

Final warming of the Southern Hemisphere polar vortex in high- and low-top CMIP5 models

Article

Accepted Version

Author final version

Wilcox, L. ORCID: <https://orcid.org/0000-0001-5691-1493> and Charlton-Perez, A. J. ORCID: <https://orcid.org/0000-0001-8179-6220> (2013) Final warming of the Southern Hemisphere polar vortex in high- and low-top CMIP5 models. *Journal of Geophysical Research - Atmospheres*, 118 (6). pp. 2535-2546. ISSN 0148-0227 doi: 10.1002/jgrd.50254 Available at <https://centaur.reading.ac.uk/30939/>

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

Published version at: <http://onlinelibrary.wiley.com/doi/10.1002/jgrd.50254/abstract>
To link to this article DOI: <http://dx.doi.org/10.1002/jgrd.50254>

Publisher: American Geophysical Union

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

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

¹ Final warming of the Southern Hemisphere polar ² vortex in high- and low-top CMIP5 models

3 L. J. Wilcox and A. J. Charlton-Perez

4 Department of Meteorology, University of Reading, Earley Gate,
5 PO Box 243, Reading, RG6 6BB, UK

Abstract

The final warming date of the polar vortex is a key component of Southern Hemisphere stratospheric and tropospheric variability in spring and summer. We examine the effect of external forcings on Southern Hemisphere final warming date, and the sensitivity of any projected changes to model representation of the stratosphere. Final warming date is calculated using a temperature-based diagnostic for ensembles of high- and low-top CMIP5 models, under the CMIP5 historical, RCP4.5, and RCP8.5 forcing scenarios. The final warming date in the models is generally too late in comparison with those from reanalyses: around two weeks too late in the low-top ensemble, and around one week too late in the high-top ensemble. Ensemble Empirical Mode Decomposition (EEMD) is used to analyse past and future change in final warming date. Both the low- and high-top ensemble show characteristic behaviour expected in response to changes in greenhouse gas and stratospheric ozone concentrations. In both ensembles, under both scenarios, an increase in final warming date is seen between 1850 and 2100, with the latest dates occurring in the early twenty-first century, associated with the minimum in stratospheric ozone concentrations in this period. However, this response is more pronounced in the high-top ensemble. The high-top models show a delay in final warming date in RCP8.5 that is not produced by the low-top models, which are shown to be less responsive to greenhouse gas forcing. This suggests that it may be necessary to use stratosphere resolving models to accurately predict Southern Hemisphere surface climate change.

30 1 Introduction

31 The Southern Hemisphere (SH) stratosphere and troposphere have been shown
32 to be coupled, with wave driving from the upward propagation of tropospheric
33 Rossby waves influencing the stratospheric zonal wind, and anomalies in the
34 stratospheric polar vortex having an impact down to the surface. This coupling
35 predominantly occurs in the late spring, or summer, when the final warming
36 of the polar vortex strongly influences both the stratospheric and tropospheric

37 circulation (Black et al., 2006), resulting in the stratospheric and tropospheric
38 annular mode having its largest variance in this season (Baldwin et al., 2003).
39 Changes in the strength of the polar vortex are associated with persistent circulation
40 anomalies in the lower stratosphere, with weaker flow resulting in negative
41 Southern Annular Mode (SAM) anomalies. Thompson et al. (2005) showed that
42 final warming events are also associated with tropospheric circulation anomalies
43 of the same sign, which can persist for in excess of two months. They found that
44 significant increases in tropospheric geopotential height over the pole and de-
45 creases in the midlatitudes, with a similar structure to the negative phase of the
46 SAM, followed major weakenings in the SH polar vortex. Coherent changes in
47 Antarctic surface temperature, with positive temperature anomalies over much
48 of the continent outside the Peninsula region, were also identified in association
49 with these changes.

50 Climate forcings have been shown to change the final warming date of the SH
51 polar vortex. In recent years, changes have been found to be strongly determined
52 by decreases in stratospheric ozone concentrations, with final warming dates
53 observed to be later in the 1990s compared to the 1980s (Waugh et al. (1999);
54 Zhou et al. (2000); Karpetchko et al. (2005); Langematz and Kunze (2006);
55 Haigh and Roscoe (2009)). Ozone depletion causes local cooling over the pole,
56 resulting in an increased temperature gradient and a stronger vortex, and hence,
57 later final warming dates.

58 Several studies have suggested that, in SH spring, the effects on surface
59 climate of ozone recovery and increasing greenhouse gases will be equal and
60 opposite, leading to a near cancellation, or even a reversal, in current trends in
61 the early twenty-first century (Arblaster et al. (2011); McLandress et al. (2011);
62 Polvani et al. (2011); Thompson et al. (2011); Wilcox et al. (2012)). Ozone
63 depletion causes a larger local decrease in temperature compared to greenhouse
64 gas increases, and has been shown to be the primary driver of recent changes
65 in final warming date (Langematz and Kunze, 2006). It is expected that ozone
66 recovery will similarly be the primary driver of near-term changes in final warm-
67 ing date, and that the vortex breakdown will become earlier. A return to later
68 dates towards the end of the twenty-first century is possible as lower strato-
69 spheric temperature trends become dominated by well-mixed greenhouse gas
70 forcing, which has been shown to result in an increased temperature gradient
71 near 100 hPa (Shindell et al. (1998); Wilcox et al. (2012)). If these changes are
72 coupled to the surface then changes in springtime Antarctic surface temperature
73 trends would be likely to occur in conjunction with these changes in the vor-
74 tex. Therefore, one important facet of the stratospheric impact on tropospheric
75 climate is how external forcings may change the final warming date.

76 The significant tropospheric circulation anomalies associated with final warm-
77 ing events demonstrate that changes in the timing of this phenomenon will play
78 a key role in future SH tropospheric circulation change (Black and McDaniel,
79 2007). Hence, understanding potential changes in final warming date, and their
80 drivers, is an important part of SH climate prediction. Several studies have
81 shown that the final warming signature in the SH propagates downwards (e.g.
82 Baldwin et al. (2003); Thompson et al. (2005)). Hardiman et al. (2010) recently

83 showed that this propagation begins at 1 hPa. As such, the representation of
84 changes in final warming date may be sensitive to the position of the model
85 top, which is often located near or below 1 hPa in models. Here, we attempt
86 to quantify the effect of external forcings on SH final warming date, and the
87 sensitivity of any projected changes to the position of the model top.

88 2 Data and Methods

89 The aim of this study is to identify robust changes in SH final warming date,
90 their drivers, and their potential sensitivity to the position of the model top.
91 The fifth Coupled Model Intercomparison Project (CMIP5) provides a unique
92 opportunity to analyse the response of a large number of models to the same
93 future greenhouse gas scenarios. CMIP5 also includes a substantial number of
94 'high-top' models, which have an explicit representation of the stratosphere.
95 High-top models have been defined here as those with model tops at pressures
96 ≤ 1 hPa, or altitudes ≥ 45 km. In addition to having a higher model top, the
97 high-top models used in this study typically have higher vertical resolution in
98 the stratosphere, and a larger proportion of model levels above 200 hPa (54%
99 of high-top model levels are in the stratosphere, compared to 36% for low-
100 top models). The models used in this study, their classification, and vertical
101 distribution of levels, are shown in Table 1. Only one model from each model
102 family is included in each classification to avoid biasing the ensemble mean.

103 We examine monthly mean data from the historical (1850-2005), Representative
104 Concentration Pathway (RCP) 4.5 (Thomson et al., 2011), and RCP8.5
105 (both 2006 to 2100) (Riahi et al., 2011) integrations. The two future pathways
106 result in a radiative forcing of 4.5 Wm^{-2} and 8.5 Wm^{-2} respectively by 2100,
107 with RCP4.5 carbon dioxide emissions peaking around 2040, and RCP8.5 emis-
108 sions peaking in 2100. The rate of change of greenhouse gas concentrations
109 stabilises by ~ 2070 in RCP4.5, and continues to increase throughout RCP8.5
110 (Figure 1(a)). The time series analysed in this paper are concatenations of the
111 historical and RCP experiments for consistent ensemble members of each model,
112 and are referred to throughout by the name of the relevant future pathway.

113 Although a recommended ozone time series was compiled for CMIP5 (Cionni
114 et al., 2011), only three of the models used in this study are forced with these
115 data. Others included modified versions of the Cionni et al. (2011) data, some
116 prescribed ozone concentrations from different data sets, and others treat ozone
117 interactively. The different representation of ozone in the subset of CMIP5
118 models used in this study is shown in Table 1, following the categorisation of
119 Eyring et al. (2012). Example time series of the September to November mean
120 75° - 90° S mean concentration at 50 hPa for each prescribed category are shown
121 in Figure 1(b), alongside the time series from models with interactive ozone.
122 Comparison of the different categories reveals a range of Antarctic stratospheric
123 ozone concentrations, with 1900 values between 2.4 ppmv and 4 ppmv. There is
124 some spread in the rate of recovery in the twenty-first century. Ozone concen-
125 trations tend to recover faster in the time series from models with interactive

126 ozone. The relative change in ozone concentrations prior to 2000 is similar in the
127 interactive and Cionni timeseries, but smaller in the other prescribed categories.
128 However, the turning points are comparable across the categories (Figure 1(b)).
129 The aim of this study is to identify the drivers of robust projections in SH final
130 warming date, which will depend on the forcings, and the response to them, hav-
131 ing the same characteristics across the model ensemble. As the turning points
132 in the ozone timeseries are comparable, it is anticipated that the qualitative
133 response of the final warming date to ozone will have similar characteristics
134 across the models. Hence, the quantitative differences in the ozone forcing are
135 not anticipated to influence our result.

