Clim Dyn
DOI 10.1007/s00382-013-1830-9

Atmospheric impacts of Arctic sea-ice loss, 1979–2009: separating forced change from atmospheric internal variability

James A. Screen · Clara Deser · Ian Simmonds · Robert Tomas

Received: 5 March 2013 / Accepted: 4 June 2013

Abstract The ongoing loss of Arctic sea-ice cover has implications for the wider climate system. The detection and importance of the atmospheric impacts of sea-ice loss depends, in part, on the relative magnitudes of the sea-ice forced change compared to natural atmospheric internal variability (AIV). This study analyses large ensembles of two independent atmospheric general circulation models in order to separate the forced response to historical Arctic sea-ice loss (1979–2009) from AIV, and to quantify signal-to-noise ratios. We also present results from a simulation with the sea-ice forcing roughly doubled in magnitude. In proximity to regions of sea-ice loss, we identify statistically significant near-surface atmospheric warming and precipitation increases, in autumn and winter in both models. In winter, both models exhibit a significant lowering of sea level pressure and geopotential height over the Arctic. All of these responses are broadly similar, but strengthened and/or more geographically extensive, when the sea-ice forcing is doubled in magnitude. Signal-to-noise ratios differ considerably between variables and locations. The temperature and precipitation responses are significantly easier to detect (higher signal-to-noise ratio) than the sea level pressure or geopotential height responses. Equally, the local response (i.e., in the vicinity of sea-ice loss) is easier to detect than the mid-latitude or upper-level responses. Based on our estimates of signal-to-noise, we conjecture that the local near-surface temperature and precipitation responses to past Arctic sea-ice loss exceed AIV and are detectable in observed records, but that the potential atmospheric circulation, upper-level and remote responses may be partially or wholly masked by AIV.

Keywords Arctic sea ice · Atmospheric modelling · Ensembles · Detection and attribution · Internal variability · Signal-to-noise ratio

J. A. Screen (✉)
College of Engineering, Mathematics and Physical Sciences, University of Exeter, Harrison Building, Streatham Campus, North Park Road, Exeter, Devon EX4 4QF, UK
e-mail: j.screen@exeter.ac.uk

J. A. Screen · I. Simmonds
School of Earth Sciences, University of Melbourne, Melbourne, VIC, Australia

C. Deser · R. Tomas
Climate and Global Dynamics, National Center for Atmospheric Research, Boulder, CO, USA

1 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 6 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 6 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,

Published online: 15 June 2013

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; Blüthgen 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 mid-latitude weather (Honda et al. 2009; Petoukhov and Semenov 2010; Liu et al. 2012; Francis and Vavrus 2012; Screen and Simmonds 2013a, b). Progress in understanding the potential large-scale or remote impacts of Arctic sea-ice loss is hampered by large uncertainties in the atmospheric circulation response to sea-ice loss. Observational studies suggest links between autumn sea-ice loss and circulation patterns in the following winter (Francis et al. 2009; Overland and Wang 2010; Wu and Zhang 2010: Strong et al. 2010; Jaiser et al. 2012), but the statistical significance of these linkages has been questioned (Hopsch et al. 2012), causality is unclear and the mechanisms are poorly understood. In model simulations, the spatial pattern, strength, statistical significance and timing of the circulation response to sea-ice loss differs considerably between studies, and can be hard to disentangle from atmospheric internal variability (AIV). AIV, also known as "climate noise", arises from non-linear dynamical processes intrinsic to the atmosphere (see, e.g., Deser et al. 2012 and references therein).

In an attempt to better separate, and quantify, the potential forced response to Arctic sea-ice loss and AIV, this manuscript presents results from large ensembles with two independent models, in which the only prescribed forcing was observed Arctic sea-ice loss. Both models have been run multiple times with identical surface boundary conditions and external forcing, with each run beginning from a different atmospheric initial state. Therefore, the differences between the simulated atmospheric states of each of the ensemble members arise only due to AIV. These ensembles are approximately a factor of ten larger than those used in Screen et al. (2013), and appreciably larger than in most of the studies mentioned above. In part, we seek to answer the question: how many ensemble members are required to detect a significant response (in a particular variable) to Arctic sea-ice loss, if indeed it is possible to detect a significant response at all? This is pertinent to assessing the strength of the forced change compared to AIV and hence, whether it may be observable in the real world.

2 Data and methods

2.1 Simulations

We utilise two independent AGCMs: the UK–Australian Unified Model (UM) version 7.3 and the US National Center for Atmospheric Research (NCAR) Community Atmosphere Model (CAM) version 3. The UM has been developed by the UK Meteorological Office Hadley Centre and is the atmospheric model used in their Global Environmental Model version 2 (HadGEM2) and in the Australian Community Climate and Earth System Simulator (ACCESS). Both HadGEM2 and ACCESS are participating models in the fifth Coupled Model Intercomparison Project (CMIP). The configuration of the UM used here has 38 vertical levels with a horizontal resolution of 1.25 degrees of latitude by 1.875 degrees of longitude. CAM is the atmospheric component of the NCAR Community Climate System Model version 3 (CCSM3), which participated in the third CMIP. The version used here has 26 vertical levels and a spectral resolution of T42, roughly equivalent to 2.8 degrees of latitude and longitude. For further details the reader is directed to Martin et al. (2011) and Bi et al. (2013) for the UM/ACCESS and Collins et al. (2006) for the CAM.

We primarily analyse two distinct simulations, performed identically with each model, termed the control (CTRL) and perturbation (PERT) simulations. In CTRL, the models were prescribed with an annually-repeating monthly cycle of climatological (CLM) sea-ice concentration (*SIC*) and sea surface temperature (*SST*). Monthly-mean *SIC* and *SST* were taken from the Hurrell et al. (2008) data set, updated to 2009, which is derived from a combination of in situ and remotely-sensed observations. In PERT, the linear trend (TRD) in *SIC* over 1979–2009 for each month was added to the climatological monthly values and these CLM + TRD values were prescribed in the models. The prescription of *SST* in PERT was based on the approach introduced by Screen et al. (2013) and was as follows. In grid-boxes and months where the *SIC* TRD is not zero, then the CLM + TRD *SST* were prescribed. Elsewhere, CLM *SST* was prescribed. This approach captures *SST* changes directly related to *SIC* changes, but does not include *SST* changes outside the sea-ice zone (see Screen et al. 2013 for further details and justification). CTRL and PERT were run for 100 years in the UM and for 60 years in CAM. Since the prescribed surface forcing

repeats annually, but the atmospheric initial conditions vary, each year is considered to be an independent ensemble member (atmospheric "memory" is negligible from year-to-year).

A further simulation has been performed with the UM, termed the PERT*2 simulation, in which the linear trends in *SIC* were doubled before being added to climatological values and these CLIM + (TRD*2) values were prescribed in the model. *SST* were prescribed as above, but this time with CLM + (TRD*2) values in place of CLM + TRD values. PERT*2 was run for 100 years in the UM only.

To isolate the atmospheric impacts of sea-ice loss, we compare the ensemble-mean of a particular variable in CTRL with the ensemble-mean in PERT or PERT*2. The ensemble-mean difference, PERT-CTRL or PERT*2-CTRL, is referred to as the "response" to Arctic sea-ice loss in the single- or double-perturbation experiments. We refer to "local" and "remote" responses, by which we mean responses that are in close proximity to sea-ice changes and those that are geographically distant from sea-ice changes, respectively.

