

Evaluation and projections of the East Asian summer monsoon in a perturbed parameter ensemble

John W. Rostron (✉ john.rostron@metoffice.gov.uk)

Met Office <https://orcid.org/0000-0002-6746-7786>

David M. H. Sexton

Met Office

Kalli Furtado

Met Office

Maria João Carvalho

Met Office

Sean F. Milton

Met Office

José M. Rodríguez

Met Office

Wenxia Zhang

Institute of Atmospheric Physics Chinese Academy of Sciences

Research Article

Keywords: Perturbed parameter ensembles, East Asian Summer Monsoon, Model evaluation, Climate projections, Sub-selection

Posted Date: March 17th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1415799/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

1 Evaluation and projections of the East
2 Asian summer monsoon in a perturbed
3 parameter ensemble

4

5 John W. Rostron¹, David M. H. Sexton¹, Kalli Furtado¹, Maria João Carvalho¹, Sean F.
6 Milton¹, José M. Rodríguez¹, Wenxia Zhang²

7

8 ¹ Met Office Hadley Centre, FitzRoy Road, Exeter, EX1 3PB, UK

9 ² LASG, Institute of Atmospheric Physics, Chinese Academy of Sciences,

10 Beijing 100029, China

11

12 Corresponding author: John W. Rostron

13 Email: john.rostron@metoffice.gov.uk

14

15 ORCID iDs:

16 John W. Rostron: 0000-0002-6746-7786

17 David M. H. Sexton: 0000-0002-0004-0518

18 Kalli Furtado: 0000-0002-5166-112X

19 Maria João Carvalho: 0000-0002-7722-4710

20 Sean F. Milton: 0000-0002-2109-9770

21 José M. Rodríguez: 0000-0001-7479-5641

22 Wenxia Zhang: 0000-0001-8614-8070

23

24

25 **Abstract**

26 The East Asian summer monsoon (EASM) is a dominant driver of East Asian climate, with variations
27 in its strength potentially impacting the livelihoods of millions of people. Understanding, predicting
28 and assessing uncertainties in these variations is therefore an important area of research. Here, we
29 present a study of the projected 21st Century changes in the EASM using a ‘perturbed parameter

30 ensemble' (PPE) of HadGEM3-GC3.05 coupled climate models, which samples uncertainties arising
31 from differences in model parameter values. We show that while the present-day performance of
32 PPE members for leading order EASM metrics is comparable to CMIP5 models in many respects, the
33 PPE also exposes model biases which exist for most, if not all, parameter combinations. These
34 'structural' model biases are found mainly in the low-level circulation over the South China Sea and
35 west Pacific. We also show that future changes in regional circulation and precipitation are projected
36 consistently across the PPE members. A more detailed moisture budget analysis of the precipitation
37 changes in a region covering the Yangtze River valley shows that both dynamic and thermodynamic
38 changes contribute to an increase in precipitation over the 21st century, but that the spread in this
39 response across the PPE is mainly due to spread in the dynamic responses. We also discuss how the
40 information provided by the PPE may be used in practice, considering the plausibility of the models,
41 and giving examples of ways to sub-select ensemble members to capture the diversity in the
42 moisture budget changes.

43

44 **Keywords:** Perturbed parameter ensembles; East Asian Summer Monsoon; Model evaluation;

45 Climate projections; Sub-selection

46

47

48 1 Introduction

49

50 Monsoon systems are a key driver of seasonal variability throughout the tropics, directly affecting
51 the livelihoods of over two-thirds of the world's population (Sperber et al. 2013). Their characteristic
52 reversal of winds in the lower troposphere, and associated variations in rainfall, are driven by
53 seasonal variations in solar insolation, with substantial differences in local influences (e.g. land-sea
54 temperature contrasts, orography) giving rise to distinct systems throughout Africa, Asia, Australia,
55 and North and South America (Wang et al. 2017).

56

57 This paper concerns one such system, the East Asian summer monsoon (EASM), which covers a
58 domain stretching over the South China Sea, East China, Japan and the Korean peninsula. It is
59 characterised by an abrupt reversal of low-level winds over the South China Sea during May, and the
60 subsequent establishment of a quasi-stationary rain band (the 'Meiyu', 'Baiu' or 'Changma') which
61 propagates northwards in distinct phases through the summer, before retreating southwards in
62 August (Ding and Chan 2005). The EASM is a highly complex system, with many components and
63 influencing factors both locally and remotely e.g. variability in the West Pacific subtropical high; the
64 south Asian High, the subtropical Eurasian jet stream and Pacific and Indian ocean SST anomalies
65 (Ding and Chan 2005).

66

67 Millions of people across East Asia are affected by the monsoon: the rainfall it brings accounts for
68 around half of annual totals over the region, and interannual variations are typically around 30%
69 (Sperber et al. 2013) of the seasonal means, with potentially serious consequences through flooding,
70 drought and impacts on water supplies, agriculture and hydroelectricity generation. Providing
71 predictions of the EASM and its variations is thus vitally important, and much work has been done,
72 particularly on seasonal timescales (Wang et al. 2015). Given the complexity in modelling the
73 system, forecasting has typically been based on statistical relationships. More recently, however,
74 dynamical modelling using the Met Office's GloSea5 system has been shown to be skilful over China,
75 and has formed the basis of a forecast for summer rainfall over the Yangtze River basin (Li et al.
76 2016; Bett et al. 2018) – a region which is particularly prone to the impacts of flooding.

77

78 On longer timescales, it is important to understand the role of a changing climate on the EASM.
79 Globally, monsoon activity, variability and the strength of teleconnections to ENSO are expected to
80 increase (Hsu et al. 2012, 2013), whilst Kitoh et al. (2013) found the EASM to respond strongly to
81 warming (compared to other monsoon regions), particularly for metrics of heavy precipitation.

82 Detailed studies looking at moisture budget decompositions over the region have revealed that the
83 precipitation increase is largely driven by moistening of the atmosphere, but that uncertainties in
84 this response are mainly due to uncertainties in circulation: both in the background state and
85 responses to climate change (Zhou et al. 2018; He and Li 2019; Zhang et al. 2021).

86

87 Given the highly complex nature of monsoon systems, including the EASM, it is crucial to test the
88 robustness of these projected changes to various sources of uncertainty e.g., modelling uncertainty,
89 scenario uncertainty and internal variability (Zhou et al. 2020). Much of the work on projected
90 changes has focussed on multi-model ensembles (e.g., CMIP3, 5 and 6), which bring together the
91 latest model configuration from different centres in ‘ensembles of opportunity’. These ensembles
92 sample uncertainties in modelling structures e.g., in their resolution, complexity, parameterisation
93 schemes.

94

95 In this paper we will evaluate the present-day performance and 21st century responses of the EASM
96 using models that sample a different source of uncertainty: that arising from uncertainties in the
97 values of model parameters. We do this using an alternative approach to ensemble creation, where
98 members share the same model structure but differ in the values they take for uncertain model
99 parameters. Such ‘perturbed parameter ensembles’ (PPEs) have been used in a variety of contexts:
100 for example, to study present-day climate performance (e.g. Yokohata et al. 2013; Sexton et al.
101 2019, 2021); climate feedbacks and sensitivities (e.g. Sanderson 2011; Collins et al. 2011; Karmalkar
102 et al. 2019; Rostron et al. 2020; Tsushima et al. 2020); emergent constraints (e.g. Wagman and
103 Jackson 2018); and aerosol forcing (Regayre et al. 2018; Johnson et al. 2018). A key strength of PPEs
104 is their ability to highlight structural behaviours of a model: behaviours (e.g., biases or responses)
105 that are common to most (if not all) parameter combinations and that cannot be removed through
106 parameter tuning. Conversely, PPEs also identify those aspects of a projections that *are* susceptible
107 to tuning.

108

109 The PPE studied here is a 20-member ensemble of a very recent configuration of the Met Office’s
110 global coupled model, referred to as HadGEM3-GC3.05 (Yamazaki et al. 2021), which fed into the
111 recent UK Climate Projections for land project (UKCP18; Murphy et al. 2018). The 20 members were
112 generated through simultaneous perturbations made to 47 model parameters (across 7 atmospheric
113 parameterisation schemes), chosen to sample key parametric uncertainties (see Section 2).

114

115 To place our PPE analysis in the context of previous studies, we also analyse a subset of CMIP5
116 models. The two ensembles provide complementary datasets, sampling different sources of
117 uncertainty, and considering both allows for a more thorough representation of the uncertainties in
118 model performances and responses. This is important for testing the robustness of future changes in
119 the EASM, and for providing more comprehensive information to users interested in regional
120 impacts assessments and adaptation work. Indeed, a similar combination of these ensembles was
121 used for the global model component of the UK Climate Projections project, UKCP18 (Murphy et al.
122 2018).

123

124 However, given the different nature of the ensembles, comparisons between them are challenging.
125 Whilst the structural choices sampled by CMIP5 make it useful for capturing a wide diversity in
126 model biases and responses, it cannot reveal the structural behaviours of any individual member.
127 Conversely, whilst the PPE will expose structural biases and responses, it only does this for a single
128 model. Consider, for example, the ensemble mean biases. A clear structural model bias in the PPE
129 will be reflected in its mean, as it would not be removed by parameter perturbations. Each CMIP5
130 model could also exhibit clear structural biases (though we wouldn't know without a PPE around
131 each of them), but the effects of these would be suppressed in the CMIP5 mean due cancellation
132 across the different structures (unless it was a bias common across the CMIP5 models e.g., the
133 double ITCZ bias). So, whilst it advantageous to consider both ensembles, we must be cautious when
134 directly comparing them.

135

136 Our analysis will focus on performance and responses for leading-order metrics of low-level winds
137 and precipitation. Whilst these metrics will not represent all the complexities of the EASM, they will
138 encompass many of its key features. For example, climatological low-level winds are crucial for
139 capturing the correct flow of moisture through the region, whilst seasonal cycles of precipitation will
140 be sensitive to the northward propagation of the Meiyu rain band. Known relationships between the
141 interannual changes in these variables will also be assessed (Wang et al. 2008). Variability in these
142 relationships on decadal timescales and longer will be of particular interest as they have been noted
143 for their potential use in Met Office seasonal forecasts for the Yangtze river basin (Martin et al.
144 2020).

145

146 The remainder of the paper is ordered as follows: Section 2 provides a summary description of the
147 design of the PPE and the sub-selection of CMIP5 members, along with the methods used to analyse
148 precipitation regionally over China (including a moisture budget analysis) and a metric used for low-

149 level circulation over east Asian. In Section 3 we analyse the present-day performance of both the
150 PPE and CMIP5, looking at mean state biases and variability of low-level winds and precipitation;
151 seasonal cycles of precipitation; and the relationships in interannual anomalies related to the EASM
152 (including ENSO). Section 4 covers the responses in these variables, with Section 4.2.1 focussing on
153 understanding the drivers of the 21st century precipitation response for the Yangtze River basin,
154 including using a detailed moisture budget analysis. We will discuss our findings in Section 5, with a
155 focus on how the information provided by the PPE may be used in practice. An overall summary is
156 given in Section 6.

157
158

159 2 Models and methods

160

161 2.1 Base model and parameter perturbations

162

163 Here we provide a summary description of the PPE. Further details, including a description of the
164 GC3.05 base model, can be found in Yamazaki et al. (2021).

165

166 The PPE studied in this paper comprises 20 variants of the UK Hadley Centre Unified Model
167 HadGEM3-GC3.05 model, which is closely related to the GC3.1 configuration submitted to CMIP6
168 (Williams et al. 2018). Each ensemble member has a horizontal resolution of approximately 60km at
169 mid-latitudes (called ‘N216’) and was run for a 200-year period from 1900-2100, using CMIP5
170 historical forcings and future scenarios consistent with RCP8.5 emissions (accounting for carbon
171 cycle uncertainties).

172

173 Each variant is distinguished by taking a unique set of values for 47 model parameters across 7
174 parameterisation schemes from the atmosphere, land and aerosol model components. The initial
175 distributions of parameter values were chosen to target key modelling uncertainties, through an
176 elicitation exercise with model experts. The parameter values ultimately used for the 20 PPE
177 members were selected through a multi-stage filtering process, based on the plausibility of their
178 representation of the climate, and on the diversity of their climate change responses.

179

180 The latter was assessed using idealised forcing experiments in atmosphere-only simulations, where
181 diversity in climate feedbacks, aerosol and CO2 forcings, and regional precipitation and temperature

182 responses were targeted (Sexton et al. 2021). The plausibility of the variants was assessed in a
183 variety of historical and present-day experiments e.g. using large-scale mean climate performance in
184 5-day and 5-year atmosphere-only experiments, as well as qualitative assessments of circulation,
185 surface air temperature and precipitation over the North Atlantic and UK (Sexton et al. 2021).
186 Further screening was applied to the variants run as fully coupled simulations. An initial ensemble of
187 25 members was reduced, first to 20 members and then to 15, based on criteria such as: numerical
188 stability; the strength of the Atlantic meridional overturning circulation (AMOC); historical trends in
189 northern hemisphere surface air temperatures; and climatological temperature biases (Yamazaki et
190 al. 2021).

191

192 The final 15 members were selected for use in the UKCP18 projections (Murphy et al. 2018), but for
193 this study we are interested in exploring a diverse range of model behaviours, with a focus on a
194 different region (East Asia). So, we choose to use the 20 members selected after the first stage of
195 coupled screening and which were run for the full 200-year period.

196

197 One of the 20 PPE members we analyse uses tuned parameter values i.e. the same parameter values
198 used for the HadGEM3-GC3 model (Williams et al. 2018). We refer to this as the 'standard' PPE
199 member.

200

201

202 2.2 CMIP5 models

203

204 We also assess a subset of CMIP5 global coupled models, which allow us to sample uncertainties in
205 model structures, in contrast to the parametric modelling uncertainties sampled by the PPE. The two
206 datasets are complementary: the CMIP5 models will provide a useful context in which to place our
207 assessment of the PPE, and consideration of the results from both ensembles is recommended for
208 users of the projections (but note the caveats regarding direct comparisons discussed in the

209 Introduction).

210

211 The CMIP5 subset we use comprises 13 members which were selected based on a qualitative
212 assessment of key aspects of global and European/UK climate, along with a screening of very closely
213 related models (Murphy et al. 2018). We use data for the same 200-year period simulated by the
214 PPE, and the same CMIP5 historical forcings. For the future period, concentrations from the RCP8.5
215 scenario are used (note the slight difference to the PPE here, as carbon cycle uncertainties are not
216 sampled for the CMIP5 models). All 13 members are re-gridded from their native resolution to that
217 of the PPE members (60km at mid-latitudes) to facilitate the comparison of the two ensembles.

218

219

220 2.3 Definition of regions

221

222 Part of this study into the East Asian summer monsoon will involve an assessment of precipitation
223 across China. As noted in the Introduction, the EASM has a complex spatial and temporal structure,
224 with the quasi-stationary Meiyu rain band influencing different regions through the season. We
225 therefore separate China into regions, based on areas that display similar characteristics of
226 precipitation variability.

227

228 We do this using a K-Means clustering algorithm (Wilks 2011), where climatological monthly
229 anomalies of precipitation are calculated for each grid box over China and boxes with similar annual
230 cycles are grouped together into 3 groups. For this analysis we used GPCP observational data
231 (covering 1980-2014; Adler et al. 2018), and considered land points only. The 3 regions resulting
232 from this clustering were largely continuous, although some grid boxes fell into a different region
233 from their neighbours, particularly near the borders between regions. Because of this, we manually
234 adjusted the regions so they were completely continuous (i.e., no 'floating' grid boxes), but still
235 reflected the broad regions selected by the K-Means algorithm.

236

237 The 3 regions selected by this analysis are shown in Fig. 1. Two of these, which we label 'Southeast'
238 (SE) and 'Central-East' (CE) China align with the north-south propagation of the EASM over eastern
239 China. The other region, which we label 'North and West' (N+W) China encompasses a larger domain
240 including the arid and semi-arid regions in northwest China and the Tibetan Plateau. We also define
241 2 further regions: one for Northeast (NE) China, covering the important maize-growing region; and

242 another for Southwest (SW) China, where previous versions of Hadley Centre models have
243 consistently shown wet biases in the summer.

