

# Uncertainty in climate change projections: the role of internal variability

Clara Deser · Adam Phillips · Vincent Bourdette · Haiyan Teng

Received: 30 July 2010 / Accepted: 18 December 2010 / Published online: 31 December 2010  
© The Author(s) 2010. This article is published with open access at Springerlink.com

**Abstract** Uncertainty in future climate change presents a key challenge for adaptation planning. In this study, uncertainty arising from internal climate variability is investigated using a new 40-member ensemble conducted with the National Center for Atmospheric Research Community Climate System Model Version 3 (CCSM3) under the SRES A1B greenhouse gas and ozone recovery forcing scenarios during 2000–2060. The contribution of intrinsic atmospheric variability to the total uncertainty is further examined using a 10,000-year control integration of the atmospheric model component of CCSM3 under fixed boundary conditions. The global climate response is characterized in terms of air temperature, precipitation, and sea level pressure during winter and summer. The dominant source of uncertainty in the simulated climate response at middle and high latitudes is internal atmospheric variability associated with the annular modes of circulation variability. Coupled ocean-atmosphere variability plays a dominant role in the tropics, with attendant effects at higher latitudes *via* atmospheric teleconnections. Uncertainties in the forced response are generally larger for sea level pressure than precipitation, and smallest for air temperature. Accordingly, forced changes in air temperature can be detected earlier and with fewer ensemble members than those in atmospheric circulation and precipitation. Implications of the results for detection and attribution of observed climate change and for multi-model climate assessments are discussed. Internal

variability is estimated to account for at least half of the inter-model spread in projected climate trends during 2005–2060 in the CMIP3 multi-model ensemble.

**Keywords** Climate change · Uncertainty · Annular modes · Coupled climate models · Climate detection and attribution

## 1 Introduction

Characterizing and quantifying uncertainty in climate change projections is of fundamental importance not only for purposes of detection and attribution, but also for strategic approaches to adaptation and mitigation. Uncertainty in future climate change derives from three main sources: forcing, model response, and internal variability (e.g., Hawkins and Sutton 2009; Tebaldi and Knutti 2007). Forcing uncertainty arises from incomplete knowledge of external factors influencing the climate system, including future trajectories of anthropogenic emissions of greenhouse gases (GHG), stratospheric ozone concentrations, land use change, etc. Model uncertainty, also termed response uncertainty, occurs because different models may yield different responses to the same external forcing as a result of differences in, for example, physical and numerical formulations. Internal variability is the natural variability of the climate system that occurs in the absence of external forcing, and includes processes intrinsic to the atmosphere, the ocean, and the coupled ocean-atmosphere system.

Internal atmospheric variability, also termed “climate noise” (e.g., Madden 1976; Schneider and Kinter 1994; Wunsch 1999; Feldstein 2000), arises from non-linear dynamical processes intrinsic to the atmosphere. Although the atmosphere contains little memory beyond a few weeks,

---

C. Deser (✉) · A. Phillips · V. Bourdette · H. Teng  
Climate and Global Dynamics Division, National Center  
for Atmospheric Research, Boulder, CO, USA  
e-mail: cdeser@ucar.edu

V. Bourdette  
Centre Européen de Recherche et de Formation Avancée en  
Calcul Scientifique (CERFACS), Toulouse, France

it exhibits long-time scale variability characteristic of a random stochastic process. Low frequency variability also arises from processes internal to the coupled ocean-atmosphere system via dynamic and thermodynamic interactions. Thermodynamic coupling between the atmosphere and upper ocean mixed layer produces slow climate fluctuations via the ocean's integration of atmospheric "white noise" turbulent heat flux forcing (e.g., Frankignoul and Hasselmann 1977; Yukimoto et al. 1996; Barsugli and Battisti 1998; Deser et al. 2003; Dommenges and Latif 2008; Clement et al. 2010). Inclusion of dynamical ocean processes produces additional types of low-frequency coupled variability including wind-driven ocean-gyre fluctuations that have been found to play a role in the "Pacific Decadal Oscillation" (Mantua et al. 1997; Schneider and Miller 2001; Schneider and Cornuelle 2005; Kwon and Deser 2007; Alexander 2010). Finally, stochastic atmospheric forcing of internal oceanic variability may contribute to low-frequency climate fluctuations: for example variations in the Atlantic thermohaline circulation may underlie the "Atlantic Multi-decadal Oscillation" (e.g., Delworth et al. 1993).

The unprecedented assemblage of climate model projections from the World Climate Research Programme's (WCRP) Coupled Model Intercomparison Project Phase 3 (CMIP3) archive (Meehl et al. 2007) provides a unique opportunity for estimating uncertainty in climate change. This archive, consisting of forced twentieth and twenty-first century integrations from 23 coupled ocean-atmosphere models, forms the basis for much of the International Panel on Climate Change (IPCC) Fourth Assessment Report (AR4) of Working Group I (Solomon et al. 2007). Uncertainty, as estimated by the spread of the responses across the CMIP3 ensemble relative to the ensemble mean response, has been assessed for a number of climate variables, including air temperature, precipitation and the large-scale atmospheric circulation (see Hegerl et al. 2007; Meehl et al. 2007; and references therein). These uncertainties, as well as those based on long model control integrations, have also been used for estimating the contribution of external forcing to observed climate changes over the twentieth century (Hegerl et al. 2007 and references therein). Recently, Hawkins and Sutton (2009) used the CMIP3 archive to quantify the relative contributions of each source of uncertainty for projected decadal-scale changes in global mean air temperature and precipitation over the twenty-first century. They found that model (forcing) uncertainty dominates before (after)  $\sim 2040$ , while internal variability plays a significant role for *interannual* air temperature changes before  $\sim 2010$ . A follow-up study by Hawkins and Sutton (2010) for regional-scale precipitation found that internal variability is the dominant source of uncertainty for decadal-scale changes in the first few decades, with model variability becoming dominant thereafter. In both studies,

internal variability was defined as the residual from a 4th order polynomial fit to the regional or global mean time series for each model.

An underlying assumption of studies based on the CMIP3 archive is that the multi-model mean response to external forcing yields a more robust estimate of the forced climate signal than the response of any single model due to the reduction in uncertainty associated with model and internal variability (e.g., Tebaldi and Knutti 2007). However, this assumption has been difficult to verify in part due to the limited number of ensemble members for any given model and external forcing scenario (most of the CMIP3 models had fewer than 3 integrations for each forcing scenario). Thus, there is merit in performing a large number of simulations with a single climate model in order to provide a robust estimate of that model's forced response in addition to its internal variability. One such ensemble is the 62-member "Dutch Challenge Project" (Selten et al. 2004) which employed Community Climate System Model Version 1 (CCSM1; Boville et al. 2001) forced by the "business-as-usual" GHG scenario (similar to the SRES A1 scenario) out to 2080. The individual ensemble members, which differ only in their atmospheric initial conditions, were found to exhibit large spread in the future state of the extra-tropical northern hemisphere wintertime atmospheric circulation (Selten et al. 2004; Branstator and Selten 2009).

Here we analyze a new 40-member ensemble for the period 2000–2060 performed with one of the CMIP3 models, Community Climate System Model Version 3 (CCSM3). Compared to the "Dutch Challenge Project", this ensemble uses an improved and higher resolution state-of-the-art climate model and also stronger (and arguably more realistic) forcing consisting primarily of the SRES "A1B" GHG emissions and stratospheric ozone recovery scenarios. We use this ensemble to characterize the forced climate response and accompanying uncertainty due to internal variability. We consider three basic parameters, surface air temperature (TS), precipitation (Precip) and sea level pressure (SLP), for a broad view of the climate response. We also examine the responses as a function of season, highlighting any differences between winter and summer.

The following questions guide our investigation. What is the geographical distribution, magnitude and seasonal dependence of the ensemble mean (e.g., forced) response relative to the internal variability? Does this signal-to-noise ratio differ among the three climate parameters? What is the minimum number of ensemble members needed to detect the forced response with 95% statistical confidence? When can the forced response be detected given an ensemble of size  $n$  where  $n < 40$ ? Is there a relationship between the patterns of the forced response and the leading patterns of internal variability? What are the sources of internal variability, and in particular, how large is the

contribution from internal atmospheric variability (the latter being assessed from a 10,000 year control integration of the atmospheric component of CCSM3)? What are the relative contributions of internal and model variability to uncertainties in climate projections in the multi-model CMIP3 ensemble? Finally, what are the implications of the results based on the 40-member CCSM3 ensemble for detection and attribution studies of observed climate change and for investigations of future climate projections based on multi-model ensembles?

The rest of the paper is outlined as follows. The models and methods are given in Sect. 2. Results are presented in Sect. 3, structured following the sequence of questions listed above. A summary and discussion is provided in Sect. 4.

## 2 Models and methods

### 2.1 Models

CCSM3, a coupled ocean-atmosphere-land-cryosphere general circulation model, has been extensively documented in the *J. Climate* CCSM3 Special Issue (2006). In general, CCSM3 realistically simulates the major patterns of internal climate variability, except for ENSO which exhibits higher regularity and frequency (2–3 year periodicity) than in nature (Deser et al. 2006; Stoner et al. 2009). The 40-member CCSM3 ensemble uses the T42 version (2.8° latitude by 2.8° longitude resolution for the atmosphere, land, and cryosphere components and nominal 1° latitude by 1° longitude resolution for the ocean model component; note that the version of CCSM3 used in the CMIP3 archive was at T85 resolution). Each ensemble member undergoes the same external forcing, the main components of which are the A1B GHG scenario in which CO<sub>2</sub> concentrations increase from approximately 380 ppm in 2000 to approximately 570 ppm in 2060 and stratospheric ozone recovery by 2060, as well as smaller contributions from sulfate aerosol and black carbon changes (see Meehl et al. 2006). It is worth noting that for the period of interest, 2000–2060, the SRES A1B and A2 scenarios are very similar, and both are approximately 30% stronger than the B1 scenario. The ocean, land, and sea ice initial conditions are identical for each ensemble member, and are taken from the conditions on January 1, 2000 from a single 20th century CCSM3 integration. The atmospheric initial conditions differ for each ensemble member, and are taken from different days during December 1999 and January 2000 from the same twentieth century CCSM3 integration. Although the use of a single ocean initial condition may potentially underestimate the true internal variability of the simulated climate system, a recent predictability study using the same 40-member ensemble shows that the effect

of ocean initial conditions is lost within 6–7 years for upper ocean (0–300 m) heat content, and even more rapidly for surface temperature (Branstator and Teng 2010). Thus, the full internal variability is likely to be sampled by perturbing only the atmospheric initial conditions.

In addition to the 40-member CCSM3 ensemble, we make use of a 10,000-year control integration of CAM3, the atmospheric component of CCSM3, at T42 resolution under present-day GHG concentrations. In this integration, sea surface temperatures (SSTs) and sea ice are prescribed to vary with a repeating seasonal cycle but no year-to-year variability. The SST and sea ice conditions are based on observations during the period 1980–2000 from the data set of Hurrell et al. (2008). As in CCSM3, CAM3 is coupled to the Community Land Model (CLM; Oleson et al. 2004).

For the purposes of this study, we form our own CMIP3 multi-model ensemble using a single integration from each of the 21 models forced with the SRES A1B forcing scenario (see Table 10.4 of the IPCC WG1 AR4 Report) excluding CCSM3 to avoid any overlap with present 40-member ensemble. Note that the ozone forcing scenario varies among the CMIP3 models, with nearly half prescribing no change over the twenty-first century (Son et al. 2008).

### 2.2 Methods

We used two methods to compute the climate response: (1) epoch differences between the last 10 years (2051–2060) and the first 10 years (2005–2014); and (2) linear least-squares trends fit to the period 2005–2060. Note that both approaches use data beginning in 2005, 6 years after the integrations start, so as to avoid any artificial reduction in ensemble spread due to the memory of ocean initial conditions (see Branstator and Teng 2010 and related discussion in the Introduction above). The two methods yield virtually identical results.

