency bins.

143 3 Results

144 Figure 2 (a-d) shows decadal-smoothed time series of winter NAO, jet latitude and jet
145 strength anomalies from all available CMIP5 realizations and ensemble means from 20CRv2c

146 reanalysis data (see Supporting Information S1 for smoothing methodology). All four observed
147 diagnostics exhibit significant increases over the 1960s-1990s period, which for the NAO indices
148 and jet latitude are the largest in the period evaluated. For jet strength there is a larger 30-year
149 increase during the 1880s-1910s. This coincides with a period of declining AMV index values
150 and is part of the multi-decadal correlation between the AMV and jet strength (Woollings et al.,
151 2015).

152 From Figure 2 (a-d) there is evidence of contrasting amplitudes of variability in 20CRv2c
153 compared to CMIP5 models. This is particularly the case for jet strength (Figure 2d), for which
154 the standard deviation of the 30-year smoothed de-trended 20CRv2c ensemble mean time series
155 (0.59 m s^{-1}) exceeds the largest standard deviation derived from any of the time series from the
156 133 CMIP5 historical realizations (0.57 m s^{-1}).

157 A similar contrast is also evident in terms of extreme 30-year linear trends (Figure 2 (e-
158 h)). The 30-year period 1965-1994 has previously been highlighted as exhibiting a large
159 observed NAO increase, so this trend length was chosen in Figure 2 to compare with the
160 frequency of occurrence in the CMIP5 models. Here maxima in rolling 30-year overlapping
161 trends (starting just one year apart) for each CMIP5 realization are calculated from the decadal-
162 smoothed time series shown in Figure 2 (a-d). The maximum of these overlapping trends is then
163 extracted for each realization and used to populate the histograms shown in Figure 2 (e-h). The
164 same procedure is applied to the decadal-smoothed reanalysis-derived time series and the
165 resulting trends are shown by the vertical lines in Figure 2 (e-h). By this measure, 30-year trends
166 matching or exceeding the largest seen in the 20CRv2c ensemble mean are less likely for jet
167 strength (3.0 standard deviations (std) from the multi-realization mean) than for the NAO indices
168 (2.4 std for the point index and 2.7 std for the EOF index) or jet latitude (1.0 std). Linear de-
169 trending of the time series brings CMIP5 and 20CRv2c maximum jet strength trends marginally
170 closer (2.9 std) due to a small positive long-term trend in the re-analysis-estimated jet strength
171 (see Figure S3).

172 A more comprehensive evaluation across different timescales is achieved by using power
173 spectra analysis (Figure 3). This allows an examination of the extent to which CMIP5 historical
174 simulations reproduce the reanalysis-derived significant peak in low-frequency spectral power in
175 jet strength identified by Woollings et al. (2014). Figure 3d shows that the observed low-
176 frequency maximum in spectral power of jet strength at approximately 70 years stands out from
177 the 133 historical CMIP5 realizations. The same conclusion, though less clear, can be reached
178 for the NAO indices (Figures 3a,b). For jet latitude however the CMIP5 models broadly overlap
179 the 20CRv2c-derived power spectrum (Figure 3c). Isolating the low-frequency part of the power
180 spectra and displaying the spectral power for each model and realization immediately illustrates
181 the large internal climate variability across different historical realizations (Figure 4). Despite
182 the fact that in many cases historical simulations from the same model produce a wide range of
183 low-frequency power, it is also clear that only a very small number of realizations get close to
184 the magnitude of the observed relative spectral power for jet strength. Note also that CMIP5-
185 reanalysis differences are even more pronounced in the absence of linear de-trending (Figures S4
186 and S5).

187 Another key characteristic of the observed NAO is that its low frequency variability is
188 more strongly related to jet strength variability than jet latitude variability (Figure 5) (Woollings
189 et al., 2015). However, here we found that the opposite is seen across the CMIP5 ensemble, with
190 on average a higher correlation between 30-yr smoothed jet latitude and NAO time series ($r =$

191 0.66 for point index and 0.68 for EOF index) than for jet strength vs NAO ($r = 0.37$ for point
192 index and 0.30 for EOF index. Note that for jet strength vs jet latitude, the CMIP5 models are
193 consistent with 20CRv2c in exhibiting no significant correlation ($r = 0.04$). With regard to NAO
194 vs jet strength, it remains unclear whether the contrast between models and reanalysis is due to
195 sampling uncertainty of the real world and/or structural biases in the CMIP5 models. The former
196 cannot be discounted since models with a large number of realizations exhibit a wide range of
197 low-frequency correlations over the historical period (Figure 5). This suggests that the high
198 correlation between jet strength and NAO indices evident in reanalyses may be specific to the
199 observational period and not necessarily representative of longer-term behavior.

200 **4 Conclusions**

201 This paper was motivated by recent studies highlighting an apparent lack of multi-
202 decadal NAO variability across contemporary climate models (Kim et al., 2018; Kravtsov, 2017)
203 and the recognition that winter multi-decadal variability appears from reanalyses/observations to
204 be expressed more strongly in North Atlantic (NA) jet strength (Woollings et al., 2015). We have
205 extended previous assessments of NAO variability in multi-model ensembles and included NA
206 westerly jet strength and latitude in a comparison between multi-decadal variability in reanalysis
207 data and historical simulations of the CMIP5 models.

208 The main result is that observed (reanalysis) post-1862 multi-decadal jet strength
209 variability falls further into the upper tail of the range spanned by CMIP5 historical simulations
210 than multi-decadal NAO variability. Only a very limited subset of CMIP5 models appear to
211 reproduce the observed peak in low-frequency variability. In contrast, jet latitude exhibits no
212 clear contrast between observed and CMIP5 variability. This picture is evident across a range of
213 different diagnostics: simple smoothed time series, maximum rolling 30-year linear trends and
214 power spectra.