244
245

246 2.4 Metrics

247 2.4.1 Reversed Wang and Fan Index (RWFI)

248 Indices are used widely in climate science to quantify features of the climate system in a simple way
249 (e.g., circulation patterns). In this study, we focus on one index used in studies of the East Asian
250 summer monsoon – the ‘reversed Wang and Fan index’ (RWFI). This is defined, using summer (JJA)
251 means, as:

252

$$253 \quad RWFI = u_{850}(22.5^{\circ} - 32.5^{\circ}N, 110^{\circ} - 140^{\circ}E) - u_{850}(5^{\circ} - 15^{\circ}N, 90^{\circ} - 130^{\circ}E)$$

254

255 (See the red boxes in Fig. 2.) This index reflects the low-level shear vorticity over the region and was
256 initially used to quantify variability in the western North Pacific summer monsoon (Wang and Fan
257 1999). The complexity of the East Asian summer monsoon precludes any index from capturing all
258 aspects of the system (Wang et al. 2008). However, Wang et al. (2008) found the RWFI correlates
259 very well with the first multivariate EOF for precipitation, surface level pressure and winds over
260 China, and thus provides a simple metric to capture some key features of East Asian summer
261 monsoon variability. We therefore use the RWFI metric as our leading-order metric to study low-
262 level circulation for the EASM.

263

264 2.4.2 Nino-3.4 index

265 In our analysis of the connections between modes of variability affecting the EASM (Section 3.3) we
266 consider the role of ENSO and its impact on the EASM circulation. We quantify ENSO using the Nino-
267 3.4 index, which is defined using the long-term anomalies of the monthly-mean SST averaged over
268 the Nino-3.4 region (5S–5N, 170W–120W), where we use a baseline climatology period from 1950-
269 2006. We then smooth the time series using a 5-month window and normalise using the standard
270 deviation of the smoothed time series over the climatological period¹.

271

272

¹ The definition of this Nino-3.4 index can be found in the ‘Technical Notes’ section of <https://climatedataguide.ucar.edu/climate-data/nino-sst-indices-nino-12-3-34-4-oni-and-tni>

273 2.5 Moisture budget analysis

274 In Section 4.2.1 we analyse 21st century changes in summer precipitation for the Central-East China
 275 region. Part of the analysis is based on changes in the components of the moisture budget for this
 276 region, which we calculate following a similar method to Seager et al. (2010). Following their
 277 Equation (1), climatological averages of the difference between precipitation (P) and
 278 evapotranspiration (E) are balanced by the convergence of the vertically integrated moisture fluxes:
 279

$$280 \quad \bar{P} - \bar{E} = -\frac{1}{g} \int_0^{p_s} \nabla \cdot \bar{q}\bar{v} dp - q_s \mathbf{v}_s \cdot \nabla p_s + RES' \quad (1)$$

$$281 \quad \bar{P} - \bar{E} = -\frac{1}{g} \int_0^{p_s} \nabla \cdot [\bar{q}\bar{v} + \overline{q'v'}] dp + RES \quad (2)$$

282

283 Here, p denotes pressure, q is the specific humidity and \mathbf{v} is the horizontal wind vector. Overbars
 284 denote monthly climatologies, while dashes denote the difference between daily mean values and
 285 the climatology (e.g., $q = \bar{q} + q'$). Surface values are denoted with subscript 's'. The term $q_s \mathbf{v}_s \cdot$
 286 ∇p_s in Equation (1) accounts for the contribution to moisture fluxes from the flow at the surface
 287 (Seager et al. 2010). We incorporate this into a residual term (RES), which we use to account for
 288 imbalances between P-E and the vertically integrated moisture fluxes. Additional contributions to
 289 this residual will result from the fact the monthly means for the cross terms are derived from daily
 290 mean data (rather than timestep data), and from errors in the divergence and integration
 291 calculations (from the use of 17 discrete pressure levels), for example. From Equation (1) to (2) we
 292 have used $\overline{q\mathbf{v}} = \bar{q}\bar{v} + \bar{q}\bar{v}' + \bar{q}'\bar{v} + \overline{q'\mathbf{v}'}$, where the cross terms in this equation are 0 (since $\bar{q}' =$
 293 $\bar{v}' = 0$). At the timescales we consider in our analysis (30-year climatologies), changes in the total
 294 water content of the column are assumed to be negligible (Rodríguez et al. 2017).

295

296 We calculate changes in the moisture budget using differences in 30-year means, for the periods
 297 2070-2099 vs 1980-2009. Denoting these differences as 'D', from Equation (2):

298

$$299 \quad D\bar{P} - D\bar{E} = -\frac{1}{g} \int_0^{p_s} \nabla \cdot [D(\bar{q}\bar{v}) + D(\overline{q'\mathbf{v}'})] dp + DRES \quad (3)$$

$$300 \quad D\bar{P} - D\bar{E} = -\frac{1}{g} \int_0^{p_s} \nabla \cdot [\bar{v}D\bar{q} + \bar{q}D\bar{v} + D\bar{q}D\bar{v} + D(\overline{q'\mathbf{v}'})] dp + DRES \quad (4)$$

301

302 We denote the change in the convergence of integrated moisture fluxes (i.e., the first term on the
 303 right-hand side of Equation 4) as $DconvQ$. Further, using the notation from Seager et al. (2010), we
 304 define the contributions to $DconvQ$ from changes in thermodynamics (TH), mean circulation
 305 dynamics (MCD), non-linear effects (NL) and transient eddies (TE) as:

$$307 \quad DTH = -\frac{1}{g} \int_0^{p_s} \nabla \cdot (\bar{\mathbf{v}} D\bar{q}) dp \quad (5)$$

$$308 \quad DMCD = -\frac{1}{g} \int_0^{p_s} \nabla \cdot (\bar{q} D\bar{\mathbf{v}}) dp \quad (6)$$

$$309 \quad DNL = -\frac{1}{g} \int_0^{p_s} \nabla \cdot (D\bar{q} D\bar{\mathbf{v}}) dp \quad (7)$$

$$310 \quad DTE = -\frac{1}{g} \int_0^{p_s} \nabla \cdot D(\bar{q}'\mathbf{v}') dp \quad (8)$$

311
 312 Conceptually, DTH is the change in moisture flux caused by changes in specific humidity carried on
 313 the present-day wind field, whilst DMCD is the contribution from changes in the wind field with
 314 specific humidity held to present-day values. DNL describes the effect of co-varying changes in the
 315 long-term averages of the humidity and the winds, whilst DTE is the contribution from changes in
 316 short-timescale co-variations in q and \mathbf{v} .

317
 318 Using these, we re-write Equation (4) as:

$$320 \quad D\bar{P} - D\bar{E} = DconvQ + DRES \quad (9)$$

$$321 \quad D\bar{P} - D\bar{E} = DTH + DMCD + DNL + DTE + DRES \quad (10)$$

322
 323 Further, we define the sum of the contributions from the mean circulation dynamics, the transient
 324 eddies and the non-linear effects to be changes in the dynamics (DDYN):

$$325 \quad DDYN = DMCD + DNL + DTE \quad (11)$$

326
 327
 328 So that:

$$329 \quad D\bar{P} - D\bar{E} = DTH + DDYN + DRES \quad (12)$$

330

331 Our moisture budget calculations produce monthly means for each grid box. For the analysis
332 presented in Section 4.2.1 we average these temporally and spatially to produce values for summer
333 (JJA) over the Central-East China region (as defined in Section 2.3).

334

335

336 **3 Present-day performance of the EASM**

337 **3.1 850hPa circulation**

338

339 We start with an assessment of the lower-level (850hPa) circulation over the region. Fig. 2 shows the
340 JJA mean climatological 850hPa circulation from reanalysis (ERA-Interim; Dee et al. 2011) for the
341 period 1980-2015, along with the bias for the PPE ensemble mean. We also show examples of the
342 ‘best’ and ‘worst’ members from the PPE and CMIP5, as measured by root-mean-squared errors of
343 the zonal wind component over the area highlighted by the red boxes (see Section 2.4.1). The JJA
344 climatology over the region is characterised by an airflow along the Somali coast, across the Arabian
345 Sea (a moisture source for the Indian monsoon) and the Bay of Bengal and into the South China Sea.
346 This circulation develops rapidly during May (Ding and Chan 2005) and persists through NH summer.
347 The latter two regions are key moisture sources for the EASM, whilst the western North Pacific
348 subtropical high (WNPSH) is another aspect of EASM interannual variability.

349

350 The PPE mean field shows substantial biases in this circulation, with a cyclonic error over Southeast
351 Asia and the west Pacific. The westerly bias centred on the South China Sea, resulting from an over-
352 extension of the Indian monsoon flow, is particularly strong. This bias is characteristic of a known
353 circulation error that has affected previous generations of Hadley Centre models (Ringer et al. 2006;
354 Bush et al. 2015) and is present in atmosphere-only (as well as coupled) simulations, and across
355 resolutions (Rodríguez et al. 2017). It is associated with biases in the representation of the WNPSH,
356 which tends to be too weak and shifted too far to the east in these models.

357

358 The cyclonic bias pattern is also seen in the ‘best’ and ‘worst’ PPE members (labelled ‘0834’ and
359 ‘1113’, respectively). In the latter case, the errors over the South China Sea reach up to 9.2 ms^{-1} but
360 the errors are clearly much reduced for member 0834. A similar bias is seen for the worst CMIP5
361 model (MRI-CGCM3), but the best model across the two ensembles (CNRM-CM5) performs well over
362 the whole domain shown in Fig. 2.

363

364 These features are seen across the PPE and CMIP5 ensembles, as shown in the 1980-2015 mean
365 climatologies and time series' of the RWFI (Fig. 3 top panel and Fig. 4, respectively), which measures
366 the 850hPa vorticity over the region (as sampled by the red boxes in Fig. 2; see Section 2.4.1). The
367 RWFI climatologies for the PPE show a structural negative bias for the PPE, with the cyclonic nature
368 highlighted by consistently negative biases for RWFI-North and positive biases for RWFI-South.

369

370 Negative RWFI biases are also seen in most of the CMIP5 models. The structural nature of the biases
371 we have shown in the PPE could indicate the CMIP5 biases may also be structural. However, without
372 PPEs based around the CMIP5 models, this is difficult to assess. There are examples of positive
373 biases for RWFI (ACCESS1-3) and negative (i.e., easterly) biases for RWFI-South (CCSM4 and MPI-
374 ESM-MR), as well as the particularly well-performing model (CNRM-CM5) highlighted in Fig. 2. In
375 both the PPE and CMIP5, the spread in RWFI values results mainly from the southern component:
376 understanding the processes and structural changes driving the spread in this region could be crucial
377 for resolving this model error (Bush et al. 2015; Martin et al. 2021).

378

379 The bottom panel of Fig. 3 shows the interannual variability in RWFI (and its components) for each
380 ensemble member. Both ensembles span a range of variability, from 1-3 ms^{-1} , which includes the
381 value from the ERA-Interim reanalysis. PPE members typically show less variability than the
382 reanalysis in all components (see also the time series for the PPE standard model in Fig. 4). CMIP5
383 models are also typically less variable than the reanalysis for RWFI-North, but there are examples of
384 enhanced variability for RWFI and RWFI-South.

385

386

387 3.2 Precipitation

388

389 As noted in the Introduction, a key characteristic of the EASM is the northward progression of the
390 Meiyu rain band across the region, from June to August. Accordingly, we start by evaluating the
391 performance of the PPE for precipitation using the annual cycles across the distinct regions defined
392 in Section 2.3, to capture the large-scale spatial and temporal characteristics of the EASM
393 precipitation (see Fig. 5).

394

395 Across all of the regions, the observed annual cycles (GPCP; Adler et al. 2018) show a continuous
396 increase in precipitation from winter to a peak in summer and a subsequent decrease after the
397 retreat of the EASM. These variations are broadly captured by the PPE, although values for MAM are

398 consistently overestimated by all members, across the regions. This bias was also noted by
399 Rodríguez et al. (2017) for a previous, atmosphere-only configuration of HadGEM3 (GA6; Walters et
400 al. 2017), and is associated with errors in the moisture convergence.

401

402 The timings of the peak of the precipitation in summer is well captured across PPE members for
403 N+W, NE and SW China, but there is a notable spread in CE and SE China. Members that peak
404 prematurely in CE China (i.e., in June rather than the observed peak in July) also tend to have a
405 premature peak in SE China (see the red curves in Fig. 5), indicating coherence in these timing errors
406 across the PPE. However, this subset tends to capture the JJA mean precipitation in SW and N+W
407 China more accurately than the members with the correct peak timing (in July) in CE China.

408

409 We show 1980-2014 mean JJA precipitation values for PPE members in the top panel of Fig. 6, along
410 with the CMIP5 ensemble and observations. (See also the mean JJA precipitation time series' in Fig.
411 7.) Clear wet biases across the PPE in southern regions and in N+W China indicate structural model
412 errors here (though the large extent of N+W China makes the interpretation of this region difficult).
413 No such structural biases are found for CE China, which covers the Yangtze River valley, and NE
414 China – an important crop growing region.

415

416 Biases in the CMIP5 models also vary spatially. Despite the differences in their constructions (as
417 noted in the Introduction), the PPE and CMIP5 ensembles are strikingly similar for the northern
418 regions, with the observed values being well captured for NE China, but systematic wet biases for
419 N+W China, suggesting an error that persists across model structures. In CE, SE and SW China the
420 CMIP5 models are more diverse, with examples of positive and negative biases in each. Here, the
421 value of combining the two ensembles to more thoroughly capture a diversity in performance is
422 clear. This is especially the case for SE China, where all but two CMIP5 models have dry biases, in
423 contrast to the wet biases seen in PPE members.

424

425 The interannual variability in JJA precipitation is shown in the bottom panel of Fig. 6 and for the
426 standard PPE member in Fig. 7 (pink lines). The variability for PPE members tends to be too high
427 across all regions, with values ranging between 0.95 and 2.09 times the observed standard deviation
428 values. CMIP5 models also tend to overestimate the variability, but there are examples of models
429 with too little variability (e.g., for CE and SE China), reflecting a more diverse sampling of
430 precipitation variability in the CMIP5 models compared to the PPE.

431

432 We have also calculated precipitation variability scaled by the climatological mean for each region
433 (not show) and found that these were also consistently overestimated in the PPE. The reasons why
434 the PPE overestimates precipitation variability is not clear, but some of this may be driven through
435 relationships with the monsoon circulation (Wang et al. 2008; also see Section 3.3). An
436 overestimation of precipitation variability could be driven by too much variability in the circulation
437 acting through this teleconnection, or from errors in the teleconnection itself. However, we do not
438 find evidence of this in the PPE. As shown in Fig. 3, the variability of the RWFI in the PPE is
439 reasonable (even slightly underestimated). We have also analysed the relationships between
440 circulation and CE China precipitation in the PPE (see Section 3.3) but find that values of the fraction
441 of variance of precipitation explained by RWFI are between 0.02 and 0.35 (the observed value is
442 0.29), indicating that other influences are driving our overestimated precipitation variability in this
443 region.

444

445

446 3.3 Variability relationships

447

448 As has been noted in many previous studies, clear relationships exist between interannual changes
449 in key EASM circulations (e.g. the WNPSH, the low-level vorticity measured by the RWFI) and JJA
450 precipitation in the region (e.g. Wang et al. 2008; Li et al. 2016) and their links to ENSO. Here, we
451 explore how these relationships are represented in the PPE by focusing on the connection between
452 the RWFI, JJA mean precipitation for CE China and ENSO.

453

454 We focus on CE China, as this coincides with the Yangtze River valley – a region where the impacts of
455 interannual rainfall changes can be great, but also where the strength of these relationships is
456 strong, opening up opportunities for exploiting predictability in models (Bett et al. 2018; Martin et
457 al. 2020). Also, as described in Section 3.2, our PPE validates reasonably well against observations in
458 this region.

459

460 We start with the relationship between JJA mean precipitation for CE China and the RWFI. Wang et
461 al. (2008) showed a strong relationship between the RWFI and the first multivariate EOF of EASM
462 variability, where positive RWFI anomalies are associated with enhanced precipitation over a region
463 which coincides with our CE China region (see Fig. 2a in Wang et al. 2008). These anomalies are
464 characteristic of an enhanced south-westerly flow over the South China Sea (and a reduction in the

465 zonal wind as measured by RWF1-S), associated with a south-westward extension of the WNPSH and
466 a weakened monsoon trough in the western North Pacific (Wang et al. 2008).

467

468 A simple way to characterise this relationship is to evaluate the slope of a simple linear fit to
469 interannual anomalies of CE China precipitation against the RWF1 index. We show these values in
470 Table 1 (column 'prC-RWF1'). The 'observed' ('OBS') value is derived from 35 years of data from GPCP
471 (for prC) and ERA-Interim (for RWF1) where these two datasets overlap (1980-2014). The observed
472 value of 0.200 ± 0.110 confirms the expected relationship between these quantities is significant.
473 (Uncertainties given are for the 95% confidence range and significance is tested at the 5% level.) The
474 remaining rows in Table 1 give the equivalent values for the PPE members. Most PPE members also
475 exhibit significant relationships, with slopes that are indistinguishable from the observed
476 relationship (at the 5% level). This can also be seen in the γ -values in Fig. 8. The exceptions to this
477 are members 0834, 2123, 2335 and 2832, for which our evidence isn't strong enough to distinguish
478 their slopes from zero; and member 2753, which has a steep slope that is not consistent with the
479 observations. Most notable amongst these is member 0834 (see point labelled 'D' in Fig. 8). This
480 member has the smallest circulation biases in the PPE (see Section 3.1), but also has one of the least
481 sensitive and least realistic prC-RWF1 relationships. Conversely, the member with the largest
482 circulation bias (member 1113; labelled 'E') has a prC-RWF1 sensitivity of 0.253 ± 0.141 , which is
483 consistent with the observed value.

