

Consensus on Twenty-First-Century Rainfall Projections in Climate Models More Widespread than Previously Thought

SCOTT B. POWER, FRANÇOIS DELAGE, ROBERT COLMAN, AND AUREL MOISE

Centre for Australian Weather and Climate Research, Bureau of Meteorology, Melbourne, Victoria, Australia

(Manuscript received 22 June 2011, in final form 6 October 2011)

ABSTRACT

Under global warming, increases in precipitation are expected at high latitudes and near major tropical convergence zones in some seasons, while decreases are expected in many subtropical and midlatitude areas in between. In many other areas there is no consensus among models on the sign of the projected change. This is often assumed to indicate that precipitation projections in these regions are highly uncertain.

Here, twenty-first century precipitation projections under the Special Report on Emissions Scenarios (SRES) A1B scenario using 24 World Climate Research Programme (WCRP)/Coupled Model Intercomparison Project phase 3 (CMIP3) climate models are examined. In areas with no consensus on the sign of projected change there are extensive subregions where the projected change is “very likely” (i.e., probability > 0.90) to be small (relative to, e.g., the size of interannual variability during the late twentieth century) or zero. The statistical significance of and interrelationships between methods used to identify model consensus on projected change in the 2007 Intergovernmental Panel on Climate Change (IPCC) report are examined, and the impact of interdependency among model projections on statistical significance is investigated. Interdependency among projections is shown to be much weaker than interdependency among simulations of climatology. The results show that there is more widespread consistency among the model projections than one might infer from the 2007 IPCC Fourth Assessment report. This discovery highlights the broader need to identify regions, variables, and phenomena that are expected to be little affected by anthropogenic climate change and to communicate this information to the wider community. This is especially important for projections of climate for the next 1–3 decades.

1. Introduction

Global temperature has increased over the past century (Trenberth et al. 2007), and emissions of greenhouse gases arising from human activities are very likely to have been responsible for most of the warming over the past 50 years (Hegerl et al. 2007). Further warming and further changes in the climate system associated with this warming appear inevitable (Meehl et al. 2007a). Precipitation is projected to change in many locations across the globe in association with global warming over the twenty-first century.

The last Intergovernmental Panel on Climate Change (IPCC) Fourth Assessment Report was released in 2007. It provided a comprehensive assessment of the current state of climate science at that time and represents a landmark in climate science. It concluded that increases in

precipitation are very likely in high latitudes and near major convergence zones in the tropics in some seasons, while decreases are likely in many subtropical land areas (Alley et al. 2007; Christensen et al. 2007; Meehl et al. 2007a; Solomon et al. 2007b). Examples of these general tendencies are evident in Fig. 1a, which shows the percentage change in June–August (JJA) precipitation during 2090–99, relative to precipitation averaged over the period 1980–99. The figures represent the multimodel ensemble mean (MEM) of 24 different models forced using the Special Report on Emissions Scenarios (SRES) A1B scenario (Nakicenovic et al. 2000; see section 2 for further details). We follow the IPCC Summary for Policymakers (Alley et al. 2007) and the Technical Summary (Solomon et al. 2007b) convention and only shade areas where 66% or more of models agree on the sign of change, and we only place dots in regions where 90% or more of the 24 models project changes in precipitation of the same sign (Fig. 1a).

Many areas, especially areas in the tropics and subtropics, are unshaded and without dots. In these regions

Corresponding author address: Dr. Scott B. Power, Bureau of Meteorology, CAWCR, GPO Box 1289, Melbourne, VIC, Australia.
E-mail: s.power@bom.gov.au