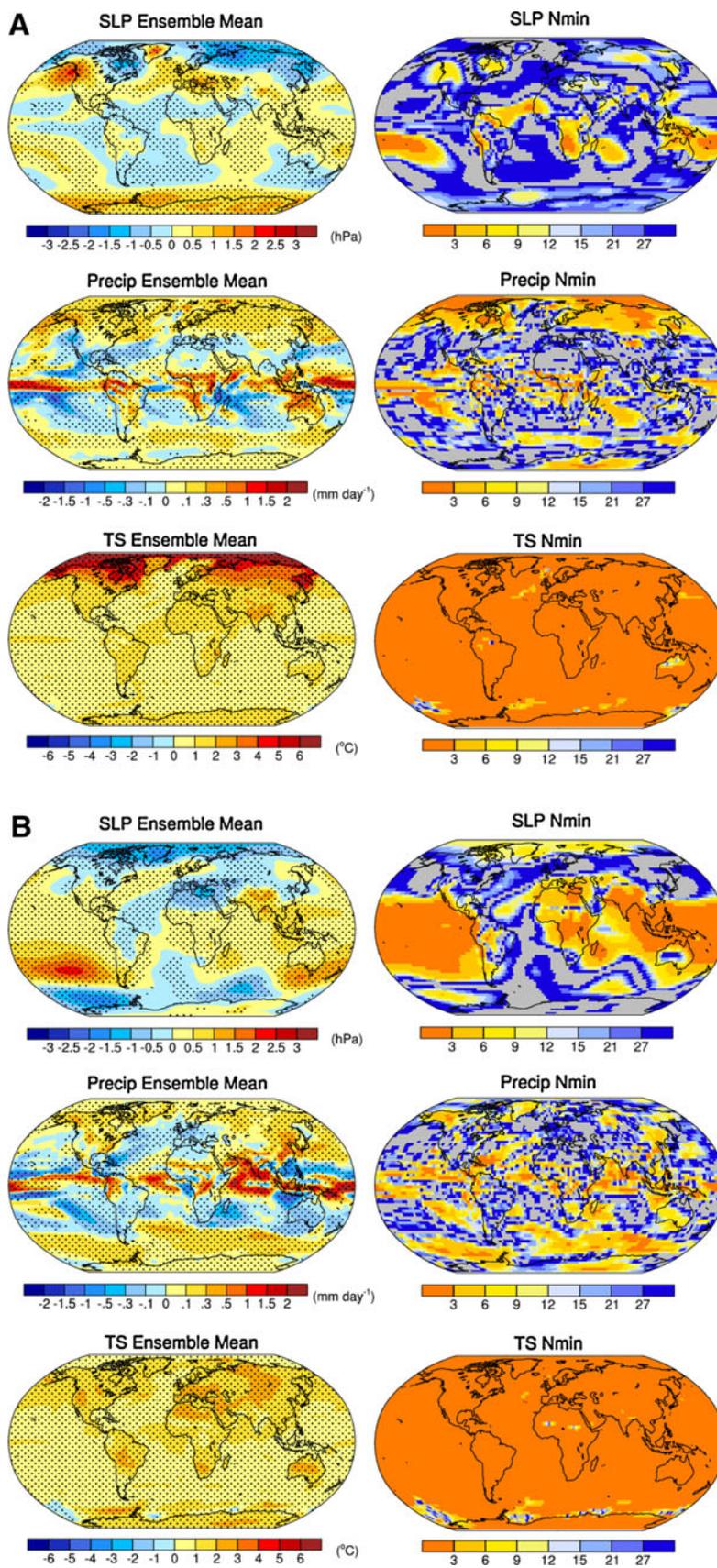
We evaluated the 95% statistical significance of the ensemble mean epoch differences and trends against a null hypothesis of zero change using a 2-sided Student's *t* test (1-sided for TS since the sign of the response is known a priori), where the spread is computed using the individual epoch difference or trend values from the 40 ensemble members. Each ensemble member's epoch difference (or trend) values are assumed to be independent.

## 3 Results

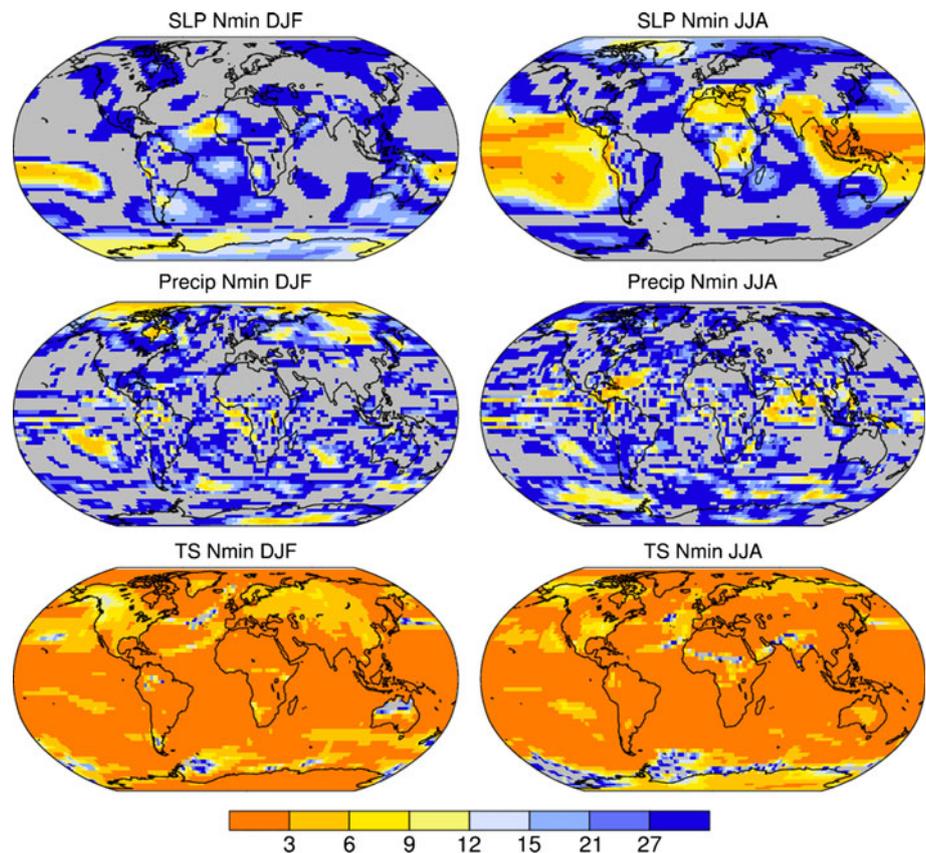
### 3.1 Ensemble mean response and minimum ensemble size requirement

The left-hand panels of Fig. 1a show ensemble-mean epoch difference maps (2051–2060 minus 2005–2014) for

**Fig. 1 a** (Left) CCSM3 40-member ensemble mean epoch differences (2051–2060 minus 2005–2014) in DJF for (top) SLP, (middle) Precip and (bottom) TS. Stippling indicates where the ensemble mean response is statistically significant at the 95% confidence level relative to the spread amongst the ensemble members. (Right) minimum number of ensemble members needed to detect a significant epoch difference response. Gray areas indicate locations where the 40-member ensemble mean response is not significant at the 95% confidence level. **b** As is in **a** but for JJA



**Fig. 2** Minimum number of ensemble members needed to detect a significant epoch difference response (2028–2037 minus 2005–2014) in (left) DJF and (right) JJA for (top) SLP, (middle) Precip and (bottom) TS. Gray areas indicate locations where the 40-member ensemble mean response is not significant at the 95% confidence level



SLP, Precip and TS during December–January–February (DJF) for the 40-member CCSM3 ensemble. Stippling indicates epoch differences that are significantly different from zero at the 95% confidence level relative to the spread of the 40 individual epoch differences, computed according to the formula for the standard error of the mean:

$$X/\sigma \geq (\pm 2/\sqrt{N-1})$$

where  $X$  is the ensemble mean epoch difference,  $\sigma$  is the standard deviation of the 40 epoch differences, and  $N$  is 40. Thus, approximately, if  $X \geq \sigma > 1/3$  then  $X$  is statistically significant at the 95% confidence level. (Note that the factor of ‘2’ in the formula above is replaced by a ‘1’ for TS due to the use of a 1-sided  $t$  test instead of a 2-sided  $t$  test.)

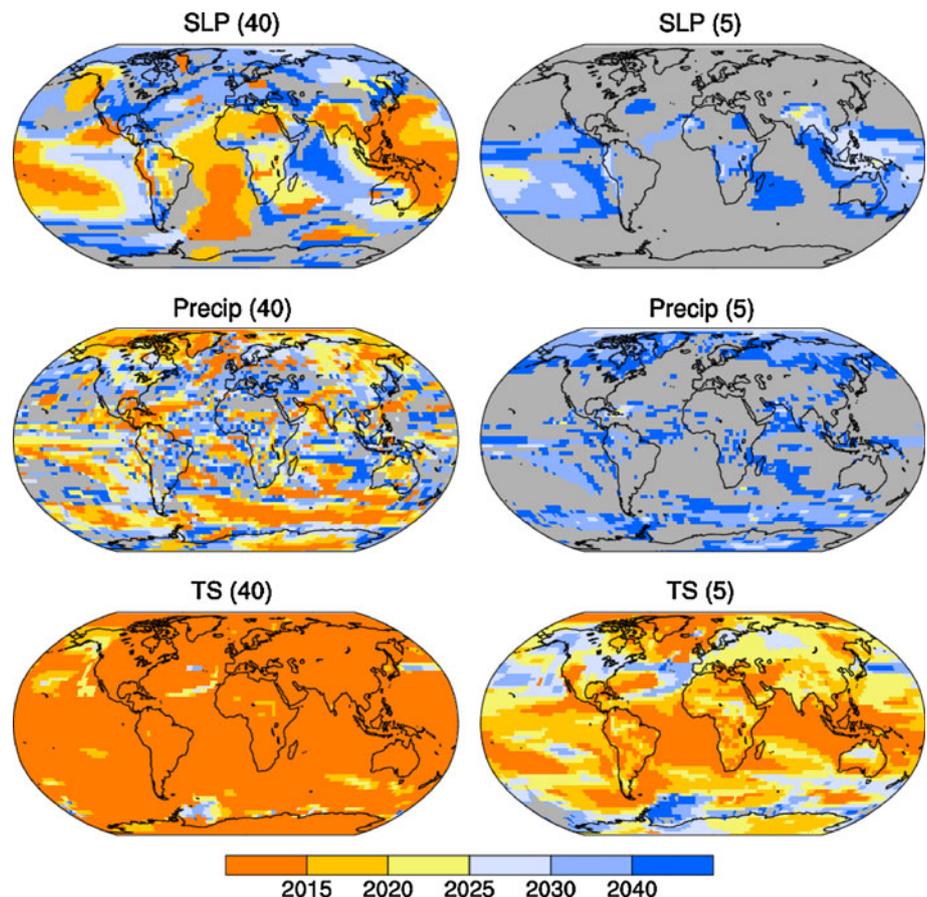
The ensemble mean response is statistically significant over most regions of the globe for all 3 variables. The large-scale SLP response over the Northern Hemisphere (NH) is characterized by generally negative (positive) values at high (middle) latitudes, with maximum amplitudes  $\sim 3$  hPa in the Gulf of Alaska and northern Eurasia. A similar pattern with reversed polarity and somewhat weaker amplitude ( $\sim 1$  hPa) is found over the Southern Hemisphere (SH). These patterns project onto the zonally-symmetric Northern and Southern Annular Modes (NAM

and SAM, respectively; e.g., Thompson and Wallace 2000). The global distribution of SLP changes is broadly consistent with that from the set of 22 CMIP3 models reported in Solomon et al. (2007). We note that the reversed polarity of the response in the SH compared to the NH is due to stratospheric ozone recovery (e.g., Son et al. 2009).

The tropical Precip response consists of mainly positive values along the equator flanked by compensating negative values, especially to the south, with maximum amplitudes  $\sim 2$  mm day $^{-1}$ . The subtropics (extra-tropics) generally exhibit reduced (enhanced) Precip, with magnitudes  $\sim < 0.5$  mm day $^{-1}$ . Surface temperature increases everywhere, with larger warming over land than ocean and maximum warming over the ice-covered Arctic Ocean and adjacent continents (maximum values  $\sim 6^\circ\text{C}$ ), the latter attributable to Arctic sea ice loss in late autumn (Deser et al. 2010). The Precip and TS responses are similar to those documented from other models (Solomon et al. 2007) and the 21 CMIP3 multi-model mean (not shown).

The right-hand panels of Fig. 1a show the minimum number of ensemble members needed to detect the forced (ensemble mean) response at the 95% significance level at each grid box, computed by inverting the formula for the

**Fig. 3** Decade when the ensemble mean change relative to the period 2005–2014 first becomes detectable at the 95% significance level for an ensemble size of 40 (left) and 5 (right), based on annual averages subject to a 10-year running mean for SLP (top), Precip (middle) and TS (bottom). Year indicated denotes the mid-point of the 10-year period. Gray areas indicate locations where the ensemble mean response is not significant at the 95% confidence level



standard error of the mean following Sardeshmukh et al. (2000):

$$N_{\min} = 8/(X/\sigma)^2$$

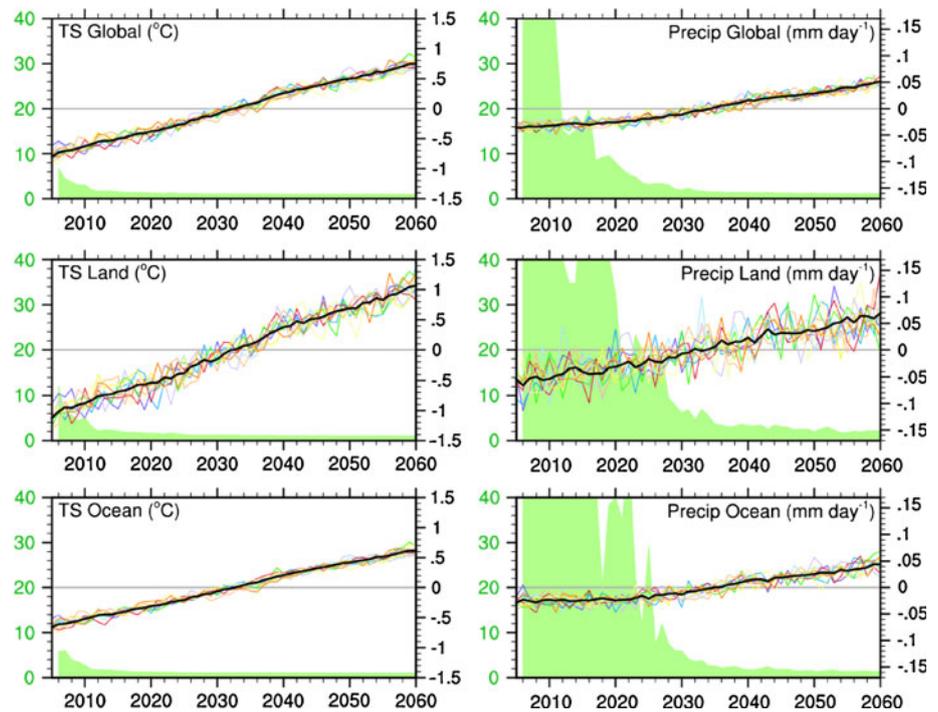
(As with the standard error formula, the factor ‘8’ in the formula for  $N_{\min}$  is replaced by a ‘4’ for TS due to the use of a 1-sided  $t$  test.) In general, SLP requires larger values of  $N_{\min}$  than Precip, and TS requires the smallest  $N_{\min}$ . Values of  $N_{\min}$  for SLP ranges from <6 in parts of the tropics (the southwestern Pacific and Indian Oceans, the northern Atlantic, and South Africa) to ~6–9 for the large-amplitude response centers in the extra-tropics (for example the west coast of Canada) and >15–21 over remaining areas, notably the Mediterranean and high latitudes of both hemispheres. Precip generally exhibits smaller values of  $N_{\min}$  than SLP, with values ~3–6 over the Arctic Ocean and northern high latitude continents, East Antarctica, and many areas of the tropics. Other regions require higher  $N_{\min}$  such as the Southern Ocean (9–12) and northern middle latitudes (>15). TS requires generally fewer than 3 ensemble members, except for isolated regions in the Southern Ocean, the eastern North Atlantic and northwestern Australia.