136 To date, different numbers of ensemble members have been provided for each
137 of the CMIP5 models. Where multi-model means have been used, they include
138 only one ensemble member for each model to avoid biasing the mean towards
139 models with a larger number of ensemble members.

140 ERA-Interim (Dee et al., 2011) and the NCEP Climate Forecast System
141 Reanalysis (CFSR) (Saha et al., 2010) were used to assess biases in the model
142 data.

143 2.1 Final warming diagnostic

144 The definition of vortex breakdown is subjective, and several approaches have
145 been used in earlier studies. These include potential vorticity-based spatial
146 diagnostics (Waugh and Randel (1999), Waugh et al. (1999), Karpetchko et al.
147 (2005), Zhou et al. (2000)), diagnostics based on wind thresholds (Black and
148 McDaniel, 2007), and temperature based diagnostics (Haigh and Roscoe, 2009).
149 However, regardless of the definition used, there is a consensus that the final
150 warming date (FWD) of the SH vortex was later in the 1990s compared to
151 the 1980s. Potential vorticity is not a standard CMIP5 output, and the coarse
152 vertical resolution of the archived data makes it difficult reliably to calculate
153 potential vorticity. Therefore, only temperature (Haigh and Roscoe, 2009) and
154 wind (Black and McDaniel, 2007) based diagnostics of the FWD have been
155 considered.

156 Black and McDaniel (2007) defined the FWD as the final time that the zonal-
157 mean zonal-wind at 60°S and 50 hPa drops below 10 ms^{-1} until the following
158 autumn. They apply the diagnostic to 5-day running averages of daily data.

159 Haigh and Roscoe (2009) define the FWD as the minimum in the second
160 time derivative of polar cap mean (90-60°S) temperature at 50 hPa. They use
161 3-day averages of daily and bi-daily data, smoothed with a 21-day triangular
162 filter. However, they found that interpolation of monthly mean data gave similar
163 fields to smoothed daily data. Here, monthly mean data is used as, at this
164 early stage in CMIP5, it facilitates the analysis of a larger number of models.
165 The sum of the first five Fourier components of the temperature time series
166 is used to produce interpolated daily data. Due to the smooth nature of the
167 evolution of the seasonal cycle in polar cap mean temperature, only negligible
168 differences were identified between FWDs calculated using this method, and
169 those calculated using daily data (see Figure 2).

170 The FWD calculated using the Haigh and Roscoe method is typically a week
171 earlier than that calculated using the Black and McDaniel diagnostic. However,
172 there is little qualitative difference between the diagnostics (Figure 2): the time
173 series are strongly correlated, with $r=0.95$ for 1950-2005 for CNRM-CM5 data.
174 The use of the Black and McDaniel (2007) threshold-based diagnostic may be
175 problematic if there are significant variations in the background state between
176 models, or under strong forcing. In some models, the use of the 10 ms^{-1} thresh-
177 old results in non-identification of a FWD for some years in the historical period.
178 As scenarios with large forcing will be considered, the Haigh and Roscoe diag-
179 nstic, from monthly mean data, will be used for the remainder of this work, in
180 order to avoid excessive non-identification of FWDS.

181 2.2 Empirical mode decomposition

182 Climate data is often non-linear and non-stationary. Deviations from monotonic
183 change are particularly apparent in the Southern Hemisphere where change is
184 governed by the competing effects of increased greenhouse gases and strato-
185 spheric ozone. Changes in FWD have been established as being strongly ozone
186 driven (Zhou et al. (2000), Karpetchko et al. (2005), Haigh and Roscoe (2009)),
187 and a better fit is found between FWD and stratospheric ozone concentrations
188 than can be achieved with linear trends for example (Haigh and Roscoe, 2009).

189 To avoid fitting extrinsic functions, which may not correspond well to the
190 non-linearity embedded in the data, or forcing data time series, which may
191 only account for changes via one of many mechanisms, Empirical Mode De-
192 composition (EMD) has been used to analyse variability in FWD. EMD is an
193 intrinsic, adaptive method for deriving the variability of a time series on vari-
194 ous timescales. EMD has successfully been applied to climate data in several
195 previous studies (e.g. Lee and Ouarda (2011), Franzke (2009), Huang and Wu
196 (2008), Wu et al. (2007), McDonald et al. (2007), and Duffy (2004)). While
197 EMD is a useful tool for analysing variability and trends in non-linear time se-
198 ries, it cannot be used to unambiguously attribute particular characteristics of
199 these trends to a given forcing mechanism. Hence, EMD is used here alongside
200 multiple linear regression analysis.

201 EMD is an algorithm used to decompose a time series into a set of Intrinsic
202 Mode Functions (IMFs), with each describing a given oscillatory mode of the
203 data. IMFs must satisfy two conditions:

- 204 • Must have a local mean of zero
- 205 • Must have a single zero crossing between two extrema

206 IMFs are extracted sequentially from a data series, from the highest frequency
207 to the lowest, until no complete oscillation can be identified. The residual from
208 this process then describes the long-term trend in the data, where the trend is
209 defined as the instantaneous mean of the time series.

210 Unlike Fourier filtering, the phase and amplitude of each IMF are time de-
211 pendent. The number of IMFs extracted from a time series is typically $\ln N$,

212 where N is the number of data points (Wu et al., 2007). There is some evi-
 213 dence of mode mixing (signals of different timescales identified in the same IMF)
 214 amongst the IMFs of FWD from EMD. To avoid this, Ensemble Empirical Mode
 215 Decomposition (EEMD) has been used. EEMD gives an ensemble mean of the
 216 IMFs for the product of FWD and a finite white noise series (Wu and Huang,
 217 2009). The inclusion of a noise series provides a uniformly distributed reference
 218 scale, which preserves the dyadic property of EMD that can fail when data is
 219 intermittent (Wu and Huang, 2009). The noise is cancelled out in the ensemble
 220 mean, so it can be used to facilitate the separation of different timescales, with-
 221 out contributing to the final IMFs. EEMD is performed here with 200 iterations
 222 and white noise with an amplitude of 0.2 times the standard deviation of the
 223 FWD series (following Wu and Huang (2009)).

224 Figure 3(a) shows a time series of FWD from MIROC-ESM-CHEM under
 225 RCP4.5, calculated using the Haigh and Roscoe (2009) method, alongside the
 226 IMFs from EEMD (Figure 3(b)). Most of the high-frequency variability in the
 227 time series, with a period of less than 3 years, is contained in the first two IMFs
 228 (not shown). The local maximum near 2000 is captured in the sixth IMF, and
 229 the increasing trend through the period shown is captured in the residual. The
 230 equivalent result using EMD is shown in Figure 3(d). In this example, it can be
 231 seen that the different frequencies have not been satisfactorily separated. This
 232 is particularly clear in the third IMF (top line of Figure 3(d)), where the period
 233 of the oscillation around the year 2000 is double that in the rest of the IMF.

234 IMFs that can be distinguished from the equivalent IMFs of a noise time
 235 series of the same length are significant, and can be taken to represent physically
 236 meaningful signals. White (Wu and Huang (2004), Wu et al. (2007)) and red
 237 (Franzke (2009)) noise have both been used in previous studies to assign signif-
 238 icance to IMFs from climate data. There is no physical reason why the FWD
 239 in one year would be dependent on the date in another year (Black et al. (2006)
 240 also considered each event as an independent sample). Therefore, a comparison
 241 with a white noise series has been used to determine when an IMF is significant,
 242 following Wu and Huang (2004).

243 A significant difference from a white noise time series is identified through
 244 analysis of the period (T) and energy density (E) of each IMF. Wu and Huang
 245 (2004) show that the probability density function for each IMF of a white noise
 246 time series is well approximated by a normal distribution, and that the prob-
 247 ability distribution of the energy of the n^{th} IMF, NE_n , is a χ^2 distribution,
 248 with $N\bar{E}_n$ degrees of freedom, where \bar{E}_n is the mean of E_n when the number
 249 of data points, N , approaches ∞ . The spread of different confidence intervals
 250 as a function of the mean energy of each IMF can then be determined. Wu and
 251 Huang (2004) define $y = \ln E$ and show that for $|y - \bar{y}| \ll 1$, the distribution of
 252 the energies is Gaussian. The spread lines can then be approximated by

$$y = -x \pm k \sqrt{\frac{2}{N}} e^{x/2} \quad (1)$$

253 where $x = \ln \bar{T}_n$, \bar{T}_n is the mean period, and k is a constant from the percentiles
 254 of the normal distribution. Example energies and periods from 1000 white noise

255 time series of 1000 data points, and the spread lines from the 95% confidence
256 interval, are shown in Figure 3(c). Energy densities from a data time series that
257 lie outside the bounds of the spread lines can be assumed to be significantly
258 different from those expected from a white noise time series, and are therefore
259 expected to contain some information at that confidence level.

260 3 Past and future trends in final warming date

261 Mean FWDs in the individual models are shown in Figure 4 for three periods:
262 1870-1900, 1979-2005, and 2070-2098. In most cases, the FWD is one to two
263 weeks later in 2070-2098 compared to 1870-1900. In the RCP4.5 experiment
264 the delay ranges from a change of 1 day in INMCM4 to 9 days in CanESM2,
265 CSIRO-Mk3.6.0, GISS-ES-R, and NorESM1-M (Figure 4(a)). In RCP8.5 the
266 delay compared to 1870-1900 ranges from 2 days in INMCM4 to 15 days in
267 CanESM2 (Figure 4(b)). With the exception of CNRM-CM5 and GISS-E2-R,
268 all models have later FWDs in 2070-2098 in the RCP8.5 experiment than in
269 RCP4.5. Figure 4(c) compares FWD from 1870-1900 to 2070-2098. There is
270 some evidence of a saturation effect here, with models with a very late historical
271 FWD appearing to show less of a change in the future.

