1

2

3 **Atmospheric impacts of Arctic sea-ice loss, 1979-2009: Separating forced change from**

4 **atmospheric internal variability**

5

6

7 James A Screen[1,2], Clara Deser[3], Ian Simmonds[2] and Robert Tomas[3]

8

9 *[1]College of Engineering, Mathematics and Physical Sciences, University of Exeter, Exeter, UK*

10 *[2]School of Earth Sciences, University of Melbourne, Melbourne, Victoria, Australia*

11 *[3]Climate and Global Dynamics, National Center for Atmospheric Research, Boulder, Colorado, US*

12

13

14

15

16 Revised manuscript for *Climate Dynamics*

17 14 May 2013

18

19

20

23

24 Author correspondence address: J. Screen, College of Engineering, Mathematics and Physical

25 Sciences, Harrison Building, Streatham Campus, University of Exeter, North Park Road, Exeter,

26 Devon, EX4 4QF, UK (j.screen@exeter.ac.uk).

27

28   **Abstract**

29

30   The ongoing loss of Arctic sea-ice cover has implications for the wider climate system. The

31   detection and importance of the atmospheric impacts of sea-ice loss depends, in part, on the relative

32   magnitudes of the sea-ice forced change compared to natural atmospheric internal variability (AIV).

33   This study analyses large ensembles of two independent atmospheric general circulation models in

34   order to separate the forced response to historical Arctic sea-ice loss (1979-2009) from AIV, and to

35   quantify signal-to-noise ratios. We also present results from a simulation with the sea-ice forcing

36   roughly doubled in magnitude. In proximity to regions of sea-ice loss, we identify statistically

37   significant near-surface atmospheric warming and precipitation increases, in autumn and winter in

38   both models. In winter, both models exhibit a significant lowering of sea level pressure and

39   geopotential height over the Arctic. All of these responses are broadly similar, but strengthened

40   and/or more geographically extensive, when the sea-ice forcing is doubled in magnitude. Signal-to-

41   noise ratios differ considerably between variables and locations. The temperature and precipitation

42   responses are significantly easier to detect (higher signal-to-noise ratio) than the sea level pressure

43   or geopotential height responses. Equally, the local response (i.e., in the vicinity of sea-ice loss) is

44   easier to detect than the mid-latitude or upper-level responses. Based on our estimates of signal-to-

45   noise, we conjecture that the local near-surface temperature and precipitation responses to past

46   Arctic sea-ice loss exceed AIV and are detectable in observed records, but that the potential

47   atmospheric circulation, upper-level and remote responses may be partially or wholly masked by

48   AIV.

Introduction

One of the clearest manifestations of recent climate change is the loss of summer and autumn sea-ice cover in the Arctic (Stroeve et al., 2011). During the 2012 melt season, the Arctic sea-ice extent shrunk to the lowest value in the satellite record, which began in 1979 (Zhang et al., 2013; Parkinson and Comiso, 2013). Especially rapid sea-ice melt occurred during August 2012 at the time of a ferocious storm (Simmonds and Rudeva, 2012), though model hindcasts suggest that a new sea-ice minimum would have been recorded even without this storm (Zhang et al., 2013). The last six years (2007-2012) have witnessed the six lowest September sea-ice extents on record, possibly suggesting a "tipping point" has been passed (Livina and Lenton, 2013). Recent dramatic sea-ice reductions augment longer-term trends, but statistically significant sea-ice extent reductions are apparent in all calendar months even if the last six years are excluded (Kay et al., 2011).

The ongoing retreat of Arctic sea-ice has implications for the climate system. In order to better understand these, a number of studies have perturbed sea-ice conditions in atmospheric general circulation models (AGCMs) and examined the atmospheric response (e.g,. Singarayer et al., 2006; Seierstad and Bader, 2009; Deser et al., 2010; Strey et al., 2010; Bluthgen et al., 2012; Orsolini et al., 2012; Ghatak et al., 2012; Porter et al., 2012; Screen et al., 2012; 2013). In a model setting, the sea-ice cover can be manipulated in a controlled manner to reveal how and by what processes it affects the wider climate system. These studies have identified some robust and reasonably well-understood features of the local atmospheric response to sea-ice loss (i.e., impacts proximate to regions of sea-ice loss). These include warming and moistening of the lower troposphere and increases in cloud cover and precipitation.

The impacts of Arctic sea-ice loss may not be limited to the high latitudes. Increasing attention is now turning to the potential remote impacts of Arctic sea-ice loss, including possible changes in

75 mid-latitude weather (Honda et al., 2009; Petoukhov and Semenov, 2012; Liu et al., 2012; Francis

76 and Vavrus, 2012; Screen and Simmonds, 2013a,b). Progress in understanding the potential large-

77 scale or remote impacts of Arctic sea-ice loss is hampered by large uncertainties in the atmospheric

78 circulation response to sea-ice loss. Observational studies suggest links between autumn sea-ice loss

79 and circulation patterns in the following winter (Francis et al., 2009; Overland and Wang, 2010; Wu

80 and Zhang, 2010: Strong et al., 2010; Jaiser et al., 2012), but the statistical significance of these

81 linkages has been questioned (Hopsch et al., 2012), causality is unclear and the mechanisms are

82 poorly understood. In model simulations, the spatial pattern, strength, statistical significance and

83 timing of the circulation response to sea-ice loss differs considerably between studies, and can be

84 hard to disentangle from atmospheric internal variability (AIV). AIV, also known as "climate

85 noise", arises from non-linear dynamical processes intrinsic to the atmosphere (see, e.g., Deser et

86 al., 2012 and references therein).

87

88 In an attempt to better separate, and quantify, the potential forced response to Arctic sea-ice loss

89 and AIV, this manuscript presents results from large ensembles with two independent models, in

90 which the only prescribed forcing was observed Arctic sea-ice loss. Both models have been run

91 multiple times with identical surface boundary conditions and external forcing, with each run

92 beginning from a different atmospheric initial state. Therefore, the differences between the

93 simulated atmospheric states of each of the ensemble members arise only due to AIV. These

94 ensembles are approximately a factor of ten larger than those used in Screen et al. (2013), and

95 appreciably larger than in most of the studies mentioned above. In part, we seek to answer the

96 question: how many ensemble members are required to detect a significant response (in a particular

97 variable) to Arctic sea-ice loss, if indeed it is possible to detect a significant response at all? This is

98 pertinent to assessing the strength of the forced change compared to AIV and hence, whether it may

99 be observable in the real world.

100

101 **Data and Methods**

102

103 *Simulations*

104

105 We utilise two independent AGCMs: the UK-Australian Unified Model (UM) version 7.3 and the

106 US National Center for Atmospheric Research (NCAR) Community Atmosphere Model (CAM)

107 version 3. The UM has been developed by the UK Meteorological Office Hadley Centre and is the

108 atmospheric model used in their Global Environmental Model version 2 (HadGEM2) and in the

109 Australian Community Climate and Earth System Simulator (ACCESS). Both HadGEM2 and

110 ACCESS are participating models in the fifth Coupled Model Intercomparison Project (CMIP).

111 The configuration of the UM used here has 38 vertical levels with a horizontal resolution of 1.25

112 degrees of latitude by 1.875 degrees of longitude. CAM is the atmospheric component of the NCAR

113 Community Climate System Model version 3 (CCSM3), which participated in the third CMIP. The

114 version used here has 26 vertical levels and a spectral resolution of T42, roughly equivalent to 2.8

115 degrees of latitude and longitude. For further details the reader is directed to Martin et al. (2011)

116 and Bi et al. (2013) for the UM/ACCESS and Collins et al. (2006) for the CAM.

117

118 We primarily analyse two distinct simulations, performed identically with each model, termed the

119 control (CTRL) and perturbation (PERT) simulations. In CTRL, the models were prescribed with

120 an annually-repeating monthly cycle of climatological (CLM) sea-ice concentration (*SIC*) and sea

121 surface temperature (*SST*). Monthly-mean *SIC* and *SST* were taken from the Hurrell et al. (2008)

122 data set, updated to 2009, which is derived from a combination of in situ and remotely-sensed

123 observations. In PERT, the linear trend (TRD) in *SIC* over 1979-2009 for each month was added to

124 the climatological monthly values and these CLM+TRD values were prescribed in the models. The

125 prescription of *SST* in PERT was based on the approach introduced by Screen et al. (2013) and was

126 as follows. In grid-boxes and months where the *SIC* TRD is not zero, then the CLM+TRD *SST* were

127   prescribed. Elsewhere, CLM *SST* was prescribed. This approach captures *SST* changes directly

128   related to *SIC* changes, but does not include *SST* changes outside the sea-ice zone (see Screen et al.,

129   2013 for further details and justification). CTRL and PERT were run for 100 years in the UM and

130   for 60 years in CAM. Since the prescribed surface forcing repeats annually, but the atmospheric

131   initial conditions vary, each year is considered to be an independent ensemble member (atmospheric

132   "memory" is negligible from year-to-year).