484

485 Interannual variability in the EASM is known to be strongly influenced by the El Nino-Southern
486 Oscillation (ENSO) and many studies have explored the potential mechanisms linking ENSO to
487 anomalies in EASM circulation. Wang et al. (2008) showed that the peak of a lead-lag correlation
488 between their first multivariate EOF, which exhibits an anomalous subtropical high in the west
489 Pacific, and the NINO3.4 index, occurs in the preceding winter. That is, the anomalous anticyclone,
490 which is associated with an enhanced south-westerly flow over the South China Sea (decreased
491 RWF1-S and increased RWF1) occurs in the summer after an El Nino. This has been linked to positive
492 feedback mechanisms in the west Pacific and the Indian ocean, where ENSO-induced SST anomalies
493 drive Rossby (west Pacific) and Kelvin (Indian ocean) waves, which reinforce the anticyclonic
494 circulation, allowing it to persist into the summer (Wang et al. 2000; Xie et al. 2016; Xie and Zhou
495 2017; Hardiman et al. 2018).

496

497 We characterise this relationship in a similar way to the prC-RWF1 relationship – using the slope of a
498 simple linear regression between the RWF1 and the NINO3.4 index (see definition in Section 2.4.2).

499 We use DJF averages for the NINO3.4 index and regress against the RWFI for the following summer,
500 to capture the peak of the lead-lag correlation described in Wang et al. (2008). As with the prC-RWFI
501 relationship, we use the period 1980-2014 to evaluate the slopes (with the data for the NINO3.4
502 index starting from December 1979). The results are shown in the third column of Table 1 ('RWFI-
503 NINO3.4'). The observed relationship is significant, as expected, with a value of 0.970 ± 0.525 .
504 However, the PPE exhibits different behaviours: for most of the members (15 out of 20) the RWFI vs
505 NINO3.4 slope is not distinguishable from zero, and 9 of these have distinctly different relationships
506 to the observations. This can be seen quite clearly in Fig. 8 (x-values), where PPE members generally
507 have smaller RWFI-NINO3.4 slope values than the observations, with some values even being
508 negative. The exceptions to this are PPE members 1113, 1554, 2089, 2491 and 2832, which are
509 indistinguishable from the observed relationship (at a 5% confidence level). Except for 2832, these
510 members all matched the observed prC-RWFI relationship too, as highlighted by the clustering of
511 these members around the observations in Fig. 8. We again highlight the cases of members 0834
512 and 1113. The former, which has the smallest circulation errors, captures neither the prC-RWFI
513 relationship nor the RWFI-NINO3.4 relationship; whilst the latter, which performed poorly in its
514 circulation, is mostly closely matched to the observations for these variability relationships (see
515 points 'D' and 'E' compared to the observations in Fig. 8).

516

517 **Table 1** Relationships between CE China precipitation (prC), EASM circulation (using the RWFI) and ENSO (using Nino3.4
518 index) for observations (top row) and the PPE (remaining rows). Values are calculated as the slope of the linear relationship
519 between interannual anomalies of these metrics for the period 1980-2014. For prC and RWFI we use JJA means, whilst for
520 the NINO3.4 index we use DJF means from the preceding winter. Values in brackets give the uncertainties capturing the
521 95% confidence range. Crosses (*) denote slopes which are not significant, whilst daggers (†) denote a significant difference
522 from the observed value. Values with neither a cross nor a dagger therefore denote significant slopes which are consistent
523 with the observed value. (Significance is assessed at the 5% level).

Member		prC-RWFI	RWFI-NINO3.4
X	OBS	0.200 [0.110]	0.970 [0.525]
A	0000	0.392 [0.262]	0.371 [0.409] ^x
B	0090	0.349 [0.232]	0.030 [0.416] ^{x†}
C	0605	0.311 [0.255]	0.262 [0.520] ^x
D	0834	0.117 [0.309] ^x	-0.193 [0.371] ^{x†}
E	1113	0.253 [0.141]	0.956 [0.633]
F	1554	0.246 [0.230]	0.538 [0.382]
G	1649	0.281 [0.170]	-0.322 [0.560] ^{x†}
H	1843	0.305 [0.234]	0.169 [0.488] ^{x†}
I	1935	0.344 [0.215]	0.205 [0.520] ^{x†}
J	2089	0.253 [0.164]	1.237 [0.467] ^x

K	2123	0.224 [0.246] ^x	-0.051 [0.525] ^{x†}
L	2242	0.257 [0.240]	-0.231 [0.614] ^{x†}
M	2305	0.236 [0.200]	0.109 [0.557] ^{x†}
N	2335	0.146 [0.212] ^x	0.339 [0.605] ^x
O	2491	0.330 [0.212]	0.730 [0.486]
P	2753	0.562 [0.270] [†]	0.073 [0.420] ^{x†}
Q	2832	0.090 [0.223] ^x	0.498 [0.431]
R	2868	0.260 [0.166]	0.474 [0.742] ^x
S	2884	0.154 [0.143]	0.423 [0.611] ^x
T	2914	0.254 [0.165]	0.400 [0.504] ^x

524

525

526

527 4 Future changes in the EASM

528

529 Each PPE member was simulated out to 2100 under the CMIP5 RCP8.5 scenario (as outlined in
530 Section 2). Here we assess how the EASM responds to this future scenario across our PPE members,
531 in terms of the mean-state and variability of the low-level circulation (Section 4.1) and precipitation
532 (Section 4.2). In Section 4.2.1 we use a more detailed moisture budget analysis for changes in
533 precipitation of the CE China region, to highlight the relative impact of thermodynamic and dynamic
534 changes on the precipitation response.

535

536

537 4.1 850hPa circulation

538

539 21st century circulation responses, based on differences in 30-year averages around 1995 (1980-
540 2009) and 2085 (2070-2099), are shown in Fig. 9. The PPE mean shows a clear anti-cyclonic response
541 over the region, with a weakened westerly flow over the SE Asian peninsula and South China Sea,
542 and increased south-westerly flow over East China, suggesting an intensification of the EASM
543 circulation. This response is seen consistently, but with varying magnitudes, across individual PPE
544 members (3 of which are shown in Fig. 9). The mean state responses in the RWFI (and its
545 components) shown in the top panel of Fig. 10 are consistent with this: westerly (easterly) responses
546 are seen in the north (south) components, and the RWFI response is consistently positive as a result.

547 This can also be seen in the time series for the PPE in Fig. 4, where RWFI values increase over the
548 21st Century This systematic response across PPE members suggests it is driven by a structural
549 response of HadGEM3-GC3.05, which the parameter perturbations do not alter significantly.

550

551 Amongst the CMIP5 models there are examples of both anti-cyclonic and cyclonic responses in the
552 region (2 examples are shown in Fig. 9). But without PPEs based on the CMIP5 models it is not
553 possible to assess whether any of these are systematic responses (as we find for HadGEM3-GC3.05).
554 Interestingly, Chen et al. (2020) found that constraining CMIP5 models, based on present-day SST
555 patterns associated with uncertainties in projections of the WNPSH, favoured models with a future
556 strengthening of the WNPSH. Such a change is consistent with the robust weakening of the cyclonic
557 circulation captured by the RWFI that we have found in the PPE. Even so, Fig. 10 shows that
558 magnitudes of the changes for the CMIP5 models are smaller than in the PPE; that is, the PPE
559 appears to have a particularly strong circulation response. We also note that the CMIP5 member
560 with the largest positive change in RWFI, HadGEM2-ES, is the most closely related to the HadGEM3-
561 GC3.05 configuration used for the PPE.

562

563 In Fig. 11 we show these circulation responses against the present-day mean values. This shows a
564 cluster of 6 PPE members with comparable performance to CMIP5 models (with present-day means
565 $> -9 \text{ ms}^{-1}$) that sample responses at the lower end of the PPE, but that are systematically higher than
566 the CMIP5 model responses. Again, this highlights the benefit of considering information from both
567 ensembles, with CMIP5 providing a wide diversity of future behaviours, and the PPE providing more
568 examples of large, but still plausible, responses. Note though, that the combination of PPE and
569 CMIP5 ensembles should not be considered as one entity, so the combined scatter should not be
570 considered as evidence of an emergent relationship without more information e.g., PPEs based on
571 each CMIP5 member.

572

573 Whilst the CMIP5 models show little relationship between present-day biases and future responses
574 ($R^2=0.13$), there is a significant relationship for the PPE (at the 5% level; $R^2=0.38$). Models with larger
575 present-day biases tend to have stronger future responses. This connection is somewhat tentative
576 and would require a more targeted set of experiments to understand the common drivers, if any,
577 behind the relationship. However, for the PPE we can start to look at the influence of model
578 parameters by regressing the present-day mean and responses for RWFI against the parameter
579 values. We find that the deep entrainment parameter ('ent_fac_dp') explains the most variance for
580 both quantities. For present-day values, deep entrainment has an R^2 of 0.60 and clearly stands out

581 compared to other parameters (the next highest is R^2 is 0.24 and could be due to chance). However,
582 for the responses the correlation is weaker, with $R^2=0.31$, and many other parameters have similar-
583 sized effects, making it difficult to pick-out the key processes here. We acknowledge that with so few
584 PPE members (20) and so many parameters (47), robust conclusions aren't possible from this
585 analysis, but suggest that a focus on the deep entrainment parameter may provide a useful starting
586 point for follow-up studies (e.g. Zou and Zhou 2011; Bush et al. 2015). Such sensitivity studies are an
587 important strength of PPEs over multi-model ensembles (like CMIP5), as they allow a deeper
588 understanding of the processes affecting a model's biases and responses.

589

590 The bottom panel of Fig. 10 shows the change in variability (the standard deviation) in the RWFI
591 components. No consistent change in the variability is seen for any component – both the PPE and
592 CMIP5 ensembles have members with increases and decreases in variability, and the magnitude of
593 these changes are similar in the two ensembles.

594

595

596 4.2 Precipitation

597

598 Responses in JJA precipitation, for 2070-2099 vs 1980-2009, are shown for our 5 regions in Fig. 12.
599 Positive changes in both the mean state and interannual variability are widespread: all PPE members
600 get wetter in all regions, as do most CMIP5 members (top panel of Fig. 12; also Fig. 7). A small
601 number of CMIP5 members show a slight drying in some regions, namely IPSL-CM5A-MR in CE, SE,
602 and SW China, and MPI-ESM-MR in CE and N+W China. For some members/regions, the responses
603 are weak and not statistically significant (at the 5% level) - for example CNRM-CM5 does not show a
604 significant change over CE, SE and SW China, whilst the weak drying seen in MPI-ESM-MR for CE and
605 N+W China is also not significant. Typically, however, the responses *are* significant, particularly in the
606 PPE. Variability also typically increases in these regions (bottom panel of Fig. 12; Zhang et al. 2021),
607 except for a handful of members from each ensemble.

608

609 As described in Section 3.2, biases vary across the different models and our chosen regions. For
610 climate service applications, users may want to apply bias corrections by analysing the percentage
611 change in summer precipitation to the modelled climatology, which we show in the middle panel of
612 Fig. 12. In both the percentage and absolute changes, the precipitation responses in the PPE are
613 typically larger than those in CMIP5: PPE values range from 0.3 to 4.5 mm/day (7 to 58 %), whilst

614 CMIP5 values range from -0.7 to 2.9 mm/day (-15 to 36 %). This is most notable in CE and SE China,
615 although there is substantial overlap between the ensembles across the regions.

616

617 We would emphasise that the structural precipitation responses seen in the HadGEM3-GC3.05 PPE
618 may also be present in CMIP5 models. But without PPEs based around these models we cannot
619 assess this. As was the case for the precipitation biases (Section 3.2), considering the two ensembles
620 will clearly be of benefit to users interested in capturing an appropriate level of plausible diversity in
621 precipitation changes over China.

622

623

624 4.2.1 Drivers of precipitation response

625

626 In the previous section we showed that PPE members consistently project increases in precipitation
627 for China over the 21st Century, but that there is a sizable amount of spread in the magnitude of this
628 change. We now look in more detail at what is driving these changes for the CE China region, starting
629 with a simple assessment of future prC-RWFI relationships, followed by a closer look at changes in
630 the moisture budget. We use this region as a case study for several reasons. As noted in the
631 Introduction it is particularly sensitive to interannual variability, with past flooding events impacting
632 the lives of 100s of millions of people. It is also an important economic region, including some of
633 China's largest cities and being a centre for key industries (e.g., agriculture and hydroelectricity).
634 Consequently, the region has been a major focus of forecasting and climate prediction research, and
635 recent work has shown skilful seasonal predictions over the region in the Met Office's HadGEM3-
636 GC2 model. Additionally, in this paper we have shown that the HadGEM3-GC3.05 PPE (based on a
637 more recent, but closely related model to HadGEM3-GC2) captures the observed summer
638 precipitation well over the region).

639

640 In Section 3.3 we showed positive relationships between anomalies in summer precipitation for CE
641 China (prC) and RWFI for the PPE, as well as observations. We have also seen increases in the mean
642 values for both prC and RWFI, so a simple question to ask is: can the PPE's prC responses be
643 explained by the RWFI responses applied to the present-day prC-RWFI relationship? This is
644 analogous to rainfall changes being driven by changes in the large-scale monsoon circulation (to the
645 extent that this is described solely by RWFI), but assuming that any adjustments in precipitable
646 water or the relationship between RWFI and prC are small. The second of these assumptions
647 (stationarity of RWFI-prC relationship throughout the 21st century) is also related to changes in the

648 large-scale circulations: for example, strengthening or weakening of the WNPSH could alter the
649 relation between RWFI and prC. We tested the stationarity of the RWFI-prC relationship by
650 calculating the gradients for the prC-RWFI linear fits in 4 50-year periods across the 21st Century
651 (1989-2039, 2009-2059, 2029-2079 and 2049-2099) and comparing to gradients for 1969-2019. The
652 results are shown in Fig. 13. This shows the prC-RWFI relationships are not constant, with gradient
653 values at the end of the 21st Century showing little correlation with 1969-2019 values. For instance,
654 member 2242 has a prC-RWFI gradient of 0.21 for 1969-2019 (closely matching the observed value
655 of 0.20 for 1980-2014), but has a quite different relationship for 2049-2099, with a slope of -0.22.
656 Conversely, member 2832 has a weak gradient of 0.04 for 1969-2019, but stronger relationships of
657 0.22 and 0.26 for 2029-2079 and 2049-2099, respectively. The variability in the prC-RWFI
658 relationship suggests that this simple framework is not sufficient for understanding prC changes, and
659 a more detailed analysis of changes in the moisture budget is required.

660

661 To do this we analysed daily means of precipitation, evaporation and variables contributing to the
662 moisture convergence - including its decomposition into thermodynamic and dynamic contributions.
663 From these, we determine changes in the components of the moisture budget for CE China,
664 averaged over summer (JJA) for two 30-year periods: 1980-2009 and 2070-2099. The details of these
665 calculations are given in Section 2.5. The results of this decomposition are shown in Fig. 14.

666

667 The precipitation changes (DP) shown in Fig. 14 are equivalent to those for CE China in Fig. 12.
668 Changes in evaporation (DE) are small, with values ranging from -0.16 to 0.39 mm/day (not shown).
669 Consequently, the precipitation changes are largely driven by changes in the moisture convergence
670 ($R^2=0.84$; see Equation 9).

671

672 The residual term (DRES) from the decomposition is also shown in Fig. 14. Whilst values are small
673 compared to DP and DconvQ, they are typically negative across the PPE, ranging from -0.52 to 0.06
674 mm/day. The main driver of this residual is not clear, but there will be contributions from the change
675 in the surface term (see Equation 1), from errors introduced through the use of daily means and
676 from errors in the divergence and integration calculations. The size of these residuals places limits on
677 how confidently we can discuss terms in the moisture budget analysis.

678

679 Results from the further decomposition of DconvQ into thermodynamic (DTH) and dynamic terms
680 (DDYN; see Equations 5-12) are also shown in Fig. 14. We find that DTH is always positive, as
681 expected from the moistening of the atmosphere in a warmer climate. DDYN also typically

682 contributes positively to DconvQ: only three members have negative dynamic changes, and their
683 magnitudes are small.