215 The reanalysis/CMIP5 differences in jet strength variability are most prominent for
216 periods of approximately 70 years (Figure 3), which is consistent with a link between multi-
217 decadal variability in jet strength and Atlantic Multi-decadal Variability (AMV) that was
218 identified previously by Woollings et al. (2015) (who found a correlation of -0.48 between AMV
219 and multi-decadal jet strength variability using a similar jet strength metric). Studies show that
220 the observed AMV is stronger than that in most CMIP5 models (Cheung et al., 2017; Zhang &
221 Wang, 2013). The comparatively prominent observed low-frequency strength variability
222 compared to CMIP5 is consistent with previous work suggesting a too-weak feedback between
223 AMV and atmospheric variability (Kim et al., 2018; Peings et al., 2016; Wang et al., 2017). For
224 example Kim et al. (2018) found that even with realistic AMV-related SST anomalies used in
225 atmosphere-only simulations, multi-decadal NAO variability in the CAM5 model remains
226 stubbornly weaker than observed. Nevertheless, certain CMIP5 models do exhibit evidence of
227 stronger feedbacks. Two CMIP5 models (GFDL-ESM2G (model 22) and HadGEM2-ES (model
228 30)) were identified by Peings et al. (2016) as producing a clear lagged signal between the NAO
229 and AMV. Both models come from families of models (i.e. GFDL models 20-23 and Met Office
230 Hadley Centre models 29-30) that emerge as producing historical realizations with relatively
231 large multi-decadal variability in jet strength (e.g. Fig. 4). It is worth noting from Figure 4 that
232 there is a large scatter between individual realizations from the same model (i.e. internal climate
233 variability). An implication of this is that for each of these and a number of other CMIP5 models,

234 the ensemble size could be too small to confidently assert how unusual the observed low-
235 frequency peak in jet strength is.

236 Overall the above results indicate that the too-weak 20th century multi-decadal variability
237 in the NA jet stream, as identified through the NAO in previous studies, is even more
238 pronounced in terms of the jet strength. In terms of climate impacts, such biases in characteristics
239 and strength of low-frequency jet variability are highly important. This is because compared to
240 jet shifting, jet pulsing (i.e. strengthening and weakening) produces contrasting surface
241 temperature correlation patterns over Europe (e.g. Woollings et al., 2015). In addition, variations
242 in the background strength of the jet would also influence shorter-term seasonal and intra-
243 seasonal jet variability (Hanna et al., 2015; Woollings et al., 2018).

244 Of further relevance to impacts of atmospheric variability, Figure 5 shows that the
245 observed stronger NAO-jet strength relationship on multi-decadal timescales is not stationary in
246 the CMIP5 models. This raises the possibility that the observed correlation is specific to the post-
247 1862 period. The simulated lack of stationarity in such correlations potentially relates to multi-
248 decadal variability of NAO centers of action that have been identified in re-analysis datasets
249 covering the modern era (Moore et al., 2013; Wang et al., 2012), longer-term proxy
250 reconstructions, and climate models (Raible et al., 2006; Raible et al., 2014). As such, multi-
251 decadal NAO variability across different centuries could exhibit contrasting characteristics and
252 impacts. For a shorter-term example, compare the observed positive NAO trends in the periods
253 1880-1910 and 1960-1990 (Figure 2). The latter of these arose from a combination of jet latitude
254 and strength trends which were not exceptional in themselves. In contrast, the earlier period was
255 linked to an exceptional jet strength trend, potentially associated with the more sustained decline
256 in AMV indices during this period (e.g. Peings et al., 2016). More broadly this study highlights
257 the importance of taking a comprehensive approach to evaluating large-scale NA atmospheric
258 variability, such as the combination of NAO and jet diagnostics presented here or more detailed
259 diagnostics of NAO structure such as the NAO angle index of Wang et al. (2012). Ultimately,
260 such analyses should be used to help inform studies aiming to identify the key physical
261 mechanisms to explain model-observation differences in low-frequency NA climate variability.

262

263 **Acknowledgments**

264 TB and HL were supported as part of the NERC ACSIS project (NE/N018028/1). TW
265 was supported by the NERC IMPETUS project (NE/L01047X/1). RE was supported by the Met
266 Office Hadley Centre Climate Programme (GA01101). We acknowledge the WCRP's Working
267 Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate
268 modeling groups (listed in Table S1 of this paper) for producing and making available their
269 model output. For CMIP the U.S. Department of Energy's Program for Climate Model Diagnosis
270 and Intercomparison provided the coordinating support and led development of software
271 infrastructure in partnership with the Global Organization for Earth System Science Portals.
272 CMIP5 data can be accessed at <https://esgf-node.llnl.gov/projects/cmip5/>. The 20CRv2c data
273 was provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA from their Web site at
274 <https://www.esrl.noaa.gov/psd/>. Support for the 20CRv2c dataset is provided by the U.S.
275 Department of Energy, Office of Science Biological and Environmental Research (BER), and by
276 the National Oceanic and Atmospheric Administration Climate Program Office.