133

134   A further simulation has been performed with the UM, termed the PERT*2 simulation, in which the

135   linear trends in *SIC* were doubled before being added to climatological values and these

136   CLIM+(TRD*2) values were prescribed in the model. *SST* were prescribed as above, but this time

137   with CLM+(TRD*2) values in place of CLM+TRD values. PERT*2 was run for 100 years in the

138   UM only.

139

140   To isolate the atmospheric impacts of sea-ice loss, we compare the ensemble-mean of a particular

141   variable in CTRL with the ensemble-mean in PERT or PERT*2. The ensemble-mean difference,

142   PERT-CTRL or PERT*2-CTRL, is referred to as the "response" to Arctic sea-ice loss in the single-

143   or double-perturbation experiments. We refer to "local" and "remote" responses, by which we mean

144   responses that are in close proximity to sea-ice changes and those that are geographically distant

145   from sea-ice changes, respectively.

146

147   *Statistical methods*

148

149   To test the statistical significance of the ensemble-mean differences we compute the Student's t-

150   statistic, *t*, using the difference of means test (Von Storch and Zwiers, 1999),

151

152 $$t = \frac{\left|\overline{x} - \overline{y}\right|}{s_p \times \sqrt{\frac{2}{N}}} \quad (1)$$

153

154 where x is the ensemble-mean from PERT (or PERT*2), y is the ensemble-mean from CTRL, N is

155 the ensemble size and s_p is the pooled standard deviation, given by,

156

157 $$s_p = \sqrt{\frac{\sum_{i=1}^{n}\left(x_i - \overline{x}\right)^2 + \sum_{i=1}^{m}\left(y_i - \overline{y}\right)^2}{n + m - 2}} \quad (2)$$

158

159 where x_i is an individual ensemble member from PERT (or PERT*2), y_i is an individual ensemble

160 member from CTRL and n and m are the respective ensemble sizes (in our cases $N=n=m$). The

161 ensemble-mean difference is considered statistically significant when $t \geq t_c$ where t_c is the cutoff

162 value of the Student's t-distribution for a two-tailed probability of 0.025 (i.e., 95% confidence

163 interval) and $n+m$-2 degrees of freedom.

164

165 To calculate the minimum ensemble-size required to detect a statistically significant ensemble-

166 mean difference, N_{min}, we re-write (1) replacing t with t_c and N with N_{min},

167

168 $$t_c = \frac{\left|\overline{x} - \overline{y}\right|}{s_p \times \sqrt{\frac{2}{N_{min}}}} \quad (3)$$

169

170 and re-arrange to give,

171

172 $$N_{min} = 2t_c^2 \times \left(\frac{s_p}{\overline{x} - \overline{y}}\right)^2 \quad (4)$$

173

It can be seen from combining (1) and (3) that when $t \geq t_c$ then $N_{min} \leq N$. N_{min} can be considered a

measure of the signal-to-noise ratio, with small values of N_{min} implying a large signal-to-noise ratio

and large values of N_{min} implying a small signal-to-noise ratio. A similar approach for computing

N_{min} was used in Deser et al. (2012) and Terray et al. (2012). Equation (4) assumes that s_p is

insensitive to the ensemble size (i.e., s_p for N_{min} is equal to s_p for N). This assumption is

approximately valid, except for small values of N_{min} (when s_p for N_{min} is generally lower than s_p for

N; not shown). When N_{min} is small however, the denominator in (4) is appreciably larger than the

numerator ($s_p << x - y$) and thus, N_{min} is relatively insensitive to discrepancies in s_p.

182

In this manuscript we focus on autumn (September-November; SON) and winter (December-

February; DJF) as the atmospheric response to sea-ice loss is largest in these two seasons (e.g.,

Deser et al., 2010; Porter et al., 2012; Screen et al., 2013) and five key atmospheric variables: near-

surface (defined as 1.5 m in the UM and 2 m in CAM) air temperature (T_{ref}), air temperature on

constant pressure levels (T), precipitation (P), sea level pressure (SLP) and geopotential height on

constant pressure levels (Z).

189

Results

191

Figure 1a shows the SON SIC differences (PERT-CTRL) in the single-perturbation experiment. SIC

is reduced over most of the Arctic marginal seas, with the greatest losses in the Beaufort, Chukchi

and East Siberian Seas. By design, this pattern closely matches the SIC trends observed over the

period 1979-2009. In DJF, SIC reductions are most pronounced over the Barents Sea, Sea of

Okhotsk, Hudson Bay and the Labrador Sea (Figure 1b). Small SIC increases are located along the

east coast of Greenland and south of the Bering Strait. The difference in sea-ice area between

CTRL and PERT is 1.73 and 0.98 million km^2 in SON and DJF, respectively, and between CTRL

199 and PERT*2 is 2.53 and 1.48 million km^2, respectively, in SON and DJF. Note that the loss of sea-

200 ice area in the double-perturbation experiment is less than twice that in the single-perturbation

201 experiment because the *SIC* in any grid-box cannot be lower than zero. Recall that the boundary

202 conditions in PERT are based on *SIC* trends from 1979 through to 2009, which was the last full

203 year of *SIC* data when the model runs were initiated. The past 3 years (2010, 2011 and 2012) have

204 had low sea-ice coverage, with summer 2012 a new record minimum (Zhang et al., 3013; Parkinson

205 and Comiso, 2013), enhancing the long-term trend. The observed sea-ice area loss from 1979 to

206 2012, based on the National Snow and Ice Data Center (NSIDC) sea-ice index

207 (http://nsidc.org/data/G02135), is 2.40 and 1.34 million km^2 in SON and DJF, respectively. Thus,

208 the single-forcing experiment represents a smaller (by 28% and 27% in SON and DJF) loss of sea-

209 ice than observed from 1979 to 2012 and the double-forcing experiment represents a slightly larger

210 (by 5% and 10% in SON and DJF) loss of sea-ice than observed from 1979 to 2012. Figure 1c and

211 d show the corresponding differences in *SST* for SON and DJF, respectively. In general, the *SST*

212 warms where *SIC* decreases, and vice versa. By design, *SST* is unchanged in regions of constant or

213 zero *SIC* change. The *SIC* and *SST* differences in the double-perturbation experiment have the same

214 spatial patterns as in Figure 1, but with differences that are larger in magnitude (not shown).

215

216 Figure 2 shows the ensemble-mean T_{ref} responses (a-c; g-i) and associated values of N_{min} (d-f; j-l),

217 with the panels arranged as follows. The first (a-c) and second (d-f) rows correspond to SON and

218 the third (g-i) and fourth (j-l) rows to DJF. The first (a, d, g, j) and second (b, e, h, k) columns are

219 for the single-perturbation experiment in the CAM and UM, respectively, and the third column (c, f,

220 i, l) is for the double-perturbation experiment.

221

222 In SON, both models show widespread and significant warming over the Arctic Ocean and adjacent

223 continents (Figure 2a, b). Unsurprisingly, warming is largest over the regions of greatest ice loss

224 (cf. Figure 1a, c). The models are in very close agreement. The most obvious difference is that the

225   warming extends further over Scandinavia and northeastern Russia in the CAM than UM. The DJF

226   responses in both models show four warming centres: the Barents Sea, Hudson Bay, northern

227   Bering Sea and the Sea of Okhotsk (Figure 2g, h). These regions correspond to areas of winter sea-

228   ice loss and associated SST warming (cf. Figure 1b, d). The atmospheric warming is largely

229   confined to maritime regions in the case of the Bering Sea and the Sea of Okhotsk, but spreads to

230   neighbouring land masses around the Barents Sea and Hudson Bay. Farther away from the regions

231   of sea-ice loss, there are very few areas of significant T_{ref} response in either model. The UM depicts

232   significant cooling over the Caspian Sea and CAM depicts warming over central Asia.

233

234   As might be expected, the local T_{ref} response is larger in the double-perturbation experiment than in

235   the single-perturbation experiment (cf. Figures 2b and c; h and i). Additionally, the T_{ref} response is

236   larger over the high-northern continents and significant T_{ref} responses are detectable at lower

237   latitudes. This suggests that if Arctic sea-ice loss continues unabated, the geographical area affected

238   by sea-ice loss induced warming will increase. Under doubled forcing, there is a weak cooling

239   response over mid-latitude Eurasia in DJF, but this is only significant over a limited area

240   surrounding the Caspian Sea (Figure 2i).