Ensemble mean epoch difference maps and ensemble size requirements for the June–July–August (JJA) season

are shown in Fig. 1b. The SLP response pattern is considerably different in JJA compared to DJF. For example, over the SH the quasi-zonally symmetric pattern in DJF (e.g., the SAM) is replaced with a regional meridional dipole over the Pacific sector. In addition, the positive SLP response centered over the Mediterranean region in DJF is replaced by negative values in JJA. SLP decreases over the Arctic Ocean in both seasons, although the maximum negative anomalies are centered over the western Arctic in JJA compared to the eastern Arctic in DJF. The Precip response pattern in JJA is similar to that in DJF except that the tropical signals are largest within the NH following the position of the sun. The biggest difference between the TS responses in JJA and DJF is the lack of northern hemisphere polar amplification in summer, consistent with the muted influence of Arctic sea ice loss during this season (Deser et al. 2010).

In general, fewer ensemble members are needed for detecting significant SLP changes in JJA compared to DJF (Fig. 1b, right). For example,  $N_{\min} < 3$  over the entire tropical Pacific, and <12 over the Arctic and portions of the Southern Ocean. On the other hand, larger (smaller) values of  $N_{\min}$  are needed to detect the enhanced Precip over the

**Fig. 4** Time series of annual mean (left) TS and (right) Precip anomalies averaged over the (top) globe, (middle) land and (bottom) ocean for the 40-member ensemble mean (thick black curve) and the first 10 ensemble members (thin colored curves). The green shaded curve shows the minimum number of ensemble members needed to detect a 95% significant change relative to 2005 as a function of time

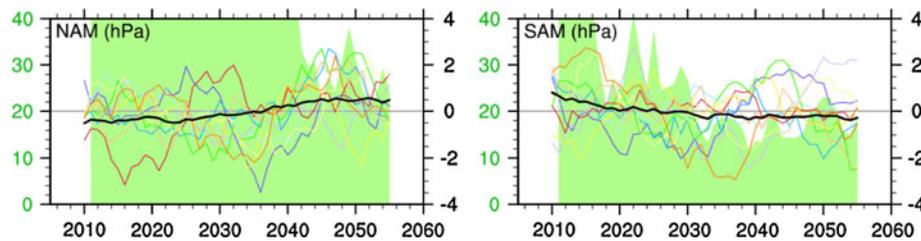


Arctic (Southern Ocean) in JJA compared to DJF, in part related to the weaker (stronger) amplitude of the signal. The ensemble size requirements for TS in JJA are similar to those in DJF.

What are the minimum ensemble size requirements for detecting the forced climate signal near the mid-point of the integration period? Figure 2 shows the distributions of  $N_{\min}$  in winter and summer for epoch differences based on the decade 2028–2037 relative to the decade 2005–2014. Many of the climate changes that become significant in 2051–2060 are not yet significant in 2028–2037 (with an ensemble of 40 members), and those that are require considerably more ensemble members to detect them. For example, the forced SLP response over the northern hemisphere in DJF for 2028–2037 is not detectable with a 40 member ensemble except in a few areas of North America and Siberia where  $N_{\min} > 27$ . SLP changes in 2028–2037 in JJA remain detectable, but  $N_{\min}$  increases by 3–9 compared to that for 2051–2060. Despite that the DJF SLP response over the NH high latitudes is largely undetectable, the DJF Precip response over the Arctic and adjacent continents is detectable by 2028–2037 albeit with a larger ensemble size ( $N_{\min} \sim 3$ –9) than for 2051–2060 ( $N_{\min} < 3$ ). Precip changes in other regions require an increase in ensemble size of  $\sim 3$ –9 members relative to that for 2051–2060, similar to the results for SLP. Finally, values of  $N_{\min}$  needed to detect significant changes in TS in 2028–2037 remain  $< 3$  over much of the tropics but increase to 3–9 over portions of North America, Eurasia, Australia and Antarctica as well as the North Pacific and

Atlantic and the Southern Ocean. Coastal regions of Antarctica in austral winter require  $N_{\min} > 21$ –27.

When does the forced signal first become detectable with an ensemble of  $n$  members (where  $n$  is  $\leq 40$ )? Here we consider 10-year running means; less (more) temporal smoothing would yield later (earlier) detection times. Figure 3 shows the 10-year period (centered in the year given) when the forced signal becomes 95% significant relative to the decade 2005–2014 for an ensemble size of 40 (left panels) and 5 (right panels) based on annual averages. With a 40-member ensemble, decadal SLP changes are detectable within approximately 5–10 years (2015–2020) over the tropical western Pacific and tropical Atlantic Oceans and around 2030 elsewhere; decadal Precip changes are detectable in the next 5–10 years over the Arctic, the Southern Ocean, and portions of the tropics, and around 2030 over Europe; and decadal TS changes are detectable within the next 5 years (e.g., by 2015) over most regions (10–15 years over Alaska and the eastern North Pacific). With a 5-member ensemble, detection of the forced SLP signal over the tropics is delayed to 2030–2040, and no detection is possible for extra-tropical SLP. Similarly, detection of the forced Precip signal with a 5-member ensemble is generally confined to polar regions, the Southern Ocean and portions of the tropical oceans, with detection times around 2030–2040. Although TS continues to be detectable with a 5-member ensemble at nearly all locations, the time of detection is delayed to 2020–2030 over much of Eurasia and North America and parts of the Southern Ocean, and



**Fig. 5** Ten-year running mean DJF time series of the NAM (left) and SAM (right), defined as the zonally-averaged SLP anomaly difference between high ( $55^{\circ}$ – $90^{\circ}$ ) and middle latitudes ( $30^{\circ}$ – $55^{\circ}$ ) of the northern and southern hemisphere, respectively. The *thick black*

*curve* denotes the 40-member ensemble mean, and the *thin colored curves* denote the first 10 ensemble members. The *green shaded curve* shows the minimum number of ensemble members needed to detect a 95% significant change relative to the decade centered on 2010

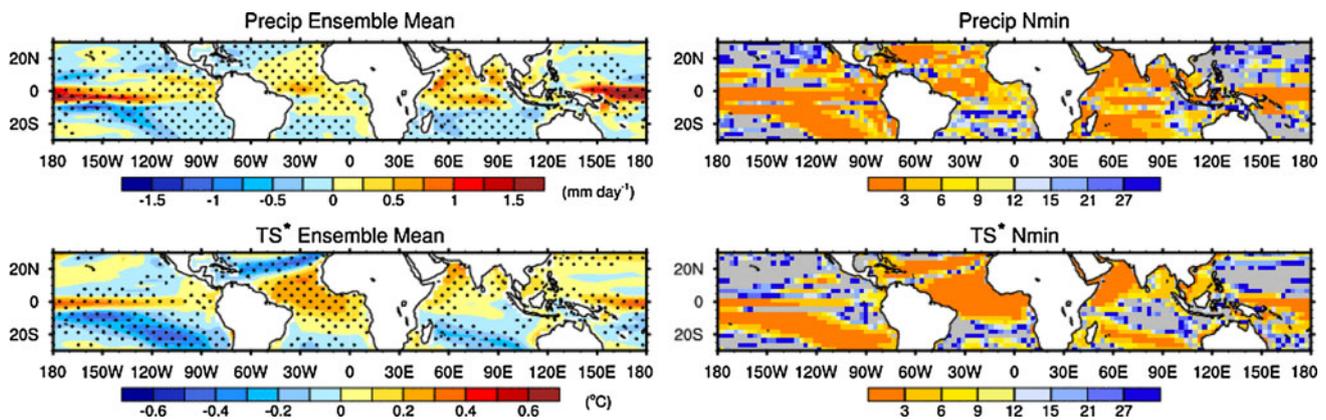
to 2015–2020 over Africa and portions of South America as well as the eastern tropical Pacific. A complementary analysis of the timing of the externally-forced global warming signal (relative to the period 1910–1959) in the TS field from observations and the CMIP3 twentieth and twenty-first century model simulations was presented in Kattsov and Sporyshev (2006).

It is instructive to view  $N_{\min}$  (a measure of the amplitude of the forced signal relative to the noise) as a function of time for a given region. Figure 4 shows the ensemble mean time series of annual mean Precip and TS averaged over the globe, and over land and ocean areas separately, along with the associated  $N_{\min}$  time series. Also shown are the Precip and TS records for each of the first 10 realizations to illustrate the ensemble member spread as a function of time. A statistically significant increase in TS relative to 2005 is detected within a year with an ensemble of 4–6 members, and within approximately 10 years with a single realization for global, land and oceanic averages. In contrast, detection of a statistically significant increase in Precip with a 40-member ensemble does not occur until 2012 for global means and approximately 2023 (2020) for ocean (land) averages. With just a few realizations, detection of a significant Precip change occurs in the mid 2020s for global averages and in the early 2030s for land and ocean averages. The similarity of the detection times for the marine and terrestrial averages is due to the compensation between the magnitude of the intra-ensemble spread and the forced signal. That is, the land records exhibit a larger spread and a larger forced signal than their oceanic counterparts.

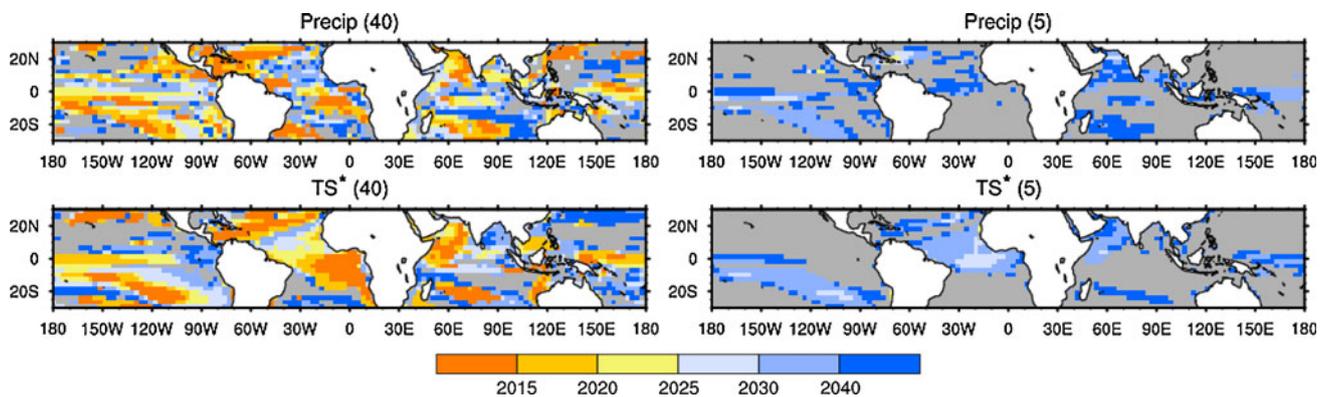
The DJF NAM and SAM indices, defined as the zonally-averaged SLP difference between high ( $55^{\circ}$ – $90^{\circ}$ ) and middle latitudes ( $30^{\circ}$ – $55^{\circ}$ ) of their respective hemispheres, are shown in Fig. 5 along with their associated  $N_{\min}$  time series. The indices have been smoothed with a 10-year running mean and are displayed as differences relative to the period 2005–2014; the calculation of  $N_{\min}$  is based on the 10-year running mean records. The ensemble-mean low-pass filtered NAM record exhibits a

monotonic upward trend, but due to the considerable spread amongst the individual ensemble members, the time of detection of the forced NAM response does not occur until 2042 with a 40-member ensemble, and a relatively large number of realizations ( $\sim 25$ ) is needed to detect the response thereafter. The downward trend in the ensemble-mean low-pass filtered SAM record is detectable by 2017 with a 40-member ensemble, and  $\sim 14$ – $18$  realizations are needed to detect the response thereafter. Thus, the signal-to-noise ratio of the forced trend is larger for the SAM than the NAM.