277 **References**

- 278 Ba, J., Keenlyside, N. S., Latif, M., Park, W., Ding, H., Lohmann, K., et al. (2014). A multi-
279 model comparison of Atlantic multidecadal variability. *Clim. Dyn.*, *43*(9-10), 2333-2348.
280 doi:10.1007/s00382-014-2056-1
- 281 Baker, H. S., Woollings, T., & Mbengue, C. (2017). Eddy-Driven Jet Sensitivity to Diabatic
282 Heating in an Idealized GCM. *J. Clim.*, *30*(16), 6413-6431. doi:10.1175/jcli-d-16-0864.1
- 283 Cheung, A. H., Mann, M. E., Steinman, B. A., Frankcombe, L. M., England, M. H., & Miller, S.
284 K. (2017). Comparison of Low-Frequency Internal Climate Variability in CMIP5 Models
285 and Observations. *J. Clim.*, *30*(12), 4763-4776. doi:10.1175/jcli-d-16-0712.1
- 286 Compo, G. P., Whitaker, J. S., Sardeshmukh, P. D., Allan, R. J., McColl, C., Yin, X., et al.
287 (2015). NOAA/CIRES Twentieth Century Global Reanalysis Version 2c. Boulder, CO:
288 Research Data Archive at the National Center for Atmospheric Research, Computational
289 and Information Systems Laboratory.
- 290 Han, Z., Luo, F. F., Li, S. L., Gao, Y. Q., Furevik, T., & Svendsen, L. (2016). Simulation by
291 CMIP5 models of the atlantic multidecadal oscillation and its climate impacts. *Adv.*
292 *Atmos. Sci.*, *33*(12), 1329-1342. doi:10.1007/s00376-016-5270-4
- 293 Hanna, E., Cropper, T. E., Jones, P. D., Scaife, A. A., & Allan, R. (2015). Recent seasonal
294 asymmetric changes in the NAO (a marked summer decline and increased winter
295 variability) and associated changes in the AO and Greenland Blocking Index. *Int. J.*
296 *Climatol.*, *35*(9), 2540-2554. doi:10.1002/joc.4157
- 297 Hurrell, J. W. (1995). Decadal trends in the North-Atlantic oscillation - regional temperatures
298 and precipitation. *Science*, *269*, 676-679.
- 299 Keeley, S. P. E., Sutton, R. T., & Shaffrey, L. C. (2012). The impact of North Atlantic sea
300 surface temperature errors on the simulation of North Atlantic European region climate.
301 *Q. J. Roy. Meteorol. Soc.*, *138*(668), 1774-1783. doi:10.1002/qj.1912
- 302 Kelley, C., Ting, M. F., Seager, R., & Kushnir, Y. (2012). The relative contributions of radiative
303 forcing and internal climate variability to the late 20th Century winter drying of the
304 Mediterranean region. *Clim. Dyn.*, *38*(9-10), 2001-2015. doi:10.1007/s00382-011-1221-z
- 305 Kim, W. M., Yeager, S., Chang, P., & Danabasoglu, G. (2018). Low-Frequency North Atlantic
306 Climate Variability in the Community Earth System Model Large Ensemble. *J. Clim.*, *31*,
307 787-813. doi:10.1175/jcli-d-17-0193.1
- 308 Kravtsov, S. (2017). Pronounced differences between observed and CMIP5-simulated
309 multidecadal climate variability in the twentieth century. *Geophys. Res. Lett.*, *44*(11),
310 5749-5757. doi:10.1002/2017gl074016
- 311 McGraw, M. C., & Barnes, E. A. (2016). Seasonal Sensitivity of the Eddy-Driven Jet to
312 Tropospheric Heating in an Idealized AGCM. *J. Clim.*, *29*(14), 5223-5240.
313 doi:10.1175/jcli-d-15-0723.1
- 314 Moore, G. W. K., Renfrew, I. A., & Pickart, R. S. (2013). Multidecadal Mobility of the North
315 Atlantic Oscillation. *J. Clim.*, *26*(8), 2453-2466. doi:10.1175/jcli-d-12-00023.1
- 316 Omrani, N. E., Keenlyside, N. S., Bader, J., & Manzini, E. (2014). Stratosphere key for
317 wintertime atmospheric response to warm Atlantic decadal conditions. *Clim. Dyn.*, *42*(3-
318 4), 649-663. doi:10.1007/s00382-013-1860-3
- 319 Osborn, T. J. (2004). Simulating the winter North Atlantic Oscillation: the roles of internal
320 variability and greenhouse gas forcing. *Clim. Dyn.*, *22*(6-7), 605-623.
321 doi:10.1007/s00382-004-0405-1

- 322 Peings, Y., Simpkins, G., & Magnusdottir, G. (2016). Multidecadal fluctuations of the North
323 Atlantic Ocean and feedback on the winter climate in CMIP5 control simulations. *J.*
324 *Geophys. Res.-Atmos.*, *121*(6), 2571-2592. doi:10.1002/2015jd024107
- 325 Raible, C. C., Casty, C., Luterbacher, J., Pauling, A., Esper, J., Frank, D. C., et al. (2006).
326 Climate variability-observations, reconstructions, and model simulations for the Atlantic-
327 European and Alpine region from 1500-2100 AD. *Clim. Change*, *79*(1-2), 9-29.
328 doi:10.1007/s10584-006-9061-2
- 329 Raible, C. C., Lehner, F., Gonzalez-Rouco, J. F., & Fernandez-Donado, L. (2014). Changing
330 correlation structures of the Northern Hemisphere atmospheric circulation from 1000 to
331 2100 AD. *Clim. Past*, *10*(2), 537-550. doi:10.5194/cp-10-537-2014
- 332 Scaife, A. A., Kucharski, F., Folland, C. K., Kinter, J., Bronnimann, S., Fereday, D., et al.
333 (2009). The CLIVAR C20C project: selected twentieth century climate events. *Clim.*
334 *Dyn.*, *33*(5), 603-614. doi:10.1007/s00382-008-0451-1
- 335 Sutton, R. T., McCarthy, G. D., Robson, J., Sinha, B., Archibald, A., & Gray, L. J. (2018).
336 Atlantic Multi-decadal Variability and the UK ACSIS programme. *Bull. Am. Meteorol.*
337 *Soc.*, in press. doi:10.1175/bams-d-16-0266.1
- 338 Taylor, K. E., Stouffer, R. J., & Meehl, G. A. (2012). An Overview of CMIP5 and the
339 Experiment Design. *Bull. Am. Meteorol. Soc.*, *93*, 485-498. doi:10.1175/bams-d-11-
340 00094.1
- 341 Wang, X. F., Li, J. P., Sun, C., & Liu, T. (2017). NAO and its relationship with the Northern
342 Hemisphere mean surface temperature in CMIP5 simulations. *J. Geophys. Res.-Atmos.*,
343 *122*(8), 4202-4227. doi:10.1002/2016jd025979
- 344 Wang, Y. H., Magnusdottir, G., Stern, H., Tian, X., & Yu, Y. (2012). Decadal variability of the
345 NAO: Introducing an augmented NAO index. *Geophys. Res. Lett.*, *39*.
346 doi:10.1029/2012gl053413
- 347 Woollings, T., Barnes, E., Hoskins, B., Kwon, Y.-O., Lee, R. W., Li, C., et al. (2018). Daily to
348 decadal modulation of jet variability. *J. Clim.*, *31*, 1297-1314. doi:10.1175/jcli-d-17-
349 0286.1
- 350 Woollings, T., Czuchnicki, C., & Franzke, C. (2014). Twentieth century North Atlantic jet
351 variability. *Q. J. Roy. Meteorol. Soc.*, *140*(680), 783-791. doi:10.1002/qj.2197
- 352 Woollings, T., Franzke, C., Hodson, D. L. R., Dong, B., Barnes, E. A., Raible, C. C., et al.
353 (2015). Contrasting interannual and multidecadal NAO variability. *Clim. Dyn.*, *45*(1-2),
354 539-556. doi:10.1007/s00382-014-2237-y
- 355 Zhang, L. P., & Wang, C. Z. (2013). Multidecadal North Atlantic sea surface temperature and
356 Atlantic meridional overturning circulation variability in CMIP5 historical simulations. *J.*
357 *Geophys. Res. - Oceans*, *118*(10), 5772-5791. doi:10.1002/jgrc.20390