241

242   In terms of N_{min}, five or fewer ensemble members are required to detect a statistically significant T_{ref}

243   response in the proximity of sea-ice loss, irrespective of the model, season or the magnitude of

244   forcing. Away from the regions of ice loss, approximately 30-50 ensemble members are required to

245   detect a significant response. The response over regions adjacent to ice loss is likely mediated by

246   horizontal advection due to synoptic systems (Deser et al., 2010), so it follows that the non-local

247   response will be weaker and subject to larger AIV (i.e., lower signal-to-noise ratio, higher N_{min})

248   than the local response that is directly driven by surface heat flux changes (e.g., Deser et al., 2010;

249   Screen et al., 2013). Under doubled forcing, the T_{ref} response over the high-northern continents is

250   easier to detect (lower N_{min} than with single forcing) and is detectable further south (cf. Figures 2e

251   and f; k and l).

252

253   Figure 3 shows the latitudinal and vertical structure of the zonal-mean T response. Consistent with

254   earlier work (Screen et al., 2012; 2013), Arctic warming due to observed sea-ice loss is strongest in

255   the lowermost atmosphere and is almost entirely confined to below 700 hPa in both seasons and

256   models (Figure 3a-b, g-h). In the double-perturbation experiment, the T response is stronger in the

257   near-surface levels, but the response remains trapped in the lower troposphere (Figure 3c, i). This

258   implies that Arctic sea-ice loss has no discernable influence on T aloft. N_{min} generally increases with

259   altitude. At the few locations where there is significant zonal-mean warming above 700 hPa, N_{min} is

260   50 or more.

261

262   Figure 4 shows the ensemble-mean P responses, arranged as in Figure 2. In SON, widespread P

263   increases are found over the Beaufort, Chukchi and East Siberia Seas in the UM (Figure 4b). In the

264   CAM, P also increases in these regions but with less spatial coherence (Figure 4a). In both models,

265   the P increases are associated with significant increases in cloud cover, principally low cloud, but

266   the cloud responses are weaker in the CAM than UM (not shown). The weaker P and cloud cover

267   responses in the CAM versus UM was previously noted by Screen et al. (2013) and appears to relate

268   to problems with the cloud cover scheme in CAM version 3. In the double-forcing experiment, the

269   P increases are stronger and are significant over most of the Arctic Ocean (Figure 4c). Away from

270   the Arctic Ocean there are isolated patches of significant P response in both models, but no large-

271   scale features even in the double-perturbation experiment. In DJF, both models show significant P

272   increases over the regions of winter sea-ice loss: the Barents Sea, Sea of Okhotsk and Hudson Bay

273   (Figures 4g, h) . Remote P decreases occur over the North Atlantic and Pacific in CAM, but they

274   are only significant in small areas (Figure 4g). In the double-forcing experiment, a similar spatial

275   pattern is found, but with increased magnitude (Figure 4i). Over regions of maximum sea-ice loss,

276   N_{min} for P is less than 10 and over other regions of sea-ice loss it is around 10-30 (Figures 4d-f, j-l).

277   The majority of grid-boxes with a significant P response have an associated N_{min} of less than 40,

278   with the main exceptions being the sporadic remote P responses.

279

280   Figure 5 shows the SLP responses. In SON, SLP decreases significantly over the Beaufort, Chukchi

281   and East Siberian Seas in the UM (Figure 5b). A second low SLP centre is located over the Baltic

282   countries. In CAM, two regions of lowered SLP are identified in broadly similar, but non-identical,

283   locations (Figure 5a). The first low is shifted to the southwest to be centred over Alaska and the

284   second low is shifted westward to be located over Scandinavia. Away from these limited regions,

285   the SLP response is statistically insignificant in both models. The spatial patterns of the SLP

286   responses are largely consistent between the single- and double-perturbation experiments, but there

287   are differences in the magnitudes and significance of the responses (cf. Figure 5b and c). In the

288   double-perturbation experiment, significantly lowered SLP is found over a larger area, including

289   most of the Arctic Ocean, the Canadian Archipelago and Hudson Bay. The low-pressure centre over

290   the Baltic countries that is significant in the single-perturbation is statistically insignificant in the

291   double-perturbation experiment. Conversely, SLP increases over Europe and East Asia become

292   significant in the double-perturbation experiment.

293

294   In DJF, significant large-scale SLP decreases are found over the Arctic Ocean, Hudson Bay and

295   eastern Canada in the UM (Figure 5h). Isolated regions of significant SLP reductions are also

296   identified over the Sea of Okhotsk and central North America. SLP is increased over Europe, but

297   this feature is not statistically significant. In the CAM, SLP decreases significantly over Hudson

298   Bay, Greenland and the Atlantic-side of the Arctic Ocean (Figure 5g). SLP increases significantly

299   over the Bering Sea. Whilst the two models exhibit broadly similar SLP responses in the Arctic and

300   sub-Arctic Canada, the mid-latitude responses are rather different (cf. Figure 5g and h). In

301   particular, CAM depicts larger SLP increases over the north Atlantic and north Pacific than does the

302   UM, and the responses over the United States are opposite in sign between the two models.

303 However, the mid-latitude responses are predominantly statistically insignificant in both models, so

304 these discrepancies can be explained by AIV. No regions show significant responses of opposite

305 sign between the models. The spatial patterns of the DJF *SLP* responses are similar in the single-

306 and double-perturbation experiments (cf. Figures 5h and i). The *SLP* decrease over the Arctic and

307 Hudson Bay is larger in magnitude in the latter, but the geographical extent of the significant *SLP*

308 response is not overly different. Three small regions show significant responses in the double-

309 perturbation experiment that are not significant in the single-perturbation experiment. These are

310 *SLP* increases over the Bering Sea, eastern Europe and eastern China. The region of weak, but

311 significant, *SLP* decrease over central North America in the single-perturbation experiment is not

312 significant in the double-perturbation experiment.

313

314 N_{min} for the *SLP* response is as low as 10 in the UM over regions of maximum ice loss, especially in

315 the double-perturbation case, but N_{min} values this low are only found in very limited geographical

316 regions (Figure 5d-f; j-l). Generally, approximately 30-50 ensemble members are required to detect

317 a significant *SLP* response, and upwards of 50 members are required to detect a significant response

318 in remote regions. It is notable that even with 100 ensemble members in the UM, very few mid-

319 latitude regions show a significant *SLP* response in the single-perturbation experiment. Further,

320 despite larger mid-latitude responses in CAM, an ensemble size of 60 is insufficient for these

321 achieve statistical significance. This implies that the remote *SLP* response to recent Arctic sea-ice

322 loss is considerably smaller than AIV.

323

324 Figure 6 shows the zonal-mean *Z* responses. In SON, the high-latitude response is baroclinic with *Z*

325 decreases in the lowermost atmosphere and *Z* increases aloft. Significant zonal-mean *Z* responses

326 are only found at 1000 hPa. The vertical profile is fairly consistent across the models and

327 experiments. Taken together, the *SLP* and *Z* responses in SON are suggestive of a shallow thermal

328 (heat) low in response to sea-ice loss. Thermal lows can occur when cold air overlies warmer water,

329 as is the case in regions of sea-ice loss (Higgins and Cassano, 2009; Deser et al., 2010; Strey et al.,

330 2010; Orsolini et al., 2012). In DJF, the vertical profile of the Z response is completely different.

331 Both models show a quasi-barotropic Z decrease over high northern latitudes. This high-latitude Z

332 decrease is significant in the UM below 500 hPa, but only at 1000 hPa in CAM. Both models show

333 Z increases over mid-latitudes. In CAM, these extend throughout the troposphere, but are only

334 significant above 700 hPa. In the UM, Z increases are found aloft but not at 1000 hPa, and are

335 shifted polewards in comparison to those in CAM. They are insignificant in the single-perturbation

336 experiment, but significant above 850 hPa in the double-perturbation experiment. In all other

337 respects, the Z responses in the single- and double-perturbation are very similar. N_{min} for Z is high,

338 typically 50 or above in the single-perturbation experiment and only slightly lower in the double-

339 perturbation experiment.

340

341 In summary, the SLP and Z responses point to rather different spatial and vertical structures to the

342 circulation responses in SON and DJF. In SON, the response is baroclinic (restricted to the near-

343 surface levels) and localised. Similar local circulation responses to sea-ice loss have been identified

344 in other simulations (Higgins and Cassano, 2009; Deser et al., 2010; Strey et al., 2010; Orsolini et

345 al., 2012). By contrast in DJF, the circulation response is fairly barotropic and more spatially

346 extensive. This seasonal transition from a local baroclinic response to a larger-scale barotropic

347 response was also noted by Deser et al. (2010), although the horizontal structure of their winter

348 responses are rather different to that found here. In our CAM simulations, the DJF responses project

349 onto the positive phase of the Arctic Oscillation (AO). This is in contrast to the negative-type AO

350 responses found in February by Deser et al. (2010) and in DJF by Liu et al. (2012), both using

351 CAM but in response to projected future and past sea-ice trends, respectively. Screen et al. (2013)

352 reported a negative North Atlantic Oscillation (NAO) response in early-winter (November-

353 December) in the CAM and UM, but cautioned that the response was weak and often exceeded by

354 AIV. The larger ensembles presented here do not support a shift towards to negative phase of the

355 NAO in response to observed sea-ice loss. Instead, in CAM the response projects onto the positive

356 NAO phase and in the UM the response is not NAO-like. Thus, the wintertime circulation responses

357 (and their interactions with the large-scale modes of atmospheric variability) are not robust across

358 simulations, even those using the same models.