2.2 Statistical methods

To test the statistical significance of the ensemble-mean differences we compute the Student's t-statistic, t, using the difference of means test (Von Storch and Zwiers 1999),

$$t = \frac{|\overline{x} - \overline{y}|}{s_p \times \sqrt{\frac{2}{N}}} \tag{1}$$

where $\overline{x}$ is the ensemble-mean from PERT (or PERT*2), $\overline{y}$ is the ensemble-mean from CTRL, N is the ensemble size and s_p is the pooled standard deviation, given by,

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

where x_i is an individual ensemble member from PERT (or PERT*2), y_i is an individual ensemble member from CTRL and n and m are the respective ensemble sizes (in our cases $N = n = m$). The ensemble-mean difference is considered statistically significant when $t \geq t_c$ where t_c is the cutoff value of the Student's t-distribution for a two-tailed probability of 0.025 (i.e., 95 % confidence interval) and $n + m - 2$ degrees of freedom.

To calculate the minimum ensemble-size required to detect a statistically significant ensemble-mean difference, N_{min}, we re-write (1) replacing t with t_c and N with N_{min},

$$t_c = \frac{|\overline{x} - \overline{y}|}{s_p \times \sqrt{\frac{2}{N_{\min}}}} \tag{3}$$

and re-arrange to give,

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

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 ($s_p \ll |\overline{x} - \overline{y}|$) and thus, N_{min} is relatively insensitive to discrepancies in s_p.

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).

3 Results

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 (Fig. 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 and PERT*2 is 2.53 and 1.48 million km^2, respectively, in SON and DJF. Note that the loss of sea-ice area in the double-perturbation experiment is less than twice that in the single-perturbation experiment because the *SIC* in any grid-box cannot be lower than zero. Recall that the boundary conditions in PERT are based on *SIC* trends from 1979 through to 2009, which was the last full year of *SIC* data when the model runs were initiated. The past 3 years (2010, 2011 and 2012) have had low sea-ice coverage, with summer 2012 a new record minimum (Zhang et al. 2013; Parkinson and Comiso 2013), enhancing the long-term trend. The observed sea-ice

area loss from 1979 to 2012, based on the National Snow and Ice Data Center (NSIDC) sea-ice index (http://nsidc.org/data/G02135), is 2.40 and 1.34 million km^2 in SON and DJF, respectively. Thus, the single-forcing experiment represents a smaller (by 28 and 27 % in SON and DJF) loss of sea-ice than observed from 1979 to 2012 and the double-forcing experiment represents a slightly larger (by 5 and 10 % in SON and DJF) loss of sea-ice than observed from 1979 to 2012. Figure 1c, d shows the corresponding differences in *SST* for SON and DJF, respectively. In general, the *SST* warms where *SIC* decreases, and vice versa. By design, *SST* is unchanged in regions of constant or zero *SIC* change. The *SIC* and *SST* differences in the double-perturbation experiment have the same spatial patterns as in Fig. 1, but with differences that are larger in magnitude (not shown).

Figure 2 shows the ensemble-mean T_{ref} responses (a–c; g–i) and associated values of N_{min} (d–f; j–l), with the panels arranged as follows. The first (a–c) and second (d–f) rows correspond to SON and 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 for the single-perturbation experiment in the CAM and UM, respectively, and the third column (c, f, i, l) is for the double-perturbation experiment.

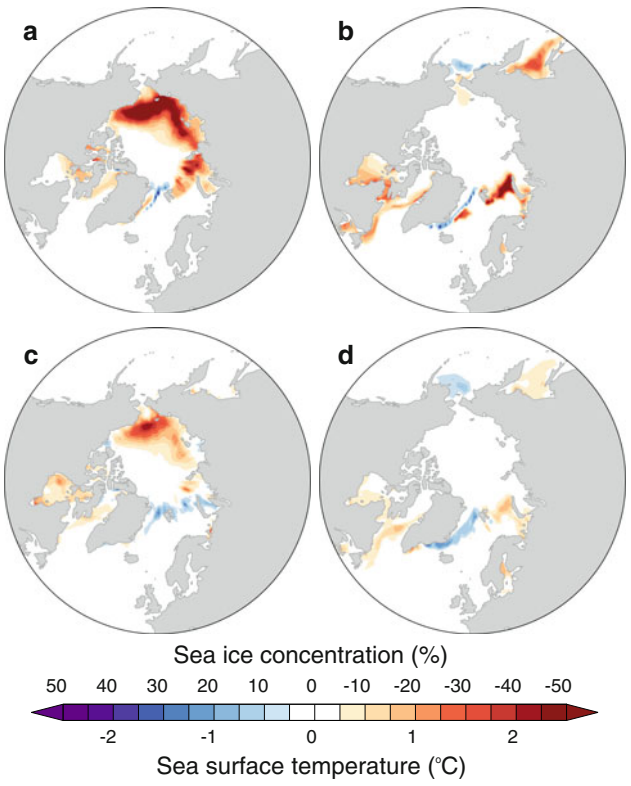

Fig. 1 Ensemble-mean differences (PERT-CTRL) in sea-ice concentration (*SIC*) for **a** autumn and **b** winter. (**c**–**d**) As (**a**–**b**), but for sea surface temperature (*SST*). Note the inverse scale for *SIC*

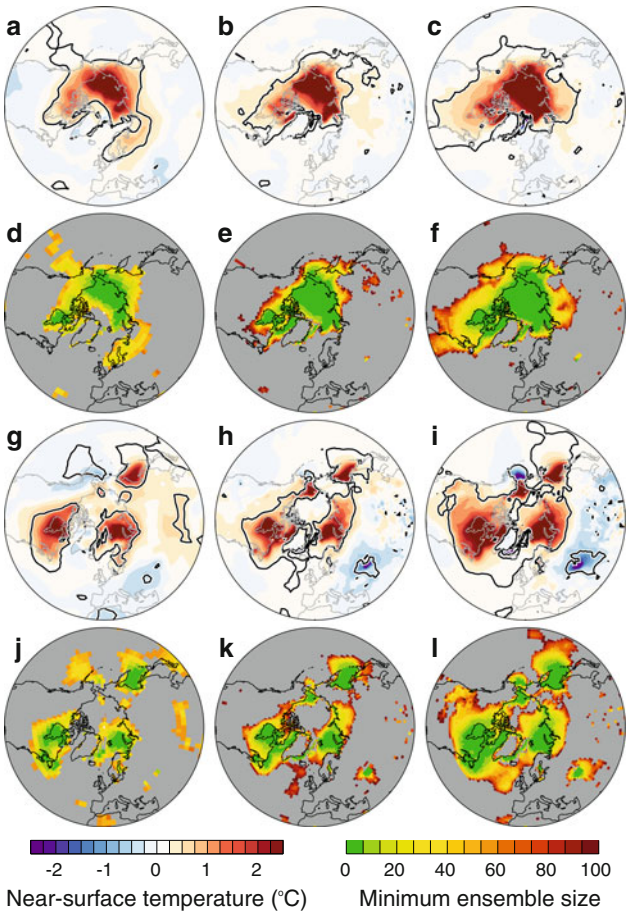

Fig. 2 Ensemble-mean differences in autumn near-surface air temperature (T_{ref}) for **a** CAM PERT-CTRL, **b** UM PERT-CTRL and **c** UM PERT*2-CTRL. Statistically significant differences (at the $p \leq 0.05$ level) are enclosed by *black* contours. **d**–**f** N_{min} for the differences shown in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