358

359

360 **Figure Captions**

361

362 **Figure 1.** Reanalysis (20CRv2c) climatological (1861-2014) winter-mean U850 (line countours)
 363 and regression slopes (color fill) of time series of winter-mean U850 at each grid point on **(a)**
 364 NAO point index, **(b)** jet latitude and **(c)** jet strength. For the U850 climatologies, solid (dashed)
 365 contour lines denote positive (negative) values with levels -9, -6, -3, 0, 3, 6, 9 m s^{-1} . All time
 366 series were linearly de-trended prior to calculation of regression slopes.
 367

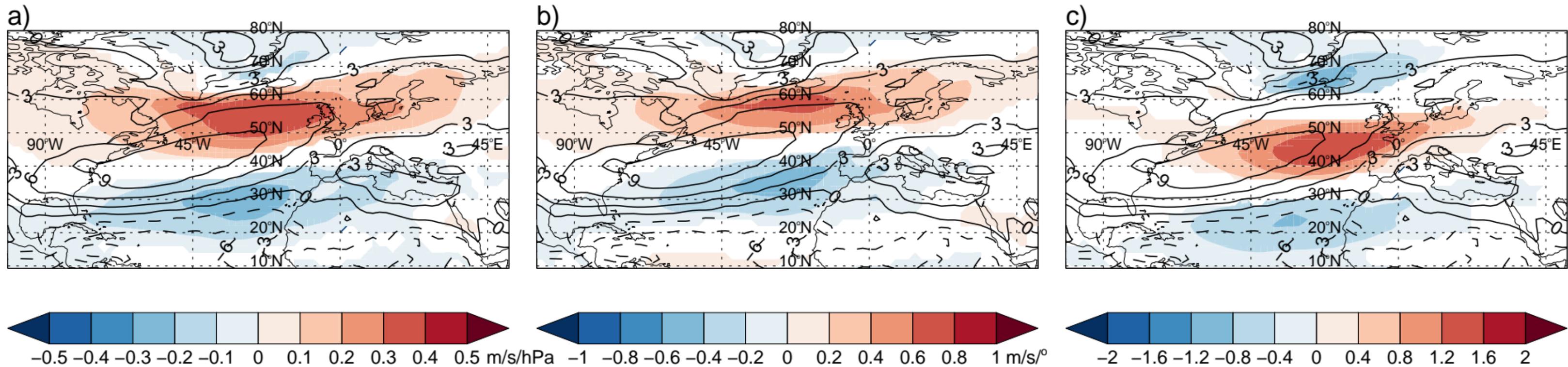
368 **Figure 2.** Panels **a-d** show smoothed time series of winter NAO point index, NAO EOF-based
 369 index, jet latitude and jet strength anomalies respectively, with CMIP5 ensemble members
 370 shown by decadal-smoothed thin grey lines and the 20CRv2c ensemble mean shown by the
 371 thick black lines (dashed shows decadal-smoothed time series and dotted shows 30-year
 372 smoothing). The solid segments of the decadal-smoothed 20CRv2c time series highlight
 373 periods containing the maximum 30-year linear trend. All available years are shown from each
 374 time series, hence the differing start and end dates in some cases. Panels **e-h** show histograms of
 375 maximum 30-year linear trends of the smoothed and de-trended CMIP5 time series shown in
 376 panels **a-d**. Vertical black lines show maximum 30-year linear trends from each of the 20CRv2c
 377 ensemble members (i.e. trends during the periods highlighted by solid black lines in **a-d**).
 378