359

360 So far we have considered a limited number of atmospheric variables. For a wider perspective,

361 Table 1 provides the mean N_{min} for a broad selection of atmospheric variables. The values given in

362 Table 1 are averages of N_{min} across all grid-points that exhibit a significant response in that variable

363 (recall N_{min} is undefined where the response is insignificant), all experiments, models (UM and

364 CAM) and seasons (SON and DJF). For example, the mean N_{min} for T_{ref} is the average of all the

365 values in Figures 2d-f and j-l. Table 1 also provides the mean percentage area of northern

366 hemisphere extra tropics (>30°N) exhibiting a significant response in each variable. Both the mean

367 N_{min} and area metrics mask substantial spatial, seasonal and inter-model variability, so the precise

368 numbers must be interpreted with caution. However, comparison of the mean N_{min} between

369 variables is insightful as it clearly demonstrates that the responses in certain variables are easier to

370 detect than others. To aid interpretation, the variables in Table 1 are listed in order of ascending

371 mean N_{min}. Recall, smaller values indicate that the response is easier to detect than larger values.

372 The ranked variables can be split into four categories of increasing mean N_{min}. This ranking is

373 largely insensitive to whether or not the double-perturbation experiment is included in the analysis

374 (not shown). The variables with smallest values ($N_{min} < 30$) are the surface heat fluxes and T_{ref}. The

375 next group ($30 < N_{min} < 50$) includes variables related to clouds, precipitation and radiation. A third

376 group ($50 < N_{min} < 60$) contains variables related to surface atmospheric circulation, including SLP

377 and near-surface wind. The hardest responses to detect ($N_{min} > 60$) are in upper-level variables, for

378 example, mid-tropospheric (500 hPa) temperature (T_{500}) and geopotential height (Z_{500}) and

379 jetstream-level (250 hPa) wind (U_{250}, V_{250}). Although we have not considered stratospheric

380 variables here, Cai et al. (2012) found that the stratospheric response to sea-ice loss is small

381   compared to the tropospheric response.

382

383   Clearly, AIV is a key source of uncertainty in the simulated atmospheric response to Arctic sea-ice

384   loss. Larger ensembles can reduce this uncertainty by averaging out, to some extent, the effects of

385   AIV. Figure 7 quantifies the reduction in uncertainty in the response to Arctic sea-ice loss, due to

386   AIV, as the ensemble size increases. To construct this figure, we have sub-sampled our large

387   ensembles into smaller sub-ensembles of varying size. For each sub-ensemble size, a large number

388   (100,000) of unique combinations are sampled to produce a large set of sub-ensemble mean

389   responses. For example, for a sub-ensemble size of 5 we sub-sampled 100,000 unique combinations

390   of 5 members from the full set. For each combination, we averaged the selected members to

391   produce a sub-ensemble mean. This results in a set of 100,000 sub-ensemble mean responses. The

392   spread (difference between maximum and minimum values) of these sub-ensemble mean responses

393   provides a measure of the uncertainty in the response due to AIV, for an ensemble of that size.

394   Figures 7a and b shows examples for the Arctic-mean (>70°N) SON T_{ref} response and DJF SLP

395   response, respectively, but qualitatively similar results are found for other seasons and variables.

396

397   Uncertainty due to AIV, as estimated by the spread of sub-ensemble mean responses, can be seen to

398   decrease almost exponentially as the ensemble size increases. This implies that to reduce

399   uncertainty by one half, the ensemble size has to be doubled. In absolute terms, uncertainty due to

400   AIV decreases rapidly as the ensemble-size increases from 5 (or fewer) to 20 members, and then

401   continues to reduce more slowly as further ensemble members are added. This behaviour is very

402   similar in the two models and in both the single- and double-perturbation experiments, however,

403   CAM has larger AIV than the UM for both T_{ref} and SLP. In the UM a point is reached, around 50-

404   60 ensemble members, where adding further ensemble members has almost no impact on the

405   uncertainty due to AIV (spread of responses). We assume a similar point would occur in the CAM,

406   but we have insufficient ensemble members to confirm this.

407

Discussion and Conclusions

409

Arguably, one of the most surprising aspects of our results is that over the Arctic we have identified

a robust (in the sense that it is statistically significant in both models, and in both the single- and

double-perturbation experiments) lowering of *SLP* in response to Arctic sea-ice loss. This is

surprising because this high-latitude DJF *SLP* response is opposite to that suggested in some

empirical studies (e.g., Francis et al., 2009; Jaiser et al., 2012) and to that found in other modelling

studies, for example, Deser et al (2010) and Liu et al. (2012). It is of especial interest to draw

comparisons with Liu et al. because that study performed very similar experiments to those

presented here. It used the same model (CAM version 3 at T42 resolution), experimental set-up and

forcing based on observed sea-ice trends. However, Liu et al. report a significant increase in DJF

SLP in response to Arctic sea-ice loss, in stark contrast to the decrease shown here (cf. our Figure

3g and their Figure 4c). One notable difference between the two studies is that Liu et al. used an

ensemble of 20 CAM simulations compared to our ensemble of 60 CAM simulations.

422

It is plausible that 20 ensemble members are insufficient to accurately separate the forced signal

from AIV (N_{min} for DJF *SLP* is generally larger than 20). To test this hypothesis, we considered if it

is possible to derive the Liu et al. result from a subset of our larger CAM ensemble. Figure 8 shows

probability distribution functions (PDFs) for the Arctic-mean DJF sub-ensemble mean *SLP*

responses (i.e., the set of 100,000 sub-sampled responses). The PDFs narrow as the ensemble size

increases, implying that larger ensembles yield more precise responses to Arctic sea-ice loss. For a

sub-ensemble size of 10, it is possible to obtain both positive and negative sub-ensemble mean *SLP*

responses. This shows that AIV influences both the magnitude and sign of *SLP* responses in small

ensembles. However for a sub-ensemble size of 20, an Arctic-mean increase in *SLP* is found in less

than 0.1% of cases. This implies that the Liu et al. result cannot be derived from a 20-member

433   subset of our larger ensemble. Thus, it is very unlikely that the discrepancy between the Arctic

434   winter SLP responses is this study and in Liu et al. is due to AIV alone. A further difference

435   between this study and Liu et al. is the magnitude of the sea-ice forcing. Although the spatial

436   pattern of the sea-ice forcing is highly similar between the two studies, our forcing is approximately

437   two-to-three times larger magnitude than that in Liu et al. (cf. our Figures 1a ,b and their Figures 4a,

438   b; note the different colour scales). Whether or not this is the cause of the opposing winter SLP

439   responses is unclear, but the discrepancy highlights that experimental differences between

440   simulations, even with the same model, can lead to fundamentally different responses.

441

442   We now return to the question posed earlier: how many ensemble members are required to detect a

443   significant response to Arctic sea-ice loss? In reality, there is no simple answer to this question as

444   N_{min} varies considerably in space and by variable. As a general rule of thumb, we suggest that

445   detection of the thermo-dynamical (e.g., SLP, wind) response requires an ensemble size

446   approximately twice as large as the thermal response (e.g., surface heat fluxes, T). The hydrological

447   response (e.g., cloud, P) lays in-between, which likely reflects influences of both thermal and

448   thermo-dynamical factors on these variables. Deser et al. (2012) reported very similar findings in

449   the context of the coupled climate response to greenhouse gas forcing, as did Wehner (2000) and

450   Taschetto and England (2008) both in the context of the atmospheric response to global SST and

451   SIC trends. The upper-level response to Arctic sea-ice loss (e.g., Z_{500}, T_{500}, U_{250}, V_{250}) is harder to

452   detect than the near-surface response as the signal-to-noise ratio decreases with altitude. Equally,

453   the remote T_{ref} or P responses (that are mediated by thermo-dynamical processes) are harder to

454   detect than the local T_{ref} and P responses (that are primarily-driven by surface fluxes). Figure 7

455   suggests large gains, in terms of reduced uncertainty, by increasing from a small (less than 20

456   members) to moderate-sized (20-50 members) ensemble. Further increases in ensemble size

457   represent a case of "diminishing returns" with smaller reductions in uncertainty per additional

458   ensemble member. On this basis, we argue that an ensemble size of around 50 members is

459   desirable. This is considerably larger than the typical ensemble size used in past studies of the

460   atmospheric response to observed Arctic sea-ice loss (e.g., 5 in Ghatak et al. (2012) and Orsolini et

461   al. (2012), 5/8 in Screen et al. (2013), 10 in Strey et al. (2010), 15 in Porter et al. (2012), 20 in Liu

462   et al. (2012)).