272 Figure 4(d) shows the 1979-2005 mean FWD for each model, compared to
273 ERA-Interim and CFSR. In all models except MIROC5, the FWD is too late
274 compared to the reanalyses, with most models having an FWD that is signifi-
275 cantly later. Such a late bias has been identified in earlier model evaluations,
276 e.g. Butchart et al. (2011). It can also be seen in Figure 4(d) that most models
277 underestimate the inter-annual variability in FWD compared to reanalyses.

278 The late bias in model FWDs is reflected in the high- and low-top ensemble
279 means, shown in Table 2, and in Figure 5 alongside those from ERA-Interim
280 and CFSR. The mean FWDs in the period 1979-2005 are day 312 and day 313
281 in ERA-Interim and CFSR respectively. The low-top mean FWD is around 2
282 weeks late, with a 1979-2005 mean of day 327. The high-top ensemble mean is
283 in better agreement with the reanalysis values, but is still late on average, with
284 a 1979-2005 mean of day 321 (Table 2). For all periods shown in Figure 4, the
285 mean FWD from the low-top ensemble is around a week later than that from
286 the high-top ensemble (Table 2).

287 The FWD from the low- and high-top ensemble is shown in Figure 6 for
288 the historical and RCP4.5 and historical and RCP8.5 experiments. There is
289 more inter-model spread and inter-annual variability in the low-top ensemble,
290 although there is still a considerable amount of inter-annual variability in the
291 FWD from the individual high-top models.

292 A marked delay in FWD can be seen in the high-top ensemble from the
293 late 1970s to the late 1990s (Figure 6). This is associated with the localised,
294 seasonal, cooling that results from ozone depletion in this period. Under RCP4.5
295 this increase in FWD is followed by a steady decrease to 2100, but in RCP8.5
296 a more modest decrease is seen, followed by a small trend towards later FWDs
297 by 2100. The large inter-model spread amongst the low-top models makes such

298 features difficult to distinguish in the low-top ensemble. However, there is some
299 sense of a shift towards later FWDs in the late twentieth century.

300 The large interannual variability and inter-model spread in FWD makes
301 it difficult to compare patterns of behaviour across the models, although the
302 spread in absolute values is important to bear in mind. The FWD in all models
303 is now adjusted to the 1860-1900 mean to assist discussion of the change in
304 FWD across the models. In Figure 7, an 11-year running mean has also been
305 applied, which removes high frequency inter-annual variability, without obscuring
306 decadal variability. The ensemble means shown in Figure 7 are calculated
307 by first finding the ensemble mean of the adjusted raw data, then calculating
308 the 11-year running mean.

309 More similarities can be seen in the behaviour of the low- and high-top
310 models in Figure 7 compared to Figure 6. A clear increase in FWD can now be
311 seen in the low-top ensemble, although the change is not as rapid, large, or as
312 consistent across models, as in the high-top ensemble. A return to earlier FWDs
313 in the twenty-first century can now be seen in the low-top ensemble mean under
314 RCP4.5, although the rate of change is still small compared to that seen in
315 the high-top ensemble. Under RCP8.5, the FWD in the low-top ensemble mean
316 shows very little change in the twenty-first century. In contrast, a clear decrease
317 can be seen in the first half of the twenty-first century in the high-top ensemble,
318 followed by an increase towards the end of the century. The large twenty-first
319 century inter-model spread in the low-top ensemble, even after adjusting to the
320 1860-1900 mean, may obscure some of this behaviour in the low-top ensemble
321 mean. However, there is no convincing evidence of such a pattern in the FWDs
322 from individual models. Such behaviour can be seen in a number of the high-top
323 models.

324 **4 Drivers of past and future trends in final warming 325 date**

326 The primary drivers of changes in FWD are anticipated to be changes in strato-
327 spheric ozone and well-mixed greenhouse gas concentrations. These changes will
328 occur on different timescales, and have different functional forms in the time-
329 series. As such, their signature can be expected to be seen in different IMFs.
330 Increasing greenhouse gases are expected to be linked to a delay in the FWD,
331 while the depletion and recovery of stratospheric ozone will produce a delay fol-
332 lowed by an advance: a signature with a period in the region of 60 years. These
333 responses are likely to be seen in the residual and the last IMF respectively. Fig-
334 ure 1(b) shows that the largest changes in stratospheric ozone concentrations at
335 southern high latitudes occur in the first half of the twenty-first century. Hence,
336 it is anticipated that changing ozone concentrations will be the primary driver
337 of FWD changes here, with greenhouse gases becoming increasingly important
338 in the second half of the century. Figure 1(a) shows that greenhouse gas con-
339 centration changes in RCP4.5 and RCP8.5 are very different in the latter half

340 of the century, with almost no change in concentrations in RCP4.5 and rapid
341 increases in RCP8.5. The potential influence of this difference on FWD was
342 hinted at in Figure 7. It is particularly clear in the comparison of the high-top
343 ensemble means for the two scenarios, where a negative trend from ~ 2070 is
344 seen in RCP4.5 and a positive trend is seen in RCP8.5.

345 The sum of the residual and the last IMF for each model, and the low-
346 and high-top ensemble means, are shown in Figure 8. The ensemble mean is
347 calculated by finding the ensemble mean of the adjusted data, then performing
348 EEMD on this mean. All models and the ensemble means show, with the
349 exception of MIROC5, later FWDs around the turn of the century, under both
350 RCP4.5 and RCP8.5. Patterns of behaviour seen in the ensemble mean are
351 similar to those seen in the running means in Figure 7: an increase then decrease
352 in FWD under RCP4.5; and an increase then decrease then increase in the high-
353 tops under RCP8.5. There is even a suggestion of this RCP8.5 response in the
354 low-top models HadGEM2-ES and CSIRO-Mk3-6-0. However, the amplitude
355 of twenty-first century changes are smaller in the the low-top ensemble than
356 the high-top case. The larger response of high-top models to greenhouse gas
357 forcing towards the end of the twenty-first century is consistent with the larger
358 temperature gradient changes at the tropopause level simulated by these models
359 (Wilcox et al., 2012).

360 Significance testing was carried out to determine which IMFs show patterns
361 significantly different to those that may be identified in a white noise time series.
362 The Wu and Huang (2004) method was used, including their assumption that
363 the energy of the first IMF comes solely from noise and can be used to re-scale
364 the energy density of the other IMFs. Figure 9 shows the sum of significant IMFs
365 (at the 5% level) with periods greater than 50 years (in order to consider only
366 inter-decadal variability) for the low- and high-top ensemble mean (Figure 9 (a)
367 and (b) respectively). The signatures of the high- and low-top significant IMFs
368 follow the patterns seen in the running means, and sums of the last two IMFs: a
369 more pronounced peak at the turn of the century in the high-top ensemble, and
370 a trend towards later FWDs at the end of the twenty-first century in RCP8.5
371 in the high-top ensemble only.

372 The spread function of the 95% and 99% confidence intervals for white noise
373 and energies of the individual IMFs are shown in Figure 10. Here, a significant
374 IMF is identified when it lies outside the inner pair of dotted lines, which indicate
375 the 5th and 95th percentile for white noise. The outer pair of dotted lines indicate
376 the 1st and 99th percentile.

377 Figure 10 shows that the residual is clearly significant for both ensembles
378 and scenarios. For the high-top ensemble, the last IMF is also significant at the
379 1% level for both scenarios. In a reflection of the larger inter-model spread, and
380 the resulting weaker peak in FWD around the turn of the century, the last IMF
381 of the low-top ensemble mean is significant at the 5% level for the historical
382 and RCP4.5 scenario, and not at all for the historical and RCP8.5 scenario
383 (Figure 10(b)). The higher energy of the last IMF in RCP8.5 in the high-top
384 mean compared to the low-top mean is not due only to a differing response to
385 ozone forcing. Analysis of the structure of the IMFs shows that in the high-top

386 RCP8.5 case the last IMF includes some response to GHG forcing, in addition
387 to the anticipated ozone response. The delay in FWD towards the end of the
388 twenty-first century is incorporated in the last IMF as the timing of the trend
389 fits with the \sim 60 year period of the response to stratospheric ozone changes.

390 Multiple linear regression analysis was also performed, regressing FWD against
391 a constant, a timeseries of September to November mean Antarctic mean ozone
392 at 50 hPa, and $\ln(\text{GHG})$, where GHG is represented by the CO_2 equivalent
393 values shown in Figure 1(a). Following Roscoe and Haigh (2007), these indices
394 are normalised to allow direct comparison of the regression coefficients. The re-
395 gression slope, Pearson correlation coefficient, and significance from a two-tailed
396 student's t-test are shown in Table 3 for the high- and low-top ensemble mean
397 for RCP4.5 and RCP8.5. For ensemble mean calculations, ozone was taken from
398 the Cionni et al. (2011) data.

399 In both scenarios, FWD has a stronger relationship with both the GHG index
400 and the ozone index in the high-top ensemble. This can be seen in the larger
401 regression slopes and linear correlations shown in Table 3, and in comparison of
402 the multiple linear correlations: 0.63 (0.64) and 0.76 (0.78) for RCP4.5 (RCP8.5)
403 for the low- and high-top ensemble mean respectively. This is a reflection of
404 the more consistent cross-model behaviour seen in the high-top models (e.g.
405 Figure 8).