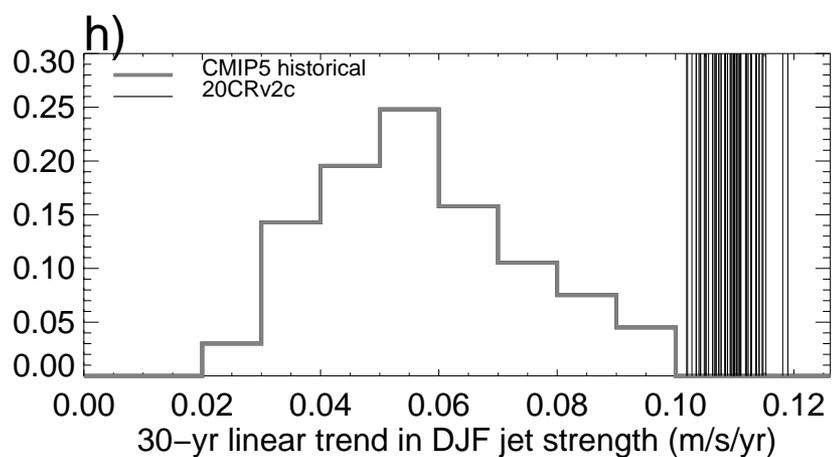
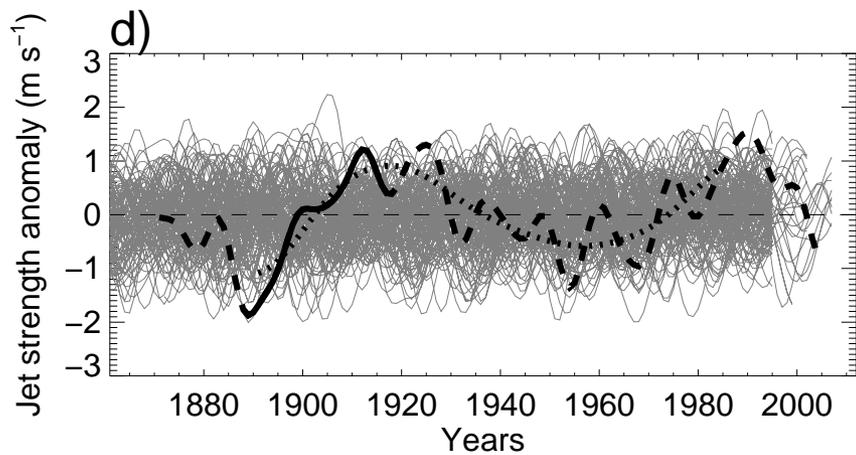
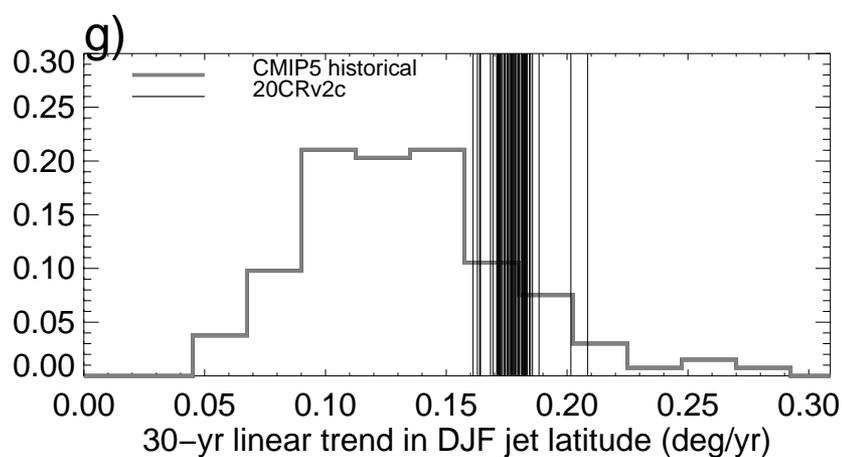
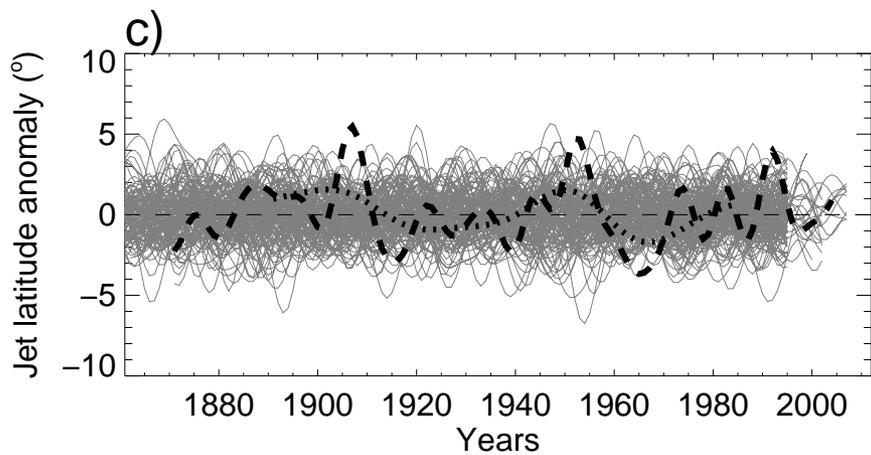
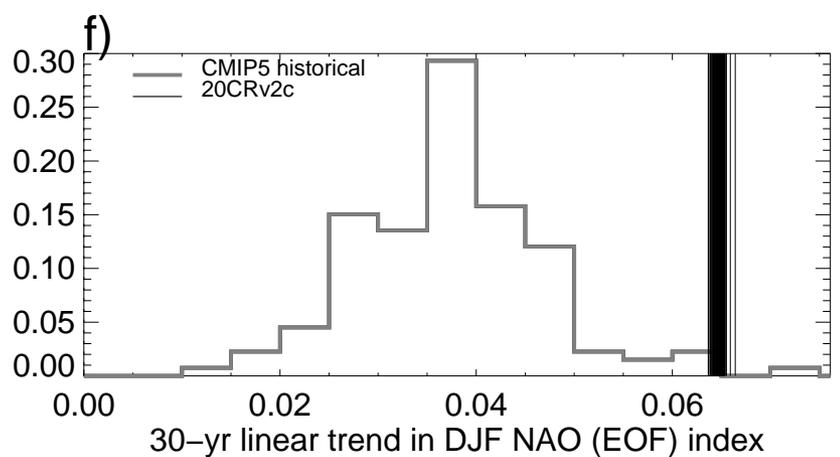
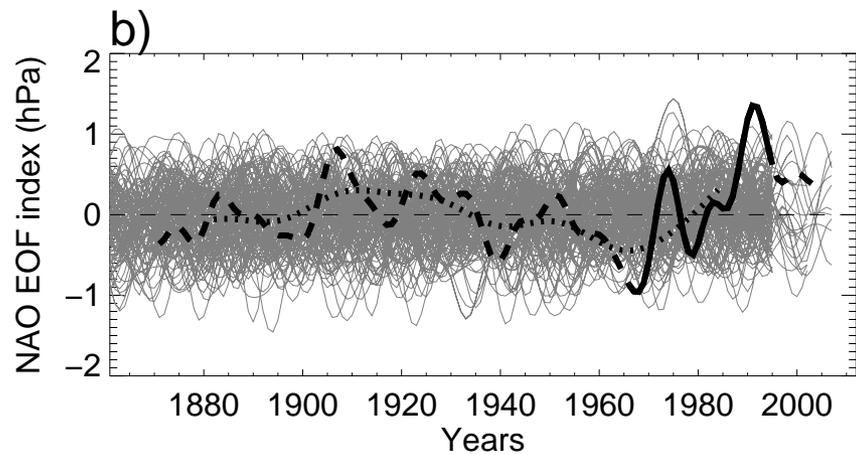
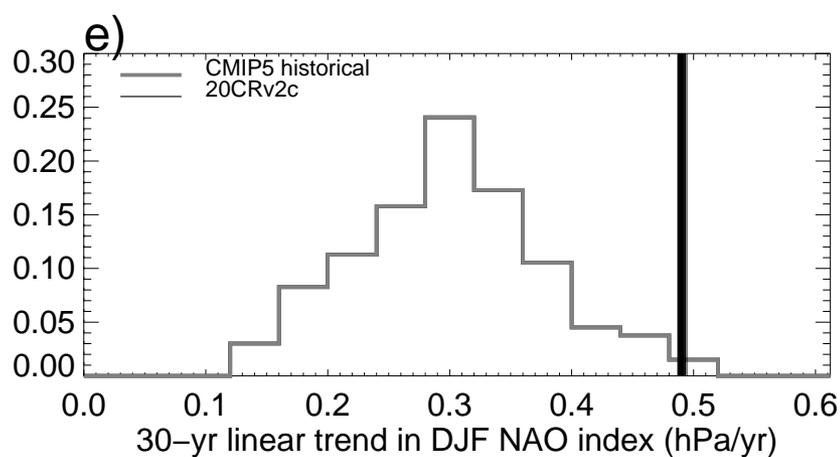
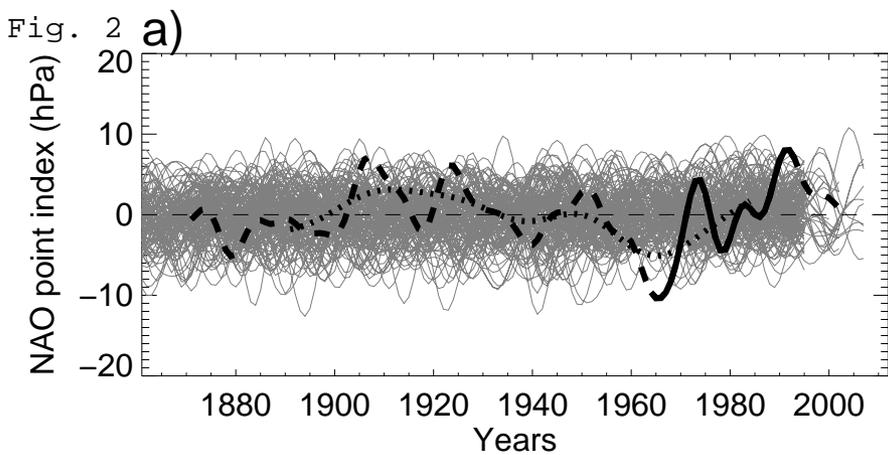
379 **Figure 3.** Proportional power spectral density (i.e. divided by the mean across all frequency
 380 components) of **(a)** NAO point index, **(b)** NAO EOF index, **(c)** jet latitude and **(d)** jet strength.
 381 The grey lines denote individual CMIP5 realizations and the black lines show 20CRv2c
 382 ensemble members. Each time series was linearly de-trended prior to the spectral analysis and
 383 the time range was restricted to the common period of 1862-2005 for comparison of the same
 384 frequency bins. The dashed line indicates the $p = 0.05$ threshold for the significance of a peak in
 385 any particular frequency bin using Fisher's g -statistic. In **(d)**, the mean p -value across the 56
 386 20CRv2c ensemble members is 0.038 in the 72-year period bin.