A recent study by Xie et al. (2010) emphasized the importance of spatial gradients in the tropical Sea Surface Temperature (SST) response to GHG forcing. In particular, they showed that the pattern of the tropical precipitation response is positively correlated with spatial deviations of the SST response from the tropical mean warming. To explore this aspect, Fig. 6 compares the annual ensemble mean epoch difference maps (2051–2060 minus 2005–2014) and associated  $N_{\min}$  distributions for Precip (top) and TS\* (bottom), defined as the residual SST response from the tropical mean ( $30^{\circ}\text{N}$ – $30^{\circ}\text{S}$ ). Consistent with Xie et al. (2010), the pattern of the tropical Precip response is similar to that of TS\*, with positive values over the equatorial Pacific, northern Indian Ocean and tropical Atlantic, and negative values elsewhere. The  $N_{\min}$  distributions associated with these responses are also similar, with values  $< 3$ – $6$  in the equatorial and southeastern Pacific, the tropical Atlantic, and the off-equatorial western Indian Ocean. Finally, the time of detection of the annual TS\* response is comparable to that of the annual Precip response, based on 10-year running means relative to 2005–2014 (Fig. 7). Detection times are approximately 2015–2020 (2030–2040) based on a 40-member (5-member ensemble) in regions where  $N_{\min} < 3$ – $6$  (Fig. 7). Thus, the spatially-varying component of the forced SST response in the tropics exhibits a similar spatial pattern and signal-to-noise ratio (as measured by  $N_{\min}$  and detection time) as the total Precip response, corroborating the results of Xie et al. (2010).



**Fig. 6** As in Fig. 1a but for tropical Precip (*top*) and TS\* (*bottom*) based on annual values. TS\* is defined as TS minus the tropical mean (30°N–30°S) TS computed from oceanic grid points only. Values over land are omitted



**Fig. 7** As in Fig. 3 but for tropical Precip (*top*) and TS\* (*bottom*). TS\* is defined as TS minus the tropical mean (30°N–30°S) TS computed from oceanic grid points only. Values over land are omitted

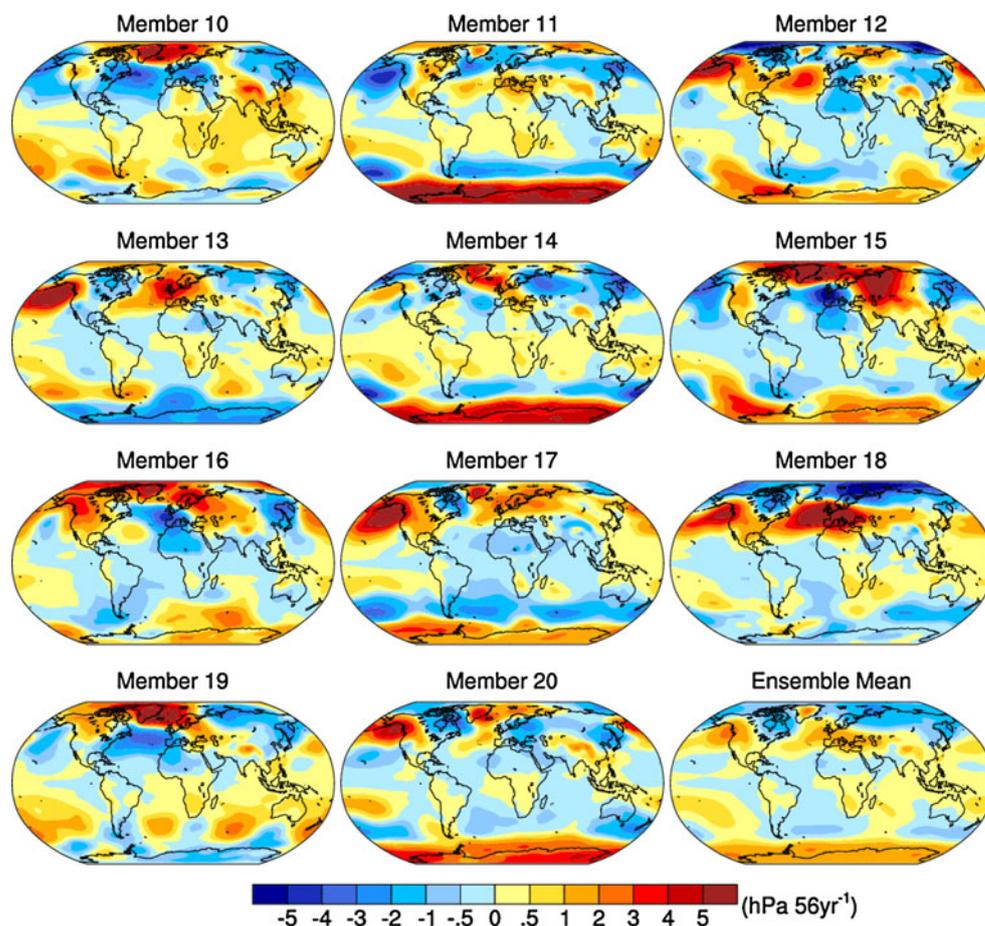
### 3.2 Characterization and mechanisms for uncertainties in future climate trends: the role of “weather noise”

To illustrate the range of uncertainty in future SLP changes, Fig. 8 shows linear trend maps for ensemble members 10–20 individually and for the 40-member ensemble mean (lower right panel) based on DJF during 2005–2060 (similar results are obtained for epoch differences; not shown). The individual realizations reveal a wide range of trend responses to the same external forcing (other ensemble member subsets show a similar range of patterns; not shown). For example, members 11 and 13 exhibit similar patterns over the extratropics but generally opposite polarity, while members 13 and 17 exhibit similar patterns and the same (opposite) sign over the NH (SH). Other members show different spatial distributions: for example member 19 exhibits a zonal wave 3 response over the southern hemisphere in contrast to the more zonally symmetric responses of members 11 and 14. The wide variety of SLP responses in individual realizations underscores the need for a large ensemble ( $\sim 20$ – $30$  members) for accurate estimation of the forced response.

What mechanisms contribute to the spread of the trends across the ensemble members? First we consider the role of internal atmospheric variability using the 10,000-year CAM3 control integration. In this integration, the specified repeating seasonal cycles of SST and sea ice are based on observations from the period 1980–2000. Ideally, the CAM3 control integration should be forced with the SST and sea ice conditions simulated by the CCSM3 40-member ensemble mean during 2000–2060 to obtain identical boundary conditions for the two sets of experiments; however, the differences between atmospheric internal variability under observed present-day (1980–2000) and simulated future (2000–2060) SST and sea ice conditions are likely to be small compared to the magnitude of the internal variability itself.

The spread of the trends across the ensemble members, assessed by the standard deviation, are compared for CCSM3 and CAM3 in Fig. 9a and b for DJF and JJA, respectively. Trends were computed for the period 2005–2060 (56 years in length) for CCSM3 and for 56-year periods from the CAM3 control integration obtained by dividing the 10,000-year record into 178 consecutive

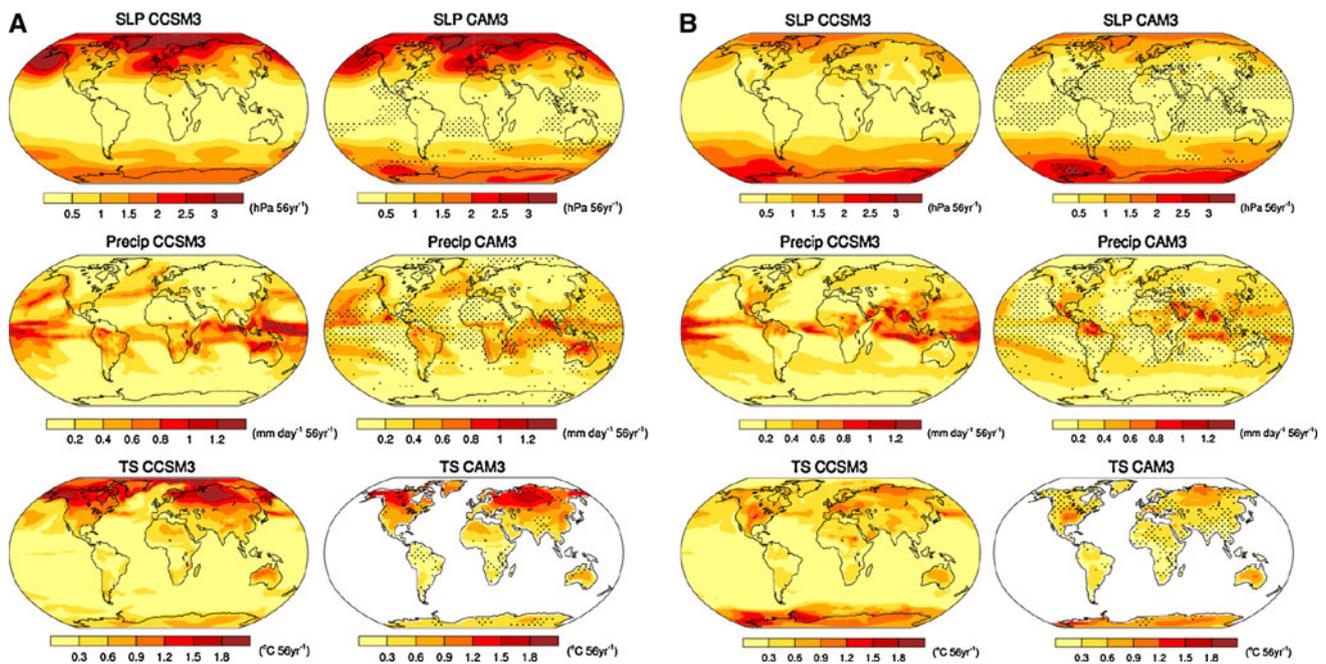
**Fig. 8** DJF SLP trends for individual ensemble members 10 through 20, computed over the period 2005–2060. The ensemble-mean trend based on all 40 members is shown in the lower right panel



segments. Standard deviations that differ significantly between the two sets of experiments, as assessed by an  $f$  test at the 95% confidence level, are indicated with stippling on the CAM3 panels. In each season, the spread of the SLP trends is remarkably similar for the two models in both pattern and amplitude, with significant differences only over the tropics and subtropics especially in JJA. In particular, the large trend standard deviations over the NH and SH extra-tropics in CCSM3 in both seasons are consistent with the null hypothesis of internal atmospheric variability. The patterns of the spread in the Precip trends are also generally similar between CCSM3 and CAM3, but the magnitudes are significantly greater (by approximately a factor of 2–3) within the tropics when the ocean is allowed to interact with the atmosphere. The terrestrial distributions of the spread in TS trends (note that only terrestrial values are shown for CAM3 due to the lack of interannual variability of specified TS values over the oceans) also show similar patterns in CCSM3 and CAM3, with maximum values of 1–2°C over the high latitude NH continents in DJF. Most terrestrial regions in the NH in DJF and in the SH in JJA show no significant differences in amplitude between CCSM3 and CAM3. In summary, internal atmospheric variability contributes substantially to

the spread of the SLP, Precip and terrestrial TS trends during 2005–2060 in the 40-member CCSM3 ensemble.

To characterize the dominant patterns of uncertainty in future climate trends, we have performed EOF analysis on the set of 40 (178) trend maps from CCSM3 (CAM3) for each variable (SLP, Precip and TS). A separate EOF analysis, based on the area-weighted covariance matrix, has been computed for each hemisphere poleward of 30°, and also for the tropics (30°N–30°S). Note that because ensemble mean trends are removed in the EOF procedure, the results characterize the dominant patterns of the “noise” component of the future trends. The leading EOF of extra-tropical SLP trends for each season (DJF and JJA) and hemisphere are shown in Fig. 10 for both the 40-member CCSM3 ensemble and the “178-member” CAM3 control integration. In both hemispheres and seasons, the leading EOF is characterized by an annular mode structure consisting of zonally symmetric anomalies of opposite sign north and south of approximately 55°–60°. These patterns, referred to as the NAM and SAM, also characterize the leading EOF of the interannual variability (not shown). The annular modes account for similar percentages of the total variance in both models, with more variance explained by the SAM compared to the NAM especially in DJF



**Fig. 9** a Trend standard deviations in DJF from the 40-member CCSM3 ensemble (*left*) and the “178-member” CAM3 control integration (*right*) for SLP (*top*), Precip (*middle*) and TS (*bottom*). Trends are computed over the period 2005–2060 for CCSM3 and for

56-year non-overlapping segments for CAM3. *Stippling* in the *right-hand* panels indicates where differences between the two models are statistically significant at the 95% confidence level. **b** As in **a** but for JJA

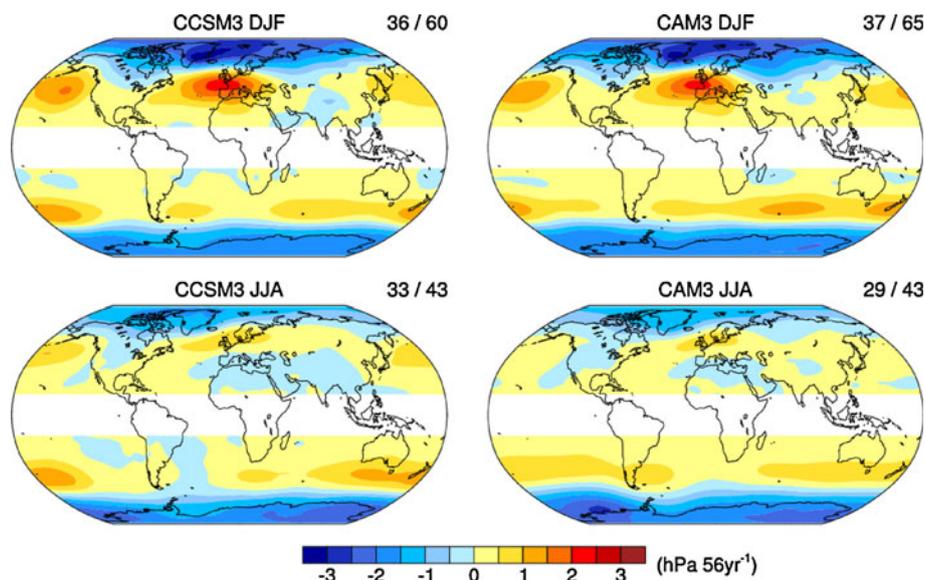
(60–65% compared to 36–37%). These hemispheric modes occur independently of one another: e.g., the correlation between the principal component (PC) records in the NH and SH is near zero in both CCSM3 and CAM3.