406 There is little difference between RCP4.5 and RCP8.5 in the statistics relat-
407 ing to the ozone index (Table 3). The more influential role of GHGs in RCP8.5
408 is reflected in the regression slopes as well as the significance. The larger regres-
409 sion slope, linear correlation, and significance associated with the GHG index
410 in RCP8.5 for the high-top ensemble compared to the low-top is likely to be
411 a reflection of the delay in FWD in the high-top ensemble mean near the end
412 of the twenty-first century in response to GHG forcing, which is not seen in
413 RCP4.5, or the low-top ensemble mean. This echoes the higher energies found
414 in the last IMF and residual of the high-top ensemble mean in RCP8.5.

415 In the illustrations of FWD in CMIP5 models shown in this study, MIROC5
416 has been a clear outlier. The model shows almost no change in FWD from 1860
417 to 2100 (Figure 4, Figure 7) and the structure of the timeseries from the sum
418 of the last IMF and the residual mirrors those from other low-top models. In
419 the high-top ensemble, there are no such striking outliers (Figure 8). However,
420 MIROC-ESM-CHEM shows larger inter-decadal variations in FWD than other
421 models in the group. While the behaviour of FWD in MIROC-ESM-CHEM is
422 not especially unusual in the context of the other models, is it possible that the
423 large changes simulated by MIROC-ESM-CHEM and the very small changes
424 from MIROC5 have enough influence on their respective ensemble means to
425 dominate the differences seen between the high- and low-top ensembles?

426 It was found that removing the MIROC models from the ensembles had no
427 effect on our conclusions from EEMD analysis at the 5% level. As one would
428 expect, there are small changes to the energies of the IMFs as a result of the
429 removal, but the IMFs identified as being significantly different to those expected
430 from white noise are the same, and their structure is qualitatively unchanged.

431 The results of the multiple linear regression analysis without the MIROC

432 models is shown alongside the results for the whole ensemble in Table 3. As
433 expected, removing MIROC5, a model that shows little change in FWD, from
434 the low top ensemble slightly increases the correlation between the FWD and
435 both ozone and $\ln(\text{GHG})$ in both the RCP4.5 and RCP8.5 case, but not to such
436 an extent that the significance level is altered. The removal of MIROC5 results
437 in an increase in the magnitude of the regression slope for the ozone index, and
438 for the GHG index in the RCP8.5 scenario. It also brings the regression slope
439 for the GHG index closer to the anticipated positive value in the RCP4.5 case.

440 MIROC-ESM-CHEM simulates a slightly larger response to stratospheric
441 ozone depletion compared to the rest of the high-top ensemble, but doesn't
442 show a delay in FWD towards the end of the twenty-first century. Thus, it is
443 anticipated that the removal of the model from the high-top ensemble will result
444 in a decrease in the magnitude of the regression slope of the ozone index and
445 correlation, and an increase in the regression slope and correlation for the GHG
446 index. Such changes can be seen in both the RCP4.5 and RCP8.5 case (Table 3).
447 These changes are marked enough to decrease the significance of the relationship
448 between stratospheric ozone and RCP8.5 FWD, and of the relationship between
449 RCP4.5 FWD and GHG.

450 As one would expect, removing the MIROC models from the analysis does
451 change the statistics. However, the conclusions drawn from the analysis are un-
452 changed. The importance of stratospheric ozone changes as a driver of changes
453 in FWD is consistent across both scenarios, with a unit change in ozone concen-
454 tration having more influence on the high-top ensemble mean than the low-top
455 ensemble mean. GHG changes play more of a role in RCP8.5 than RCP4.5, and,
456 as for ozone changes, result in a larger change in FWD in the high-top ensemble
457 mean than the low-top mean. The larger values of the regression coefficients in
458 the high-top case reflect the higher energies of the residual and last IMF seen
459 in Figure 10, and the more consistent behaviour of the models seen in Figure 8.

460 5 Conclusions

461 Changes in final warming date are known to drive persistent tropospheric anom-
462 alies with a similar structure to the southern annular mode (Thompson et al.
463 (2005), Black et al. (2006)). Such changes are sensitive to external forcing
464 from greenhouse gases and, in particular, stratospheric ozone. This results in
465 pronounced changes in Southern Hemisphere final warming date, with a peak
466 around the year 2000, which can be expected to influence spring and summer-
467 time trends in high-latitude surface climate.

468 The Southern Hemisphere final warming date is around one week too late
469 in CMIP5 high-top models, and two weeks too late in low-top models com-
470 pared to ERA-Interim and the Climate Forecast System Reanalysis (1979-2005).
471 The high-top models show more consistent absolute values and changes in final
472 warming date in both the historical and future periods than low-top models

473 After adjustment to the 1860-1900 mean, similar behaviour can be seen in
474 both the high- and low-top ensembles. A shift to later final warming dates

475 is seen in the historical period as a response to stratospheric ozone depletion,
476 and a return to earlier final warming dates occurs as ozone recovers. In the
477 high-top ensemble, there is also a shift towards later final warming dates in the
478 latter half of the twenty-first century in RCP8.5, which is consistent with the
479 larger meridional temperature gradient identified in high-top models by Wilcox
480 et al. (2012). The high-top models show a more consistent pattern of change,
481 and larger changes, in response to forcing compared to the low-top models.
482 This difference is apparent in both the comparison of significant IMFs, and the
483 coefficients from multiple linear regression.

484 Further investigations with larger ensembles of high- and low-top models,
485 with consistent ozone concentrations, are required. Simpson et al. (2011) showed
486 that the late bias in final warming date contributes to too-persistent southern
487 annular mode anomalies in summer, and may cause models to respond too
488 strongly to anthropogenic forcing in this season. Hence, the difference between
489 the high- and low-top ensemble mean results, the large spread in the low-top
490 ensemble, and the more pronounced late bias in final warming date in the low-
491 top ensemble, suggest that high-top models are likely to be required to produce
492 accurate projections of future Southern Hemisphere surface climate.

493 6 Acknowledgments

494 This work was funded by the National Centre for Atmospheric Science (NCAS)-
495 Climate via a CMIP5 grant.

496 We acknowledge the World Climate Research Programme's Working Group
497 on Coupled Modelling, which is responsible for CMIP, and we thank the cli-
498 mate modelling groups for producing and making available the model output
499 listed in Table 1. For CMIP the U.S. Department of Energy's Program for
500 Climate Model Diagnosis and Intercomparison provides coordinating support
501 and led development of software infrastructure in partnership with the Global
502 Organization for Earth System Science Portals.

503 We thank the British Atmospheric Data Centre (BADC) for providing access
504 to their CMIP5 data archive, and acknowledge the use of ERA data made
505 available by the BADC and NCAS-Climate. NCEP/NCAR data was provided
506 by NOAA/OAR/ESRL PSD, and CFSR data was provided by NCAR, via the
507 Research Data Archive (RDA).

508 We also thank three anonymous reviewers for their helpful suggestions.

509 **References**

- 510 Arblaster, J. M., G. A. Meehl and D. J. Karoly, 2011: Future climate change in
511 the Southern Hemisphere: Competing effects of ozone and greenhouse gases.
512 *Geophys. Res. Lett.*, **38**, L02701, doi:10.1029/2010GL045384.
- 513 Baldwin, M. P., D. B. Stephenson, D. W. J. Thompson, T. J. Dunkerton, A. J.
514 Charlton, and A. O'Neill, 2003: Stratospheric memory and skill of extended-
515 range weather forecasts. *Science*, **301**, 636–640.
- 516 Black, R. X. and B. A. McDaniel, 2007: Interannual variability in the Southern
517 Hemisphere circulation organized by stratospheric final warming events. *J.
518 Atmos. Sci.*, **64**, 2968–2974.
- 519 Black, R. X., B. A. McDaniel, and W. A. Robinson, 2006: Stratosphere-
520 troposphere coupling during spring onset. *J. Clim.*, **19**, 4891–4901.
- 521 Butchart, N., et al., 2011: Multimodel climate and variability of the strato-
522 sphere. *J. Geophys. Res.*, **116**, D05102.
- 523 Cionni, I., et al., 2011: Ozone database in support of CMIP5 simulations: results
524 and corresponding radiative forcing. *Atmos. Chem. Phys. Discuss.*, **11** (4),
525 10 875–10 933.
- 526 Dee, D. P., et al., 2011: The ERA-Interim reanalysis: configuration and per-
527 formance of the data assimilation system. *Quarterly Journal of the Royal
528 Meteorological Society*, **137** (656), 553–597, doi:10.1002/qj.828.
- 529 Duffy, D. G., 2004: The application of Hilbert-Huang transforms to meteoro-
530 logical datasets. *J. Atmos. Oceanic Technol.*, **21**, 599–611.
- 531 Eyring, V., et al., 2012: Long-term changes in tropospheric and stratospheric
532 ozone and associated impacts in CMIP5 simulations. *J. Geophys. Res.*, sub-
533 mitted.
- 534 Franzke, C., 2009: Multi-scale analysis of teleconnection indices: climate noise
535 and nonlinear trend analysis. *Nonlin. Process Geophys.*, **16**, 65–76.
- 536 Haigh, J. D. and H. K. Roscoe, 2009: The final warming date of the Antarctic
537 polar vortex and influences on its interannual variability. *J. Clim.*, **22**, 5809–
538 5819.
- 539 Hardiman, S. C., N. Butchart, S. M. Osprey, L. J. Gray, A. C. Bushell, and
540 T. J. Hinton, 2010: The climatology of the middle atmosphere in a vertically
541 extended version of the Met Office's climate model. part I: Mean state. *J.
542 Atmos. Sci.*, **67**, 1509–1525.
- 543 Huang, N. E. and Z. Wu, 2008: A review on Hilbert-Huang transform: Method
544 and its applications to geophysical studies. *Rev. Geophys.*, **46**, RG2006.