463

464   The values of N_{min} have implications not only for modelling studies, but also for what aspects of the

465   simulated Arctic sea-ice response may be observable in the real world. Since each ensemble

466   member is one-year of simulation, N_{min} can also be thought of as approximate measure of the

467   minimum number of years required to detect a significant response due to Arctic sea-ice loss,

468   assuming the rate of loss is linear. The differences in prescribed boundary conditions between

469   CTRL and PERT have, in reality, occurred over a 31-year period (1979-2009). Assuming that the

470   models are realistic in their depiction of the forced response and AIV, a N_{min} of 31 or less suggests

471   that the simulated response should be observable in nature over the period 1979-2009. Conversely,

472   a N_{min} of greater than 31 suggests that more than 31 years are required to separate the forced

473   response from AIV and therefore, the response to past sea-ice loss would not be expected to be

474   detectable in observed records. Accordingly, we argue that the simulated local T_{ref} and P responses

475   to Arctic sea-ice loss should be detectable, but that the atmospheric circulation (e.g., SLP, 10 m

476   wind), upper-level (e.g., Z_{500}, T_{500}, U_{250}, V_{250}) and remote responses may be partially or wholly

477   masked by AIV. In practise, the detection and importance of the atmospheric impacts of sea-ice loss

478   not only depend on the relative magnitudes of the sea-ice forced change compared to AIV, but also

479   on the relative magnitudes of sea-ice forced response to other forced responses.

480

481   Figures 9a and d show observed (from ERA-Interim; Dee et al., 2011) trends in T_{ref} over the period

482   1979-2009 for SON and DJF, respectively. These are highly similar to the simulated T_{ref} responses

483   to Arctic sea-ice loss (Figure 2), suggesting that the T_{ref} response is indeed detectable, consistent

484   with previous studies (Screen and Simmonds, 2010a; 2010b; 2012). Figures 9b and e show

485   observed trends in P for SON and DJF, respectively, taken from the Global Precipitation

486   Climatology Project (GPCP) data set (Adler et al., 2003). These can be compared to the simulated P

487   responses in Figure 4. Although the moderate values of N_{min} in Figure 4 suggested that the local P

488   response may be detectable, the observed trends are not in agreement with the simulated P

489   responses. We propose that there are two likely reasons for this apparent disparity. One reason is

490   that detection of the atmospheric impacts of sea-ice loss not only depend on the relative magnitudes

491   of the sea-ice forced change compared to AIV, but also on the relative magnitudes of sea-ice forced

492   response to other forced responses. It is likely that the observed P trends are forced by factors other

493   than, or in addition to, sea-ice loss. Secondly, there is considerable uncertainty as to the sign and

494   magnitude of observed P trends over the poorly observed Arctic region. P trends from alternative

495   observationally constrained data sources – for example the CPC Merged Analysis of Precipitation

496   (CMAP; Xie and Arkin, 1997) or ERA-Interim - depict rather different patterns of P change over

497   the Arctic Ocean (not shown). It is possible therefore, that the simulated P response to sea-ice loss

498   is undetectable in observations because of considerable observational uncertainty. Observed (ERA-

499   Interim) SON and DJF SLP trends are shown in Figures 9c and f. Neither resemble the simulated

500   SLP responses in any of the models/experiments (Figure 5), which is consistent with the conclusion

501   that the SLP response to Arctic sea-ice loss is masked by AIV, or SLP trends due to other forcing

502   factors. We note that the observed trends are only one realisation (effectively one ensemble

503   member) and likely contain a sizeable component of natural (unforced) variability. Thus, the

504   observed trends would not be expected to match the ensemble-mean simulated responses.

505

506   We close by emphasising two arguably obvious, but nonetheless important, considerations. Firstly,

507   this study has only considered the first-order "direct" atmospheric response to Arctic sea-ice loss.

508   The fully coupled climate system response to Arctic sea-ice loss may be different to that shown

509   here. Secondly, Arctic sea-ice loss is only one forcing factor that may be relevant to northern

510   hemisphere climate variability and change. Observed trends reflect changes in multiple forcing

511   factors and the complex interactions between them.

512

520

521   **References**

522

523   Adler, S.G., et al., 2003: The Version 2 Global Precipitation Climatology (GPCP) Monthly

524   Precipitation Analysis (1979-present), *J. Hydrometeor., 4,* 1147-1167.

525   Bi, D, et al., 2013: The ACCESS coupled model: Description, control climate and evaluation, *Aus.*

526   *Met. Oceanog. J.,* in press.

527   Bluthgen, J., R. Gerdes, and M. Werner, 2012: Atmospheric response to the extreme Arctic sea-ice

528   conditions in 2007, *Geophys. Res. Lett.*, *39*, L02707, doi:10.1029/2011GL050486.

529   Cai, D., M. Dameris, H. Garny and T. Runde, 2012: Implications of all season Arctic sea-ice

530   anomalies on the stratosphere, *Atmos. Chem. Phys.*, *12*, 11819-11831.

531   Collins, W., et al., 2006: The formulation and atmospheric simulation of the Community

532   Atmosphere Model Version 3 (CAM3), *J. Climate*, *19*, 2144–2161.

533   Dee, D.P., et al., 2011: The ERA-Interim reanalysis: configuration and performance of the data

534   assimilation system, *Q. J. R. Meteorol. Soc., 137*, 553-597.

535   Deser, C., R. Tomas, M. Alexander, and D. Lawrence, 2010: The seasonal atmospheric response to

536   projected Arctic sea-ice loss in the late twenty-first century, *J. Climate*, *23*, 333–351.

537   Deser, C., A. Phillips, V. Bourdette and H. Teng, 2012: Uncertainty in climate change projections:

538   the role of internal variability, *Climate Dynamics*, *38*, 527-546.

539   Francis, J., W. Chen, D. Leathers, J. Miller, and D. Veron, 2009: Winter northern hemisphere

540   weather patterns remember summer Arctic sea-ice extent, *Geophys. Res. Lett.*, *36*, L07503,

541   doi:10.1029/2009GL037274.

542   Francis, J. and S. Vavrus, 2012: Evidence linking Arctic amplification to extreme weather in mid-

543   latitudes, *Geophys. Res. Lett.*, *39*, L06801, doi:10.1029/2012GL051000.

544   Ghatak, D., C. Deser, A. Frei, G. Gong, A. Phillips, D. Robinson and J. Stroeve, 2012: Simulated

545   Siberian snow cover response to observed Arctic sea-ice loss, 1979-2008, *J. Geophys. Res.*,

546   *117*, D23108, doi:10.1029/2012JD018047.

547   Higgins, M.E. and J.J Cassano, 2009: Impacts of reduced on winter Arctic atmospheric circulation,

548   precipitation, and temperature, *J. Geophys. Res.*, *114*, D16107, doi:10.1029/2009JD011884.

549   Honda, M., J. Inoue, and S. Yamane, 2009: Influence of low Arctic sea-ice minima on anomalously

550   cold Eurasian winters, *Geophys. Res. Lett.*, *36*, L08707, doi:10.1029/2008GL037079.

551   Hopsch, S., J. Cohen and K. Dethloff, 2012: Analysis of a link between fall Arctic sea-ice

552   concentration and atmospheric patterns in the following winter, *Tellus*, *64A*, 18624,

553   doi:10.3402/tellusa.v64i0.18624.

554   Hurrell, J., J. Hack, D. Shea, J. Caron, and J. Rosinski, 2008: A new sea surface temperature and

555   sea-ice boundary dataset for the Community Atmosphere Model, *J. Climate*, *21*, 5145–5153.

556   Jaiser, R., K. Dethloff, D. Handorf, A. Rinke, and J. Cohen, 2012: Impact of sea-ice cover changes

557   on the northern hemisphere atmospheric winter circulation, *Tellus*, *64A*, 11595,

558   doi:10.3402/tellusa.v64i0.11595.

559   Kay, J., M. Holland and A. Jahn, 2011: Inter-annual to multi-decadal Arctic sea ice extent trends in

560   a warming world, *Geophys. Res. Lett.*, *38*, L15708, doi:10.1029/2011GL048008.

561   Liu, J., J. Curry, H. Wang, M. Song, and R. Horton, 2012: Impact of declining Arctic sea-ice on

562   winter snowfall, *Proc. Natl. Acad. Sci. USA*, *109*, 4074–4079.

563   Livina, V. and T. Lenton, 2013: A recent tipping point in Arctic sea-ice over: abrupt and persistent

564   increase in the seasonal cycle since 2007, *Cryosphere*, *7*, 275-286.

565   Martin, G. M., et al., 2011: The HadGEM2 family of Met Office Unified Model climate

566   configurations, *Geoscientific Model Development*, *4*, 723–757.