387 **Figure 4.** Low-frequency proportional power (mean over periods of between 40 to 100 years of
 388 results shown in Figure 3) for **(a)** NAO point index, **(b)** NAO EOF index, **(c)** jet latitude and **(d)**
 389 jet strength. The grey horizontal lines show the values calculated from 20CRv2c ensemble
 390 members. Each asterisk represents one historical CMIP5 realization arranged into columns for
 391 each model (see supporting Table S1 for model names).

392 **Figure 5.** Correlation coefficients of 30-year smoothed time series of NAO point index vs **(a)** jet
 393 latitude and **(b)** jet strength. As in Figure 4 each CMIP5 historical realization is indicated by an
 394 asterisk and the 20CRv2c reanalysis ensemble members by the horizontal solid grey lines. See
 395 Figure S6 for a version based on the EOF-based NAO index.
 396
 397

Fig. 1





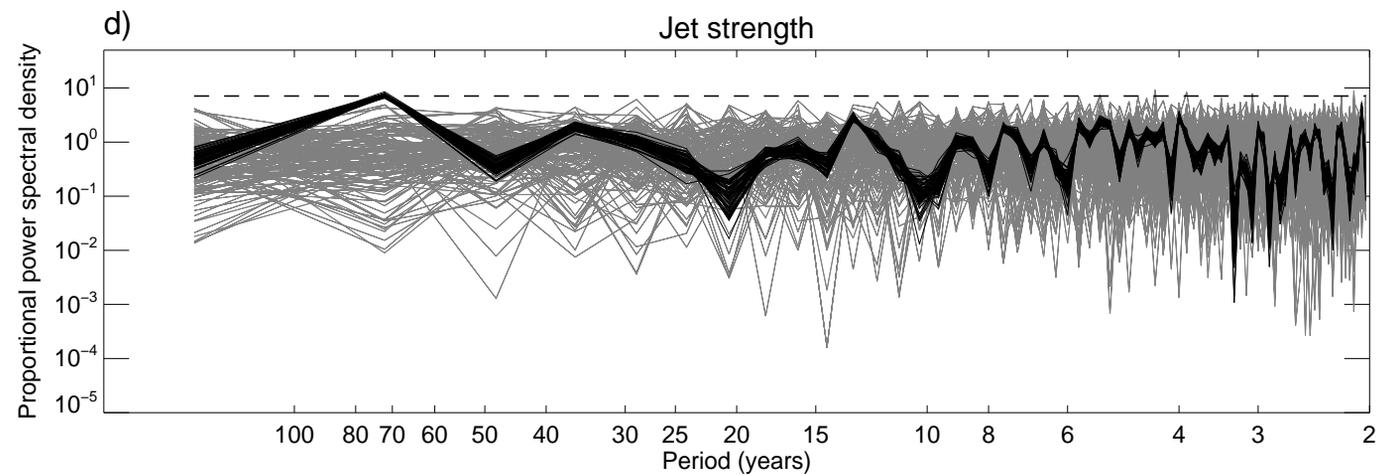
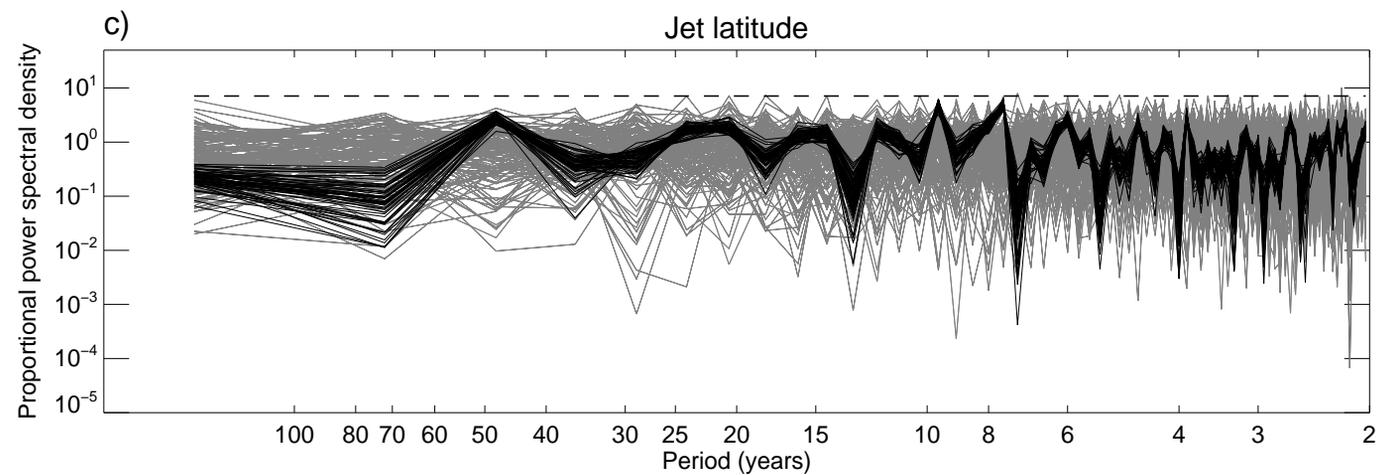
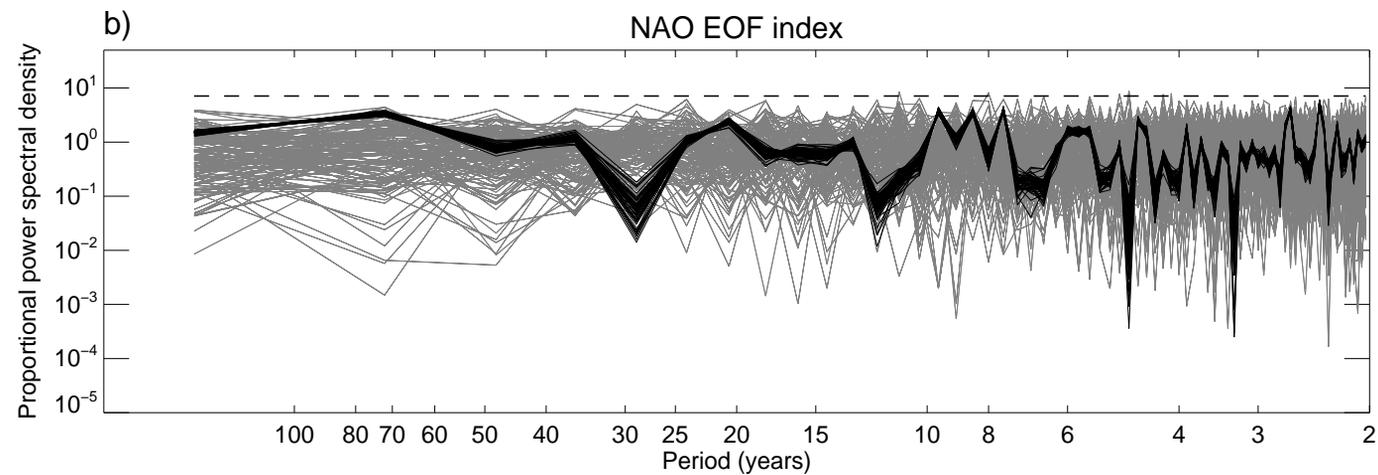
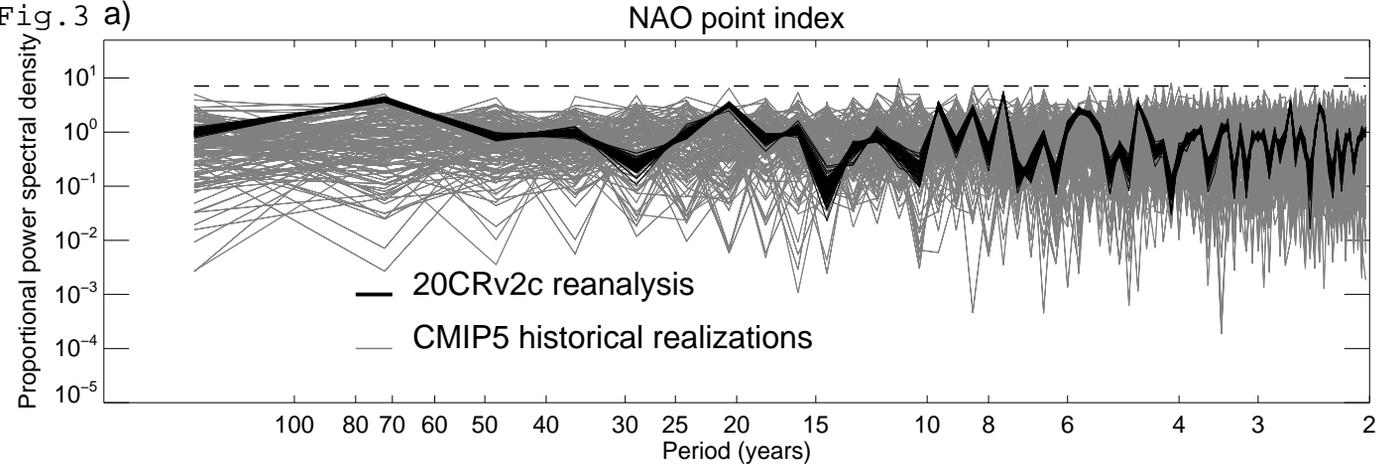