In SON, both models show widespread and significant warming over the Arctic Ocean and adjacent continents (Fig. 2a, b). Unsurprisingly, warming is largest over the regions of greatest ice loss (cf. Fig. 1a). The models are in very close agreement. The most obvious difference is that the warming extends further over Scandinavia and north-eastern Russia in the CAM than UM. The DJF responses in both models show four warming centres: the Barents Sea, Hudson Bay, northern Bering Sea and the Sea of Okhotsk (Fig. 2g, h). These regions correspond to areas of winter sea-ice loss and associated *SST* warming (cf. Fig. 1b, d). The atmospheric warming is largely confined to maritime regions in the case of the Bering Sea and the Sea of Okhotsk, but spreads to neighbouring land masses around the Barents Sea and Hudson Bay. Farther away from the regions of sea-ice loss, there are very few areas of significant T_{ref} response in either model. The UM depicts significant cooling over the Caspian Sea and CAM depicts warming over central Asia.

As might be expected, the local T_{ref} response is larger in the double-perturbation experiment than in the single-perturbation experiment (cf. Fig. 2b, c, h, i). Additionally, the T_{ref} response is larger over the high-northern continents and significant T_{ref} responses are detectable at lower latitudes. This suggests that if Arctic sea-ice loss continues unabated, the geographical area affected by sea-ice loss induced warming will increase. Under doubled forcing, there is a weak cooling response over mid-latitude Eurasia in DJF, but this is only significant over a limited area surrounding the Caspian Sea (Fig. 2i).

In terms of N_{min}, five or fewer ensemble members are required to detect a statistically significant T_{ref} response in the proximity of sea-ice loss, irrespective of the model, season or the magnitude of forcing. Away from the regions of ice loss, approximately 30–50 ensemble members are required to detect a significant response. The response over regions adjacent to ice loss is likely mediated by horizontal advection due to synoptic systems (Deser et al. 2010), so it follows that the non-local response will be weaker and subject to larger AIV (i.e., lower signal-to-noise ratio, higher N_{min}) than the local response that is directly driven by surface heat flux changes (e.g., Deser et al. 2010; Screen et al. 2013). Under doubled forcing, the T_{ref} response over the high-northern continents is easier to detect (lower N_{min} than with single forcing) and is detectable further south (cf. Fig. 2e, f, k, l).

Figure 3 shows the latitudinal and vertical structure of the zonal-mean T response. Consistent with earlier work (Screen et al. 2012, 2013), Arctic warming due to observed sea-ice loss is strongest in the lowermost atmosphere and is almost entirely confined to below 700 hPa in both seasons and models (Fig. 3a, b, g, h). In the double-perturbation experiment, the T response is stronger in the near-surface levels, but the response remains trapped in the lower troposphere (Fig. 3c, i). This implies that Arctic sea-ice loss has no discernable influence on T aloft. N_{min} generally increases with altitude. At the few locations where there is significant zonal-mean warming above 700 hPa, N_{min} is 50 or more.

Figure 4 shows the ensemble-mean P responses, arranged as in Fig. 2. In SON, widespread P increases are found over the Beaufort, Chukchi and East Siberia Seas in the UM (Fig. 4b). In the CAM, P also increases in these regions but with less spatial coherence (Fig. 4a). In both models, the P increases are associated with significant increases in cloud cover, principally low cloud, but the cloud responses are weaker in the CAM than UM (not shown). The weaker P and cloud cover responses in the CAM versus UM were previously noted by Screen et al. (2013) and appears to relate to problems with the cloud cover scheme in CAM version 3. In the double-forcing experiment, the P increases are stronger and are significant

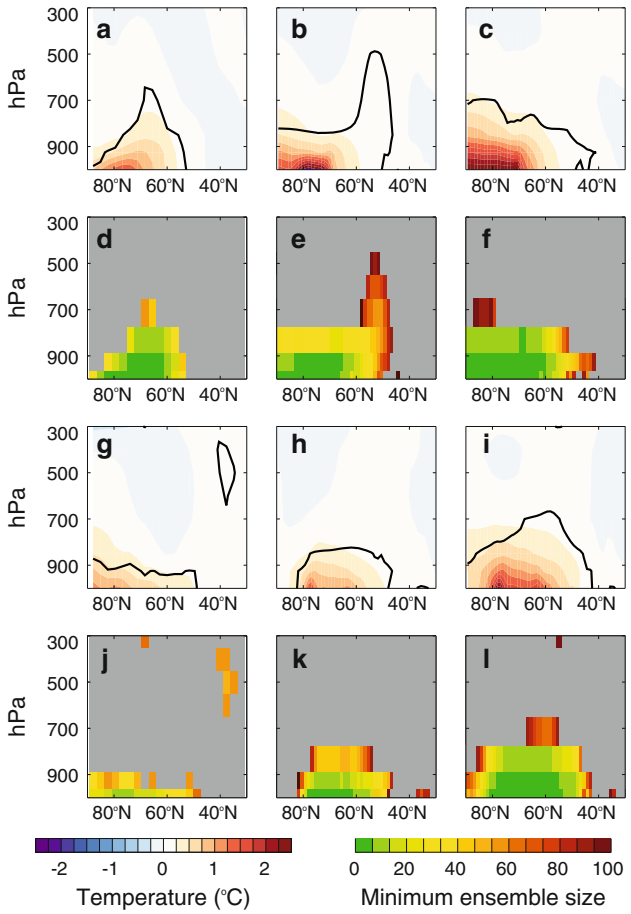

Fig. 3 Zonal-mean ensemble-mean differences in autumn air temperature (T) for **a** CAM PERT-CTRL, **b** UM PERT-CTRL and **c** UM PERT*2-CTRL. Statistically significant differences are enclosed by *black* contours. **d–f** N_{min} for the differences shown in (**a–c**), respectively. *Grey shading*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

over most of the Arctic Ocean (Fig. 4c). Away from the Arctic Ocean there are isolated patches of significant P response in both models, but no large-scale features even in the double-perturbation experiment. In DJF, both models show significant P increases over the regions of winter sea-ice loss: the Barents Sea, Sea of Okhotsk and Hudson Bay (Fig. 4g, h). Remote P decreases occur over the North Atlantic and Pacific in CAM, but they are only significant in small areas (Fig. 4g). In the double-forcing experiment, a similar spatial pattern is found, but with increased magnitude (Fig. 4i). Over regions of maximum sea-ice loss, N_{min} for P is less than 10 and over other regions of sea-ice loss it is around 10–30 (Fig. 4d–f, j–l). The majority of grid-boxes with a significant P response have an associated N_{min} of less than 40, with the main exceptions being the sporadic remote P responses.