567   Orsolini, Y., R. Senan, R. Benestad, and A. Melsom, 2012: Autumn atmospheric response to the

568   2007 low Arctic sea-ice extent in coupled ocean-atmosphere hindcasts, *Climate Dynamics*,

569   *38*, 2437–2448.

570   Overland, J. and M. Wang, 2010: Large-scale atmospheric circulation changes are associated with

571   the recent loss of Arctic sea-ice, *Tellus*, *62A*, 1–9.

572   Parkinson, C.L., and J.C. Comiso, 2013: On the 2012 record low Arctic sea ice cover: Combined

573   impact of preconditioning and an August storm, *Geophys. Res. Lett., 40*, 1356-1361.

574   Petoukhov, V. and V. Semenov, 2010: A link between reduced Barent-Kara sea-ice and cold winter

575   extremes over northern continents, *J. Geophys. Res.*, *115*, D21111,

576   doi:10.1029/2009JD013568.

577   Porter, D., J. Cassano, and M. Serreze, 2012: Local and large-scale atmospheric responses to

578   reduced Arctic sea-ice and ocean warming in the WRF model, *J. Geophys. Res.*, *117*,

579   D11115, doi:10.1029/2011JD016969.

580   Screen, J.A. and I. Simmonds, 2010a: The central role of diminishing sea ice in recent Arctic

581   temperature amplification, *Nature*, *464*, 1334-1337.

582   Screen, J.A. and I. Simmonds, 2010b: Increasing fall-winter energy loss from the Arctic Ocean and

583   its role in Arctic temperature amplification, *Geophys. Res. Lett.*, *37*, L16797,

584   doi:10.1029/2010GL044136.

585   Screen, J.A. and I. Simmonds, 2012: Declining summer snowfall in the Arctic: causes, impacts and

586   feedbacks, *Climate Dynamics*, *38*, 2243-2256.

587   Screen, J.A., C. Deser, I. Simmonds, 2012: Local and remote controls on observed Arctic

588   warming, *Geophys. Res. Lett.*, *39*, L10709, doi:10.1029/2012GL051598.

589   Screen, J.A., I. Simmonds, C. Deser and R. Tomas, 2013: The atmospheric response to three

590   decades of observed Arctic sea-ice loss, *J. Climate*, *26*, 1230-1248.

591   Screen, J.A. and I. Simmonds, 2013a: Exploring links between Arctic amplification and mid-

592   latitude weather, *Geophys. Res. Lett.*, *40*, 959-964.

593   Screen, J.A. and I. Simmonds, 2013b: Caution needed when linking weather extremes to amplified

594   planetary waves, *Proc. Natl. Acad. Sci. USA*, in press.

595   Seierstad, I. and J. Bader, 2009: Impact of projected future Arctic sea-ice reduction on extratropical

596   storminess and the NAO, *Climate Dynamics*, *33*, 937–943.

597   Simmonds, I. and I. Rudeva, 2012: The Great Arctic Cyclone of August 2012, *Geophys. Res. Lett.*,

598   *39*, L23709, doi:10.1029/2012GL054259.

599   Singarayer, J., J. Bamber, and P. Valdes, 2006: Twenty-first-century climate impacts from a

600   declining Arctic sea-ice cover, *J. Climate*, *19*, 1109–1125.

601   Strey, S., W. Chapman, and J. Walsh, 2010: The 2007 sea-ice minimum: impacts on the Northern

602   Hemisphere atmosphere in late autumn and early winter, *J. Geophys. Res.*, *115*, D23103,

603   doi:10.1029/2009JD013294.

604   Stroeve, J., M. Serreze, M. Holland, J. Kay, J. Maslanik, and A. Barrett, 2011: The Arctic's rapidly

605   shrinking sea-ice cover: A research synthesis, *Clim. Change*, *110*, 1005–1027.

606   Strong, C., G. Magnusdottir, and H. Stern, 2010: Observed feedback between winter sea-ice and the

607   North Atlantic Oscillation, *J. Climate*, *22*, 6021–6032.

608   Taschetto, A.S. and M.H. England, 2008: Estimating ensemble size requirements of AGCM

609   simulations, *Meteorol. Atmos. Phys.*, 100, 23-36.

610   Terray, L., L. Corre, S. Cravatte, T. Delcroix, G. Reverdin and A. Ribes, 2012: Near-surface

611   salinity as Nature's rain gauge to detect human influence on the tropical water cycle, *J.*

612   *Climate*, *25*, 958-977.

613   Von Storch, H. and F.W. Zwiers, 1999: Statistical analysis in climate research, Cambridge
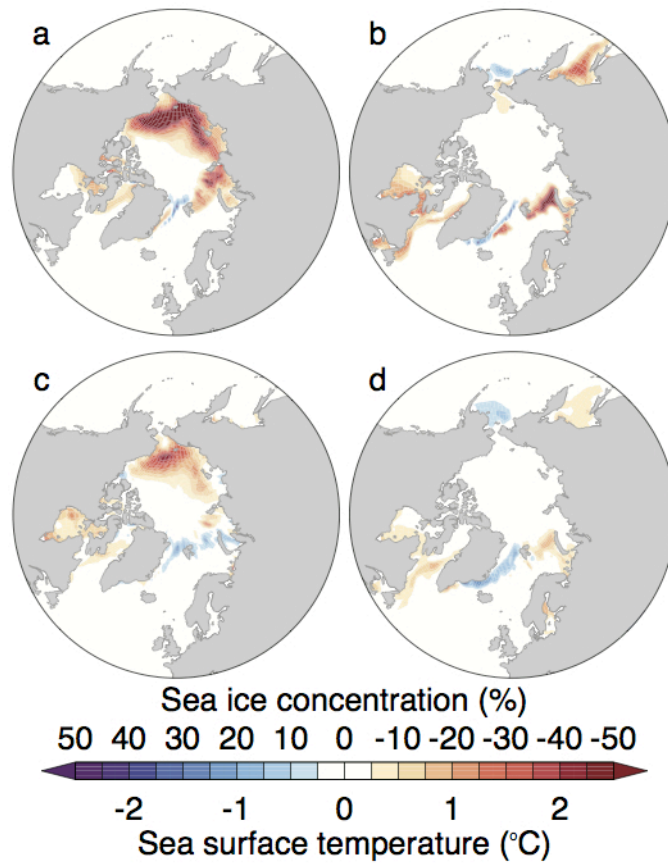
614   University Press, Cambridge, UK.

615   Wehner, M.F., 2000: A method to aid in the determination of the sampling size of AGCM ensemble

616   simulations, *Climate Dynamics*, 16, 321-331.

617   Wu, Q. and X. Zhang, 2010: Observed forcing-feedback processes between northern hemisphere

618   atmospheric circulation and Arctic sea-ice coverage, *J. Geophys. Res.*, *115*, D14199,

619   doi:10.1029/2009JD013574.

620   Xie, P. and P.A. Arkin, 1997: Global precipitation: a 17-year monthly analysis based on gauge

621   observations, satellite estimates, and numerical model outputs, *Bull. Amer. Meteorol. Soc.,*

622   *78*, 2539-2558.

623   Zhang. J., R. Lindsay, A. Schweiger and M. Steele, 2013: The impact of an intense summer cyclone

624   on 2012 Arctic sea-ice retreat, *Geophys. Res. Lett.*, doi:10.1002/grl.50190.

625

626

Variable	Mean N_{min}	Mean area
Surface sensible heat flux	21.9	19.1
Surface latent heat flux	24.5	19.7
1.5 m air temperature (T_{ref})	26.0	28.9
Low cloud cover	29.9	17.5
925 hPa air temperature (T_{925})	31.1	25.7
Net surface short-wave radiation	32.6	12.9
Net surface long-wave radiation	32.8	15.2
Total cloud cover	33.3	16.5
Precipitation	40.3	11.4
Sea level pressure (SLP)	50.3	11.9
10 m meridional wind speed	50.7	9.0
10 m zonal wind speed	50.8	9.4
500-1000 hPa thickness	56.1	7.0
250 hPa zonal wind speed (U_{250})	65.4	5.5
250-1000 hPa thickness	67.6	4.7
250 hPa meridional wind speed (V_{250})	68.0	3.2
500 hPa air temperature (T_{500})	70.0	4.3
500 hPa geopotential height (Z_{500})	71.9	5.0
250 hPa geopotential height (Z_{250})	73.3	4.8

627

628 **Table 1**: Mean N_{min} for a selection of atmospheric variables. For each variable, the value of N_{min}

629 given is the average over all grid-boxes (with a significant response in that variable), all models and

630 experiments, and both autumn and winter. The right-hand column shows the mean percentage area

631 of northern hemisphere extra tropics (>30°N) exhibiting a significant response in that variable

632 (again averaged across models, experiments and seasons). To aid interpretation, the variables are

633 listed in order of ascending mean N_{min}. Note that low values imply a response that is easier to detect

634 than high values.