- 545 Karpetchko, A., E. Kyrö, and B. M. Knudsen, 2005: Arctic and Antarctic polar
546 vortices 1957-2002 as seen from the ERA-40 reanalyses. *J. Geophys. Res.*,
547 **110**, D21109.
- 548 Kawase, H., T. Nagashima, K. Sudo, and T. Nozawa, 2011. Future changes in
549 tropospheric ozone under Representative Concentration Pathways (RCPs).
550 *Geophys. Res. Lett.*, **38**, L05801.
- 551 Lamarque, J. F., et al., 2010: Historical (1850-2000) gridded anthropogenic and
552 biomass burning emissions of reactive gases and aerosols: methodology and
553 application. *Atmos. Chem. Phys.*, **10** (15), 7017–7039.
- 554 Lamarque, J. F., G. P. Kyle, M. Meinshausen, K. Riahi, S. J. Smith, D. P.
555 van Vuuren, A. J. Conley, and F. Vitt, 2011: Global and regional evalua-
556 tion of short-lived radiatively-active gases and aerosols in the Representative
557 Concentration Pathways. *Climatic Change*, **109** (1-2), 191–212.
- 558 Langematz, U. and M. Kunze, 2006: An update on dynamical changes in the
559 Arctic and Antarctic stratospheric polar vortices. *Climate Dynamics*, **27** (6),
560 647–660.
- 561 Lee, T. and T. B. M. J. Ouarda, 2011: Prediction of climate nonstationary
562 oscillation processes with empirical mode decomposition. *J. Geophys. Res.*,
563 **116**, D06107.
- 564 McDonald, A. J., A. J. G. Baumgaertner, G. J. Fraser, S. E. George, and
565 S. Marsh, 2007: Empirical mode decomposition of the atmospheric wave field.
566 *Ann. Geophys.*, **25**, 375–384.
- 567 McLandress, C., T. G. Shepherd, J. F. Scinocca, D. A. Plummer, M. Sigmond,
568 A. I. Jonsson, M. C. Reader, 2011: Separating the Dynamical Effects of
569 Climate Change and Ozone Depletion. Part II: Southern Hemisphere Tropo-
570 sphere. *J. Climate*, **24**, 1850–1868.
- 571 Polvani, L. M., M. Previdi and C. Deser, 2011: Large cancellation, due to
572 ozone recovery, of future Southern Hemisphere atmospheric circulation trends.
573 *Geophys. Res. Lett.*, **38**, L04707, doi:10.1029/2011GL046712.
- 574 Riahi, K., et al., 2011: RCP 8.5—a scenario of comparatively high greenhouse
575 gas emissions. *Climatic Change*, **109**, 33–57, 10.1007/s10584-011-0149-y.
- 576 Roscoe, H. K., and J. D. Haigh, 2007: Influences of ozone depletion, the solar
577 cycle and the QBO on the Southern Annular Mode. *Q. J. R. Meteorol. Soc.*,
578 **133**, 1855–1864.
- 579 Saha, S., et al., 2010: The NCEP Climate Forecast System Reanalysis. *Bulletin*
580 *of the American Meteorological Society*, **91**, 1015–1057.
- 581 Shindell, D..T., D. Rind, and P. Lonergan, 1998: Increased polar stratospheric
582 ozone losses and delayed eventual recovery owing to increasing greenhouse-gas
583 concentrations. *Nature*, **392**, 589-592, doi:10.1038/33385.

- 584 Simpson, I. R., P. Hitchcock, T. G. Shepherd, J. F. Scinocca, 2011: Stratospheric
585 variability and tropospheric annular mode timescales. *Geophys. Res. Lett.*, **38**,
586 L20806.
- 587 Thompson, D. W. J., M. P. Baldwin, and S. Solomon, 2005: Stratosphere–
588 troposphere coupling in the southern hemisphere. *J. Atmos. Sci.*, **62**, 708–
589 715.
- 590 Thompson, D. W. J., S. Solomon, P. J. Kushner, M. H. England, K. M.
591 Grise and D. J. Karoly, 2011: Signatures of the Antarctic ozone hole in
592 Southern Hemisphere surface climate change. *Nature Geoscience*, **4**, 741–749,
593 doi:10.1038/ngeo1296.
- 594 Thomson, A., et al., 2011: RCP4.5: a pathway for stabilization of radiative
595 forcing by 2100. *Climatic Change*, **109**, 77–94, 10.1007/s10584-011-0151-4.
- 596 Waugh, D. W. and W. J. Randel, 1999: Climatology of Arctic and Antarctic
597 polar vortices using elliptical diagnostics. *J. Atmos. Sci.*, **56**, 1594–1613.
- 598 Waugh, D. W., W. J. Randel, S. Pawson, P. A. Newman, and E. R. Nash,
599 1999: Persistence of the lower stratospheric polar vortices. *J. Geophys. Res.*,
600 **104 (D22)**, 27191–27201.
- 601 Wilcox, L. J., A. J. Charlton-Perez, and L. J. Gray, 2012: Trends in Austral jet
602 position in ensembles of high- and low-top CMIP5 models. *J. Geophys. Res.*,
603 **117**, D13115.
- 604 Wu, Z. and N. E. Huang, 2004: A study of the characteristics of white noise
605 using the empirical mode decomposition method. *Proc. R. Soc. Lond. A*, **460**,
606 1597–1611.
- 607 Wu, Z. and N. E. Huang, 2009: Ensemble Empirical Mode Decomposition:
608 A noise-assisted data analysis method. *Advances in Adaptive Data Analysis*,
609 **1 (1)**, 1–41.
- 610 Wu, Z., N. E. Huang, S. R. Long, and C.-H. Peng, 2007: On the trend, de-
611 trending, and variability of nonlinear and nonstationary time series. *PNAS*,
612 **104 (38)**, 14 889–14 894.
- 613 Zhou, S., M. E. Gelman, A. J. Miller, and J. P. McCormack, 2000: An inter-
614 hemisphere comparison of the persistent stratospheric polar vortex. *Geophys.*
615 *Res. Lett.*, **27 (8)**, 1123–1126.

Table 1: CMIP5 models used in this study. High top models are denoted by *. C¹: Cionni et al. (2011); C²: Modified Cionni et al. (2011), with a solar cycle added in future; C³: Modified Cionni et al. (2011), with zonal averages in troposphere, and future concentrations in the stratosphere determined by combining two terms in a multiple linear regression analysis; P¹: Lamarque et al. [2010, 2011]; P²: Kawase et al. (2011); S¹: Ozone concentrations from a chemistry climate model, used offline.

Model	Model top	Number of levels	Number of levels above 200 hPa	% of levels above 200 hPa	Ozone
BCC-CSM1.1	2.917 hPa	26	13	50	C ¹
CNRM-CM5	10 hPa	31	9	29	Interactive
CSIRO-Mk3.6.0	4.52 hPa	31	9	29	C ¹
HadGEM2-ES	40 km (~2.3 hPa)	38	15	39	C ²
INMCM4	10 hPa	21	8	38	C ¹
NorESM1-M	3.54 hPa	26	13	50	P ¹
MIROC5	3 hPa	56	17	30	P ²
CanESM2*	1 hPa	35	10	29	C ³
GISS-E2-R*	0.1 hPa	40	19	48	Interactive
HadGEM2-CC*	85 km (~0.01 hPa)	60	37	62	C ²
IPSL-CM5A-LR*	0.04 hPa	39	22	56	S ¹
MIROC-ESM-CHEM*	0.0036 hPa	80	63	79	Interactive
MPI-ESM-LR*	0.01 hPa	47	25	53	C ²
MRI-CGCM3*	0.01 hPa	48	20	42	C ²

Table 2: Final warming date in the high- and low-top ensemble mean, and from reanalyses

	1870-1900	1979-2005	2070-2098 (RCP4.5)	2070-2098 (RCP8.5)
High-top	310	322	317	322
Low-top	318	327	323	325
ERA-Interim/CFSR	-	312/313	-	-

Table 3: Results from multiple linear regression analysis. Significance is from a 2-tailed t-test. Values in brackets show the equivalent values when the MIROC models are excluded from the ensemble mean.