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

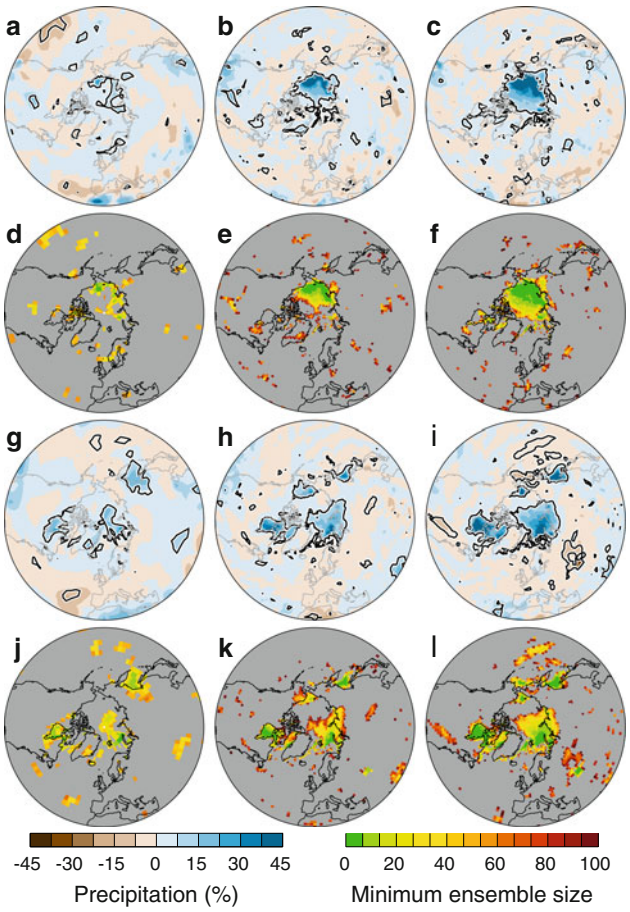

Fig. 4 As Fig. 2, but for precipitation (*P*). The *P* differences are expressed as percentages relative to the ensemble-means in CTRL

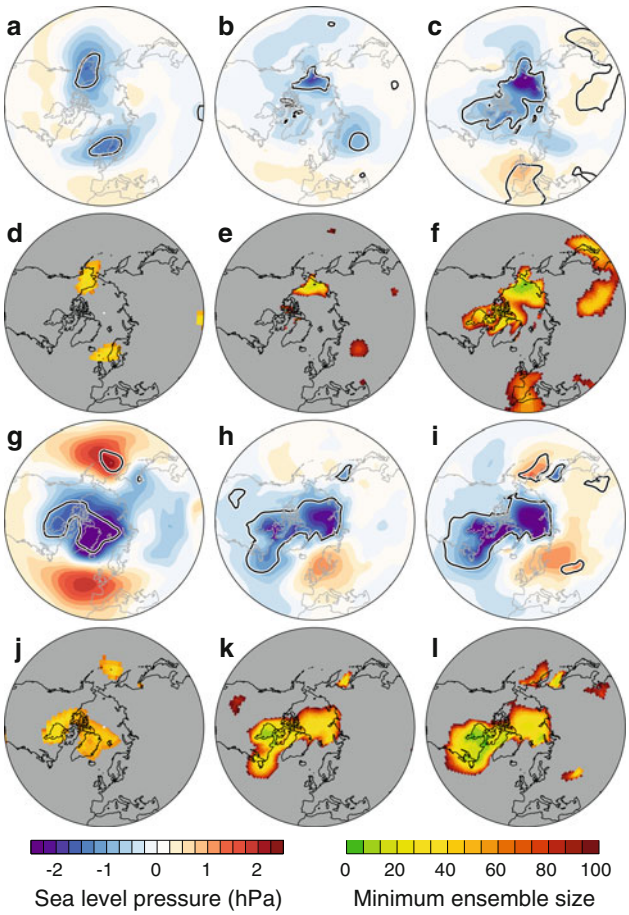

Fig. 5 As Fig. 2, but for sea level pressure (*SLP*)

Siberian Seas in the UM (Fig. 5b). A second low *SLP* centre is located over the Baltic countries. In CAM, two regions of lowered *SLP* are identified in broadly similar, but non-identical, locations (Fig. 5a). The first low is shifted to the southwest to be centred over Alaska and the second low is shifted westward to be located over Scandinavia. Away from these limited regions, the *SLP* response is statistically insignificant in both models. The spatial patterns of the *SLP* responses are largely consistent between the single- and double-perturbation experiments, but there are differences in the magnitudes and significance of the responses (cf. Fig. 5b, c). In the double-perturbation experiment, significantly lowered *SLP* is found over a larger area, including most of the Arctic Ocean, the Canadian Archipelago and Hudson Bay. The low-pressure centre over the Baltic countries that is significant in the single-perturbation is statistically insignificant in the double-perturbation experiment. Conversely, *SLP* increases over Europe and East Asia become significant in the double-perturbation experiment.

In DJF, significant large-scale *SLP* decreases are found over the Arctic Ocean, Hudson Bay and eastern Canada in the UM (Fig. 5h). Isolated regions of significant *SLP* reductions are also identified over the Sea of Okhotsk and central North America. *SLP* is increased over Europe, but this feature is not statistically significant. In the CAM, *SLP* decreases significantly over Hudson Bay, Greenland and the Atlantic-side of the Arctic Ocean (Fig. 5g). *SLP* increases significantly over the Bering Sea. Whilst the two models exhibit broadly similar *SLP* responses in the Arctic and sub-Arctic Canada, the mid-latitude responses are rather different (cf. Fig. 5g, h). In particular, CAM depicts larger *SLP* increases over the north Atlantic and north Pacific than does the UM, and the responses over the United States are opposite in sign between the two models. However, the mid-latitude responses are predominantly statistically insignificant in both models, so these discrepancies can be explained by AIV. No regions show significant responses of opposite sign between the models. The spatial patterns of the DJF *SLP* responses are similar in the single- and double-perturbation experiments (cf. Fig. 5h, i). The *SLP* decrease over the Arctic and Hudson Bay is larger in magnitude in the latter, but the geographical extent of the significant *SLP* response is not overly different. Three

small regions show significant responses in the double-perturbation experiment that are not significant in the single-perturbation experiment. These are *SLP* increases over the Bering Sea, Eastern Europe and eastern China. The region of weak, but significant, *SLP* decrease over central North America in the single-perturbation experiment is not significant in the double-perturbation experiment.

N_{min} for the *SLP* response is as low as 10 in the UM over regions of maximum ice loss, especially in the double-perturbation case, but N_{min} values this low are only found in very limited geographical regions (Fig. 5d–f, j–l). Generally, approximately 30–50 ensemble members are required to detect a significant *SLP* response, and upwards of 50 members are required to detect a significant response in remote regions. It is notable that even with 100 ensemble members in the UM, very few mid-latitude regions show a significant *SLP* response in the single-perturbation experiment. Further, despite larger mid-latitude responses in CAM, an ensemble size of 60 is insufficient for these to achieve statistical significance. This implies that the remote *SLP* response to recent Arctic sea-ice loss is considerably smaller than AIV.