635

Sea ice concentration (%)

50 40 30 20 10 0 -10 -20 -30 -40 -50

-2   -1   0   1   2
Sea surface temperature (°C)

Figure 1: Ensemble-mean differences (PERT-CTRL) in sea-ice concentration (*SIC*) for (a) autumn and (b) winter. (c-d) As (a-b), but for sea surface temperature (*SST*). Note the inverse scale for *SIC*.
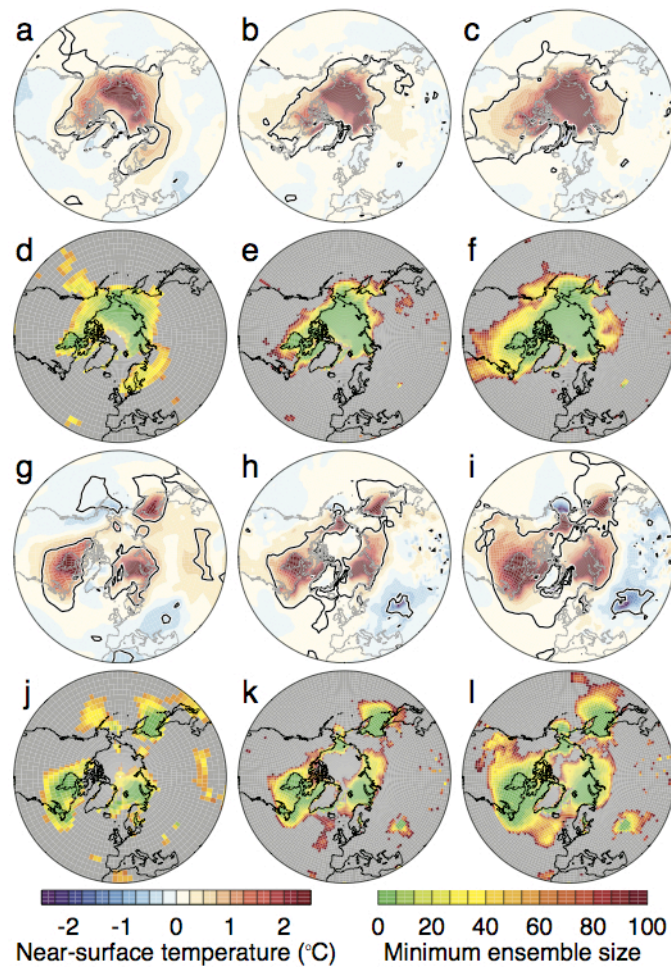
Figure 2: Ensemble-mean differences in autumn near-surface air temperature (T_{ref}) for (a) CAM
PERT-CTRL, (b) UM PERT-CTRL and (c) UM PERT*2-CTRL. Statistically significant
differences (at the $p \leq 0.05$ level) are enclosed by black contours. (d-f) N_{min} for the differences
shown in (a-c), respectively. Grey shading denotes an insignificant ensemble-mean difference. (g-l)
As (a-f), but for winter.

645
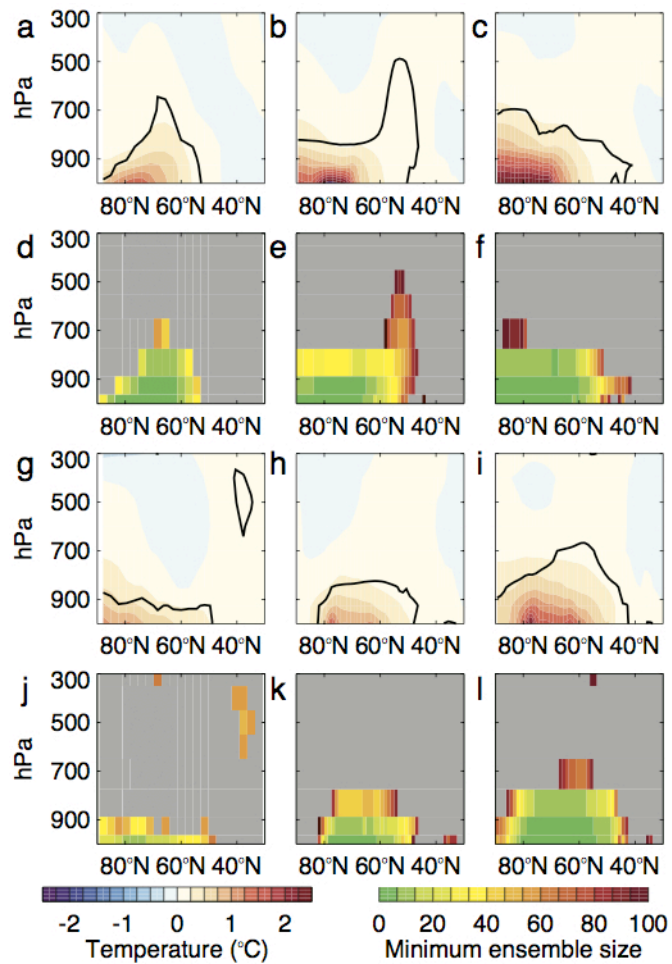
Figure 3: Zonal-mean ensemble-mean differences in autumn air temperature (T) for (a) CAM
PERT-CTRL, (b) UM PERT-CTRL and (c) UM PERT*2-CTRL. Statistically significant
differences are enclosed by black contours. (d-f) N_{min} for the differences shown in (a-c),
respectively. Grey shading denotes an insignificant ensemble-mean difference. (g-l) As (a-f), but for
winter.

652
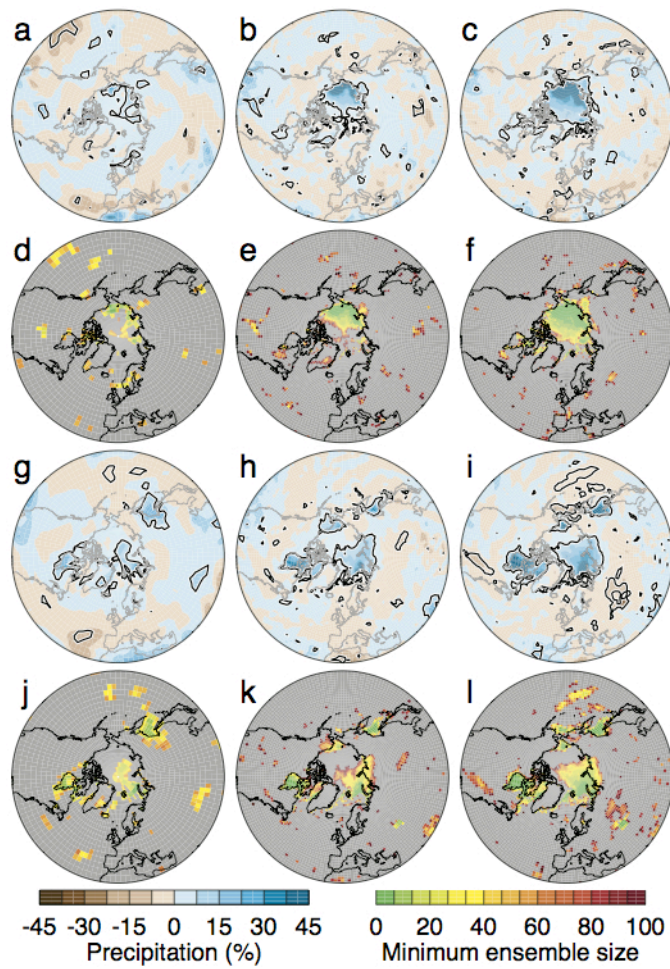
Figure 4: As Figure 2, but for precipitation (*P*). The *P* differences are expressed as percentages relative to the ensemble-means in CTRL.
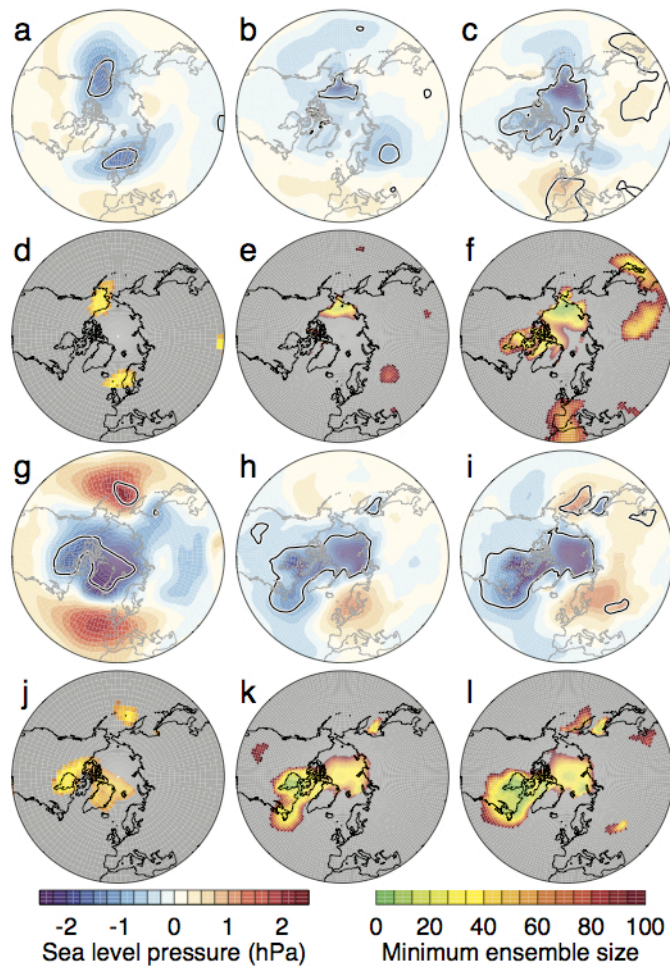
Figure 5: As Figure 2, but for sea level pressure (*SLP*).