The spatial pattern of the ensemble mean SLP response (Fig. 1) bears some similarity to the leading EOF of the trends in CCSM3 and the CAM3 control integration in both seasons. In particular, the SH response in DJF exhibits a high spatial correlation (0.88) with CAM3 EOF1 in the SH; there is also some correspondence between the NH response in DJF and the NH EOF1 from CAM3 especially over the Atlantic-Eurasian sector (spatial correlation of 0.64 for all longitudes, and 0.82 in the longitude band 20°W–140°E). In JJA, the NH ensemble mean response resembles the NH EOF1 from CAM3 (spatial correlation of 0.79), while the SH response bears some similarity to the SH EOF1 from CAM3 especially over the Pacific sector (spatial correlation of 0.61 over all longitudes and 0.84 in the longitude band 135°E–45°W). The spatial correspondence between the annular modes of atmospheric circulation variability in twentieth century coupled model integrations and the forced response to increasing concentrations of greenhouse gases and tropospheric sulfate aerosols has been documented by Miller et al. (2006) for the models in the CMIP3 archive.

Given the significant spatial projection of the ensemble mean response upon the leading EOF of the internal atmospheric variability, it is relevant to compare the

distributions of the annular mode trends from the individual ensemble members of CCSM3 and CAM3. Figure 11 shows histograms of the annular mode trends in DJF and JJA for both hemispheres, obtained by projecting the trends from each ensemble member onto EOF1 from CAM3; similar results are obtained using zonally averaged SLP differences between middle (30°–55°) and high (55°–90°) latitudes in place of the projection time series (not shown). To increase the sample size, individual trends from each month (December, January, February; June, July, August) are used, resulting in  $3 \times 40$  ( $3 \times 178$ ) samples for CCSM3 (CAM3). The spread of the trends in the annular modes is comparable in both sets of model integrations, for both hemispheres and seasons. These results indicate that internal atmospheric variability accounts for much of the spread in the future projections of atmospheric circulation trends associated with the annular modes in the 40-member CCSM3 ensemble. There is also an overall shift in the mean distribution of the annular mode trends in CCSM3 compared to CAM3, reflecting mainly the forced response. Although small ( $\sim 0.5$ – $0.75$  standard deviations for DJF in both hemispheres and for JJA in the NH), this shift is significantly different from the approximately zero mean value in the CAM3 control integration.

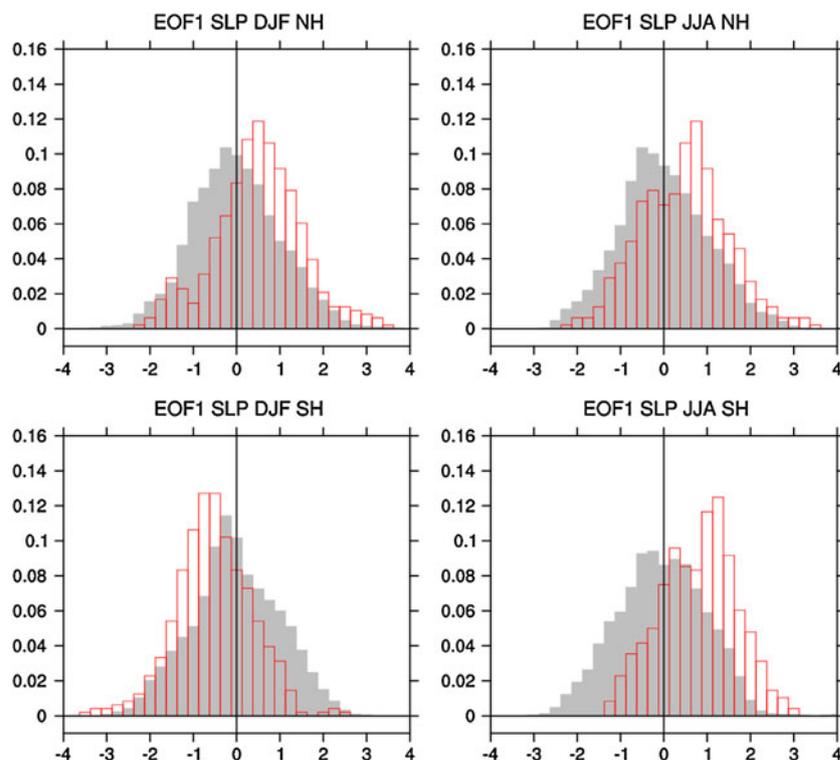
The leading EOF of SLP trends is associated with Precip and TS trend anomalies in both seasons and hemispheres as illustrated in Fig. 12 for CCSM3. In particular, the leading SLP trend EOF in both seasons is accompanied by out-of-



**Fig. 10** The leading EOF of extra-tropical SLP trends from (left) the 40-member CCSM3 ensemble and (right) the “178-member” CAM3 control integration in (top) DJF and (bottom) JJA. Trends are computed over the period 2005–2060 for CCSM3 and for 56-year non-overlapping segments for CAM3. EOF analysis is performed for each hemisphere separately but plotted on a single map for

conciseness. The percent variance explained by each EOF is given in the upper right corner of each panel, with the first number denoting the NH and the second number the SH (for example, for CCSM3 in DJF, EOF1 accounts for 36% of the variance in the NH and 60% of the variance in the SH)

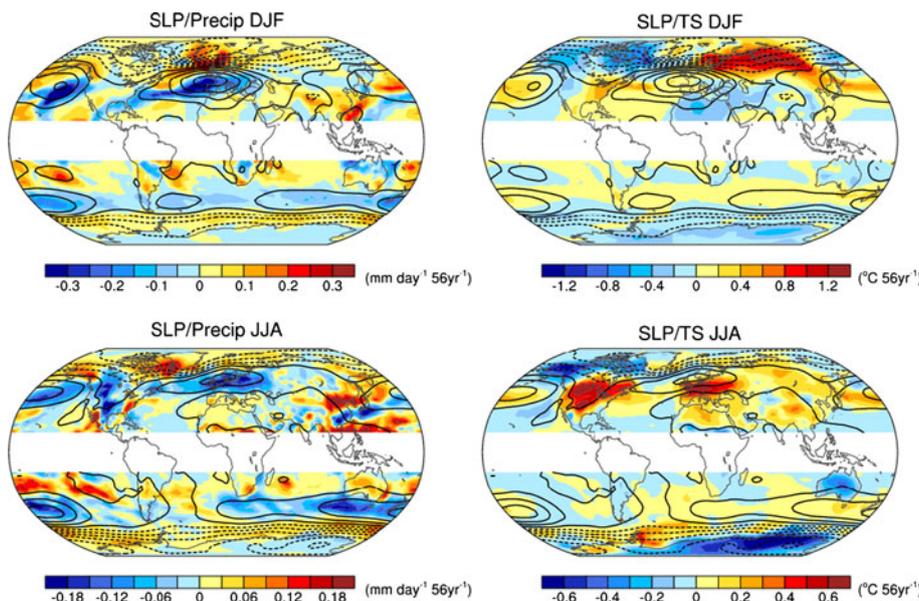
**Fig. 11** Histograms of the SLP 2005–2060 trend projections onto EOF1 from the CAM3 control integration for the (top) NH and (bottom) SH in (left) DJF and (right) JJA. The red open bars show results from the 40-member CCSM3 and the grey filled bars from the 178-member CAM3 control. The x axis is in units of standard deviations of the CAM3 control integration, and the y axis is frequency (number of ensemble members divided by the total number of ensemble members)



phase Precip trend anomalies between high and middle latitudes of the North Atlantic and Pacific and over the Southern Ocean, with positive SLP anomalies generally collocated with negative Precip anomalies. Similarly, the

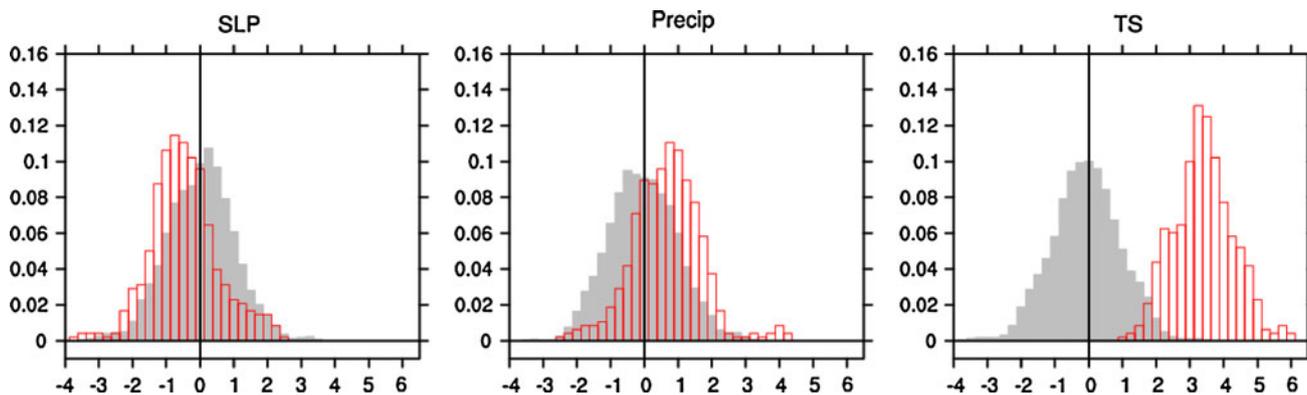
positive phase of the NH annular mode trend EOF is accompanied by positive air temperature trend anomalies over Eurasia (central Europe and the United States) and negative temperature trend anomalies over Canada and the

**Fig. 12** (Left) Precip and (right) TS trend regressions (shading) associated with the leading EOF of extra-tropical SLP trends from the 40-member CCSM3 ensemble in (top) DJF and (bottom) JJA. Contours show the SLP trend EOF (contour interval is 0.6 hPa 56 year<sup>-1</sup> for DJF, and 0.4 hPa 56 year<sup>-1</sup> for JJA; negative values are dashed). Trends are computed over the period 2005–2060. EOF and regression analyses are performed for each hemisphere separately but plotted on a single map for conciseness



Labrador Sea (Canada) in DJF (JJA). It is also interesting to note the association between the NH annular mode trend EOF in JJA with Precip trends over the Sahel and the western tropical Pacific. Over the SH, the main TS trend signal associated with the positive phase of the annular mode trend EOF is that of cooling over Antarctica and Australia, especially in JJA. The leading EOF of extra-tropical TS and Precip trends for each season and hemisphere based on the 40-member CCSM3 ensemble (not shown) are very similar to the Precip and TS trend regression patterns associated with SLP trend EOF1, with pattern correlations ranging from 0.87 to 0.97 (except for the SH in JJA which exhibits lower pattern correlations of 0.51 for Precip and 0.67 for TS). That is, the dominant trend EOF within each field is linked by virtue of a common atmospheric circulation-driven mode of variability.