Figure 6 shows the zonal-mean *Z* responses. In SON, the high-latitude response is baroclinic with *Z* decreases in the lowermost atmosphere and *Z* increases aloft. Significant zonal-mean *Z* responses are only found at 1,000 hPa. The vertical profile is fairly consistent across the models and experiments. Taken together, the *SLP* and *Z* responses in SON are suggestive of a shallow thermal (heat) low in response to sea-ice loss. Thermal lows can occur when cold air overlies warmer water, as is the case in regions of sea-ice loss (Higgins and Cassano 2009; Deser et al. 2010; Strey et al. 2010; Orsolini et al. 2012). In DJF, the vertical profile of the *Z* response is completely different. Both models show a quasi-barotropic *Z* decrease over high northern latitudes. This high-latitude *Z* decrease is significant in the UM below 500 hPa, but only at 1,000 hPa in CAM. Both models show *Z* increases over mid-latitudes. In CAM, these extend throughout the troposphere, but are only significant above 700 hPa. In the UM, *Z* increases are found aloft but not at 1,000 hPa, and are shifted polewards in comparison to those in CAM. They are insignificant in the single-perturbation experiment, but significant above 850 hPa in the double-perturbation experiment. In all other respects, the *Z* responses in the single- and double-perturbation are very similar. N_{min} for *Z* is high, typically 50 or above in the single-perturbation experiment and only slightly lower in the double-perturbation experiment.

In summary, the *SLP* and *Z* responses point to rather different spatial and vertical structures to the circulation responses in SON and DJF. In SON, the response is baroclinic (restricted to the near-surface levels) and localised. Similar local circulation responses to sea-ice loss have

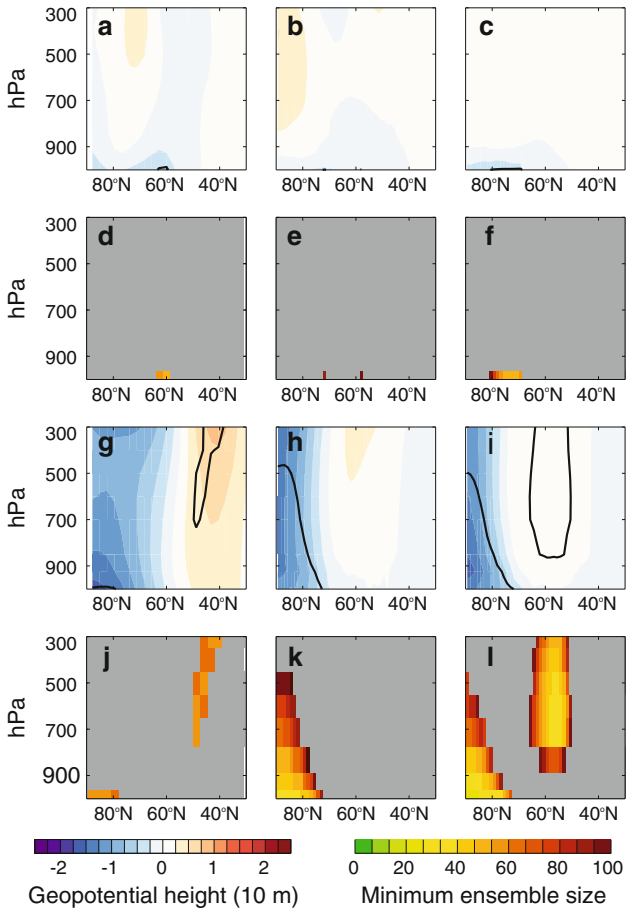

Fig. 6 As Fig. 3, but for geopotential height (*Z*)

been identified in other simulations (Higgins and Cassano 2009; Deser et al. 2010; Strey et al. 2010; Orsolini et al. 2012). By contrast in DJF, the circulation response is fairly barotropic and more spatially extensive. This seasonal transition from a local baroclinic response to a larger-scale barotropic response was also noted by Deser et al. (2010), although the horizontal structure of their winter responses are rather different to that found here. In our CAM simulations, the DJF responses project onto the positive phase of the Arctic Oscillation (AO). This is in contrast to the negative-type AO responses found in February by Deser et al. (2010) and in DJF by Liu et al. (2012), both using CAM but in response to projected future and past sea-ice trends, respectively. Screen et al. (2013) reported a negative North Atlantic Oscillation (NAO) response in early-winter (November–December) in the CAM and UM, but cautioned that the response was weak and often exceeded by AIV. The larger ensembles presented here do not support a shift towards to negative phase of the NAO in response to observed sea-ice loss. Instead, in CAM the response projects onto the positive NAO phase and in the UM the response is not NAO-like. Thus, the wintertime circulation responses (and their interactions with the large-

scale modes of atmospheric variability) are not robust across simulations, even those using the same models.

So far we have considered a limited number of atmospheric variables. For a wider perspective, Table 1 provides the mean N_{min} for a broad selection of atmospheric variables. The values given in Table 1 are averages of N_{min} across all grid-points that exhibit a significant response in that variable (recall N_{min} is undefined where the response is insignificant), all experiments, models (UM and CAM) and seasons (SON and DJF). For example, the mean N_{min} for T_{ref} is the average of all the values in Fig. 2d–f, j–l. Table 1 also provides the mean percentage area of northern hemisphere extra tropics (>30°N) exhibiting a significant response in each variable. Both the mean N_{min} and area metrics mask substantial spatial, seasonal and inter-model variability, so the precise numbers must be interpreted with caution. However, comparison of the mean N_{min} between variables is insightful as it clearly demonstrates that the responses in certain variables are easier to detect than others. To aid interpretation, the variables in Table 1 are listed in order of ascending mean N_{min}. Recall, smaller values indicate that the response is easier to detect than larger values. The ranked variables can be split into four categories of increasing mean N_{min}. This ranking is largely insensitive to whether or not the double-perturbation experiment is included in the analysis (not shown). The variables with smallest values ($N_{min} < 30$) are the surface heat fluxes and T_{ref}. The next group ($30 < N_{min} < 50$) includes variables related to clouds, precipitation and radiation. A third group ($50 < N_{min} < 60$) contains variables related to surface atmospheric circulation, including *SLP* and near-surface wind. The hardest responses to detect ($N_{min} > 60$) are in upper-level variables, for example, mid-tropospheric (500 hPa) temperature (T_{500}) and geopotential height (Z_{500}) and jetstream-level (250 hPa) wind (U_{250}, V_{250}). Although we have not considered stratospheric variables here, Cai et al. (2012) found that the stratospheric response to sea-ice loss is small compared to the tropospheric response.