684

685 In the top two panels of Fig. 15 we show correlations between DP-DE and these components. These
686 show that the spread in DP-DE is driven mainly by the dynamic changes, which explain 60% of the
687 variance across PPE members. In contrast, the thermodynamic changes show little correlation with
688 DP-DE ($R^2=0.01$), and typical values for DRES (indicated by the error bars) are comparatively small.

689

690 The dynamic component of this decomposition (DDYN) is a sum of contributions from changes in the
691 mean circulation, transient eddies, and a non-linear term (DMCD, DTE and DNL, respectively; see
692 Equations 6-8). The DMCD component, which describes moisture convergence changes resulting
693 from changes in the mean circulation with the humidity held fixed (at present-day values), accounts
694 for 24% of the variance in DP-DE (bottom-left panel of Fig. 15). This relationship is clearly degraded
695 compared to DDYN and suggests the other dynamical components (DTE, DNL and covariance terms)
696 also contribute.

697

698 The contributions from all these components highlights the need for the in-depth moisture budget
699 analysis over the simpler RWFI-based analysis we presented above. Like the DMCD component, the
700 RWFI-based analysis attempted to capture the effect of changes in the mean circulation. However,
701 they represent quite different ways to capture the effects of the changing circulation: the DMCD
702 component describes moisture convergence changes resulting from changes in the mean circulation
703 with the humidity held fixed, whilst our RWFI-based analysis estimated the effects of holding the
704 present-day relationship between precipitation and circulation anomalies fixed. Whilst we might
705 expect some level relationship between DMCD and DRWFI (indeed they are correlated with
706 $R^2=0.53$), the latter clearly misses several aspects of the drivers of the precipitation change i.e., the
707 remaining components of DDYN, as well as the contributions from DTH and DE. The bottom-right
708 panel of Fig. 15 highlights this, which shows that DRWFI only explains a similar amount of the
709 variance in DP-DE as DMCD.

710

711

712 5 Discussion on robustness and sub-selection of PPE members

713

714 The 20-member HadGEM3-GC3.05 PPE was developed to provide users with raw global model
715 output, suitable for use in regional impacts and adaptation studies. However, a dataset of this size
716 may not be practical or desirable to use, for example due to human or computational resource
717 limitations. In such cases a subset of members could be chosen, which were deemed to be plausible
718 but still representative of the future changes explored by the full ensemble (McSweeney et al. 2015).
719 The details of this sub-selection will depend on the application, but here we discuss some
720 considerations for sub-selecting the PPE in the context of the EASM.

721

722 A key consideration will be the robustness of the information provided by the models i.e., are they
723 plausible? This could be based on the global and/or regional performance of the models. For
724 example, Yamazaki et al. (2021) describe how members of the GC3.05 PPE were selected for
725 inclusion in the UKCP18 project, based on: the performance of regional SSTs over the globe; the
726 Atlantic Meridional Overturning Circulation (a key driver of climate and variability for the North
727 Atlantic and Europe); 20th century NH temperature trends; and present-day climatologies of
728 temperature and precipitation over Europe. Using these, the PPE was reduced from 25 to 20
729 members (which we have used in this paper) in a first round of filtering, and to 15 members after a
730 second round of filtering.

731

732 In this paper we have assessed leading-order metrics for the EASM, to give a broad indication of the
733 performance of the PPE for this key mode of climate variability for China. For precipitation we have
734 shown the PPE has a reasonable performance (compared to CMIP5 models) in the mean states and
735 variability. The performance does depend on the region (see Fig. 6), but the PPE has notably good
736 performance for the key CE China region, covering the Yangtze River basin.

737

738 For low-level circulation we find a structural bias in the PPE, where all members are found to have a
739 cyclonic bias in JJA over the East Asia/West Pacific region. In comparison, the CMIP5 models do
740 encompass the observed mean state for the RWFI (Fig. 3). However, since CMIP5 is comprised of
741 single variants of different model structures, we cannot tell whether these models are themselves
742 structurally biased. We note that several PPE models have a comparable level of performance to
743 CMIP5 models for the RWFI, and that PPE models compare well to CMIP5 in terms of RWFI variability
744 (Fig. 3). Of course, the tolerance level on the mean state could be set such that no PPE members are
745 accepted, but this will depend on the user and could have implications for the level of diversity if set
746 too strictly.

747

748 Once a plausible subset of models has been identified, users may want to sub-select further in a way
749 that still captures the diversity in the projected changes being studied. As an example of this, we
750 consider a simple sub-selection of the PPE for projections of CE China precipitation, based on the
751 moisture budget analysis shown in Section 4.2.1. We have shown a range of changes for CE China
752 precipitation across the PPE, and that changes in the dynamics (DDYN) are a key driver of this. In this
753 context, a representative subset would target high and low values of changes in CE China
754 precipitation and DDYN. We have highlighted 2 PPE members which do this in Fig. 14 - Fig. 17.
755 Member 2242 (marked by a filled red circle) has a low precipitation change relative to other PPE
756 members, driven in part by a slightly negative contribution from dynamical changes. In contrast,
757 member 2884 (filled blue circle) has a relatively high change in CE China precipitation, with a large
758 positive contribution from DDYN.

759
760 These two models sample high and low values in other metrics too, for example for changes in the
761 mean circulation metrics (shown in the bottom panels of Fig. 15), and in their thermodynamic
762 changes (top-left panel of Fig. 15). The thermodynamic contributions of these two models partially
763 offsets some of the differences from the dynamics i.e., DTH is large for in our low precipitation
764 change scenario (member 2242), whilst DTH is small for our high precipitation change scenario
765 (member 2884). This diversity is shown clearly in Fig. 16 and Fig. 17. In Fig. 16 our 2 example
766 members lie in opposing corners of the DTH and DDYN values sampled by the full PPE and represent
767 high and low precipitation changes due to the key role of the DDYN component. In Fig. 17 we show
768 DTH fields overlaid with the mean 850hPa wind field for 1980-2009 (top row) and DDYN fields
769 overlaid with changes in the 850hPa winds (2070-2099 vs 1980-2009; bottom row) for the two
770 members. The 850hpa winds are included to give an indication of the contribution of the mean
771 circulation to both DTH and DDYN (see Equations 5 and 6). The differences in the DDYN field, and the
772 associated mean circulation changes, are striking and clearly affect the CE China region (covered by
773 the red box), with large circulation changes bringing additional moisture into the region from the
774 south for member 2884, but much weaker circulation changes for member 2242. Meanwhile,
775 differences in DTH fields, and in the present-day circulation, are much more modest between the
776 members – reflecting the smaller spread in DTH compared to DDYN.

777

778 Here we have covered one simple example of sub-selection, based on capturing diversity in the
779 physical processes underlying 21st century precipitation changes. But there are many other ways a
780 subset could be chosen. Even within the same framework of our moisture budget decomposition,
781 other choices could be made. For example, we have highlighted 2 alternative (or additional)

782 members in Fig. 14 - Fig. 16 (see the cyan and pink circles), which also sample high and low DDYN
783 values, although these members have very similar thermodynamic changes, in contrast to our earlier
784 example. These members also have the highest/lowest CE China precipitation changes and could
785 provide useful scenarios for studies of flooding and/or drought in the region. Alternative ways to
786 sub-select might focus on other variables (e.g. temperatures or heat stress metrics for heat waves),
787 or consider a wider set of metrics in multi-sector studies or more generic assessments (McSweeney
788 et al. 2015; Palmer et al. 2021).

789
790

791 6 Summary

792

793 In this paper we have evaluated the simulation of the East Asian Summer Monsoon in a perturbed
794 parameter ensemble of HadGEM3-GC3.05 coupled models by analysing their performance with
795 respect to the observed climatology, and their projected changes. We focussed on leading-order
796 metrics of the low-level (850hPa) circulation and precipitation. In summary:

797

- 798 • For low-level circulation we find a structural bias in the PPE, where all members are found to
799 have a cyclonic bias over the East Asia/West Pacific region (for JJA means). This bias is known
800 to have affected previous Hadley Centre models, and the structural nature of the bias
801 revealed by the PPE suggests it cannot be easily corrected by model parameter choices.
802 Using the reversed Wang and Fan index (RWFI) we find negative biases are typical in both
803 the PPE and in CMIP5 models, but there are examples of much lower (and even positive)
804 biases in CMIP5.
- 805 • For precipitation we find the performance of PPE members varies spatially and temporally,
806 with good performance for JJA climatologies in CE and NE China, but wet biases over
807 southern China. The performance of CMIP5 models also varies by region, with differences
808 between the performance of the PPE and CMIP5 models highlighting the benefits of
809 considering both ensembles. Interannual variability is typically overestimated by both
810 ensembles, but especially by the PPE. For seasonal cycles, we have indicated a split in the
811 spatial and temporal modes of variability in the PPE, where members with smaller biases in
812 southern China tend to show a seasonal cycle that peaks too early over CE China.
- 813 • The observed relationship between the RWFI and precipitation for CE China is captured in
814 most PPE members. The PPE does not perform as well for the relationship between RWFI

815 and NINO3.4, with most members having a circulation response that is too weak. We note
816 the example of 2 PPE members: member 0834 had the smallest circulation bias, but failed to
817 capture both of the variability relationships, whilst member 1113 showed poor performance
818 in its mean circulation but was well matched to the observed variability relationships.

- 819 • Responses for the 21st century for low-level circulation over the region are consistently anti-
820 cyclonic in the PPE, suggesting a structural response in the HadGEM3-GC3.05 model. There
821 are examples of both cyclonic and anti-cyclonic circulation responses in CMIP5, but without
822 PPEs based on these models the structural nature of these biases is not known.
- 823 • Increases in mean precipitation are projected for the 21st century across China for all PPE
824 members, and most CMIP5 members, with increases in variability also projected for all but a
825 handful of members. For the CE China region, we find that both thermodynamic
826 (atmospheric moistening) and dynamic (circulation) changes contribute to the increased
827 precipitation, with the spread amongst PPE members largely caused by differences in the
828 dynamic response.

829

830 We have also discussed how these projections may be used in practice, where considerations about
831 the plausibility and usability of the models will be important, giving a simple example of sub-
832 selecting PPE members aimed at capturing diversity in their precipitation responses. Users should
833 also be aware of the limitations of these models in relation to structural biases which, as we have
834 shown, are readily exposed by PPEs. Of course, these limitations apply to each of the CMIP5
835 members but as we have highlighted through the text, one needs a PPE about each CMIP5 member
836 to properly understand their individual structural biases. Indeed, as shown by Rostron et al. (2020),
837 the errors of the tuned variant of HadGEM3-GA7.05 are not indicative of the structural model bias
838 for all variables. Therefore, we encourage wider use of the PPE.

839

840

841 References

842

843 Adler R, Sapiano M, Huffman G, et al (2018) The Global Precipitation Climatology Project (GPCP)
844 Monthly Analysis (New Version 2.3) and a Review of 2017 Global Precipitation. Atmosphere
845 (Basel) 9:138. doi: 10.3390/atmos9040138

846 Bett PE, Scaife AA, Li C, et al (2018) Seasonal Forecasts of the Summer 2016 Yangtze River Basin
847 Rainfall. Adv Atmos Sci 35:918–926. doi: 10.1007/s00376-018-7210-y

848 Bush SJ, Turner AG, Woolnough SJ, et al (2015) The effect of increased convective entrainment on
849 Asian monsoon biases in the MetUM general circulation model. *Q J R Meteorol Soc* 141:311–
850 326. doi: 10.1002/qj.2371

851 Chen X, Zhou T, Wu P, et al (2020) Emergent constraints on future projections of the western North
852 Pacific Subtropical High. *Nat Commun* 11:2802. doi: 10.1038/s41467-020-16631-9

853 Collins M, Booth BBB, Bhaskaran B, et al (2011) Climate model errors, feedbacks and forcings: a
854 comparison of perturbed physics and multi-model ensembles. *Clim Dyn* 36:1737–1766. doi:
855 10.1007/s00382-010-0808-0

856 Dee DP, Uppala SM, Simmons AJ, et al (2011) The ERA-Interim reanalysis: configuration and
857 performance of the data assimilation system. *Q J R Meteorol Soc* 137:553–597. doi:
858 10.1002/qj.828

859 Ding Y, Chan JCL (2005) The East Asian summer monsoon: an overview. *Meteorol Atmos Phys*
860 89:117–142

861 Hardiman SC, Dunstone NJ, Scaife AA, et al (2018) The asymmetric response of Yangtze river basin
862 summer rainfall to El Niño/La Niña. *Environ Res Lett* 13:. doi: 10.1088/1748-9326/aaa172

863 He C, Li T (2019) Does global warming amplify interannual climate variability? *Clim Dyn* 52:2667–
864 2684. doi: 10.1007/s00382-018-4286-0

865 Hsu P, Li T, Luo J-J, et al (2012) Increase of global monsoon area and precipitation under global
866 warming: A robust signal? *Geophys Res Lett* 39:n/a-n/a. doi: 10.1029/2012GL051037

867 Hsu P, Li T, Murakami H, Kitoh A (2013) Future change of the global monsoon revealed from 19
868 CMIP5 models. *J Geophys Res Atmos* 118:1247–1260. doi: 10.1002/jgrd.50145

869 Johnson JS, Regayre LA, Yoshioka M, et al (2018) The importance of comprehensive parameter
870 sampling and multiple observations for robust constraint of aerosol radiative forcing. *Atmos*
871 *Chem Phys* 18:13031–13053. doi: 10.5194/acp-18-13031-2018

872 Karmalkar A V., Sexton DMH, Murphy JM, et al (2019) Finding plausible and diverse variants of a
873 climate model. Part II: development and validation of methodology. *Clim Dyn* 53:847–877. doi:
874 10.1007/s00382-019-04617-3

875 Kitoh A, Endo H, Krishna Kumar K, et al (2013) Monsoons in a changing world: A regional perspective
876 in a global context. *J Geophys Res Atmos* 118:3053–3065. doi: 10.1002/jgrd.50258

877 Li C, Scaife AA, Lu R, et al (2016) Skillful seasonal prediction of Yangtze river valley summer rainfall.
878 *Environ Res Lett* 11:. doi: 10.1088/1748-9326/11/9/094002

879 Martin GM, Dunstone NJ, Scaife AA, Bett PE (2020) Predicting June Mean Rainfall in the
880 Middle/Lower Yangtze River Basin. *Adv Atmos Sci* 37:29–41. doi: 10.1007/s00376-019-9051-8

881 Martin GM, Levine RC, Rodriguez JM, Vellinga M (2021) Understanding the development of

882 systematic errors in the Asian summer monsoon. *Geosci Model Dev* 14:1007–1035. doi:
883 10.5194/gmd-14-1007-2021

884 McSweeney CF, Jones RG, Lee RW, Rowell DP (2015) Selecting CMIP5 GCMs for downscaling over
885 multiple regions. *Clim Dyn* 44:3237–3260. doi: 10.1007/s00382-014-2418-8

886 Murphy JM, Harris GR, Sexton DMH, et al (2018) UKCP18 Land Projections: Science Report

887 Palmer TE, Booth BBB, McSweeney CF (2021) How does the CMIP6 ensemble change the picture for
888 European climate projections? *Environ Res Lett* 16:094042. doi: 10.1088/1748-9326/ac1ed9

889 Regayre LA, Johnson JS, Yoshioka M, et al (2018) Aerosol and physical atmosphere model
890 parameters are both important sources of uncertainty in aerosol ERF. *Atmos Chem Phys*
891 18:9975–10006. doi: 10.5194/acp-18-9975-2018

892 Ringer MA, Martin GM, Greeves CZ, et al (2006) The Physical Properties of the Atmosphere in the
893 New Hadley Centre Global Environmental Model (HadGEM1). Part II: Aspects of Variability and
894 Regional Climate. *J Clim* 19:1302–1326. doi: 10.1175/JCLI3713.1

895 Rodríguez JM, Milton SF, Marzin C (2017) The East Asian Atmospheric Water Cycle and Monsoon
896 Circulation in the Met Office Unified Model. *J Geophys Res Atmos* 122:10,246–10,265. doi:
897 10.1002/2016JD025460

898 Rostron JW, Sexton DMH, McSweeney CF, et al (2020) The impact of performance filtering on
899 climate feedbacks in a perturbed parameter ensemble. *Clim Dyn*. doi: 10.1007/s00382-020-
900 05281-8

901 Sanderson BM (2011) A multimodel study of parametric uncertainty in predictions of climate
902 response to rising greenhouse gas concentrations. *J Clim* 24:1362–1377. doi:
903 10.1175/2010JCLI3498.1

904 Seager R, Naik N, Vecchi GA (2010) Thermodynamic and Dynamic Mechanisms for Large-Scale
905 Changes in the Hydrological Cycle in Response to Global Warming*. *J Clim* 23:4651–4668. doi:
906 10.1175/2010JCLI3655.1