Histograms of regional SLP, Precip and TS trends over the North Atlantic/Eurasian sector in DJF are shown in Fig. 13 based on the 40-member CCSM3 ensemble and the “178”-member CAM3 control integration. The regions used are those affected by the NAM (recall top panels of Fig. 10) as follows: (57°–90°N, 20°–120°E) for SLP; (57°–72°N, 25°W–25°E) for Precip; and (50°–75°N, 0°–125°E) for TS. For each parameter, the spread of the trends across the 40 members of the coupled model ensemble is comparable to that across the 178 members of the atmospheric control integration, indicating that internal atmospheric variability controls the trend uncertainties. The forced (ensemble mean) trend in CCSM3 is considerably smaller for SLP and Precip (~0.5 and 0.9 standard deviations, respectively) than for TS (~3.5 standard deviations), although all are significantly different from



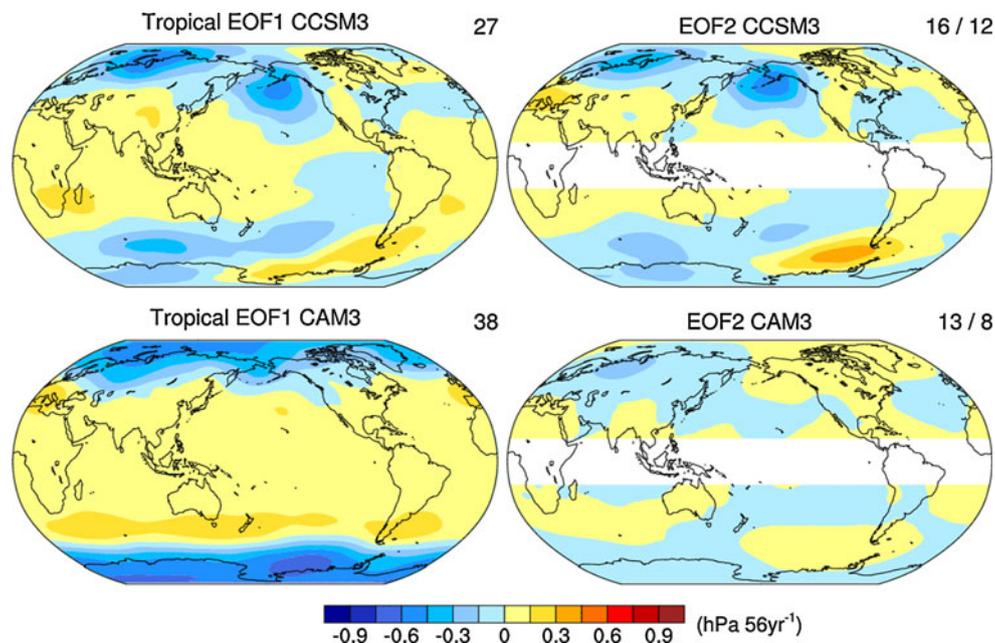
**Fig. 13** Histograms of regionally-averaged trends over the Eurasian–North Atlantic sector in DJF for SLP (left), Precip (middle) and TS (right) in DJF from the 40-member CCSM3 ensemble (open red bars) and the “178-member” CAM3 control integration (grey filled bars). Trends are computed over the period 2005–2060 for CCSM3 and for

56-year non-overlapping segments for CAM3. For all panels, the *x* axis is in units of standard deviation based on CAM3, and the *y* axis is in units of the number of ensemble members divided by the total number of ensemble members

zero at the 95% level. This result is in keeping with the fact that the amplitude of the dominant extra-tropical pattern of the noise component of the trend (Fig. 12) relative to the amplitude of the forced component of the trend (estimated by the ensemble mean; Fig. 1) is large ( $\sim 100\%$ ) for SLP and small ( $\sim 10\%$ ) for TS. The histograms also indicate that the forced component of the NAM circulation trend makes a negligible contribution to the forced component of the TS trend over Eurasia in DJF.

We have shown that internal atmospheric variability accounts for the dominant pattern of the “noise” component of extra-tropical SLP trends, which in turn drives the dominant pattern of the “noise” component of the extra-tropical Precip and air temperature trends in the 40-member CCSM3 ensemble. However, the extra-tropical atmospheric circulation is also known to be sensitive to conditions in the tropics, particularly over the Indo-Pacific sector as occurs during El Niño and La Niña events (e.g., Trenberth et al. 1998; Alexander et al. 2002). What is the role of internal variability of the tropical coupled ocean-atmosphere system in the inter-ensemble spread of future SLP trends over the extra-tropics? Figure 14 (upper left) shows the global distribution of SLP trend anomalies regressed upon the leading PC of tropical SLP trends based

on annual means from the 40-member CCSM3 ensemble. Although the EOF analysis was restricted to the tropics, the largest regression coefficient amplitudes occur over middle and high latitudes in both hemispheres. Within the tropics, EOF1 is reminiscent of the Southern Oscillation (SO), with negative anomalies over the eastern Pacific and positive anomalies over the western Pacific/Indian Ocean. This pattern must be a result of ocean-atmosphere coupling since internal atmospheric variability from the CAM3 control integration yields a very different EOF1 pattern, namely a zonally-symmetric structure with one sign throughout the tropics and middle latitudes and opposite sign at high latitudes (Fig. 14, lower left). The leading EOF of tropical SLP trends in CCSM3 is accompanied by increased Precip over the western and central equatorial Pacific and decreased Precip over the eastern Indian Ocean (not shown). These Precip anomalies in turn force global atmospheric teleconnections including a deepening of the Aleutian Low and a Rossby-like wave train over the South Pacific, similar to those which occur in association with interannual ENSO events (see Deser et al. 2006 for a description of ENSO teleconnections in CCSM3). Note that although the extra-tropical teleconnections are maximized in the winter hemisphere (not shown), the use of



**Fig. 14** (Left) The leading tropical EOF of annual SLP trends from (top) the 40-member CCSM3 ensemble and (bottom) the “178-member” CAM3 control integration. Trends are computed over the period 2005–2060 for CCSM3 and for 56-year non-overlapping segments for CAM3. The domain used for the EOF analysis is confined to the tropics, but the results are displayed for the entire globe by regressing the SLP trends at each grid point upon the tropical PC1 record. The percent variance explained by each EOF is given in the upper right corner. (Right) The second EOF of extra-tropical

annual SLP trends from (top) the 40-member CCSM3 ensemble and (bottom) the “178-member” CAM3 control integration. EOF analysis is performed for each hemisphere separately but plotted on a single map for conciseness. The percent variance explained by each EOF is given in the upper right corner of each panel, with the first number denoting the NH and the second number the SH. For each model, EOF2 is scaled by the correlation coefficient between its PC and tropical PC1

annual mean data in the EOF analysis brings out the connection to both hemispheres simultaneously.

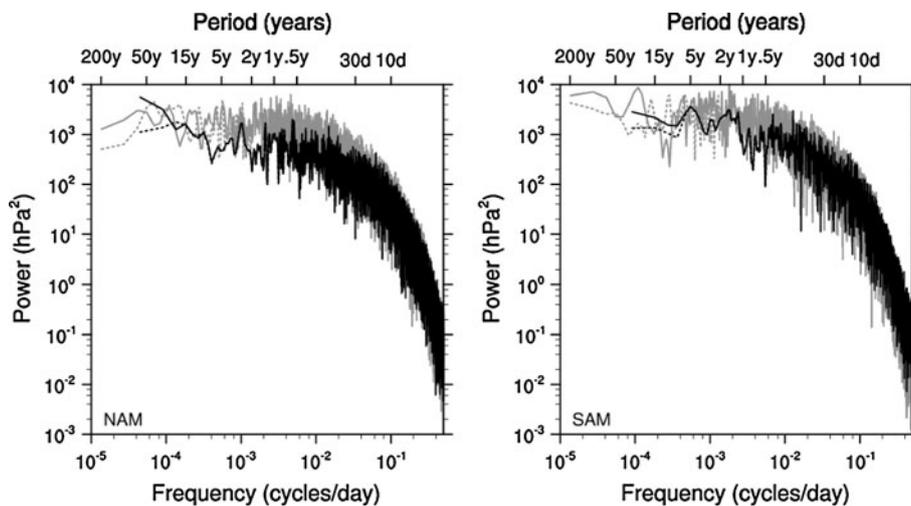
The second EOFs of annual SLP trends over the NH and SH (Fig. 14, upper right) bear a close resemblance to the SLP regressions associated with the leading EOF of tropical SLP trends in their respective hemispheres, suggesting that they are due at least in part to coupled ocean-atmosphere variability within the tropics. For example, EOF2 of the NH exhibits negative SLP trend anomalies over the North Pacific and over northern Eurasia in the vicinity of the Arctic coastline, similar albeit not identical to the NH teleconnection pattern associated with tropical EOF1. The second EOF of the SH exhibits a NE-SW oriented dipole over the South Pacific and negative SLP trend anomalies over the Indian Ocean sector of the Southern Ocean and Antarctica, generally consistent with the SH teleconnection pattern associated with tropical EOF1. There are some differences in the shapes and relative amplitudes of the centers of action of the extra-tropical EOF2 patterns and those associated with tropical EOF1, most notably over the South Pacific. These differences may indicate that internal atmospheric variability also contributes to the former. Indeed, EOF2 in the NH and SH from the CAM3 control integration (Fig. 14, lower right) exhibits centers of action over northern Eurasia and in the South Pacific north of West Antarctica, respectively.

### 3.3 Comparison with nature

The annular modes of extra-tropical atmospheric circulation variability play an important role not only in the forced climate response but also in the noise component of the

response. The variability in the noise component of the annular mode response was in turn shown to be primarily a result of processes intrinsic to the atmosphere. Thus, a natural question to address is, how realistically does CAM3 depict the temporal behavior of the annular modes? Figure 15 compares the power spectra of the observed and simulated annular mode indices, defined as the zonally averaged SLP anomaly difference between middle (30°–55°) and high (55°–90°) latitudes in each hemisphere based on daily data. The CAM3 spectra are based on a 200-year segment of the 10,000-year control integration (solid gray curves), and the observed spectra (solid black curves for raw data, dashed black curves for detrended data) are based on the NCEP/NCAR Reanalyses for 1948–2008 over the NH (similar results are obtained for 1979–2008; not shown), and 1979–2008 over the SH in view of the limited spatial coverage before the incorporation of satellite data in 1979. Due to the broad range of frequencies spanned, the spectra are displayed in a log frequency—log power format and thus do not preserve variance (e.g., the relative amount of variance in each frequency band is not a simple integral under the power spectrum curve). Note also that the spectra have a smoother appearance at low frequencies compared to high frequencies due to the higher spectral resolution at shorter periods.

The overall shape and magnitude of the observed and simulated spectra are similar, with a rapid increase in power with decreasing frequency for periods shorter than a few months, and approximately constant or slightly increasing power for periods longer than about 1 year. CAM3 overestimates the power in the NAM for periods between about 30 days and 10 years. The daily annular



**Fig. 15** Power spectra of the daily NAM (left) and SAM (right) indices, defined as the zonally-averaged SLP anomaly difference between high (55°–90°) and middle latitudes (30°–55°) of the northern and southern hemisphere, respectively, from the NCEP/NCAR Reanalysis (solid black curve; detrended version depicted by

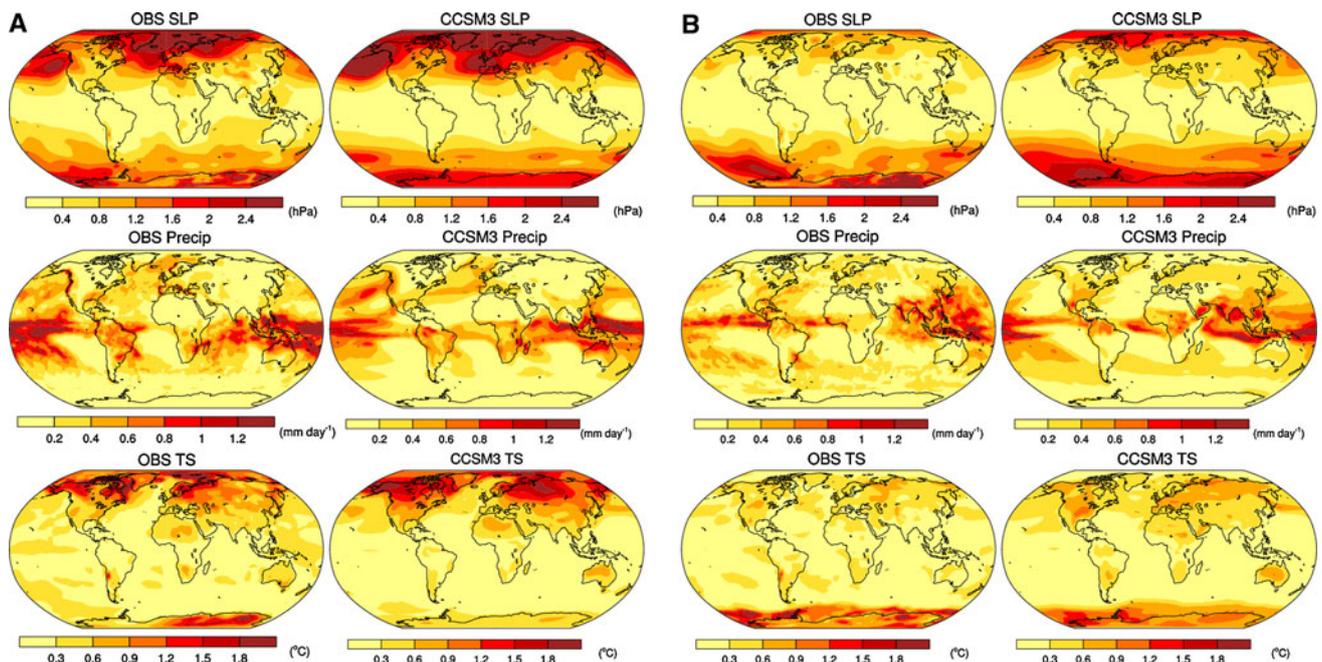
the dashed black curve), a 200-year segment of the CAM3 control integration (solid gray curve), and a 200-year segment of the CCSM3 control integration (dashed gray curve). The period 1979–2008 (1948–2008) was used for the Reanalysis in the SH (NH)

mode power spectra from a 200-year segment of a CCSM3 pre-industrial control integration, indicated by the dashed gray curves in Fig. 15 (only periods longer than 2 years are plotted for clarity), does not differ significantly from CAM3 over the range of periods relevant for this study (<60 years), confirming that interannual-to-decadal variability of the simulated annular modes is predominantly due to processes internal to the atmosphere. Thus, the null hypothesis of intrinsic atmospheric variability is a useful benchmark against which to test for the presence of externally forced trends in the annular modes in both coupled models and nature.