RCP 4.5				
	Regression slope	Pearson correlation coefficient	Significance	
Low-top Ozone	-11.00 (-12.89)	-0.63 (-0.69)	<0.1% (<0.1%)	
ln(GHG)	-2.68 (-0.38)	0.36 (0.41)	>5% (>5%)	
High-top Ozone	-14.65 (-12.69)	-0.75 (-0.71)	<0.1% (<0.1%)	
ln(GHG)	9.69 (17.12)	0.50 (0.51)	>5% (<5%)	

RCP 8.5				
	Regression slope	Pearson correlation coefficient	Significance	
Low-top Ozone	-9.94 (-12.39)	-0.63 (-0.69)	<0.1% (<0.1%)	
ln(GHG)	11.01 (12.64)	0.39 (0.42)	<1% (<1%)	
High-top Ozone	-14.51 (-12.75)	-0.76 (-0.72)	<0.1% (<1%)	
ln(GHG)	17.27 (22.11)	0.49 (0.51)	<0.1% (<0.1%)	

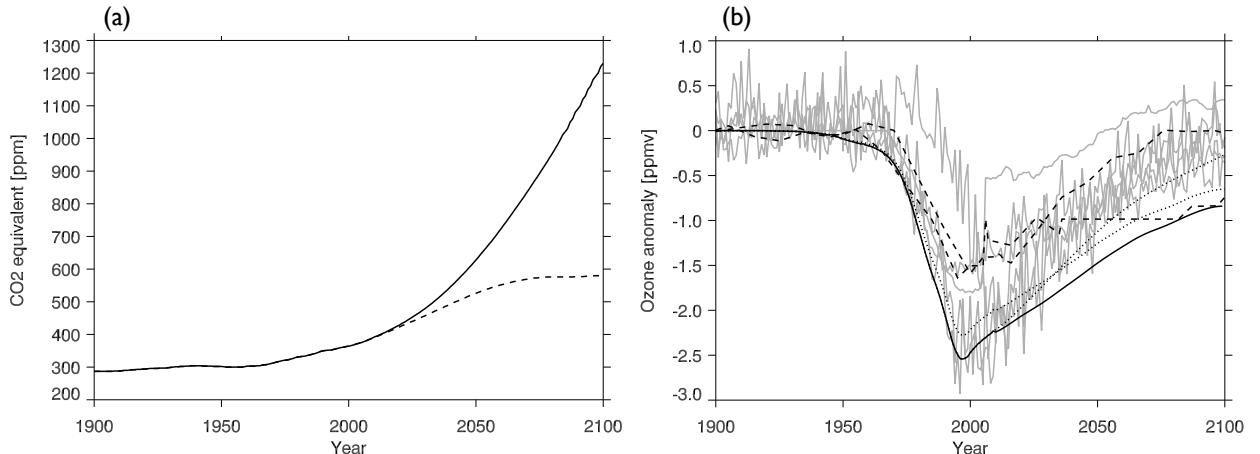


Figure 1: (a): Global-mean annual-mean greenhouse gas concentration (CO₂ equivalent) for RCP4.5 (dashed) and RCP8.5 (solid). (b): Antarctic mean (75-90°S) ozone concentrations at 50 hPa, relative to 1900 values, from Cionni et al. (2011) (black), modified versions of Cionni et al. (2011) (dotted), prescribed ozone from other sources (dashed), and from models with interactive stratospheric chemistry or those using independent chemistry climate models (grey).

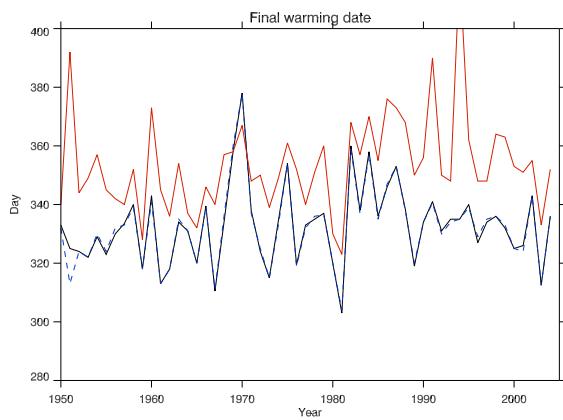


Figure 2: Historical final warming date calculated from CNRM-CM5 data using the Black and McDaniel method (red), the Haigh and Roscoe method using daily data (blue dashed), and the Haigh and Roscoe method from monthly data using interpolation (black).

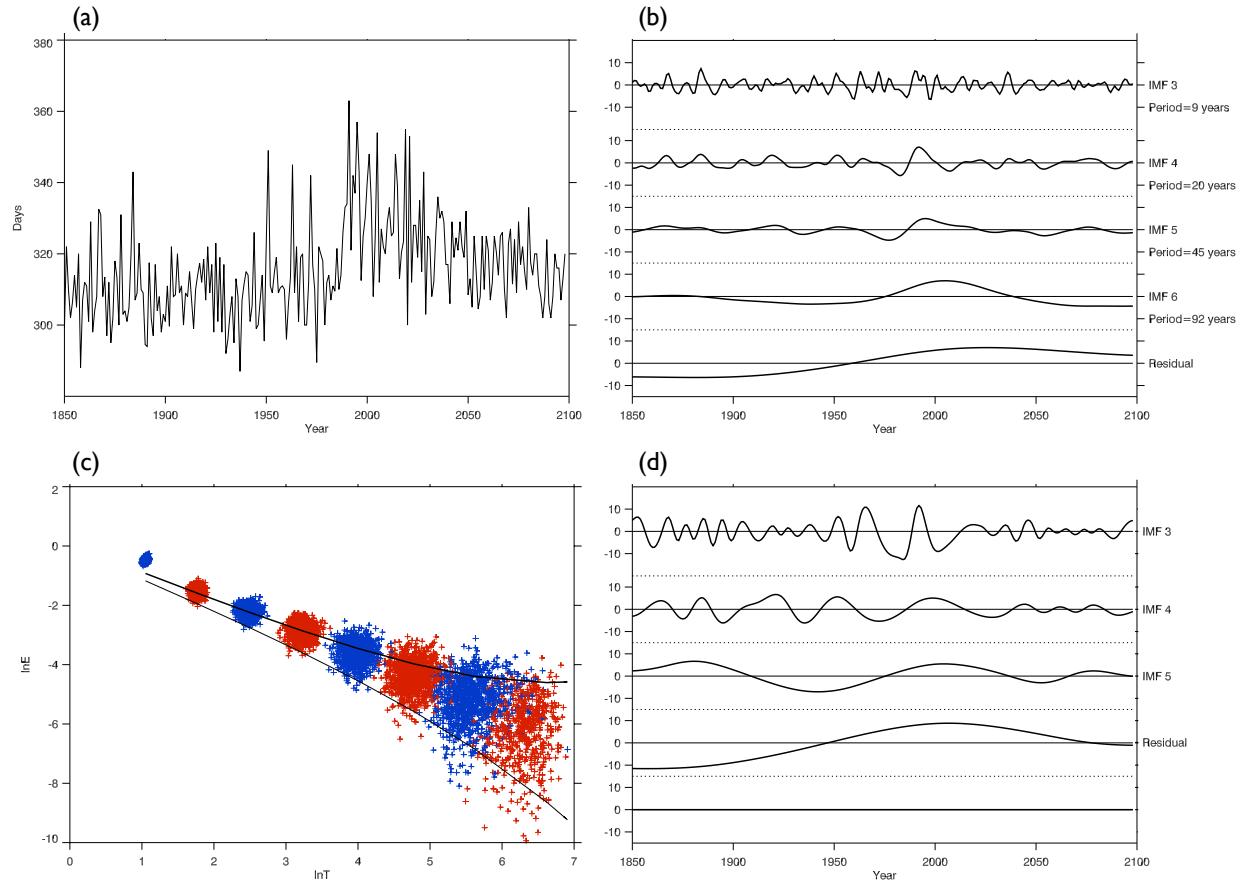


Figure 3: (a): Final warming dates from MIROC-ESM-CHEM for the historical and RCP4.5 experiments, calculated using the Haigh and Roscoe method. (b): The associated high order IMFs from EEMD. (c): The distribution of the energy and period of IMFs from 1000 white noise time series, each containing 1000 data points, and the spread function of the 95% confidence interval. (d): The associated high order IMFs from EMD, showing evidence of mode mixing.