Table 1 Mean N_{min} for a selection of atmospheric variables

Variable	Mean N_{min}	Mean area
Surface sensible heat flux	21.9	19.1
Surface latent heat flux	24.5	19.7
Near-surface 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
500–1,000 hPa thickness	56.1	7.0
250 hPa zonal wind speed (U_{250})	65.4	5.5
250–1,000 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

For each variable, the value of N_{min} given is the average over all grid-boxes (with a significant response in that variable), all models and experiments, and both autumn and winter. The right-hand column shows the mean percentage area of northern hemisphere extra tropics (>30°N) exhibiting a significant response in that variable (again averaged across models, experiments and seasons). To aid interpretation, the variables are listed in order of ascending mean N_{min}. Note that low values imply a response that is easier to detect than high values

Clearly, AIV is a key source of uncertainty in the simulated atmospheric response to Arctic sea-ice loss. Larger ensembles can reduce this uncertainty by averaging out, to some extent, the effects of AIV. Figure 7 quantifies the reduction in uncertainty in the response to Arctic sea-ice loss, due to AIV, as the ensemble size increases. To construct this figure, we have sub-sampled our large ensembles into smaller sub-ensembles of varying size. For each sub-ensemble size, a large number (100,000) of unique combinations are sampled to produce a large set of sub-ensemble mean responses. For example, for a sub-ensemble size of 5 we sub-sampled 100,000 unique combinations of 5 members from the full set. For each combination, we averaged the selected members to produce a sub-ensemble mean. This results in a set of 100,000 sub-ensemble mean responses. The spread (difference between maximum and minimum values) of these sub-ensemble mean responses provides a measure of the uncertainty in the response due to AIV, for an ensemble of that size. Figure 7a, b shows examples for the Arctic-mean (>70°N) SON T_{ref} response and DJF *SLP* response, respectively, but qualitatively similar results are found for other seasons and variables.

Uncertainty due to AIV, as estimated by the spread of sub-ensemble mean responses, can be seen to decrease almost exponentially as the ensemble size increases. This implies that to reduce uncertainty by one half, the ensemble size has to be doubled. In absolute terms, uncertainty due to AIV decreases rapidly as the ensemble-size increases from 5 (or fewer) to 20 members, and then continues to reduce more slowly as further ensemble members are added. This behaviour is very similar in the two models and in both the single- and double-perturbation experiments, however, CAM has larger AIV than the UM for both T_{ref} and *SLP*. In the UM a point is reached, around 50–60 ensemble members, where adding further ensemble members has almost

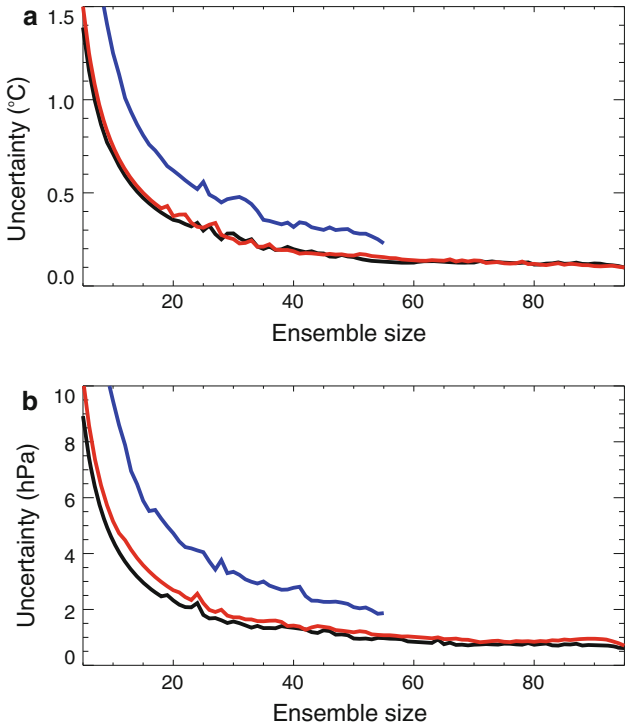

Fig. 7 Uncertainty due to atmospheric internal variability [see text for details] as a function of ensemble size for the Arctic-mean **a** autumn near-surface air temperature (T_{ref}) response and **b** winter sea level pressure (*SLP*) response. *Black* and *blue lines* correspond to PERT-CTRL in the UM and CAM respectively, and the *red lines* to PERT*2-CTRL

no impact on the uncertainty due to AIV (spread of responses). We assume a similar point would occur in the CAM, but we have insufficient ensemble members to confirm this.

4 Discussion and conclusions

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 Fig. 5g and their Fig. 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.

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 subset of our larger ensemble. Thus, it is very unlikely that the discrepancy between the Arctic winter *SLP* responses is this study and in Liu et al. is due to AIV alone. A further difference between this study and Liu et al. is the magnitude of the sea-ice forcing. Although the spatial pattern of the sea-ice forcing is highly similar between the two studies, our forcing is approximately two-to-three times larger magnitude than that in Liu et al. (cf. our Fig. 1a, b and their Fig. 4a, b; note the different colour scales). Whether or not this is the cause of the opposing winter *SLP* responses is unclear, but the discrepancy highlights that experimental differences between simulations, even with the same model, can lead to fundamentally different responses.

We now return to the question posed earlier: how many ensemble members are required to detect a significant response to Arctic sea-ice loss? In reality, there is no simple answer to this question as N_{min} varies considerably in space and by variable. As a general rule of thumb, we suggest that detection of the thermo-dynamical (e.g., *SLP*, wind) response requires an ensemble size approximately twice as large as the thermal response (e.g., surface heat fluxes, *T*). The hydrological response (e.g., cloud, *P*) lays in-between, which likely reflects influences of both thermal and thermo-dynamical factors on these variables. Deser et al. (2012) reported very similar findings in the context of the coupled climate response to greenhouse gas forcing, as did Wehner (2000) and Taschetto and England (2008) both in the context of the atmospheric response to global *SST* and *SIC* trends. The upper-level response to Arctic sea-ice loss (e.g., Z_{500}, T_{500}, U_{250}, V_{250}) is harder to detect than the near-surface response as the signal-to-noise ratio decreases

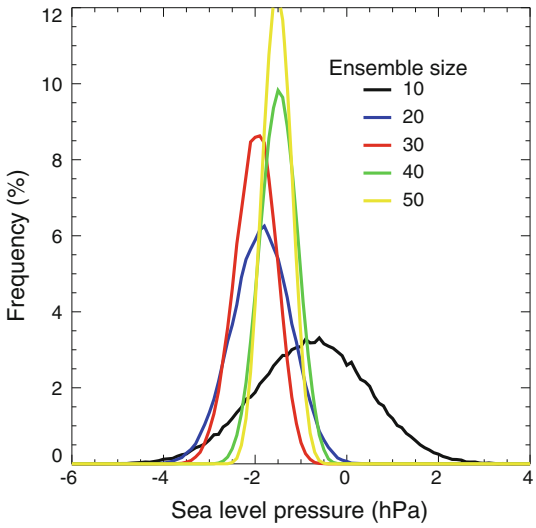

Fig. 8 Probability distribution functions (PDF) for the winter Arctic-mean sea level pressure (*SLP*) responses in sub-ensembles of varying size. Each PDF is constructed from 100,000 unique combinations sub-sampled from the 60-member CAM ensemble. For example, the *blue line* represents sub-ensemble means for 100,000 unique combinations of 20 CAM members sampled from the full set of 60 CAM members [see text for further details]

with altitude. Equally, the remote T_{ref} or P responses (that are mediated by thermo-dynamical processes) are harder to detect than the local T_{ref} and P responses (that are primarily-driven by surface fluxes). Figure 7 suggests large gains, in terms of reduced uncertainty, by increasing from a small (less than 20 members) to moderate-sized (20–50 members) ensemble. Further increases in ensemble size represent a case of "diminishing returns" with smaller reductions in uncertainty per additional ensemble member. On this basis, we argue that an ensemble size of around 50 members is desirable. This is considerably larger than the typical ensemble size used in past studies of the atmospheric response to observed Arctic sea-ice loss (e.g., 5 in Ghatak et al. (2012) and Orsolini et al. (2012), 5 or 8 in Screen et al. (2013), 10 in Strey et al. (2010), 15 in Porter et al. (2012), 20 in Liu et al. (2012)).