907 Sexton DMH, Karmalkar A V., Murphy JM, et al (2019) Finding plausible and diverse variants of a
908 climate model. Part 1: establishing the relationship between errors at weather and climate
909 time scales. *Clim Dyn* 53:989–1022. doi: 10.1007/s00382-019-04625-3

910 Sexton DMH, McSweeney CF, Rostron JW, et al (2021) A perturbed parameter ensemble of
911 HadGEM3-GC3.05 coupled model projections: part 1: selecting the parameter combinations.
912 *Clim Dyn*. doi: 10.1007/s00382-021-05709-9

913 Sperber KR, Annamalai H, Kang I-S, et al (2013) The Asian summer monsoon: an intercomparison of
914 CMIP5 vs. CMIP3 simulations of the late 20th century. *Clim Dyn* 41:2711–2744. doi:
915 10.1007/s00382-012-1607-6

916 Tsushima Y, Ringer MA, Martin GM, et al (2020) Investigating physical constraints on climate
917 feedbacks using a perturbed parameter ensemble. *Clim Dyn* 1. doi: 10.1007/s00382-020-
918 05318-y

919 Wagman BM, Jackson CS (2018) A Test of Emergent Constraints on Cloud Feedback and Climate
920 Sensitivity Using a Calibrated Single-Model Ensemble. *J Clim* 31:7515–7532. doi: 10.1175/JCLI-
921 D-17-0682.1

922 Walters D, Boutle I, Brooks M, et al (2017) The Met Office Unified Model Global Atmosphere 6.0/6.1
923 and JULES Global Land 6.0/6.1 configurations. *Geosci Model Dev* 10:1487–1520. doi:
924 10.5194/gmd-10-1487-2017

925 Wang B, Fan Z (1999) Choice of South Asian Summer Monsoon Indices. *Bull Am Meteorol Soc*
926 80:629–638. doi: 10.1175/1520-0477(1999)080<0629:COSASM>2.0.CO;2

927 Wang B, Wu R, Fu X (2000) Pacific–East Asian Teleconnection: How Does ENSO Affect East Asian
928 Climate? *J Clim* 13:1517–1536. doi: 10.1175/1520-0442(2000)013<1517:PEATHD>2.0.CO;2

929 Wang B, Wu Z, Li J, et al (2008) How to Measure the Strength of the East Asian Summer Monsoon. *J*
930 *Clim* 21:4449–4463. doi: 10.1175/2008JCLI2183.1

931 Wang H, Fan K, Sun J, et al (2015) A review of seasonal climate prediction research in China. *Adv*
932 *Atmos Sci* 32:149–168. doi: 10.1007/s00376-014-0016-7

933 Wang PX, Wang B, Cheng H, et al (2017) The global monsoon across time scales: Mechanisms and
934 outstanding issues. *Earth-Science Rev* 174:84–121. doi: 10.1016/j.earscirev.2017.07.006

935 Wilks DS (2011) *Statistical Methods in the Atmospheric Sciences*

936 Williams KD, Copsey D, Blockley EW, et al (2018) The Met Office Global Coupled Model 3.0 and 3.1
937 (GC3.0 and GC3.1) Configurations. *J Adv Model Earth Syst* 10:357–380. doi:
938 10.1002/2017MS001115

939 Xie S-P, Kosaka Y, Du Y, et al (2016) Indo-western Pacific ocean capacitor and coherent climate
940 anomalies in post-ENSO summer: A review. *Adv Atmos Sci* 33:411–432. doi: 10.1007/s00376-
941 015-5192-6

942 Xie SP, Zhou ZQ (2017) Seasonal modulations of El Niño-related atmospheric variability: Indo-
943 Western Pacific Ocean feedback. *J Clim* 30:3461–3472. doi: 10.1175/JCLI-D-16-0713.1

944 Yamazaki K, Sexton DMH, Rostron JW, et al (2021) A perturbed parameter ensemble of HadGEM3-
945 GC3.05 coupled model projections: part 2: global performance and future changes. *Clim Dyn*.
946 doi: 10.1007/s00382-020-05608-5

947 Yokohata T, Annan JD, Collins M, et al (2013) Reliability and importance of structural diversity of
948 climate model ensembles. *Clim Dyn* 41:2745–2763. doi: 10.1007/s00382-013-1733-9

949 Zhang W, Furtado K, Wu P, et al (2021) Increasing precipitation variability on daily-to-multiyear time

950 scales in a warmer world. *Sci Adv* 7:1–12. doi: 10.1126/sciadv.abf8021

951 Zhou S, Huang G, Huang P (2018) Changes in the East Asian summer monsoon rainfall under global
952 warming: moisture budget decompositions and the sources of uncertainty. *Clim Dyn* 51:1363–
953 1373. doi: 10.1007/s00382-017-3959-4

954 Zhou T, Lu J, Zhang W, Chen Z (2020) The Sources of Uncertainty in the Projection of Global Land
955 Monsoon Precipitation. *Geophys Res Lett* 47:. doi: 10.1029/2020GL088415

956 Zou L, Zhou T (2011) Sensitivity of a regional ocean-atmosphere coupled model to convection
957 parameterization over western North Pacific. *J Geophys Res Atmos* 116:. doi:
958 <https://doi.org/10.1029/2011JD015844>

959

960

961 **Statements and Declarations**

962

963 **Funding**

964 This work and its contributors (John W. Rostron, David M. H. Sexton, Kalli Furtado, Maria João
965 Carvalho, Sean F. Milton, and José M. Rodríguez) were supported by the UK-China Research &
966 Innovation Partnership Fund through the Met Office Climate Science for Service Partnership (CSSP)
967 China as part of the Newton Fund. Wenxia Zhang was supported by the National Natural Science
968 Foundation of China (41905064).

969

970 **Competing Interests**

971 The authors have no relevant financial or non-financial interests to disclose.

972

973 **Author Contributions**

974 John W. Rostron, David M. H. Sexton, Kalli Furtado, Maria João Carvalho and Sean F. Milton
975 contributed to the study conception and design. John W. Rostron and David M. H. Sexton
976 contributed the model data and performed the analyses. The first draft of the manuscript was
977 written by John W. Rostron. All authors contributed to the interpretation of the results and provided
978 comments and improvements to previous versions of the manuscript.

979

980 **Data Availability**

981 The datasets generated during and/or analysed during the current study are available from the
982 corresponding author on reasonable request.

983

984

985 Figures

986

Fig. 1 Regions of China used in our analysis. The shaded regions were selected using the K-means clustering algorithm described in Section 2.3. These are North and West China (N+W China; orange); Central-East China (CE China; purple) and Southeast China (SE China; green). Additional regions used in our analysis, covering Northeast China (NE China; red box) and Southwest China (SW China; blue box), are also shown.

987

988

Fig. 2 JJA mean 850hPa wind fields. Top-left panel shows the 1980-2015 ERA-Interim climatology, with green shading showing the wind speeds (magnitude of wind vectors). The remaining panels show model biases with respect to ERA-Interim for the same period. The arrows represent the zonal and meridional components of the wind biases, while the shading shows the wind speed bias. The top-right panel shows the mean bias across PPE members. The remaining panels show biases for the best (left column) and worst (right column) models from the PPE (middle row) and CMIP5 (bottom row). These were selected using RSME values for the regions used to calculate the RWFI (red boxes).

989

990

Fig. 3 (a) RWFI mean climatologies (1980-2015) for RWFI and its north and south components (see definition in Section 2.4.1). PPE members (green points) are shown alongside CMIP5 models (orange points) and ERA-Interim values (black crosses). Values for the PPE standard member are shown by the pink points. Panel (b) shows the interannual variability (after removal of a linear trend) over the same period.

991

992

Fig. 4 Time series of the RWFI. The ERA-Interim time series (black line) covers the period 1980-2015, whilst the PPE (green) and CMIP5 (orange) ensembles cover 1900-2100. The shaded envelopes show the ensemble minima and maxima for each year. The ensemble means are shown by darker lines. We also include the time series for the PPE standard model (pink line) to highlight the typical interannual variability in the PPE.

993

994

Fig. 5 Annual cycles of precipitation for regions in China. Monthly climatologies (based on 1980-2014 means) are shown for observations from GPCP (black lines) and for each PPE member. The colours used for each PPE member indicate the month when CE China precipitation reaches its maximum: red indicates a maximum in June; grey, a maximum in July and blue, a maximum in August.

995

996

Fig. 6 (a) JJA mean precipitation climatologies (1980-2014) for regions in China (see Section 2.3 for definitions). PPE members (blue points) are shown alongside CMIP5 models (orange points) and observations from GPCP (black crosses). Values for the PPE standard member are shown by the pink points. Panel (b) shows the interannual variability (after removal of a linear trend) over the same period.

997

998

Fig. 7 Time series of JJA mean precipitation for regions in China. Observations (black lines) are from GPCP and cover the period 1980-2014. The PPE (blue) and CMIP5 (orange) ensembles cover 1900-2100. The shaded envelopes show the ensemble minima and maxima for each year. The ensemble means are shown by darker lines. We also include the time series for the PPE standard model (pink line) to highlight the typical interannual variability in the PPE in each region.

Fig. 8 Relationships between interannual anomalies in JJA mean precipitation (for CE China), RWFI and Niño3.4, based on the period 1980-2014. Values for the gradient of the relationship between CE China precipitation and RWFI are plotted on the y-axis, whilst RWFI-Niño3.4 gradient values are plotted on the x-axis. The black cross is for the observed/reanalysis values GPCP and Era-Interim, while the PPE members are shown as blue points (with labels for each PPE member shown in the legend). The black error bars show the uncertainties capturing the 95% confidence range on these gradients for observations/reanalysis. The blue error bars show the equivalent for the mean uncertainty across the PPE members.

999

1000

Fig. 9 Changes in JJA mean 850hPa winds for 2070-2099 vs 1980-2009. The arrows show the changes in the zonal and meridional components of the wind, while the shading shows the change in the wind speed. The top four panels are for the PPE, showing changes for the PPE mean (top-left) and members with the best circulation performance (0834, top-right); the smallest RWFI change (2242, middle-left) and the largest RWFI change (1113, middle-right). The latter member also has the worst circulation performance. The bottom two panels are for the CMIP5 members with the smallest and largest RWFI changes (MPI-ESM-MR and HadGEM2-ES, respectively).

1001

1002

Fig. 10 (a) Change in the mean for RWFI and its north and south components, for 2070-2099 vs 1980-2009. PPE members are shown in green, while CMIP5 models are shown in orange. The PPE standard member is shown in pink. (b) Equivalent to (a), but for the change in the interannual variability of RWFI and its components (after the removal of linear trends for the two periods).

1003

1004

Fig. 11 The change in mean RWFI (for 2070-2099 vs 1980-2009) vs present-day mean RWFI values (1980-2015) for the PPE (green points) and CMIP5 models (orange points). The present-day value from the ERA-Interim reanalysis is shown with a vertical black line.

1005

1006

Fig. 12 (a) Change in JJA mean precipitation for regions in China (for 2070-2099 vs 1980-2009) for the PPE (blue points) and CMIP5 models (orange points). The PPE standard member is shown in pink. (b) Equivalent to (a), but with the change expressed as a percentage change from the reference period (1980-2009). (c) Change in the interannual variability of precipitation (after the removal of linear trends for the two periods).

1007

1008

Fig. 13 Relationships between JJA mean precipitation in CE China (prC) and RWFI, compared for different time periods. Gradients of prC vs RWFI are shown for 4 50-year time periods: (a) 1989-2039, (b) 2009-2059, (c) 2029-2079 and (d) 2049-2099, and each are plotted against the gradient for 1969-2019. Each point represents a single PPE member. The gradients were evaluated using detrended data, where 35-year running means were first subtracted from the prC and RWFI time series data. Detrending was applied separately for each PPE member.

1009

1010

Fig. 14 Changes in moisture budget components for CE China, based on JJA means for 2070-2099 vs 1980-2009. Values are shown for changes in precipitation (DP); precipitation minus evaporation (DP-DE); moisture convergence (DconvQ); the thermodynamic (DTH) and 'dynamic' contributions (DDYN); and a residual term (DRES) for the discrepancy between DP-DE and DTH+DDYN. The details of the calculation of these components are given in Section 2.5. Each grey point represents a single PPE member. Examples from the discussion on sub-selection of the PPE (Section 5) are highlighted in colour.

1011

Fig. 15 Changes in precipitation minus evaporation (DP-DE) for CE China are plotted against changes in the thermodynamic (a); 'dynamic' (b) and mean-circulation dynamics (c) components of the moisture budget for CE China. (d) shows the changes against the RWFI. Changes are based on JJA means for 2070-2099 vs 1980-2009. The details of these calculations are given in Sections 2.4.1 and 2.5. The black point and error bar shows the mean and standard deviation of the residuals from the moisture budget analysis (DRES) to give an indication of the limit of confidence in the moisture budget component values. The remaining data are as described for Fig. 14. An estimate of the fraction of the variance in DP-DE explained is given in each case using the square of the Pearson correlation coefficient (R^2).

1012

1013

Fig. 16 Changes in the thermodynamic (DTH) vs the dynamic (DDYN) components of the moisture budget for CE China (using JJA means for the periods 2070-2099 vs 1980-2009; see Section 2.5 for details). The data shown are as described for Fig. 15. Grey diagonals are lines of constant DTH+DDYN. Given small values for DRES and DE (see Fig. 14), these provides an estimate for changes in the moisture convergence (DconvQ) and precipitation (DP). The red line indicates where DTH and DDYN contribute equally to the moisture budget changes.

1014

1015

Fig. 17 (a) and (b) show maps of the changes in the thermodynamic (DTH) component of the moisture budget, overlaid with climatological 850hPa winds, for members 2242 and 2884, respectively. (c) and (d) show changes in the dynamic component of the moisture budget overlaid with changes in the 850hPa winds for the same members. JJA means are shown in each case, with climatologies based on the period 1980-2009 and changes based on 2070-2099 vs 1980-2009. (See Section 2.5 for details of the moisture budget calculations.)

1016

Figures

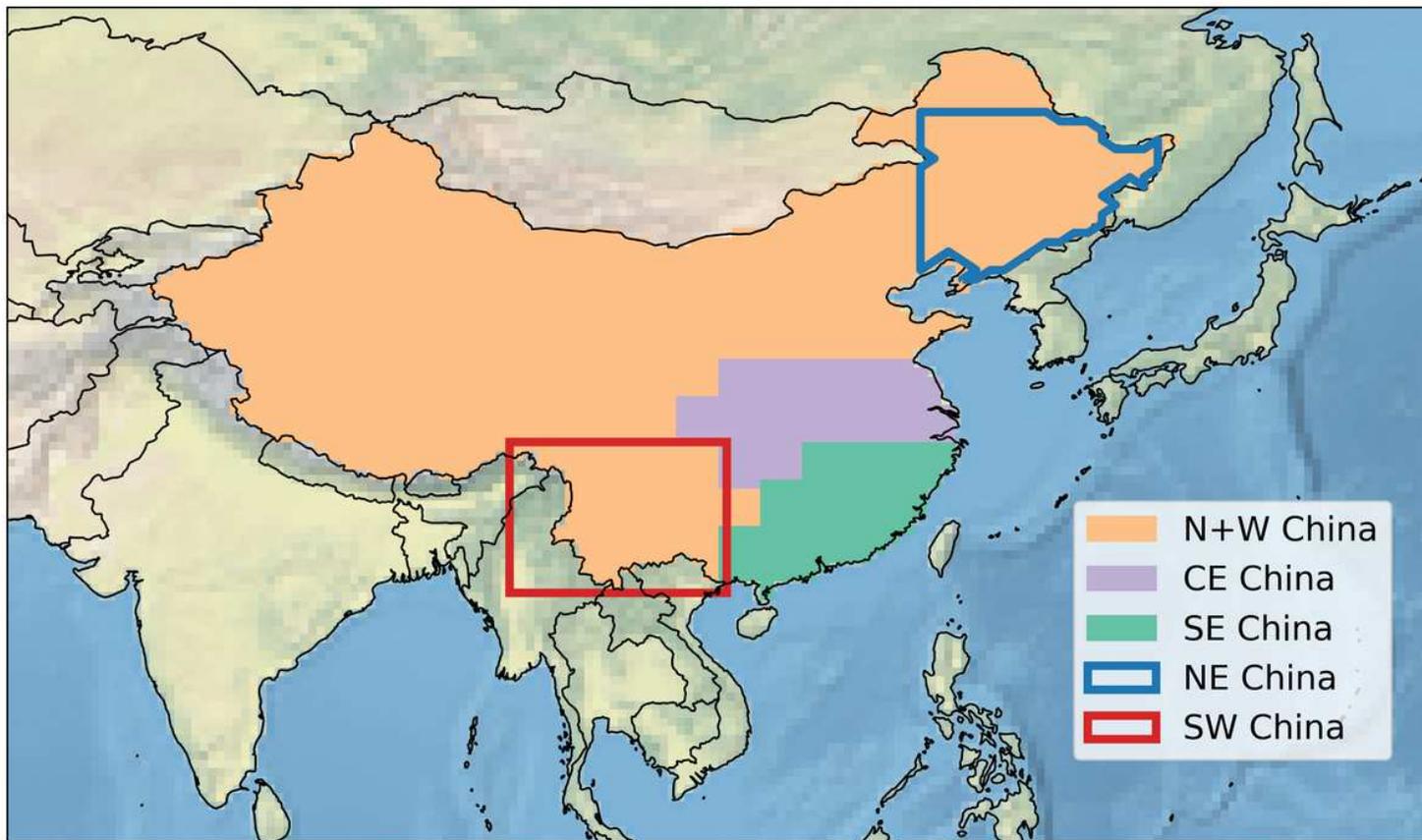


Figure 1

Regions of China used in our analysis. The shaded regions were selected using the K-means clustering algorithm described in Section 2.3. These are North and West China (N+W China; orange); Central-East China (CE China; purple) and Southeast China (SE China; green). Additional regions used in our analysis, covering Northeast China (NE China; red box) and Southwest China (SW China; blue box), are also shown.