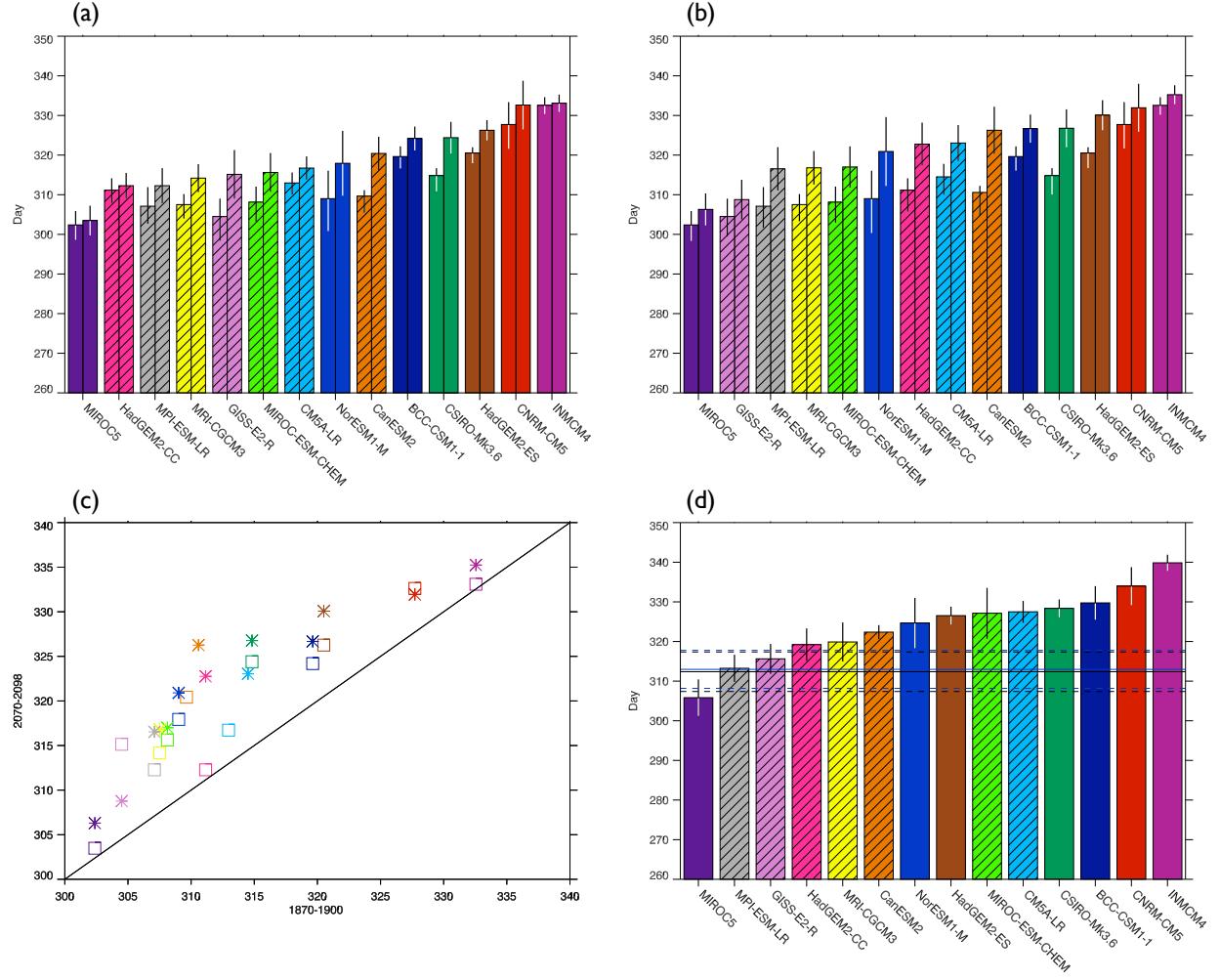


Figure 4: Mean final warming dates for each model for (a):1870-1900 (left bars) and 2070-2098 (right bars) in RCP4.5, (b): 1870-1900 (left bars) and 2070-2098 (right bars) in RCP8.5, (d): 1979-2005. Whiskers show ± 2 standard errors. High-top models are indicated by hatching. In panel (d), the horizontal solid lines show the mean final warming date from ERA-Interim (black) and CFSR (blue), with dashed lines indicating ± 2 standard errors in each case. The relationship between 1870-1900 and 2070-2098 final warming date is shown in panel (c) for RCP4.5 (squares) RCP8.5 (stars).

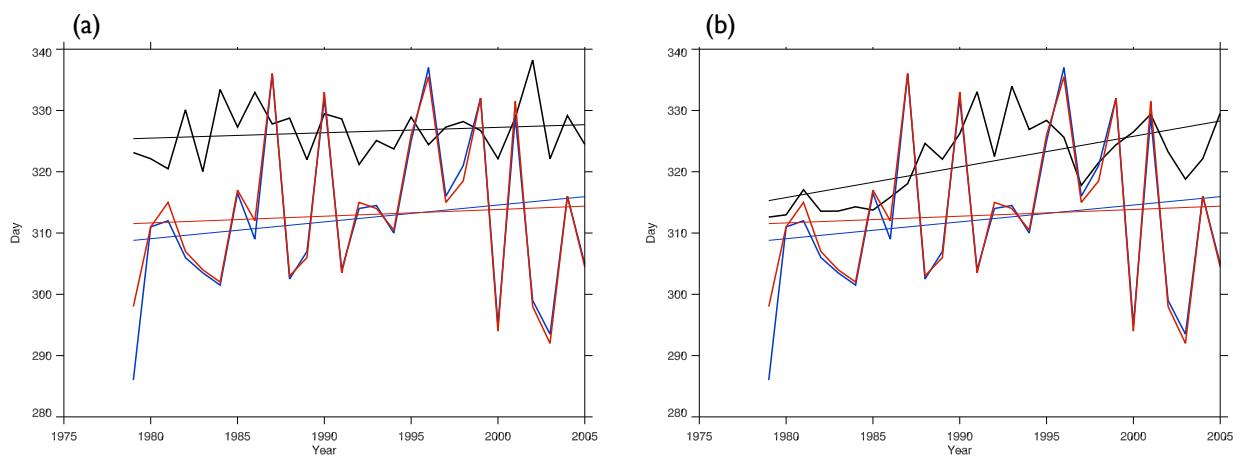


Figure 5: Final warming date from ERA-Interim (blue) and CFSR (red) with (a): the low-top ensemble mean final warming date (black), (b): the high-top ensemble mean final warming date (black).

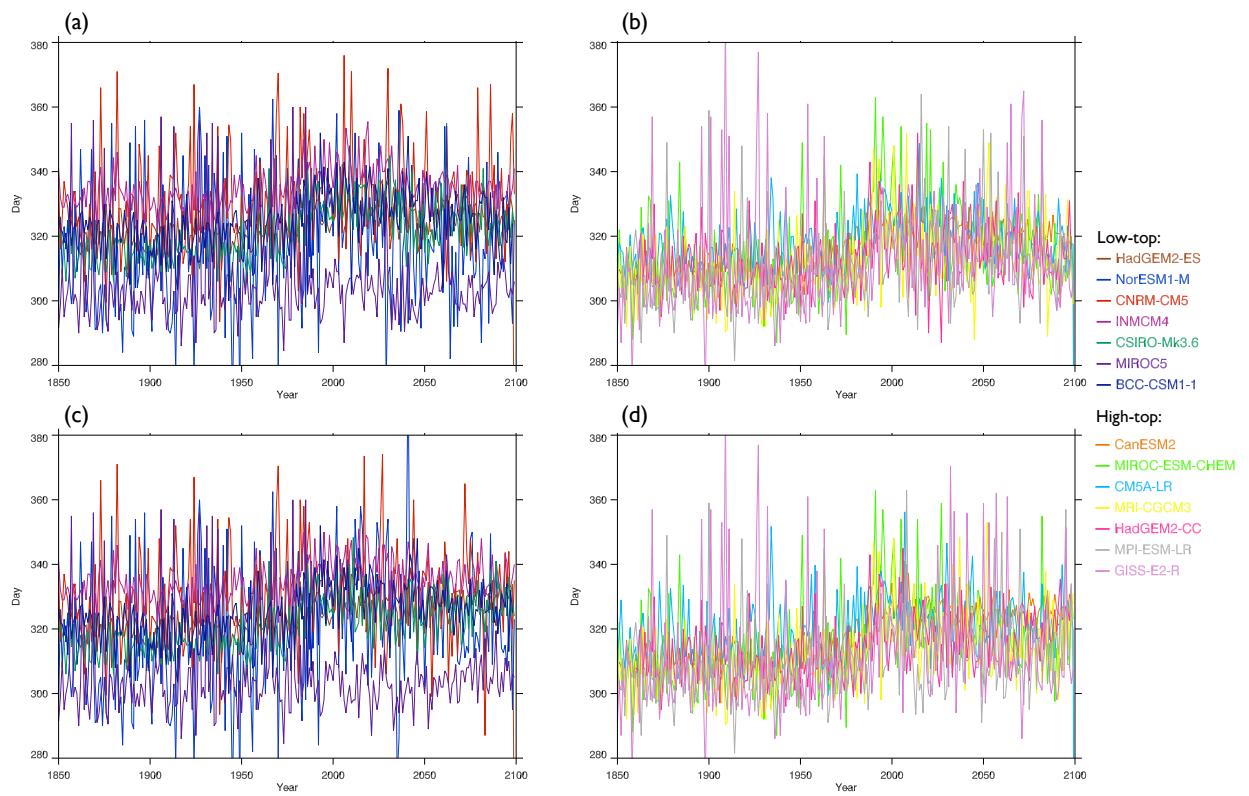


Figure 6: Final warming date for low-top (left column) and high-top (right column) models. (a,b): historical and RCP4.5, (c,d): historical and RCP8.5.