The values of N_{min} have implications not only for modelling studies, but also for what aspects of the simulated Arctic sea-ice response may be observable in the real world. Since each ensemble member is 1-year of simulation, N_{min} can also be thought of as approximate measure of the minimum number of years required to detect a significant response due to Arctic sea-ice loss, assuming the rate of loss is linear. The differences in prescribed boundary conditions between CTRL and PERT have, in reality, occurred over a 31-year period (1979–2009). Assuming that the models are realistic in their depiction of the forced response and AIV, a N_{min} of 31 or less suggests that the simulated response should be observable in nature over the period 1979–2009. Conversely, a N_{min} of greater than 31 suggests that more than 31 years are required to separate the forced response from AIV and therefore, the response to past sea-ice loss would not be expected to be detectable in observed records. Accordingly, we argue that the simulated local T_{ref} and P responses to Arctic sea-ice loss should be detectable, but that the atmospheric circulation (e.g., *SLP*, 10 m wind), upper-level (e.g., Z_{500}, T_{500}, U_{250}, V_{250}) and remote responses may be partially or wholly masked by AIV. In practise, the detection and importance of the atmospheric impacts of sea-ice loss not only depend on the relative magnitudes of the sea-ice forced change compared to AIV, but also on the relative magnitudes of sea-ice forced response to other forced responses.

Figure 9a, d shows observed (from ERA-Interim; Dee et al. 2011) trends in T_{ref} over the period 1979–2009 for SON and DJF, respectively. These are highly similar to the simulated T_{ref} responses to Arctic sea-ice loss (Fig. 2), suggesting that the T_{ref} response is indeed detectable, consistent with previous studies (Screen and Simmonds 2010a, b; 2012). Figure 9b, e shows observed trends in P for SON and DJF, respectively, taken from the Global Precipitation Climatology Project (GPCP) data set (Adler et al. 2003). These can be compared to the simulated P responses in Fig. 4. Although the moderate values of N_{min} in Fig. 4 suggested that the local P response may be detectable, the observed trends are not in agreement with the simulated P responses. We propose that there are two likely reasons for this apparent disparity. One reason is that detection of the atmospheric impacts of sea-ice loss not

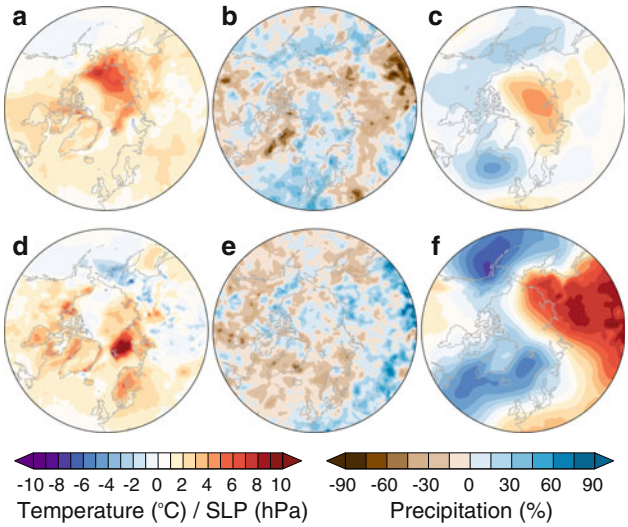

Fig. 9 Observed trends in autumn **a** near-surface temperature (T_{ref}), **b** precipitation (P) and **c** sea level pressure (*SLP*) for the period 1979–2009. Precipitation trends are expressed as percentages relative to the climatological-means. **d–f** As (**a–c**), but for winter. T_{ref} and *SLP*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

only depend on the relative magnitudes of the sea-ice forced change compared to AIV, but also on the relative magnitudes of sea-ice forced response to other forced responses. It is likely that the observed *P* trends are forced by factors other than, or in addition to, sea-ice loss. Secondly, there is considerable uncertainty as to the sign and magnitude of observed *P* trends over the poorly observed Arctic region. *P* trends from alternative observationally constrained data sources—for example the CPC Merged Analysis of Precipitation (CMAP; Xie and Arkin 1997) or ERA-Interim–depict rather different patterns of *P* change over the Arctic Ocean (not shown). It is possible therefore, that the simulated *P* response to sea-ice loss is undetectable in observations because of considerable observational uncertainty. Observed (ERA-Interim) SON and DJF *SLP* trends are shown in Fig. 9c, f. Neither resemble the simulated *SLP* responses in any of the models/experiments (Fig. 5), which is consistent with the conclusion that the *SLP* response to Arctic sea-ice loss is masked by AIV, or *SLP* trends due to other forcing factors. We note that the observed trends are only one realisation (effectively one ensemble member) and likely contain a sizeable component of natural (unforced) variability. Thus, the observed trends would not be expected to match the ensemble-mean simulated responses.

We close by emphasising two arguably obvious, but nonetheless important, considerations. Firstly, this study has only considered the first-order "direct" atmospheric response to Arctic sea-ice loss. The fully coupled climate system response to Arctic sea-ice loss may be different to that shown here. Secondly, Arctic sea-ice loss is only one forcing factor that may be relevant to northern hemisphere climate variability and change. Observed trends reflect changes in multiple forcing factors and the complex interactions between them.

Acknowledgments Laurent Terray is thanked for useful discussions on the statistical methods. Two anonymous reviewers are thanked for the constructive comments. Parts of this research were funded by the Australian Research Council, the Merit Allocation Scheme on the Australian National Computational Infrastructure, the US National Science Foundation (NSF) Office of Polar Programs, and the UK Natural Environment Research Council grant NE/J019585/1. NCAR is sponsored by the NSF.

References

Adler SG et al (2003) The version 2 global precipitation climatology (GPCP) monthly precipitation analysis (1979-present). J Hydrometeor 4:1147–1167

Bi D, Dix M, Marsl SJ, O'Farrell S, Rashid H, Uotila P, Hirst AC, Kowalczyk E, Golebiewski M, Sullivan A, Yan H, Hannah N, Franklin C, Sun Z, Vohralik P, Watterson I, Zhou X, Fiedler R, Collier M, Ma Y, Noonan J, Stevens L, Uhe P, Zhu H, Griffies SM, Hill R, Harris C, Puri K (2013) The ACCESS coupled model: description, control climate and evaluation. Aust Met Oceanog J 63:9–32

Blüthgen J, Gerdes R, Werner M (2012) Atmospheric response to the extreme Arctic sea-ice conditions in 2007. Geophys Res Lett 39:L02707. doi:10.1029/2011GL050486

Cai D, Dameris M, Garny H, Runde T (2012) Implications of all season Arctic sea-ice anomalies on the stratosphere. Atmos Chem Phys 12:11819–11831

Collins W et al (2006) The formulation and atmospheric simulation of the community atmosphere model version 3 (CAM3). J Clim 19:2144–2161