How realistic is the simulation of internal climate variability on decadal time scales in CCSM3? Traditionally, the evaluation of internally-generated climate variability in coupled models has been accomplished using long (several hundred—1,000 year) control simulations (e.g., Karoly and Wu 2005). Here we use the set of 40 CCSM3 integrations during 2005–2060, a total of 2,280 years, to evaluate internal decadal variability. To provide a baseline comparison of the amount of variance at periods of a decade and longer in nature and as simulated by the 40-member CCSM3 ensemble, we compare maps of the standard deviation of 8-year low-pass filtered data in DJF and JJA (Fig. 16a, b, respectively). The 8-year low-pass filter was achieved by smoothing the data with a 5-point binomial filter (weights 1-3-4-3-1) for each season separately

(Trenberth et al. 2007). To reduce the influence of externally-forced signals in the low-pass filtered data (e.g., to isolate the internally-generated component of decadal variability), we have removed the linear trend from the observed records, and removed the ensemble mean from each ensemble member at each time step from the CCSM3 output. The standard deviations of the resulting low-pass filtered model output were then averaged across the 40 ensemble members. We use 2 m air temperature and SLP observations from the NCEP/NCAR Reanalyses (Kistler et al. 2001) and precipitation from the Global Precipitation Climatology Project (Huffman et al. 2001) for the period 1979–2008. Generally similar results are found for other data sets and longer periods of record (not shown).

The spatial distributions of the standard deviations of the 8-year low-pass filtered data from observations (left) and CCSM3 (right) are similar for each variable and season, and the magnitudes are of the same order. For example, the standard deviations of SLP are largest at high latitudes of the winter hemisphere, with values < 0.4 hPa within the tropics increasing to ~2 hPa and greater in polar regions. The model tends to overestimate low-frequency SLP variability over the extra-tropical NH by approximately 30% in DJF and 50% in JJA. Like SLP, Precip low-frequency variability is comparable in the model and observations except for the double ITCZ-bias over the western two-thirds of the tropical Pacific in the model that is reflected in the



**Fig. 16** a Standard deviation maps of 8-year low-pass filtered SLP (top), Precip (middle) and TS (bottom) anomalies in DJF from observations (left) and the 40-member CCSM3 ensemble (right). For observations, the linear trend over the period 1979–2008 was removed before filtering. For the model, the ensemble mean was

removed from each ensemble member at each time step, and the standard deviations averaged across the 40 ensemble members. SLP and TS observations are from the NCEP/NCAR Reanalysis, and Precip observations are from the Global Precipitation Climatology Project. b As in a but for JJA

pattern of simulated variability especially in JJA. Simulated near-surface air temperature also exhibits realistic patterns and magnitudes of low-frequency variability, with larger values over land and the marginal sea ice zones compared to ocean. The highest amplitude variability occurs the NH continents in winter, with values  $\sim 1.2\text{--}1.5^\circ\text{C}$  in nature compared to  $1.5\text{--}1.8^\circ\text{C}$  in the model. The overestimate of low-frequency wintertime air temperature variability over Eurasia and Alaska in the model may be partly due to the stronger-than-observed atmospheric circulation (e.g., SLP) variability. In summary, the 40-member CCSM3 ensemble generally simulates a realistic order-of-magnitude for low-frequency ( $>8$  years) variability in near-surface air temperature, Precip and SLP.

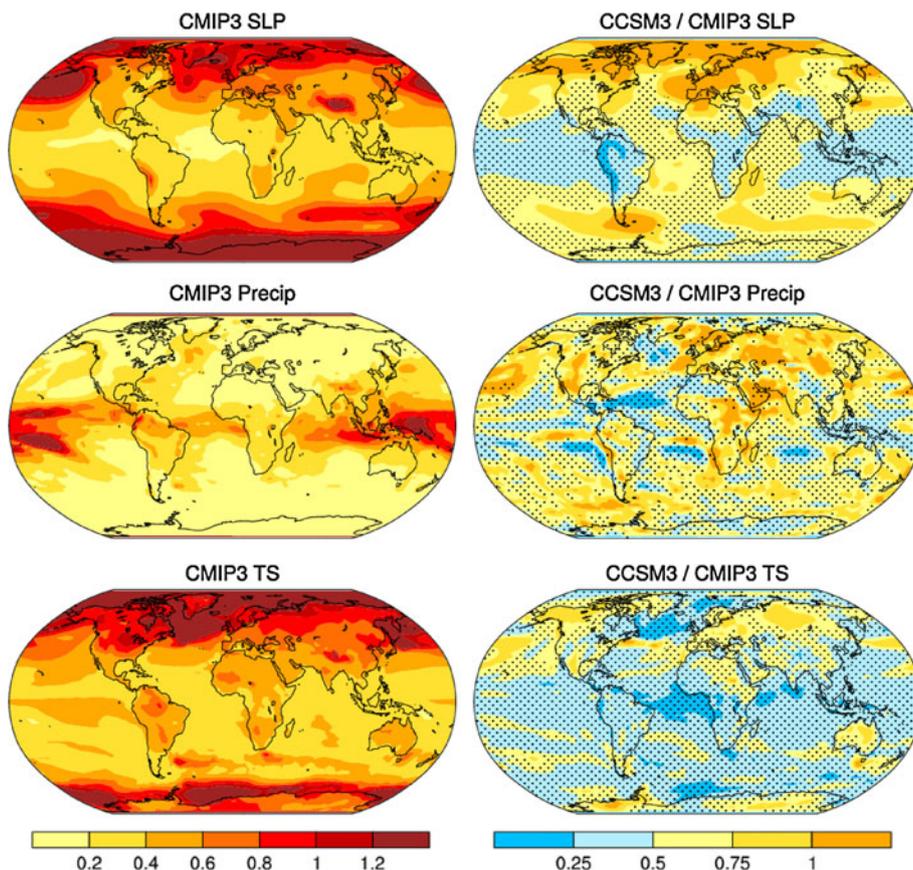
### 3.4 Contribution of internal variability to the CMIP3 multi-model ensemble

As mentioned in the Introduction, uncertainties in the forced climate change signals simulated by the multi-model mean in the IPCC WG1 4th Assessment Report (Solomon et al. 2007) contains contributions from model uncertainty and internal variability. As a first step in separating the two contributions, we have compared the internal variability of trends during 2005–2060 from the

40-member CCSM3 ensemble with the model-plus-internal variability of similarly-computed trends from the 21-model CMIP3 ensemble forced by the SRES A1B GHG scenario (see Table 8.1 in Solomon et al. 2007 for a list of models). To help mitigate seasonal biases between different models, we have used annual mean values in our trend calculations. While the internal variability estimated from one model does not necessarily represent the internal variability averaged across all models, our estimate of the contribution of internal variability to the spread of trends within the CMIP3 ensemble is intended to serve as a benchmark until sufficiently large ensembles are completed for the other models.

The spatial patterns of the standard deviations of annual-mean trends computed over the period 2005–2060 are similar for both model ensembles, with larger magnitudes for the CMIP3 ensemble (left-hand panels of Fig. 17) compared to CCSM3 (not shown but recall the seasonal trend distributions in Fig. 9) in keeping with the notion that the former contains contributions from both inter-model and internal-climate variability. The ratio of the standard deviation of annual-mean trends (CCSM3 divided by CMIP3) is shown in the right-hand panels of Fig. 17, with stippling indicating where the two sets of standard deviations differ significantly at the 95% confidence level. Note

**Fig. 17** (Left) Trend standard deviation maps from the 21-member CMIP3 ensemble and (right) the ratio of the trend standard deviations from the 40-member CCSM3 ensemble and the 21-member CMIP3 ensemble, based on annually-averaged data for (top) SLP, (middle) Precip, and (bottom) TS. Trends are computed over the period 2005–2060 for both model ensembles. Stippling indicates where the ratios are significantly different from one at the 95% confidence level. Units for the plots in the left column are  $\text{hPa } 56 \text{ year}^{-1}$  (SLP),  $\text{mm day}^{-1} 56 \text{ year}^{-1}$  (Precip), and  $^\circ\text{C } 56 \text{ year}^{-1}$  (TS)



that a ratio greater (smaller) than 0.5 indicates a larger (smaller) contribution from internal variability compared to model variability. In general, ratios  $> 0.75$  correspond to areas where the spread in the CCSM3 trends does not differ significantly from the spread in the CMIP3 trends. For SLP, internal variability within the CCSM3 ensemble generally accounts for 25–50% of the total variability within the CMIP3 ensemble over much of the tropics, and approximately 50–75% ( $>75\%$ ) over middle and high latitudes of the southern (northern) hemisphere. For Precip, ratios are  $\sim 25\text{--}50\%$  in the tropics and  $\sim 50\text{--}100\%$  at higher latitudes. And for TS, ratios are  $\sim 25\text{--}50\%$  over most areas except Eurasia, Europe and western North America where they are  $\sim 50\text{--}100\%$ . In summary, internal climate variability is more important than (comparable to) model variability for uncertainties in forced annual-mean extra-tropical SLP and Precip trends (TS trends over North America, Eurasia and Antarctica) during 2005–2060. Elsewhere, the contribution of internal variability is less than that of model variability, but rarely below half.

#### 4 Summary and discussion

We have investigated the forced climate response and associated uncertainties from a new 40-member ensemble of CCSM3 simulations forced with the SRES A1B GHG and ozone recovery scenarios during 2000–2060. The large ensemble size has enabled not only a robust estimate of the model's forced response, but also an evaluation of the spread in the response due to internal (natural) variability of the climate system. The contribution of intrinsic atmospheric variability to uncertainty in the forced response was assessed using a long (10,000-year) control integration of the atmospheric model component of CCSM3. The response was characterized for 3 basic climate parameters (surface air temperature, precipitation, and sea level pressure) and two seasons (DJF and JJA). The main results are summarized below.

Similar to the average response of the 21 models in the CMIP3 archive (Hegerl et al. 2007), the 40-member CCSM3 ensemble mean response is characterized by: increased Precip along the equator and at high latitudes, and decreased Precip within the tropics and subtropics; more warming over land than ocean, and strongest warming over the Arctic and adjacent high latitude continents in winter; and a general pattern of decreased SLP at high latitudes and increased SLP in middle latitudes, except for the SH in summer which exhibits a response of the opposite sign as a result of prescribed recovery of the stratospheric ozone hole.

Due to the relative amplitudes of the forced response and natural variability, fewer ensemble members are needed to detect a significant response in TS compared to

either Precip or SLP. More specifically, only 1 realization is needed to detect a significant (at the 95% confidence level) warming in the 2050s decade compared to the 2010s at nearly all locations, compared to approximately 3–6 ( $>15$ ) ensemble members for tropical and high latitude (middle latitude) Precip, and approximately 3–6 (9–30) members for tropical (extra-tropical) SLP, depending on location and season. Larger ensemble sizes are needed to detect a significant response in the 2030s, even for TS at middle and high latitudes where 3–12 members are required. With a 40-member ensemble, significant decadal TS changes are detectable within the next few years over most regions, while decadal SLP and Precip changes are detectable within approximately 5–10 years over portions of the tropics (and the Arctic and Southern Ocean for Precip) and around 2030 elsewhere. With a 5-member ensemble, detection of the forced signal is delayed to 2025–2040 for tropical SLP and for Precip over the Arctic, equatorial and Southern Oceans, and 2020–2030 for TS over Eurasia and North America; no detection is possible for extra-tropical SLP and middle latitude Precip. Although the spatially-uniform component of the forced tropical SST warming emerges within a few years and with a small number of realizations ( $<3$ ), the spatially-varying component is subject to a lower signal-to-noise ratio that is commensurate with the characteristics of the tropical precipitation response. The forced decadal-scale responses of the NAM and SAM require a relatively large number of realizations for detection ( $\sim 25$  and 15, respectively, in DJF) and a time horizon of detection of 2–3 decades ( $\sim 2040$  and 2030, respectively, in DJF), underscoring the low signal-to-noise ratios in even the large-scale patterns of extra-tropical atmospheric circulation response.

The leading pattern of uncertainty in the extra-tropical responses of TS, Precip and SLP, as determined from EOF analysis of the 40 individual responses, is associated with the annular mode of atmospheric circulation variability in both seasons and hemispheres. This mode, in turn, is primarily due to intrinsic atmospheric dynamics (e.g., “weather noise”: Madden 1976) and, as such, contains no predictability beyond a few months. The leading mode of uncertainty in the extra-tropical SLP response bears some resemblance to the forced (ensemble mean) SLP response, especially in the summer hemisphere. The leading pattern of uncertainty in the tropics displays a spatial structure reminiscent of the “Inter-decadal Pacific Oscillation” (Power et al. 1999) or “Pacific Decadal Oscillation” (e.g., Zhang et al. 1997). This pattern does not occur in the atmospheric control integration of CAM3, and thus owes its existence to ocean-atmosphere coupling. We note that thermodynamic coupling between the global atmosphere-ocean mixed layer system is sufficient to produce much of the spatial structure of this pattern (e.g., Yukimoto et al.