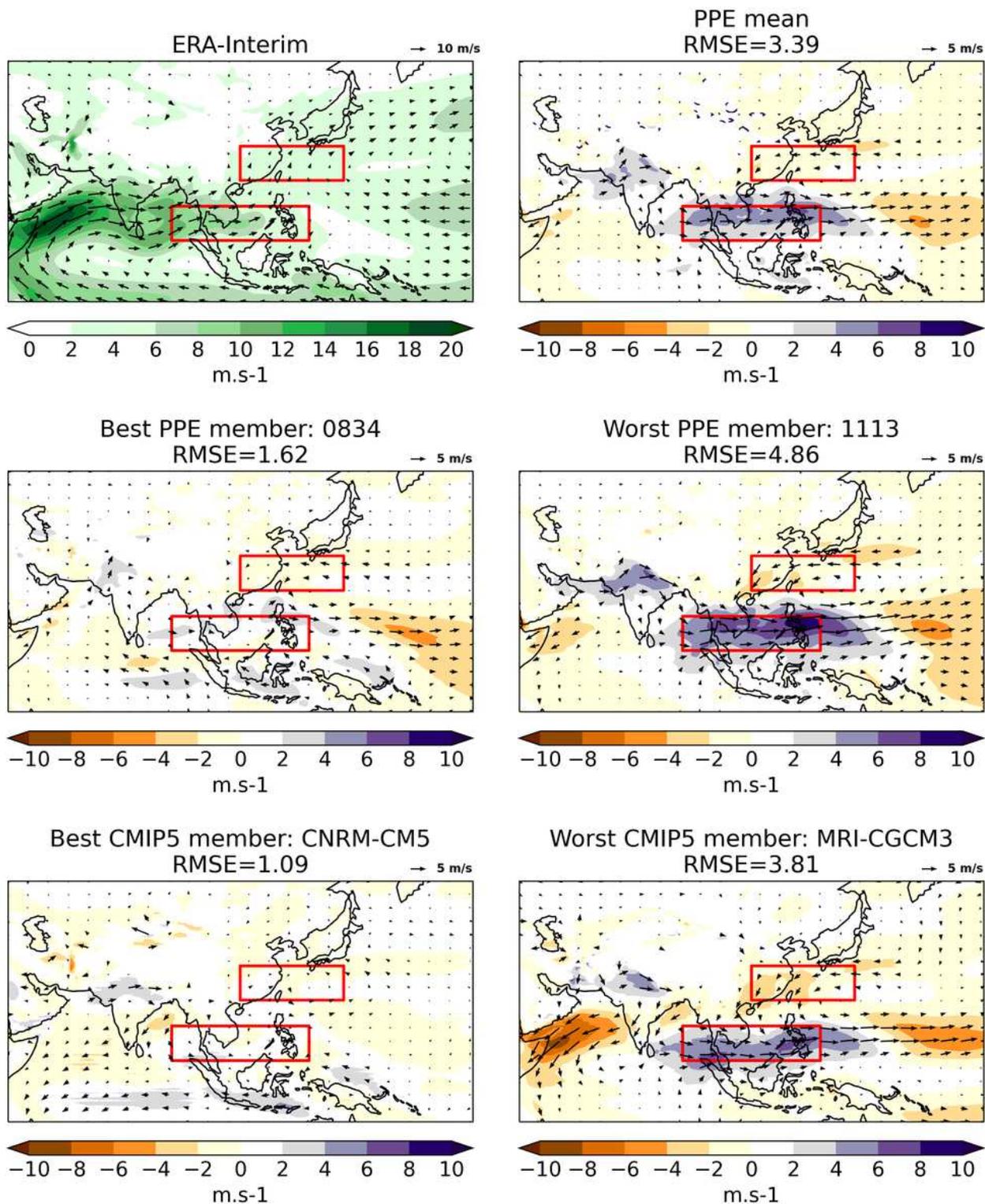


Figure 2

JJA mean 850hPa wind fields. Top-left panel shows the 1980-2015 ERA-Interim climatology, with green shading showing the wind speeds (magnitude of wind vectors). The remaining panels show model biases with respect to ERA-Interim for the same period. The arrows represent the zonal and meridional components of the wind biases, while the shading shows the wind speed bias. The top-right panel shows the mean bias across PPE members. The remaining panels show biases for the best (left column) and

worst (right column) models from the PPE (middle row) and CMIP5 (bottom row). These were selected using RSME values for the regions used to calculate the RWFI (red boxes).

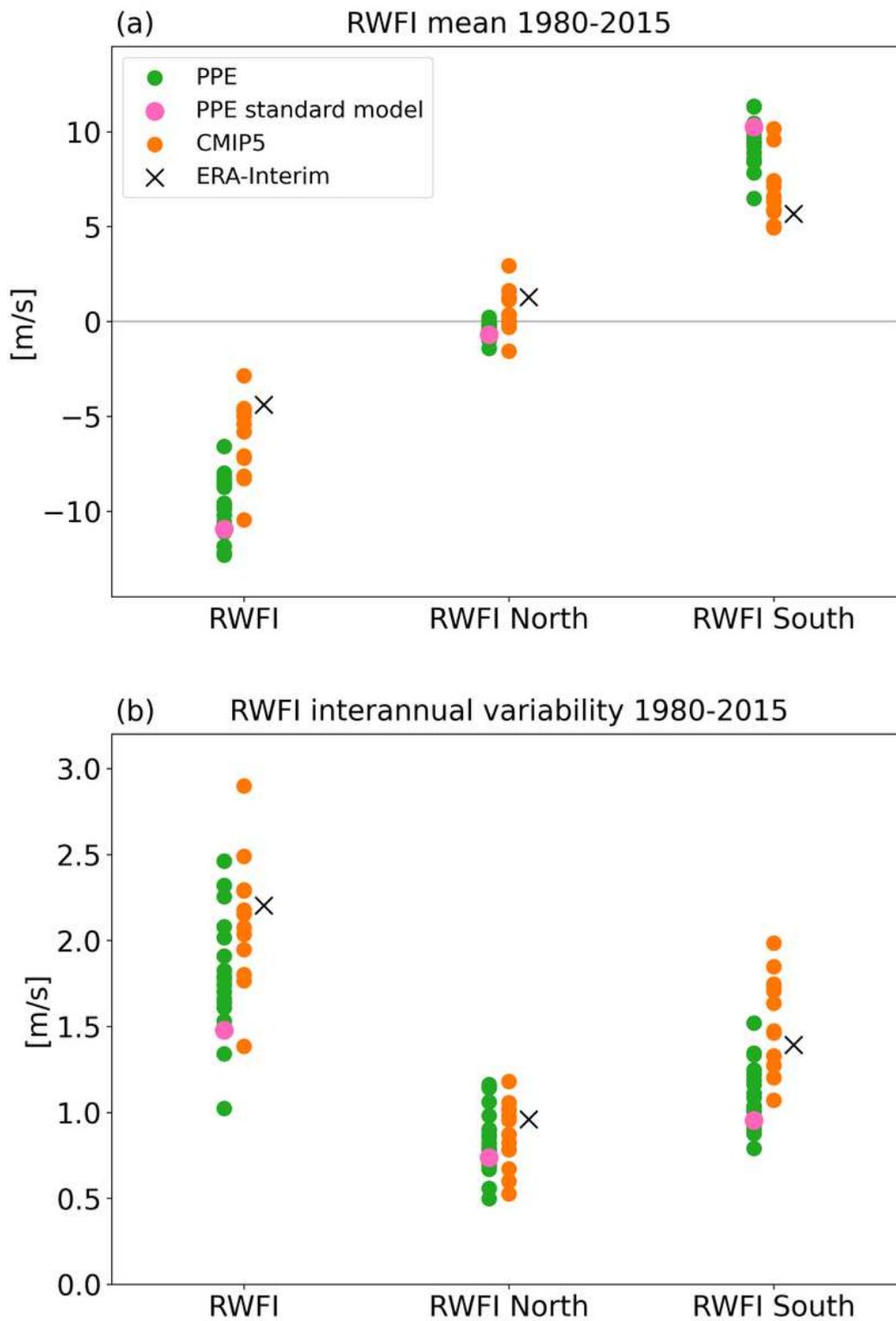


Figure 3

(a) RWFI mean climatologies (1980-2015) for RWFI and its north and south components (see definition in Section 2.4.1). PPE members (green points) are shown alongside CMIP5 models (orange points) and

ERA-Interim values (black crosses). Values for the PPE standard member are shown by the pink points. Panel (b) shows the interannual variability (after removal of a linear trend) over the same period.

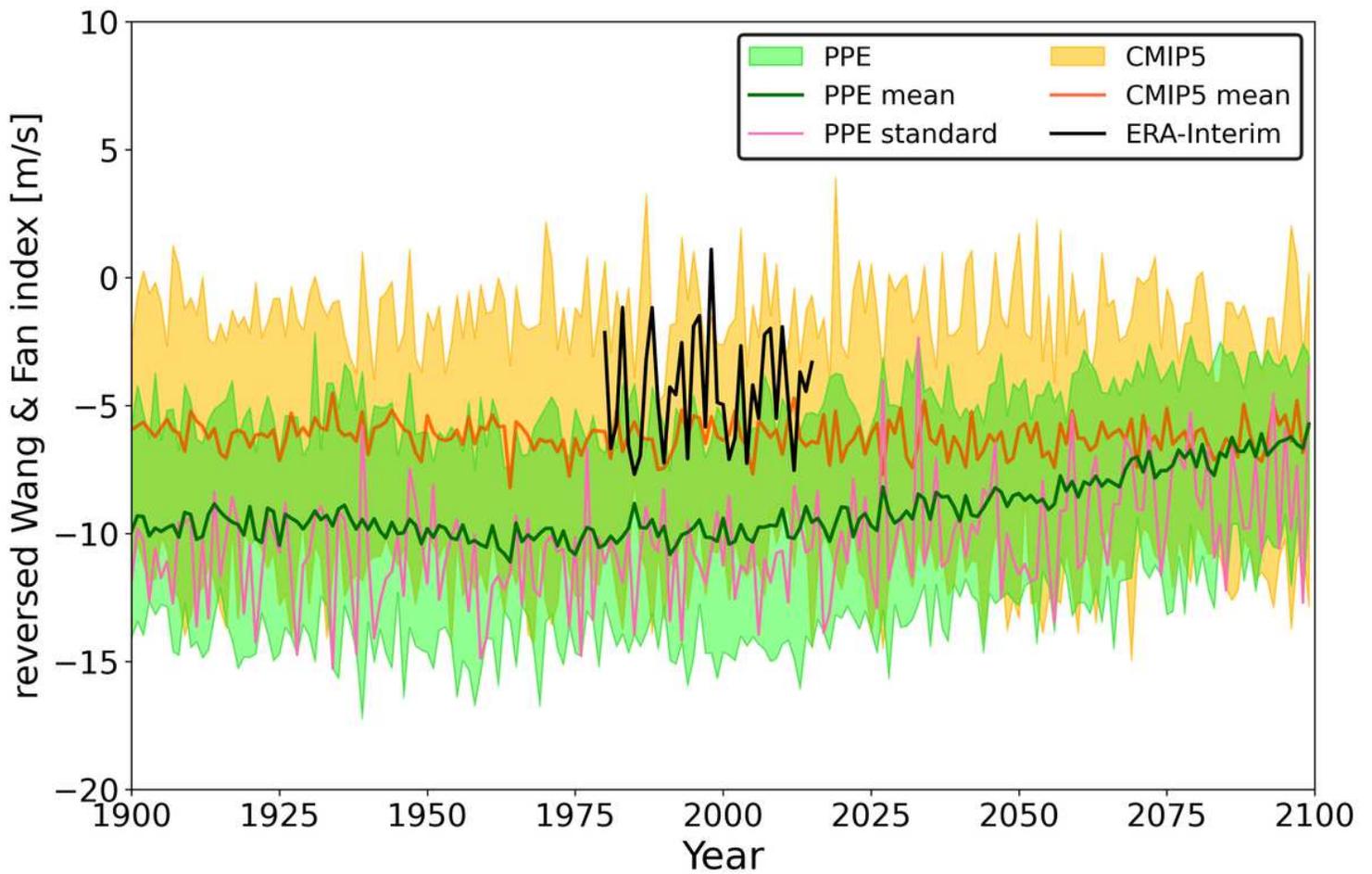


Figure 4

Time series of the RWFI. The ERA-Interim time series (black line) covers the period 1980-2015, whilst the PPE (green) and CMIP5 (orange) ensembles cover 1900-2100. The shaded envelopes show the ensemble minima and maxima for each year. The ensemble means are shown by darker lines. We also include the time series for the PPE standard model (pink line) to highlight the typical interannual variability in the PPE.

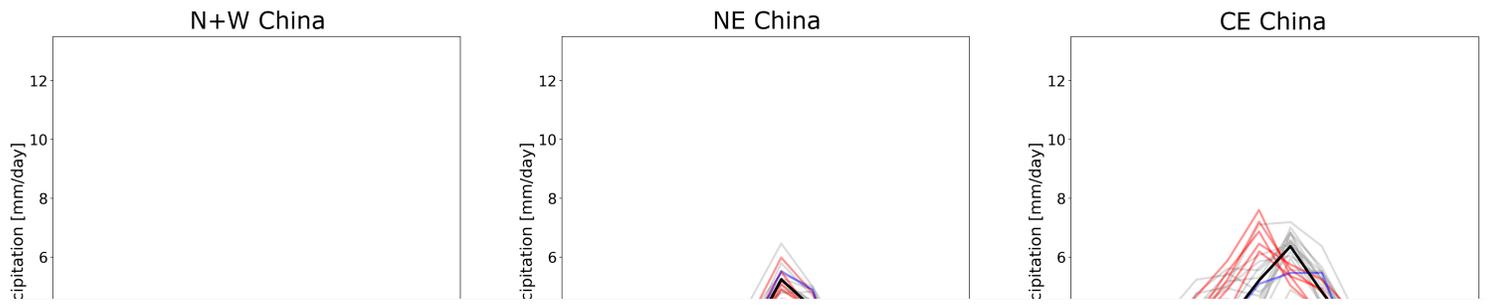


Figure 5

Annual cycles of precipitation for regions in China. Monthly climatologies (based on 1980-2014 means) are shown for observations from GPCP (black lines) and for each PPE member. The colours used for each PPE member indicate the month when CE China precipitation reaches its maximum: red indicates a maximum in June; grey, a maximum in July and blue, a maximum in August.