657
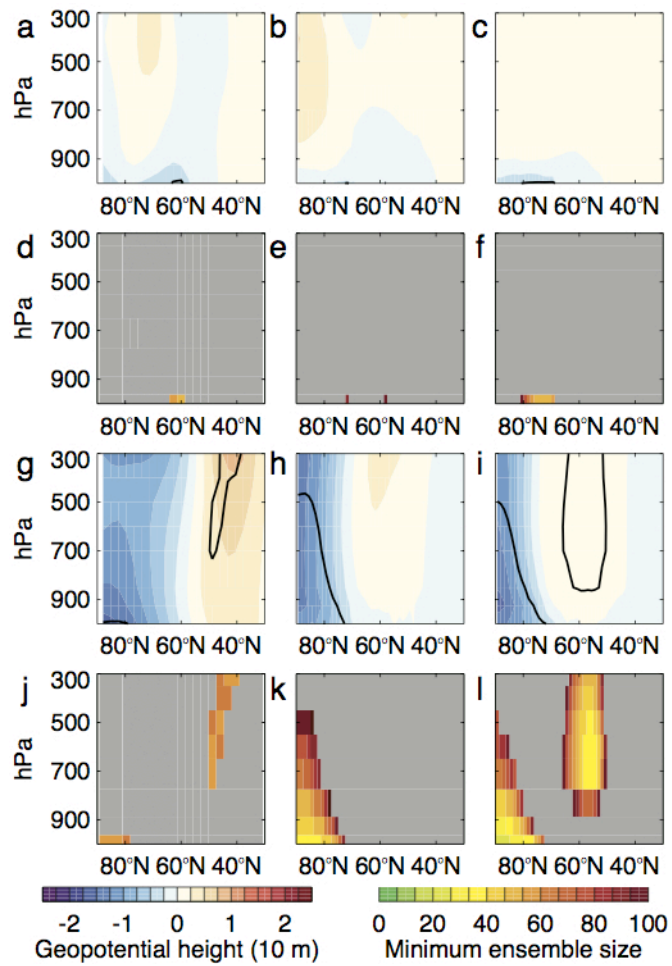
**Figure 6**: As Figure 3, but for geopotential height (Z).

659

660

661
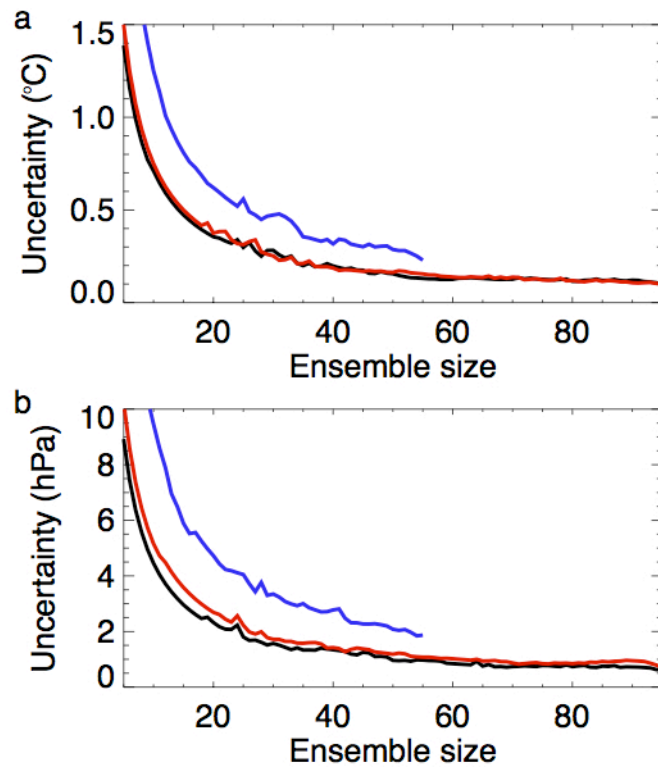
Figure 7: Uncertainty due to atmospheric internal variability [see text for details] as a function of

ensemble size for the Arctic-mean (a) autumn near-surface air temperature (T_{ref}) response and (b)

winter sea level pressure (*SLP*) response. Black and blue lines correspond to PERT-CTRL in the

UM and CAM respectively, and the red lines to PERT*2-CTRL.

666

667

668
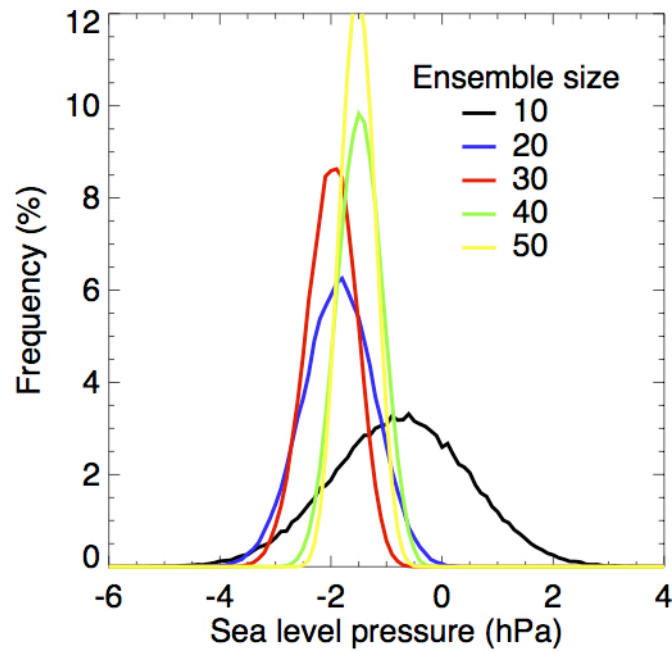
Figure 8: Probability distribution functions (PDF) for the winter Arctic-mean sea level pressure

(*SLP*) responses in sub-ensembles of varying size. Each PDF is constructed from 100,000 unique

combinations sub-sampled from the 60-member CAM ensemble. For example, the blue line

represents sub-ensemble means for 100,000 unique combinations of 20 CAM members sampled

from the full set of 60 CAM members [see text for further details].

674

675

676
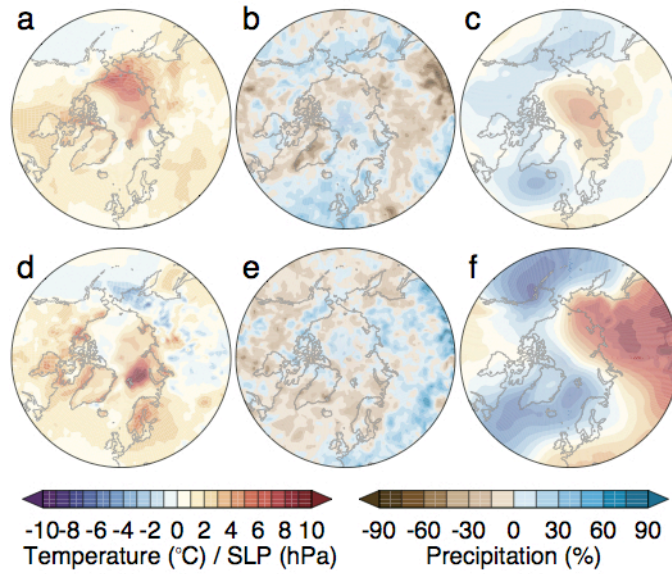
Figure 9: Observed trends in autumn (a) near-surface temperature (T_{ref}), (b) precipitation (P) and (c) sea level pressure (SLP) for the period 1979-2009. Precipitation trends are expressed as percentages relative to the climatological-means. (d-f) As (a-c), but for winter. T_{ref} and SLP data are from the ERA-Interim reanalysis and P data are from the GPCP product.