1996; Dommenges and Latif 2008; Clement et al. 2010). This tropical mode affects the extra-tropical atmospheric circulation *via* precipitation-induced teleconnections which in turn impact TS and Precip over middle and high latitudes. Indeed, the second EOF of the extra-tropical SLP response is linked in part to the leading EOF of the tropical SLP response.

The fact that forced changes in TS are more readily detectable than those in SLP over the extra-tropics by the middle of the twenty-first century indicates that the thermodynamically-induced signal in the TS response is larger than the dynamically-induced (*via* atmospheric circulation changes) uncertainty in the TS response (similar comments apply to the winter precipitation response over the northern high latitudes). For changes in earlier decades, the effect of circulation uncertainty on detection of the forced TS response is more evident (e.g., the minimum number of ensemble members needed to detect a significant TS response in the 2030s is markedly larger over the NH continents and Antarctica, regions affected by the annular modes, than other areas; recall Fig. 3). These results are in keeping with the studies of Yiou et al. (2007), Boé et al. (2009) and Vautard and Yiou (2009) focused on western Europe.

Our results have implications for detection and attribution of the twenty-first century climate response to anthropogenic forcing in nature and in the multi-model CMIP3 archive used in the AR4 IPCC assessment, and may also be useful for strategic guidance of the upcoming CMIP5 protocol of model experiments in support of the AR5. To the extent that the 40-member CCSM3 ensemble exhibits generally realistic levels of decadal variability as estimated from observational data sets spanning the past 30–50 years, attribution of observed future decadal changes in TS, Precip and SLP to anthropogenic forcing will be subject to similar levels of uncertainty reported here, taking into account any differences in climate sensitivity and forcing amplitude between CCSM3 and nature. The observed atmospheric circulation response over the extra-tropical NH may exhibit less uncertainty than indicated by CCSM3 due to the model's overestimation of decadal SLP variability in this region (by approximately 30% in DJF). We have shown that internal variability as estimated from the 40-member CCSM3 ensemble makes an appreciable contribution to the total (model plus internal) uncertainty in the future climate response simulated by the 21-model mean from the CMIP3 archive. In particular, internal variability was shown to be more important than model variability for annual-mean SLP and Precip responses in the extra-tropics, while the two sources of uncertainty are of the same order for the annual-mean TS responses over North America, Eurasia and Antarctica. The magnitude of uncertainty due to internal variability is rarely less than

half that due to model variability for forced linear climate trends during 2005–2060. Given our results, the planned set of CMIP5 model projections of twenty-first century climate in support of the AR5 IPCC Assessment should take into account the relatively high levels of uncertainty due to internal climate variability (of which internal atmospheric variability is an important component) by running enough ensemble members to provide robust assessments of the forced response in each model, perhaps by taking an adaptive approach based on the time horizon and climate parameter of interest. Similarly, given the inevitable competition between ensemble size and model resolution for a fixed level of computational resources, the former should not be sacrificed at the expense of the latter.

We have shown that the response to anthropogenic forcing is more detectable in surface temperature than in precipitation or atmospheric circulation. Thus, monitoring of observed climate change may be best served by focusing on thermodynamic components of the climate system such as air temperature and integrated quantities such as top-of-atmosphere radiation or ocean heat storage, rather than on dynamical components related to atmospheric and oceanic circulation changes.

**Acknowledgments** We thank Dr. Grant Branstator for useful discussions during the course of this work and for providing the CAM3 control integration. Dr. Michael Alexander provided helpful comments on the original manuscript. We also thank the two anonymous reviewers, Dr. Laurent Terray and Dr. Shang-Ping Xie for constructive suggestions. NCAR is sponsored by the National Science Foundation. Haiyan Teng is supported by the Department of Energy under Cooperative Agreement No. DE-FC02-97ER62402. VB's was supported by Meteo-France.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

## References

- Alexander MA (2010) Extratropical air–sea interaction, SST variability and the Pacific Decadal Oscillation (PDO). In: Sun D, Bryan F (eds) Climate dynamics: why does climate vary. AGU Monograph
- Alexander MA, Bladé I, Newman M, Lanzante JR, Lau N-C, Scott JD (2002) The atmospheric bridge: the influence of ENSO teleconnections on air–sea interaction over the global oceans. *J Climate* 15:2205–2231
- Barsugli JJ, Battisti DS (1998) The basic effects of atmosphere–ocean thermal coupling on midlatitude variability. *J Atmos Sci* 55:477–493
- Boé J, Terray L, Cassou C, Najac J (2009) Uncertainties in European summer precipitation changes: role of large scale circulation. *Clim Dyn* 33(2–3):265–276. doi:10.1007/s00382-008-0474-7
- Boville BA, Kiehl JT, Rasch PJ, Bryan FO (2001) Improvements to the NCAR CSM-1 for transient climate simulations. *J Clim* 14:164–179

- Branstator G, Selten F (2009) Modes of variability and climate change. *J Clim* 22:2639–2658. doi:[10.1175/2008JCLI2517.1](https://doi.org/10.1175/2008JCLI2517.1)
- Branstator G, Teng H (2010) Two limits of initial-value Decadal predictability in a CGCM. *J Clim* (in press)
- CCSM3 special issue 2006. *J Clim* 19:2121–2630
- Clement A, DiNezio P, Deser C (2010) Rethinking the ocean's role in the Southern Oscillation. *J Clim* (accepted with minor revisions)
- Delworth T, Manabe S, Stouffer R (1993) Interdecadal variations of the thermohaline circulation in a coupled ocean-atmosphere model. *J Clim* 6:1993–2011
- Deser C, Alexander MA, Timlin MS (2003) Understanding the persistence of sea surface temperature anomalies in midlatitudes. *J Clim* 16:57–72
- Deser C, Capotondi A, Saravanan R, Phillips AS (2006) Tropical Pacific and Atlantic climate variability in CCSM3. *J Clim* 19:2451–2481
- Deser C, Phillips AS, Alexander MA (2010) Twentieth century tropical sea surface temperature trends revisited. *Geophys Res Lett* 37:L10701. doi:[10.1029/2010GL043321](https://doi.org/10.1029/2010GL043321)
- Dommenget D, Latif M (2008) Generation of hyper climate modes. *Geophys Res Lett* 35:L02706. doi:[10.1029/2007GL031087](https://doi.org/10.1029/2007GL031087)
- Feldstein SB (2000) The timescale, power spectra, and climate noise properties of teleconnection patterns. *J Clim* 13:4430–4440
- Frankignoul C, Hasselmann K (1977) Stochastic climate models. Part II: Application to sea-surface temperature anomalies and thermocline variability. *Tellus* 29:284–305
- Hawkins E, Sutton R (2009) The potential to narrow uncertainty in regional climate predictions. *Bull Am Meteorol Soc* 90:1095. doi:[10.1175/2009BAMS2607.1](https://doi.org/10.1175/2009BAMS2607.1)
- Hawkins E, Sutton R (2010) The potential to narrow uncertainty in projections of regional precipitation change. *Clim Dyn*. doi:[10.1007/s00382-010-0810-6](https://doi.org/10.1007/s00382-010-0810-6)
- Hegerl GC et al (2007) Understanding and attributing climate change. In: Solomon S et al (eds) *Climate change 2007: the physical science basis*. Cambridge University Press, Cambridge, pp 663–745
- Huffman GJ, Adler RF, Morrissey MM, Curtis S, Joyce R, McGavock B, Susskind J (2001) Global precipitation at one-degree daily resolution from multi-satellite observations. *J Hydrometeor* 2:36–50
- Hurrell JW, Hack JJ, Shea D, Caron JM, Rosinski J (2008) A new sea surface temperature and sea ice boundary dataset for the community atmosphere model. *J Clim* 21:5145–5153
- Karoly DJ, Wu Q (2005) Detection of regional surface temperature trends. *J Clim* 18:4337–4343
- Kattsov VM, Sporyshev PV (2006) Timing of global warming in IPCC AR4 AOGCM simulations. *Geophys Res Lett* 33:L23707. doi:[10.1029/2006GL024746](https://doi.org/10.1029/2006GL024746)
- Kistler R et al (2001) The NCEP-NCAR 50-year reanalysis: monthly means CD-ROM and documentation. *Bull Am Meteorol Soc* 82:247–267
- Kwon YO, Deser C (2007) North Pacific decadal variability in the Community Climate System Model version 2. *J Clim* 20:2416–2433
- Madden RA (1976) Estimates of the natural variability of time-averaged sea-level pressure. *Mon Weather Rev* 104:942–952
- Mantua NJ, Hare SR, Zhang Y, Wallace JM, Francis R (1997) A Pacific interdecadal climate oscillation with impacts on salmon production. *Bull Am Meteorol Soc* 78(6):1069–1079
- Meehl GA et al (2006) Climate change projections for the twenty-first century and climate change commitment in the CCSM3. *J Clim* 19:2597–2616
- Meehl GA et al (2007) Global climate projections. In: Solomon S et al (eds) *Climate change 2007: the physical science basis*. Cambridge University Press, Cambridge, pp 747–845
- Miller RL, Schmidt GA, Shindell DT (2006) Forced variations in the annular modes in the 20th century IPCC AR4 simulations. *J Geophys Res* 111:D18101. doi:[10.1029/2005JD006323](https://doi.org/10.1029/2005JD006323)
- Oleson KW et al (2004) Technical description of the community land model (CLM). NCAR Technical Note NCAR/TN-461+STR, 173 pp
- Power SB, Casey T, Folland C, Colman A, Mehta V (1999) Interdecadal modulation of the impact of ENSO on Australia. *Clim Dyn* 15:319–324
- Sardeshmukh PD, Compo GP, Penland C (2000) Changes of probability associated with El Niño. *J Clim* 13:4268–4286
- Schneider N, Cornuelle BD (2005) The forcing of the Pacific Decadal Oscillation. *J Clim* 18:4355–4373
- Schneider EK, Kinter JL III (1994) An examination of internally generated variability in long climate simulations. *Clim Dyn* 10:181–204
- Schneider N, Miller AJ (2001) Predicting North Pacific Ocean climate. *J Clim* 14:3997–4002
- Selten FM, Branstator GW, Dijkstra HA, Kliphuis M (2004) Tropical originals for recent and future northern hemisphere climate change. *Geophys Res Lett* 31:L21205. doi:[10.1019/2004GL020739](https://doi.org/10.1019/2004GL020739)
- Solomon S, Qin D, Manning M, Chen Z, Marquis M et al (2007) The physical science basis. Contribution of Working Group I to the 4th Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge
- Son S-W et al (2008) The impact of stratospheric ozone recovery on the southern hemisphere westerly jet. *Science* 320:1486–1489
- Son S-W, Tandon NF, Polvani LM, Waugh DW (2009) Ozone hold and southern hemisphere climate change. *Geophys Res Lett* 36:L15705. doi:[10.1029/2009GL038671](https://doi.org/10.1029/2009GL038671)
- Stoner AM, Hayhoe K, Wuebbles DJ (2009) Assessing general circulation model simulations of atmospheric teleconnection patterns. *J Clim* 22:4348–4372
- Tebaldi C, Knutti R (2007) The use of the multimodel ensemble in probabilistic climate projections. *Phil Trans R Soc A* 365:2053–2075
- Thompson DWJ, Wallace JM (2000) Annular modes in the extratropical circulation. Part I: month-to-month variability. *J Clim* 13:1000–1016
- Trenberth KE, Branstator GW, Karoly D, Kumar A, Lau N-C, Ropelewski C (1998) Progress during TOGA in understanding and modeling global teleconnections associated with tropical sea surface temperatures. *J Geophys Res* 103:14291–14324
- Trenberth KE et al (2007) Observations: surface and atmospheric climate change. In: Solomon S et al (eds) *Climate change 2007: the physical science basis*. Cambridge University Press, Cambridge, pp 235–336
- Vautard R, Yiou P (2009) Control of recent European surface climate change by atmospheric flow. *Geophys Res Lett* 36:L22702. doi:[10.1029/2009GL040480](https://doi.org/10.1029/2009GL040480)
- Wunsch C (1999) The interpretation of short climate records, with comments on the North Atlantic and Southern Oscillations. *Bull Am Meteorol Soc* 80:245–255
- Xie S-P, Deser C, Vecchi GA, Ma J, Teng H, Wittenberg AT (2010) Global warming pattern formation: sea surface temperature and rainfall. *J Clim* 23:966–986. doi:[10.1175/2009JCLI3329.1](https://doi.org/10.1175/2009JCLI3329.1)
- Yiou P, Vautard R, Naveau P, Cassou C (2007) Inconsistency between atmospheric dynamics and temperatures during the exceptional 2006/2007 fall/winter and recent warming in Europe. *Geophys Res Lett* 34:L21808. doi:[10.1029/2007GL031981](https://doi.org/10.1029/2007GL031981)
- Yukimoto S, Endoh M, Kitamura Y, Kitoh A, Motoi T, Noda A, Tokioka T (1996) Interannual and interdecadal variabilities in the Pacific in an MRI coupled GCM. *Clim Dyn* 12:667–683. doi:[10.1007/s003820050135](https://doi.org/10.1007/s003820050135)
- Zhang Y, Wallace JM, Battisti DS (1997) ENSO-like interdecadal variability. *J Clim* 10:1004–1020