Dee DP et al (2011) The ERA-interim reanalysis: configuration and performance of the data assimilation system. Q J R Meteorol Soc 137:553–597

Deser C, Tomas R, Alexander M, Lawrence D (2010) The seasonal atmospheric response to projected Arctic sea-ice loss in the late twenty-first century. J Clim 23:333–351

Deser C, Phillips A, Bourdette V, Teng H (2012) Uncertainty in climate change projections: the role of internal variability. Clim Dyn 38:527–546

Francis J, Vavrus S (2012) Evidence linking Arctic amplification to extreme weather in mid- latitudes. Geophys Res Lett 39:L06801. doi:10.1029/2012GL051000

Francis J, Chen W, Leathers D, Miller J, Veron D (2009) Winter northern hemisphere weather patterns remember summer Arctic sea-ice extent. Geophys Res Lett 36:L07503. doi:10.1029/2009GL037274

Ghatak D, Deser C, Frei A, Gong G, Phillips A, Robinson D, Stroeve J (2012) Simulated Siberian snow cover response to observed Arctic sea-ice loss, 1979–2008. J Geophys Res 117:D23108. doi:10.1029/2012JD018047

Higgins ME, Cassano JJ (2009) Impacts of reduced on winter Arctic atmospheric circulation, precipitation, and temperature. J Geophys Res 114:D16107. doi:10.1029/2009JD011884

Honda M, Inoue J, Yamane S (2009) Influence of low Arctic sea-ice minima on anomalously cold Eurasian winters. Geophys Res Lett 36:L08707. doi:10.1029/2008GL037079

Hopsch S, Cohen J, Dethloff K (2012) Analysis of a link between fall Arctic sea-ice concentration and atmospheric patterns in the following winter. Tellus 64A:18624. doi:10.3402/tellusa.v64i0.18624

Hurrell J, Hack J, Shea D, Caron J, Rosinski J (2008) A new sea surface temperature and sea-ice boundary dataset for the community atmosphere model. J Clim 21:5145–5153

Jaiser R, Dethloff K, Handorf D, Rinke A, Cohen J (2012) Impact of sea-ice cover changes on the northern hemisphere atmospheric winter circulation. Tellus 64A:11595. doi:10.3402/tellusa.v64i0.11595

Kay J, Holland M, Jahn A (2011) Inter-annual to multi-decadal Arctic sea ice extent trends in a warming world. Geophys Res Lett 38:L15708. doi:10.1029/2011GL048008

Liu J, Curry J, Wang H, Song M, Horton R (2012) Impact of declining Arctic sea-ice on winter snowfall. Proc Natl Acad Sci USA 109:4074–4079

Livina V, Lenton T (2013) A recent tipping point in Arctic sea-ice over: abrupt and persistent increase in the seasonal cycle since 2007. Cryosphere 7:275–286

Martin GM et al (2011) The HadGEM2 family of met office unified model climate configurations. Geosci Model Dev 4:723–757

Orsolini Y, Senan R, Benestad R, Melsom A (2012) Autumn atmospheric response to the 2007 low Arctic sea-ice extent in coupled ocean-atmosphere hindcasts. Clim Dyn 38:2437–2448

Overland J, Wang M (2010) Large-scale atmospheric circulation changes are associated with the recent loss of Arctic sea-ice. Tellus 62A:1–9

Parkinson CL, Comiso JC (2013) On the 2012 record low Arctic sea ice cover: combined impact of preconditioning and an August storm. Geophys Res Lett 40:1356–1361

Petoukhov V, Semenov V (2010) A link between reduced Barent-Kara sea-ice and cold winter extremes over northern continents. J Geophys Res 115:D21111. doi:10.1029/2009JD013568

Porter D, Cassano J, Serreze M (2012) Local and large-scale atmospheric responses to reduced Arctic sea-ice and ocean warming in the WRF model. J Geophys Res 117:D11115. doi:10.1029/2011JD016969

Screen JA, Simmonds I (2010a) The central role of diminishing sea ice in recent Arctic temperature amplification. Nature 464:1334–1337

Screen JA, Simmonds I (2010b) Increasing fall-winter energy loss from the Arctic Ocean and its role in Arctic temperature amplification. Geophys Res Lett 37:L16797. doi:10.1029/2010GL044136

Screen JA, Simmonds I (2012) Declining summer snowfall in the Arctic: causes, impacts and feedbacks. Clim Dyn 38:2243–2256

Screen JA, Simmonds I (2013a) Exploring links between Arctic amplification and mid-latitude weather. Geophys Res Lett 40:959–964

Screen JA, Simmonds I (2013b) Caution needed when linking weather extremes to amplified planetary waves. Proc Natl Acad Sci USA. doi:10.1073/pnas.1304867110

Screen JA, Deser C, Simmonds I (2012) Local and remote controls on observed Arctic warming. Geophys Res Lett 39:L10709. doi:10.1029/2012GL051598

Screen JA, Simmonds I, Deser C, Tomas R (2013) The atmospheric response to three decades of observed Arctic sea-ice loss. J Clim 26:1230–1248

Seierstad I, Bader J (2009) Impact of projected future Arctic sea-ice reduction on extratropical storminess and the NAO. Clim Dyn 33:937–943

Simmonds I, Rudeva I (2012) The great Arctic cyclone of August 2012. Geophys Res Lett 39:L23709. doi:10.1029/2012GL054259

Singarayer J, Bamber J, Valdes P (2006) Twenty-first-century climate impacts from a declining Arctic sea-ice cover. J Clim 19:1109–1125

Strey S, Chapman W, Walsh J (2010) The 2007 sea-ice minimum: impacts on the Northern Hemisphere atmosphere in late autumn and early winter. J Geophys Res 115:D23103. doi:10.1029/2009JD013294

Stroeve J, Serreze M, Holland M, Kay J, Maslanik J, Barrett A (2011) The Arctic's rapidly shrinking sea-ice cover: a research synthesis. Clim Chang 110:1005–1027

Strong C, Magnusdottir G, Stern H (2010) Observed feedback between winter sea-ice and the North Atlantic oscillation. J Clim 22:6021–6032

Taschetto AS, England MH (2008) Estimating ensemble size requirements of AGCM simulations. Meteorol Atmos Phys 100:23–36

Terray L, Corre L, Cravatte S, Delcroix T, Reverdin G, Ribes A (2012) Near-surface salinity as nature's rain gauge to detect human influence on the tropical water cycle. J Clim 25:958–977

Von Storch H, Zwiers FW (1999) Statistical analysis in climate research. Cambridge University Press, Cambridge

Wehner MF (2000) A method to aid in the determination of the sampling size of AGCM ensemble simulations. Clim Dyn 16:321–331

Wu Q, Zhang X (2010) Observed forcing-feedback processes between northern hemisphere atmospheric circulation and Arctic sea-ice coverage. J Geophys Res 115:D14199. doi:10.1029/2009JD013574

Xie P, Arkin PA (1997) Global precipitation: a 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs. Bull Am Meteorol Soc 78:2539–2558

Zhang J, Lindsay R, Schweiger A, Steele M (2013) The impact of an intense summer cyclone on 2012 Arctic sea-ice retreat. Geophys Res Lett. doi:10.1002/grl.50190