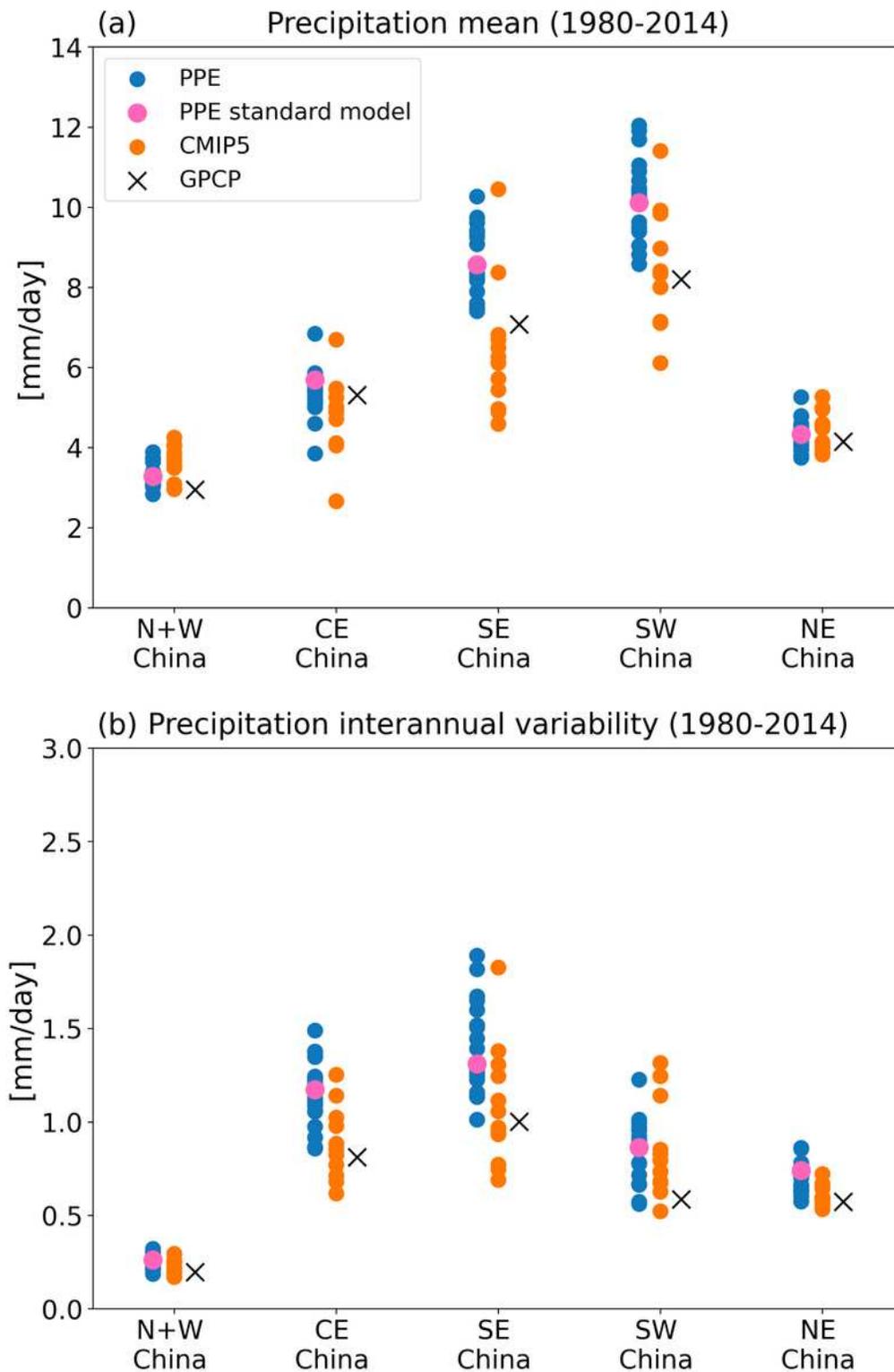


Figure 6

(a) JJA mean precipitation climatologies (1980-2014) for regions in China (see Section 2.3 for definitions). PPE members (blue points) are shown alongside CMIP5 models (orange points) and observations from GPCP (black crosses). Values for the PPE standard member are shown by the pink points. Panel (b) shows the interannual variability (after removal of a linear trend) over the same period.

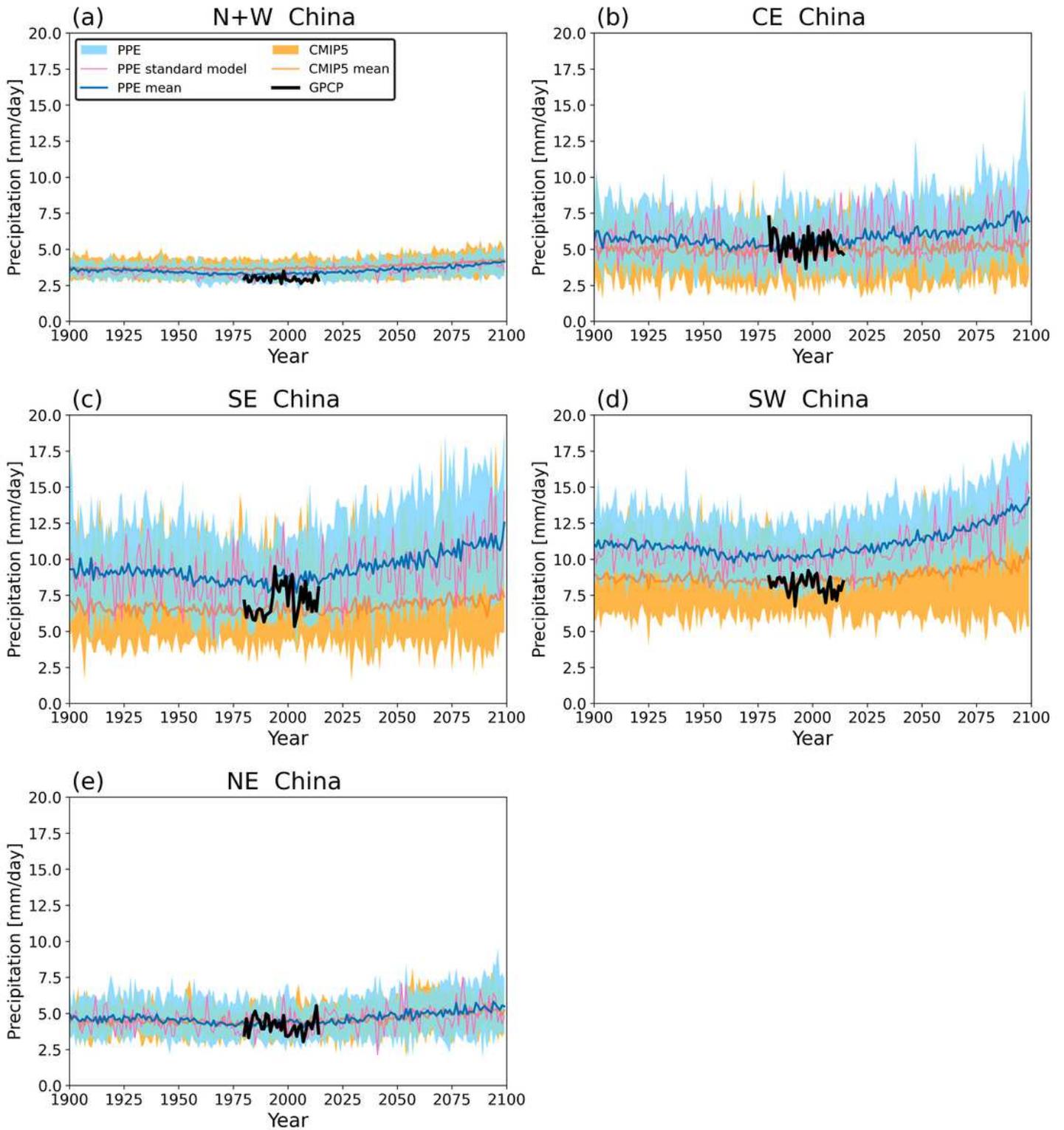


Figure 7

Time series of JJA mean precipitation for regions in China. Observations (black lines) are from GPCP and cover the period 1980-2014. The PPE (blue) and CMIP5 (orange) ensembles cover 1900-2100. The shaded envelopes show the ensemble minima and maxima for each year. The ensemble means are shown by darker lines. We also include the time series for the PPE standard model (pink line) to highlight the typical interannual variability in the PPE in each region.

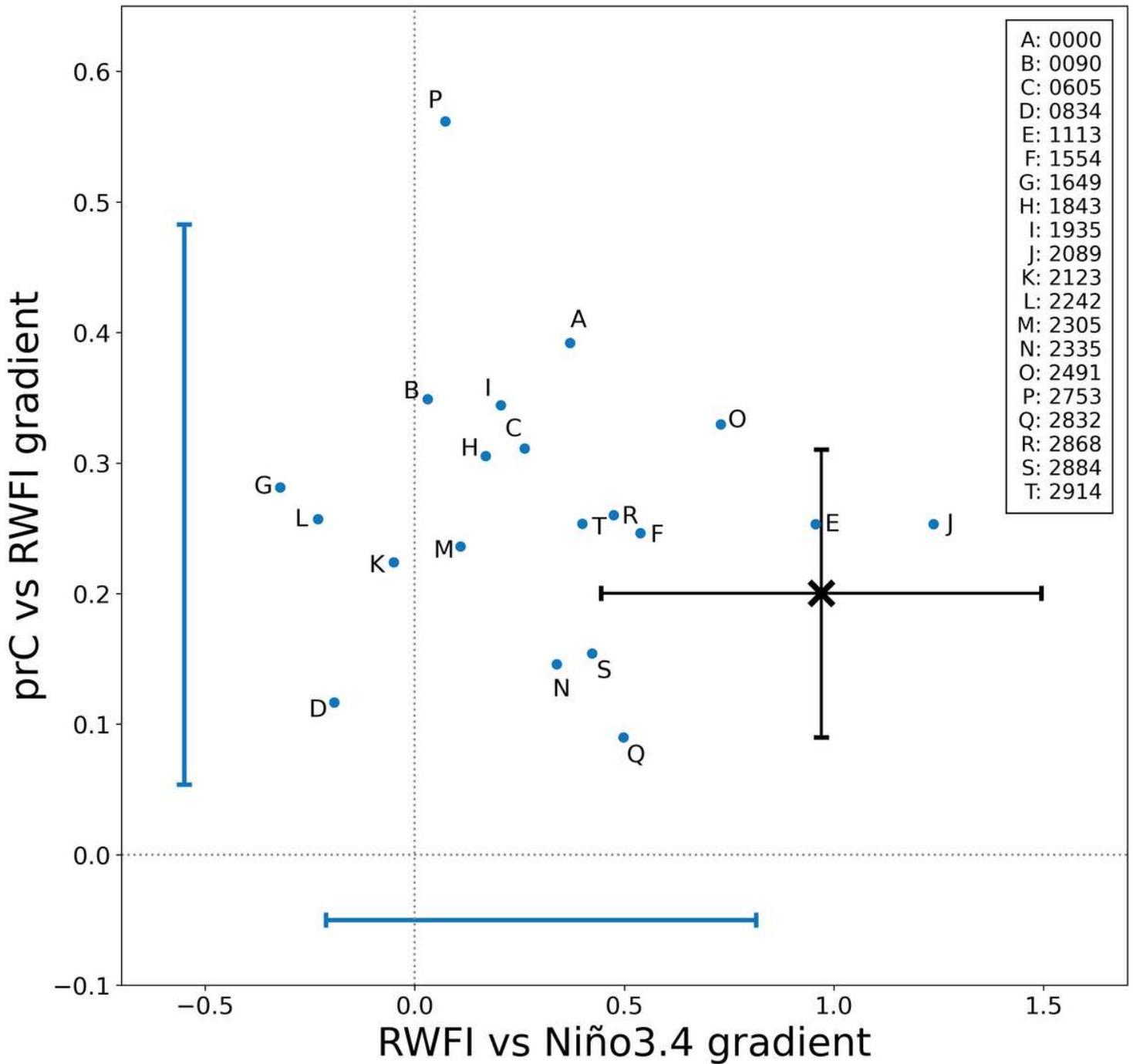


Figure 8

Relationships between interannual anomalies in JJA mean precipitation (for CE China), RWFI and Niño3.4, based on the period 1980-2014. Values for the gradient of the relationship between CE China precipitation and RWFI are plotted on the y-axis, whilst RWFI-Niño3.4 gradient values are plotted on the x-axis. The black cross is for the observed/reanalysis values GPCP and Era-Interim, while the PPE members are shown as blue points (with labels for each PPE member shown in the legend). The black error bars show the uncertainties capturing the 95% confidence range on these gradients for observations/reanalysis. The blue error bars show the equivalent for the mean uncertainty across the PPE members.

Figure 9

Changes in JJA mean 850hPa winds for 2070-2099 vs 1980-2009. The arrows show the changes in the zonal and meridional components of the wind, while the shading shows the change in the wind speed. The top four panels are for the PPE, showing changes for the PPE mean (top-left) and members with the best circulation performance (0834, top-right); the smallest RWFI change (2242, middle-left) and the largest RWFI change (1113, middle-right). The latter member also has the worst circulation performance. The bottom two panels are for the CMIP5 members with the smallest and largest RWFI changes (MPI-ESM-MR and HadGEM2-ES, respectively).

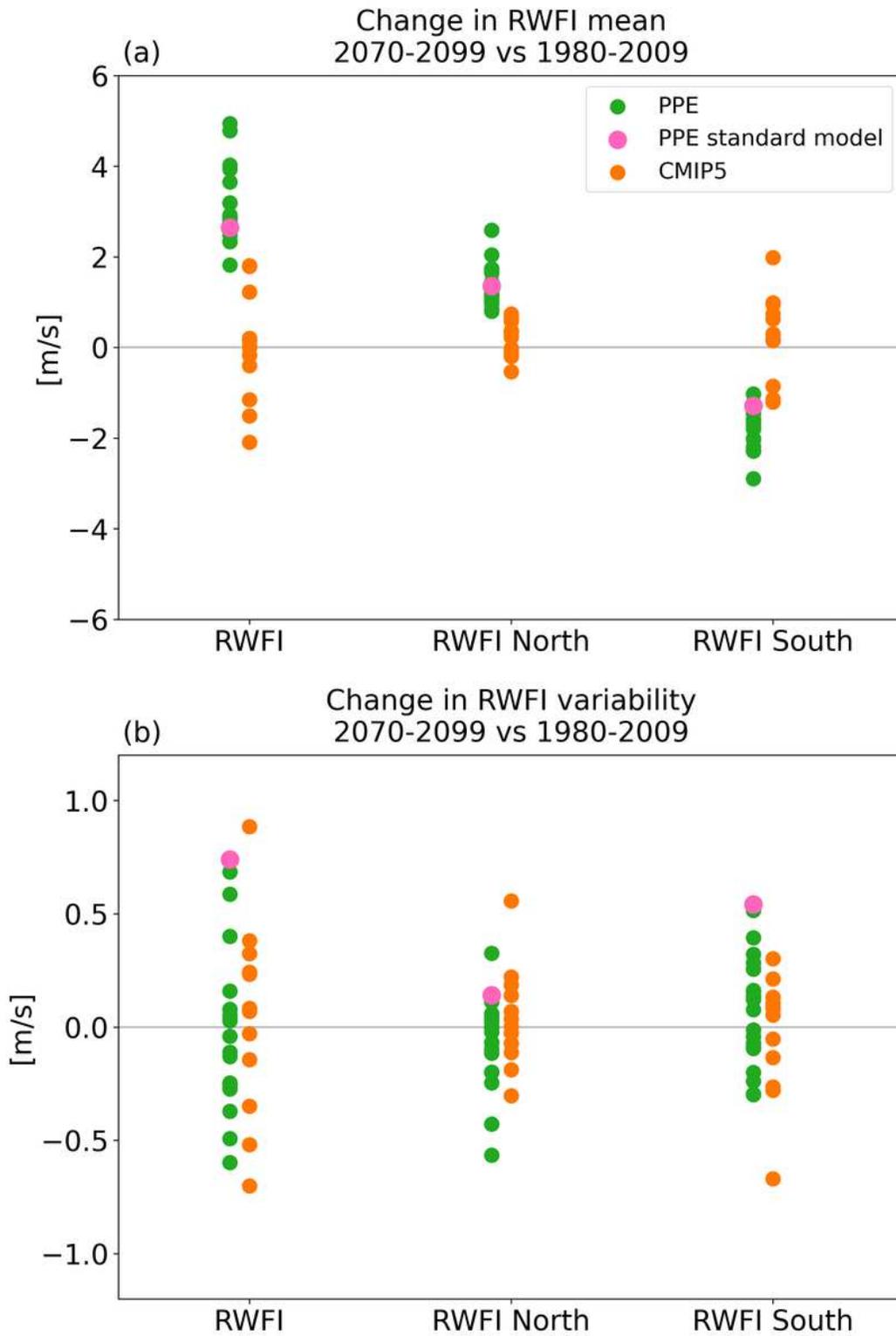


Figure 10

(a) Change in the mean for RWFI and its north and south components, for 2070-2099 vs 1980-2009. PPE members are shown in green, while CMIP5 models are shown in orange. The PPE standard member is shown in pink. (b) Equivalent to (a), but for the change in the interannual variability of RWFI and its components (after the removal of linear trends for the two periods).