Fig. 4

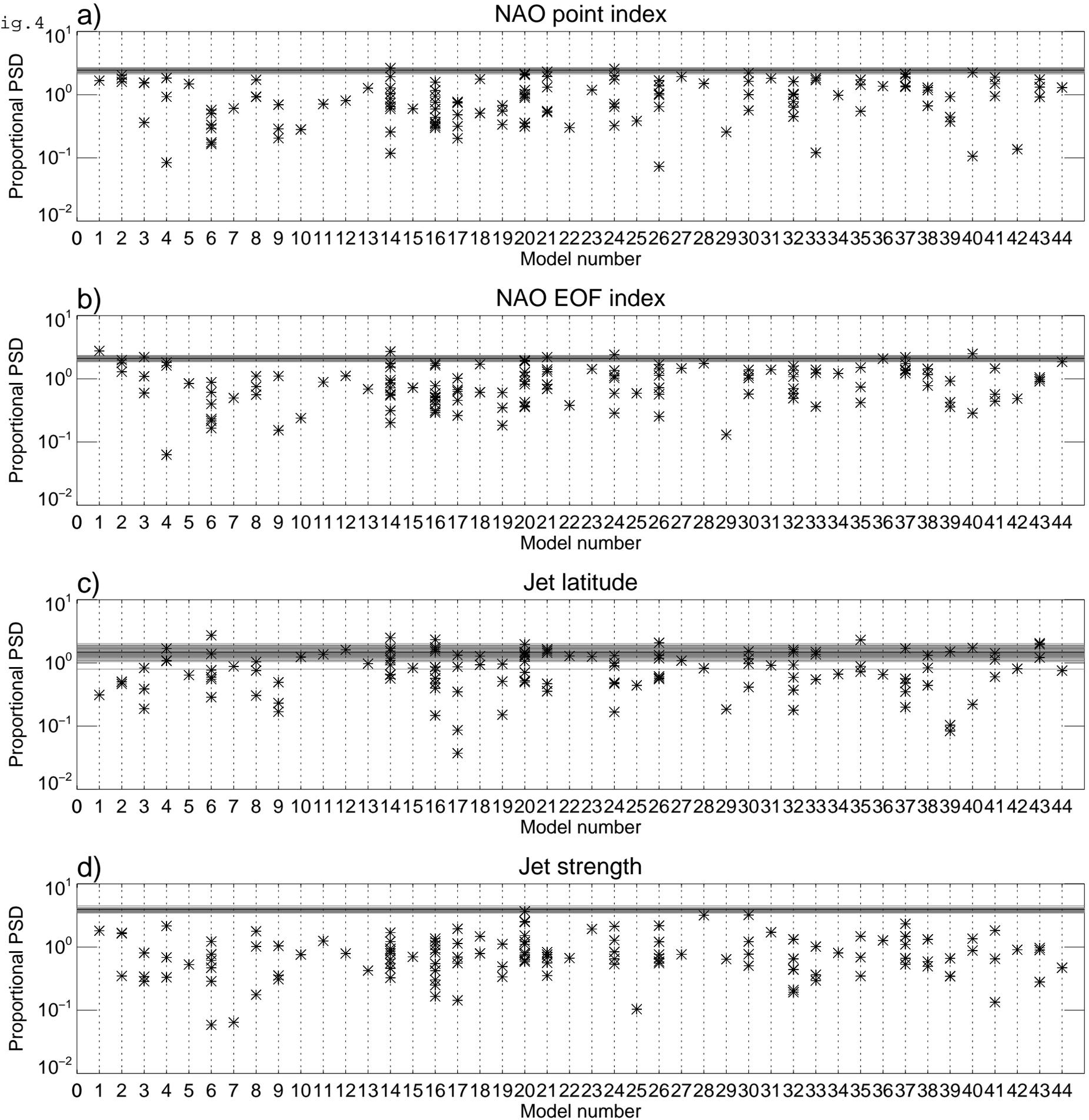


Fig. 5