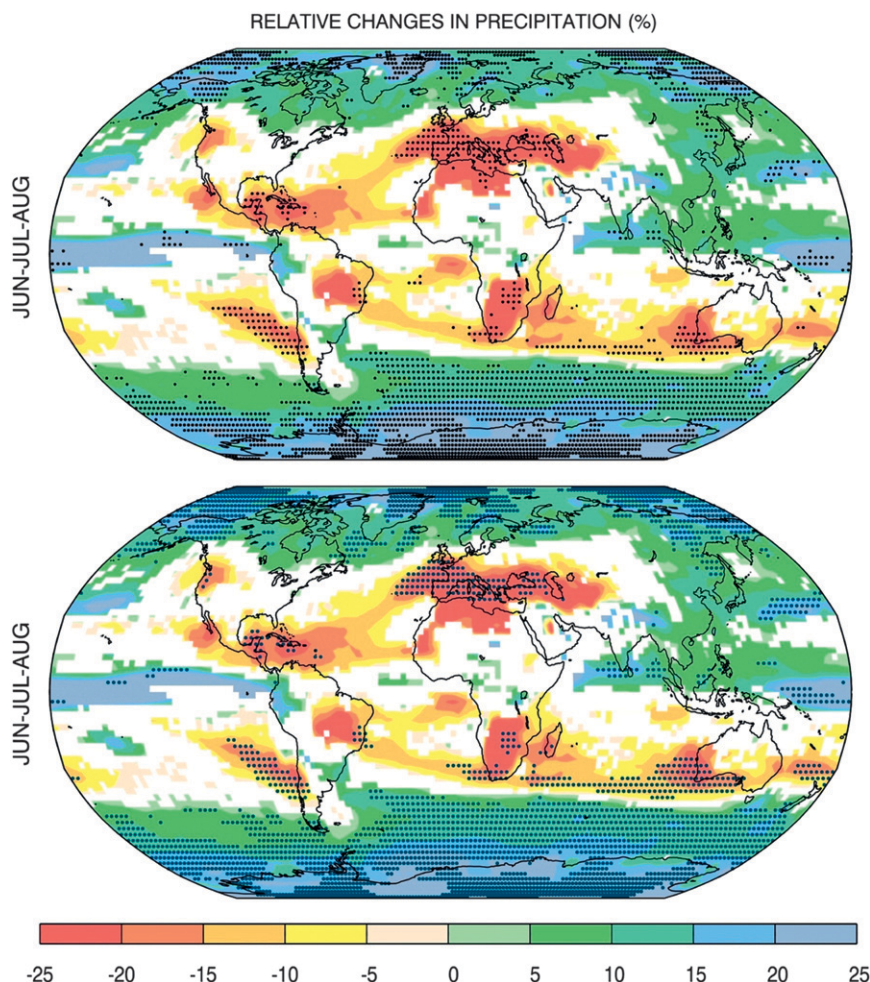


FIG. 1. Percentage (%) change in JJA precipitation during 2090–99, relative to precipitation averaged over the period 1980–99. The figures represent the MMEM of 24 different models forced using the SRES A1B scenario (see section 2 for further details). We follow the IPCC Summary for Policymakers (Alley et al. 2007) and the Technical Summary (Solomon et al. 2007b) convention and only shade (a) areas where 66% or more of the models agree on the sign of change and we only place dots in regions where 90% or more of the 24 models project changes in precipitation of the same sign or (b) regions where the MMEM change is larger in magnitude than the intermodel standard deviation of the change. Minor differences between (a) and the corresponding plot in the IPCC (2007) report arise because we use only one run per model, whereas the IPCC report uses more than one run for some of the models.

fewer than 66% of models agree on the sign of change. The absence of shading can occur, for example, if some models project a “large” increase while other models project a “large” decrease, making projected changes in such regions very uncertain. This was noted in the 2007 IPCC report by Christensen et al. (2007) who concluded that “for some regions, there are grounds for stating that the projected precipitation changes are likely or very likely. For other regions, confidence in the projected change remains weak”. Christensen et al. (2007) also noted that “it is unclear how rainfall in the Sahel, the Guinean coast and the southern Sahara will evolve”,

that “it is uncertain how annual and seasonal rainfall will change over northern South America, including the Amazon forest” and that “changes in rainfall in northern and central Australia remain uncertain.” While some of this uncertainty stems from known model deficiencies and imperfect simulations of twentieth-century regional climate (e.g., Randall et al. 2007; Brown et al. 2010; Irving et al. 2011) much of this uncertainty arises because the models do not exhibit a consensus on the sign of projected change in these regions. In fact Africa, Europe, North America, Central and South America, Australia, New Zealand, and small island regions were all found to

have subregions in which no clear consensus on the sign of the projected change was evident. Christensen et al. (2007) also noted that the regions of “large uncertainty” often lie near the boundaries between robust increases at high latitudes and the robust declines in the subtropics, and between these same robust declines in the subtropics and robust increases near major convergence zones at low latitudes.