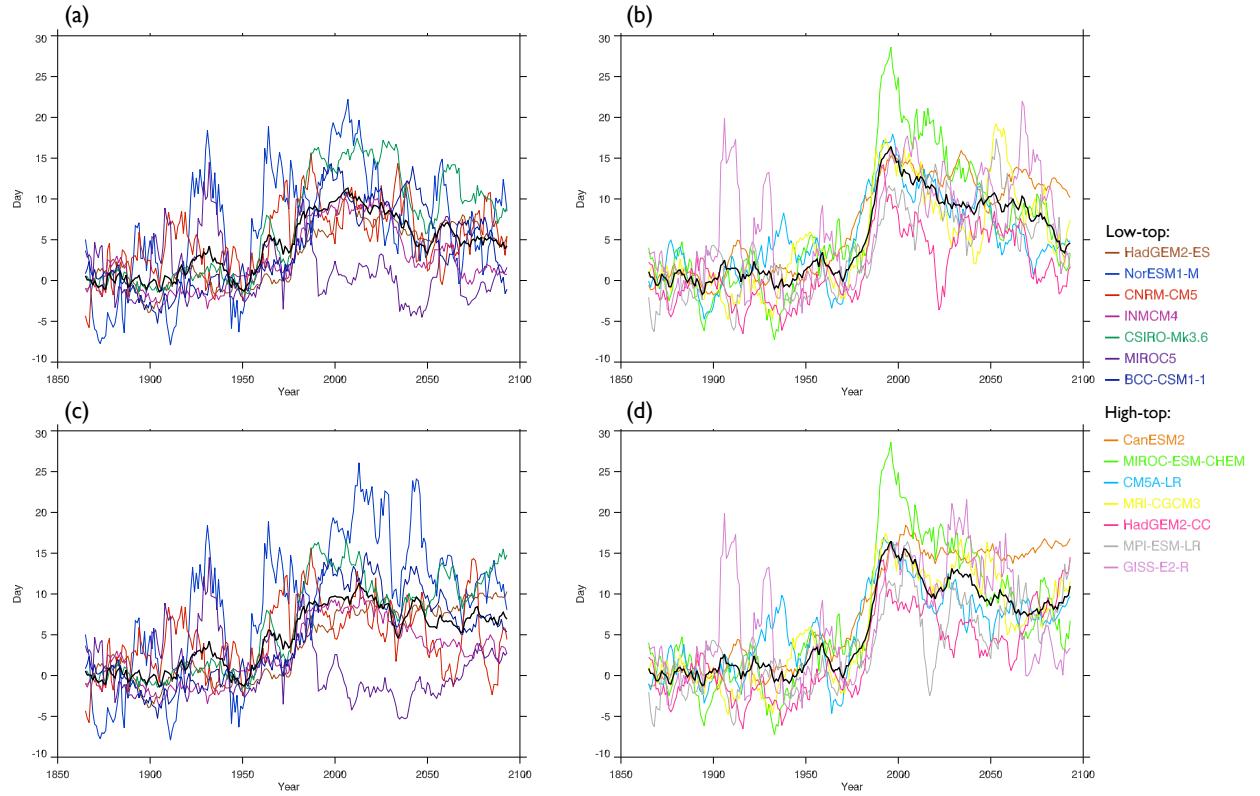


Figure 7: 11-year running mean final warming date for low-top (left column) and high-top (right column) models, with the ensemble mean (thick black line). (a,b): historical and RCP4.5, (c,d): historical and RCP8.5. Raw data is adjusted to the 1860-1900 mean.

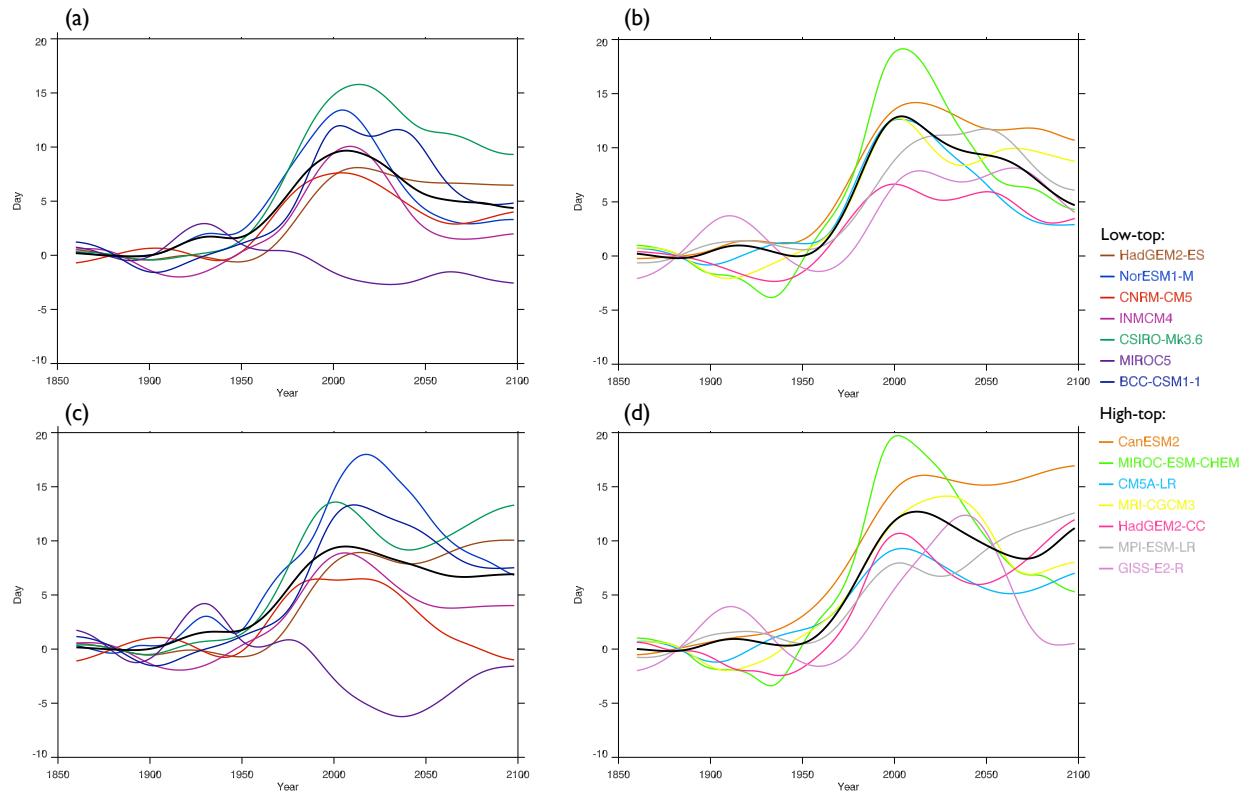


Figure 8: Sum of the residual and last IMF of final warming date for low-top (left column) and high-top (right column) models, with the ensemble mean (thick black line). (a,b): historical and RCP4.5, (c,d): historical and RCP8.5. Raw data is adjusted to the 1860-1900 mean.

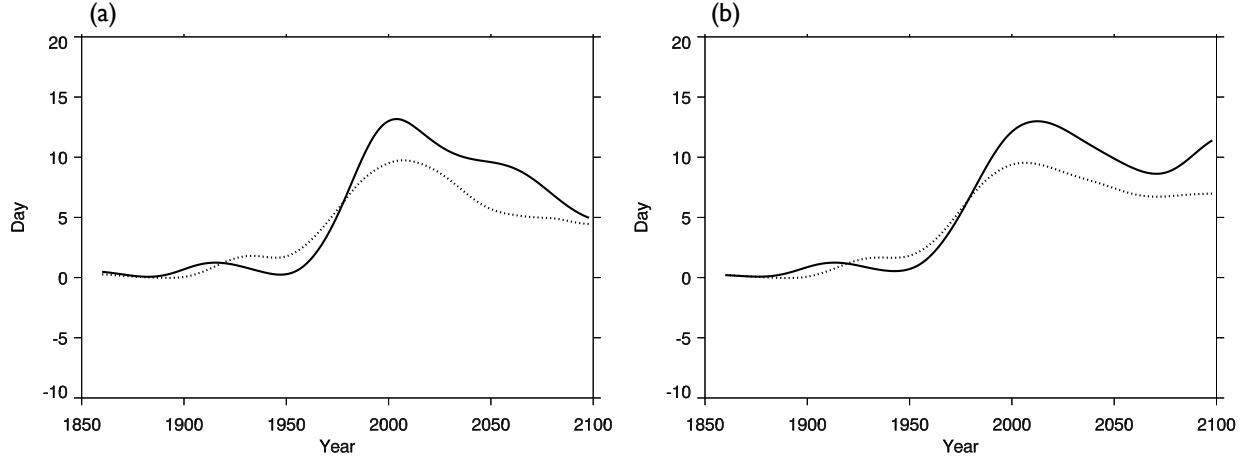


Figure 9: Sum of the significant IMFs of final warming date for low-top (dotted) and high-top (solid) ensemble means. (a): historical and RCP4.5, (b): historical and RCP8.5. Raw data is adjusted to the 1860-1900 mean.

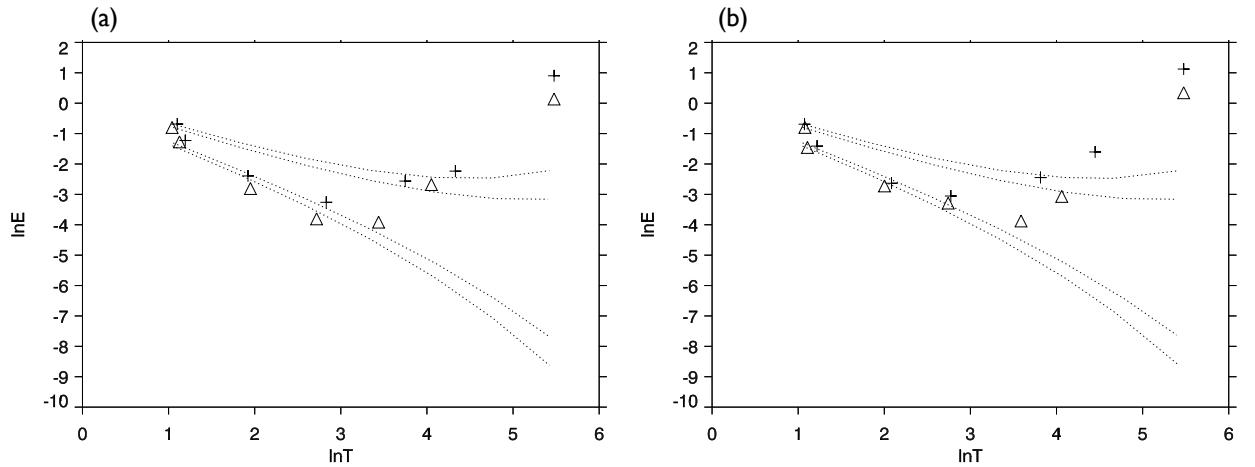


Figure 10: Spread function (dotted lines) and energies of individual IMFs for the low-top (triangles) and high-top (crosses) ensemble means. (a): historical and RCP4.5, (b): historical and RCP8.5. The inner pair of dotted lines show the 95% confidence interval, the outer pair show the 99% confidence interval.