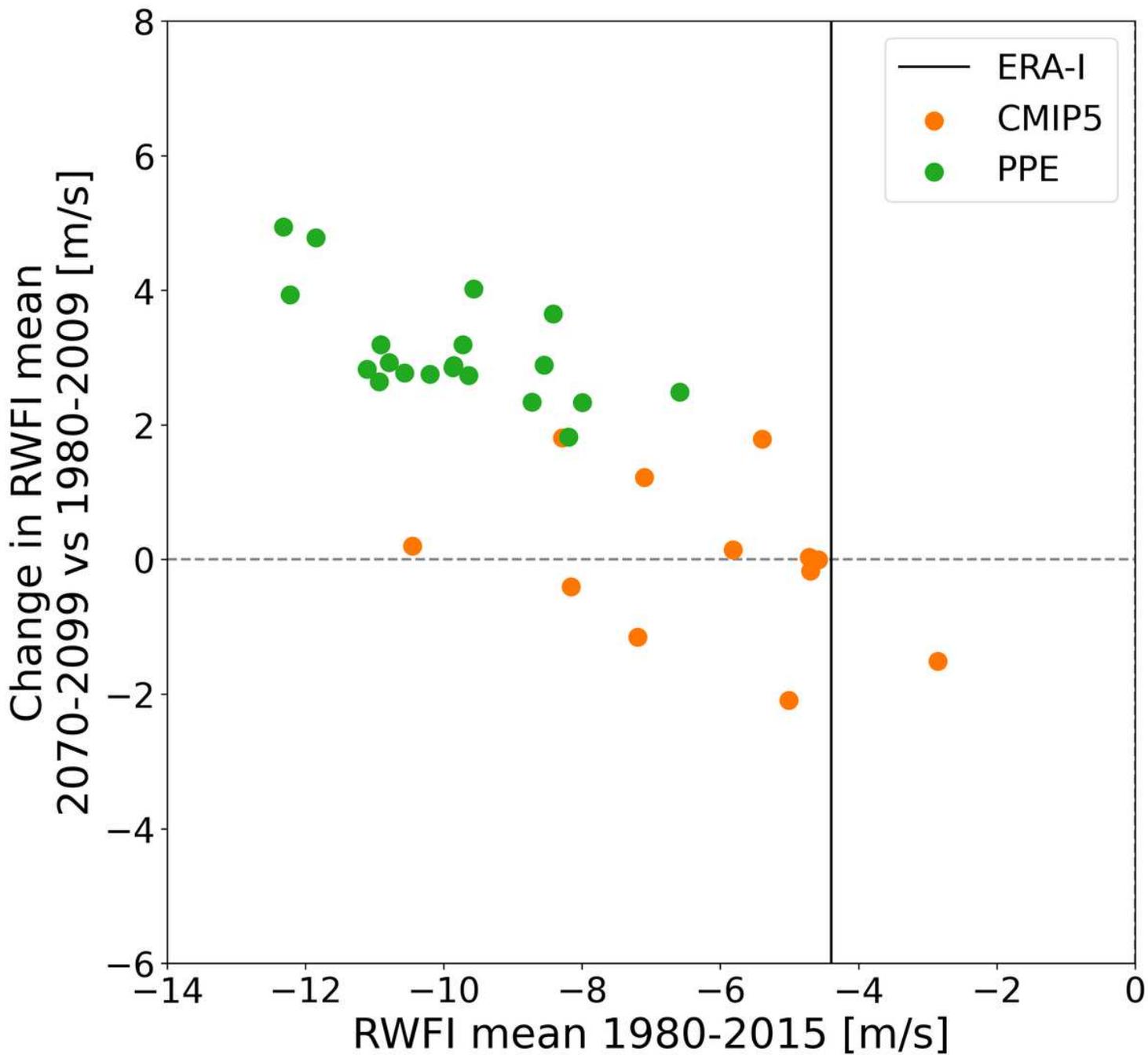


Figure 11

The change in mean RWFI (for 2070-2099 vs 1980-2009) vs present-day mean RWFI values (1980-2015) for the PPE (green points) and CMIP5 models (orange points). The present-day value from the ERA-Interim reanalysis is shown with a vertical black line.

Figure 12

(a) Change in JJA mean precipitation for regions in China (for 2070-2099 vs 1980-2009) for the PPE (blue points) and CMIP5 models (orange points). The PPE standard member is shown in pink. (b) Equivalent to (a), but with the change expressed as a percentage change from the reference period (1980-2009). (c) Change in the interannual variability of precipitation (after the removal of linear trends for the two periods).

Figure 13

Relationships between JJA mean precipitation in CE China (prC) and RWF, compared for different time periods. Gradients of prC vs RWF are shown for 4 50-year time periods: (a) 1989-2039, (b) 2009-2059, (c) 2029-2079 and (d) 2049-2099, and each are plotted against the gradient for 1969-2019. Each point represents a single PPE member. The gradients were evaluated using detrended data, where 35-year running means were first subtracted from the prC and RWF time series data. Detrending was applied separately for each PPE member.

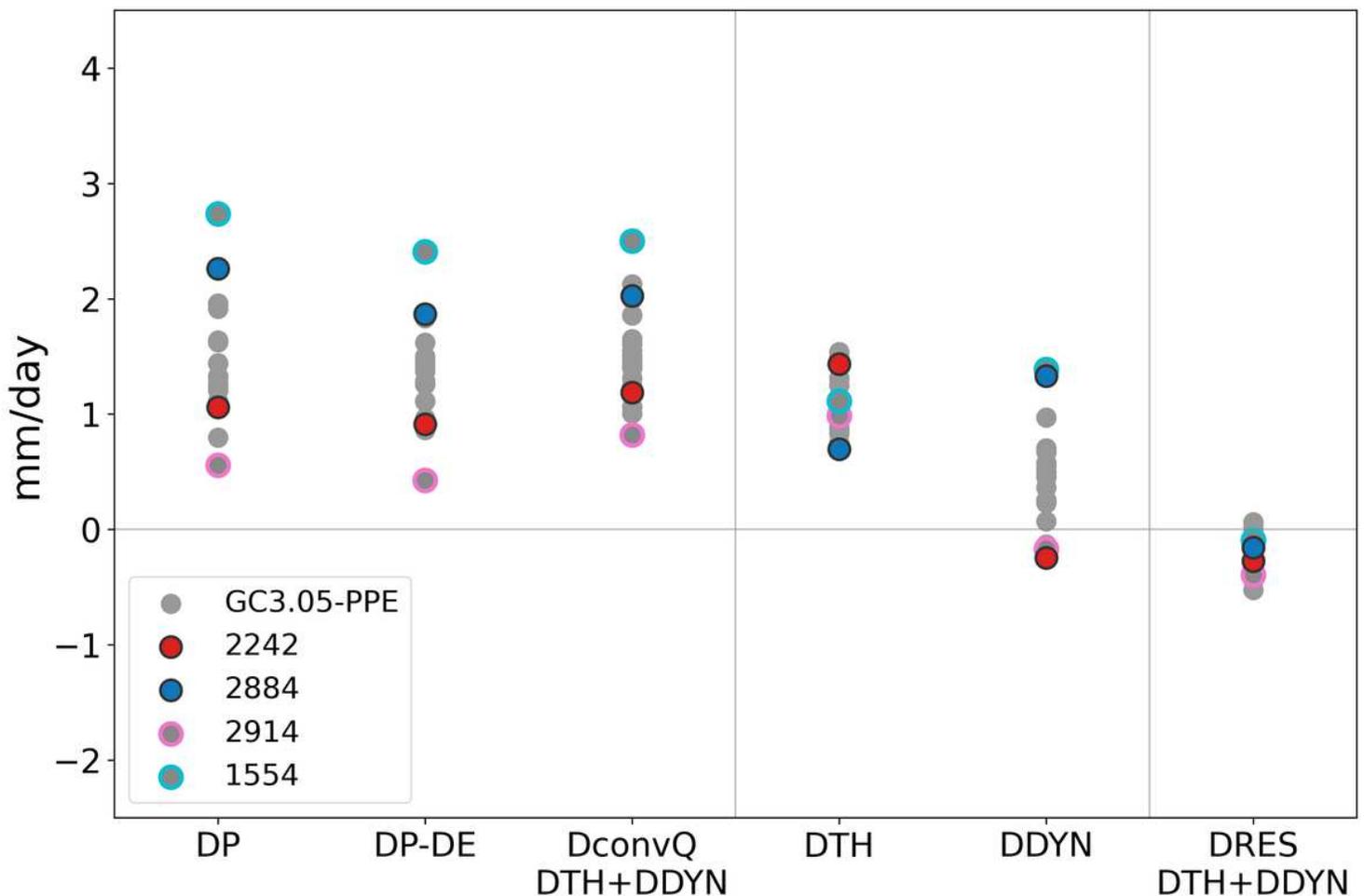


Figure 14

Changes in moisture budget components for CE China, based on JJA means for 2070-2099 vs 1980-2009. Values are shown for changes in precipitation (DP); precipitation minus evaporation (DP-DE); moisture convergence (DconvQ); the thermodynamic (DTH) and 'dynamic' contributions (DDYN); and a residual term (DRES) for the discrepancy between DP-DE and DTH+DDYN. The details of the calculation of these components are given in Section 2.5. Each grey point represents a single PPE member. Examples from the discussion on sub-selection of the PPE (Section 5) are highlighted in colour.

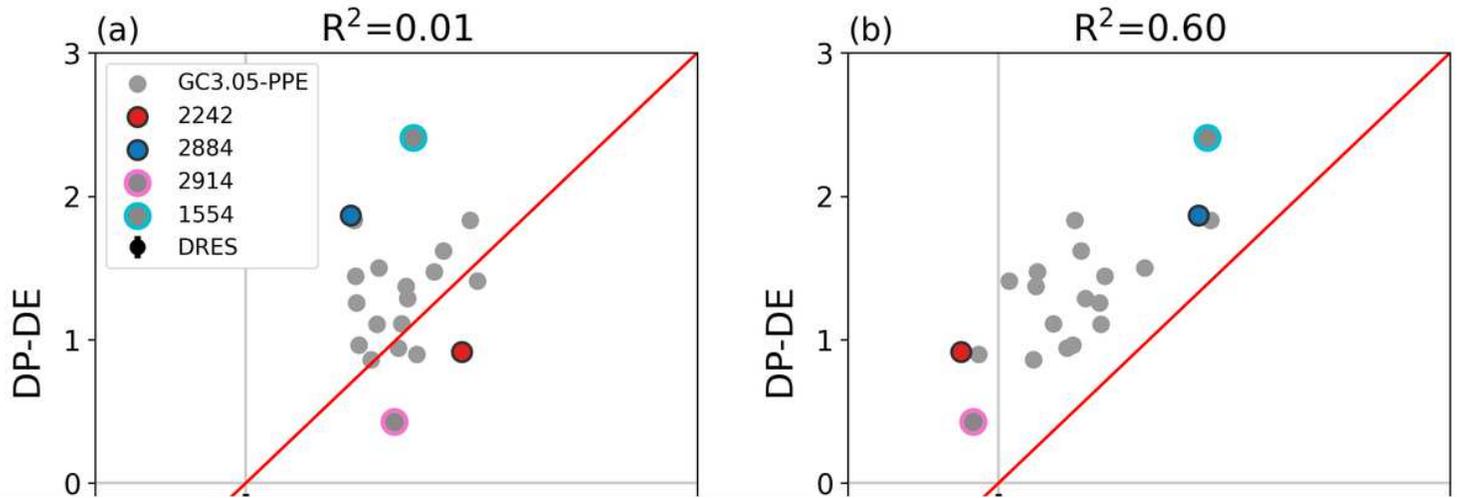


Figure 15

Changes in precipitation minus evaporation (DP-DE) for CE China are plotted against changes in the thermodynamic (a); 'dynamic' (b) and mean-circulation dynamics (c) components of the moisture budget for CE China. (d) shows the changes against the RWFI. Changes are based on JJA means for 2070-2099 vs 1980-2009. The details of these calculations are given in Sections 2.4.1 and 2.5. The black point and error bar shows the mean and standard deviation of the residuals from the moisture budget analysis (DRES) to give an indication of the limit of confidence in the moisture budget component values. The remaining data are as described for Fig. 14. An estimate of the fraction of the variance in DP-DE explained is given in each case using the square of the Pearson correlation coefficient (R^2).

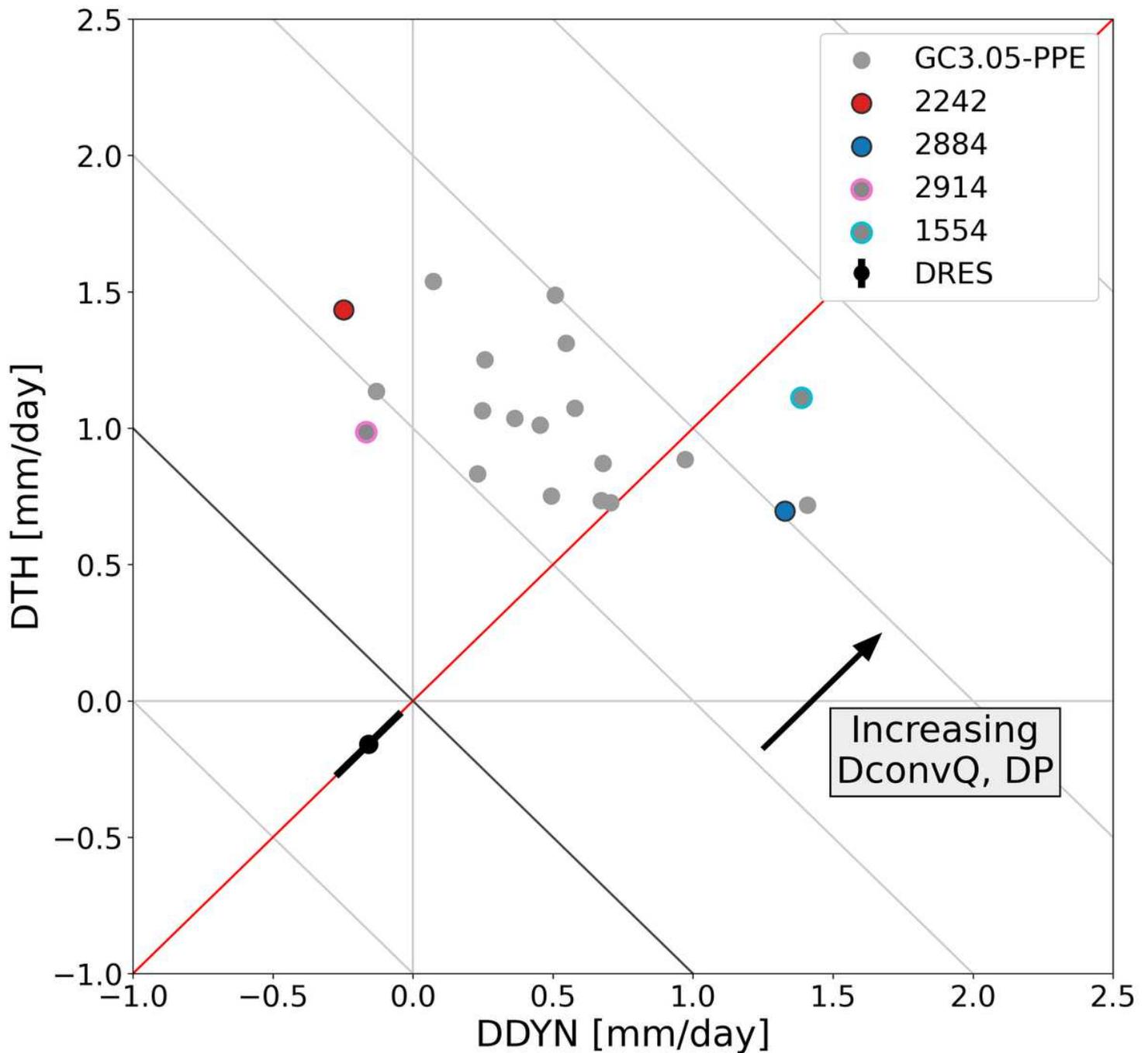


Figure 16

Changes in the thermodynamic (DTH) vs the dynamic (DDYN) components of the moisture budget for CE China (using JJA means for the periods 2070-2099 vs 1980-2009; see Section 2.5 for details). The data shown are as described for Fig. 15. Grey diagonals are lines of constant DTH+DDYN. Given small values for DRES and DE (see Fig. 14), these provides an estimate for changes in the moisture convergence (DconvQ) and precipitation (DP). The red line indicates where DTH and DDYN contribute equally to the moisture budget changes.

Figure 17

(a) and (b) show maps of the changes in the thermodynamic (DTH) component of the moisture budget, overlaid with climatological 850hPa winds, for members 2242 and 2884, respectively. (c) and (d) show changes in the dynamic component of the moisture budget overlaid with changes in the 850hPa winds for the same members. JJA means are shown in each case, with climatologies based on the period 1980-2009 and changes based on 2070-2099 vs 1980-2009. (See Section 2.5 for details of the moisture budget calculations.)