Large uncertainty is a very undesirable situation because it impedes the ability of the scientific community to provide guidance on future climate in the affected regions. This, in turn, limits the ability of decision makers in the wider community to optimize their mitigation and adaptation plans.

But does a lack of consensus on the sign of change definitely indicate greater uncertainty? Suppose, for example, that the externally forced change in every model at a particular grid point in the unshaded region is actually zero. In other words, if each model is run infinitely many times the ensemble mean (EM) of each model is zero. As we actually only have a finite number of integrations for each model, sampling error arising from essentially random internally generated variability will typically cause the sign of projected change in each run to differ in a random fashion. Some runs will exhibit positive changes, some runs negative. Some individual model ensemble means will therefore give positive projections, with others negative. In other words, we would not expect to see any consensus on the sign of the projections even if the models all agree that the externally forced response is zero.

This highlights the fact that using the degree of consensus on the sign of change will not identify regions where the models agree that the externally forced change is either absent or small compared to internal variability—if such regions actually exist. It is therefore possible that a consensus among the models exists in parts of the unshaded areas in Fig. 1a and in similar plots in the 2007 IPCC report, but it is not a consensus on the sign of change.

This motivates the primary purpose of this investigation: to determine if there are subregions in the unshaded area in which there is a strong consensus on a projected change that is small in comparison with the standard deviation of year-to-year variability. We will also determine if there are subregions in the unshaded area in which *it is very likely* that the projected change in precipitation is small in comparison with (i) year-to-year precipitation variability in comparison with (ii) twentieth-century-average precipitation in the same location. In such subregions there is compelling evidence that if there is any externally forced signal at all, it is very likely small. This is not an uncertain projection. This is a projection of little

or no change in which we have a degree of confidence. Our confidence in projections will therefore extend over a greater area than one might infer from Fig. 1 or from the 2007 IPCC report.

We will find that a subset of the unshaded area in Fig. 1a contains such regions. In the remainder of the unshaded region there tends to be genuine disagreement among the models on the sign of the change, and projected changes in some of the models are not small. Projections here *are* uncertain. Importantly, however, only a subset of the unshaded region is actually occupied by regions in which projections are uncertain. While there is some awareness of this issue in the community (e.g., E. Hawkins 2011, personal communication) it doesn't seem widely appreciated. It is often incorrectly assumed that projections are uncertain over the entire unshaded region.

Additional questions arise from the use of 66% and 90% levels of agreement (i.e., degree of consensus) in the IPCC report (or the 80% level used in other parts of the report). These values were chosen to identify robust changes and to facilitate and underpin the adoption of a formal calibrated language in which the likelihoods greater than 66% and 90% correspond to “likely” and “very likely,” respectively (Solomon et al. 2007a). Consensus on the sign of change in 90% or more of models, for example, seems “high”. But what does high actually mean? Can we assign a level of statistical significance to this value? What level of statistical significance does this level of consensus actually correspond to? Are we entitled to be confident that a 90% level of consensus is very unlikely to occur by chance? Or conversely, does 90% agreement correspond to an exceptionally high level of statistical confidence? If the answer to this last question is yes then we could use a reduced level of consensus to identify regions where we have confidence in the sign of the projected change. If this is the case then consensus over an even larger proportion of the globe might be inferred. This is again useful because it would help to provide clearer guidance to the wider community.

The third issue we will address arises from the fact that confidence in projected changes in the 2007 IPCC report was also indicated by locating regions where the MMEM change is larger in magnitude than the intermodel standard deviation of the change. Such regions are identified in Fig. 1b, again for projected change in JJA precipitation with gray-blue dots. It is again of interest to know much confidence we should have that projected changes in such regions are robust in the sense that they are unlikely to be due to chance, that is, internally generated variability. It is also of interest to know how this measure of robustness relates to the degree of consensus on the sign of change. Does a projected change that seems robust using one approach seem robust using the other approach?

The fourth issue we will address relates to statistical significance. The statistical significance discussed herein is a measure of the probability of reaching a given level of consensus on the sign of change among the models through sampling error, under the null hypothesis that there is in fact no change. This is different to the likelihood that there is in fact a change—a difference that was recently highlighted by Ambaum (2010). We will examine the relationship between these two quantities (i.e., the likelihood of change and the statistical significance of apparent change).

The fifth and final issue we examine is the impact of model codependency—through the sharing of systematic biases—on the statistical significance of projected changes. Previous research has shown that systematic biases in climate models are not independent (Masson and Knutti 2011; Pennell and Reichler 2011). We will test to see if dependency also occurs in climate change projections and not just in the simulation of past climatology. We will then examine the implications the results obtained have for assessing the statistical significance of projected changes.

To address all these issues we will examine projected precipitation changes in the late twenty-first-century World Climate Research Programme (WCRP)/Coupled Model Intercomparison Project phase 3 (CMIP3) climate model integrations forced using the SRES A1B emissions scenario. The climate model database used in this investigation is described in section 2. Regions where the signal is very likely to be small are identified in sections 3 and 4 using two different definitions of “small.” The interrelationship between the degree of consensus on the sign of the projected change, the magnitude of the standardized change and other indicators of robust change are determined in section 5. The relationship between the likelihood of change and the statistical significance of apparent change is determined in the same section. The probability of consensus on the sign of projected changes among models is considered in section 6. The results are used in section 7 to reproduce Fig. 1a, identifying in particular where the models agree that change is very likely to be small. The issue of model interdependency is addressed in section 8. Results are summarized in section 9 and discussed in section 10. Recommendations are given in section 11.

2. Climate models

We analyze both twentieth- and twenty-first-century integrations from numerous different coupled general circulation models (CGCMs) available from the WCRP/Climate Variability and Predictability (CLIVAR)/Working Group for Coupled Modelling (WGCM) CMIP3 (Meehl et al. 2007b). We will examine twenty-first-century changes

using integrations forced using the SRES (Nakicenovic et al. 2000) A1B scenario. This scenario has a 2100 CO₂ concentration of approximately 710 ppm. Twenty-four models are available that have both twentieth-century runs and twenty-first-century results for this scenario. Projected changes are calculated using data averaged over the two periods 2090–99 and 1980–99.

We include one run only for each model and so in this case the MEM corresponds to the MM average. The MEM is our model-based estimate of the externally forced response in precipitation, as the internally generated variability will tend to be very largely cancelled out from such large samples.

3. Where are projected changes very likely to be small or absent?

Suppose that Δ_j is the change in precipitation at a particular grid point in model j , from the late twentieth century to the late twenty-first century, using the periods identified in the previous section. We will use σ_j to refer to the standard deviation of temporal variability in precipitation during the twentieth century in the same model at the same grid point (during 1980–99). The MEM of Δ_j/σ_j ($=\Delta/\sigma$ say) for each season is shown in Fig. 2. The quantity Δ/σ tends to exceed 0.5 poleward of 50° in both hemispheres (not shown) and in the equatorial Pacific in all seasons. The Δ/σ drops below -0.2 in scattered but extensive regions in the subtropics, especially in the Southern Hemisphere. Such changes occur in most or all seasons in the following: the southwest of (i) Western Australia, (ii) South America, and (iii) South Africa; in and around the Mediterranean; Central America, and Mexico; and the North and South Atlantic.

The magnitude of Δ/σ tends to be less than 0.2 over scattered regions in the subtropics. Of course Δ/σ is an estimate of its true value, $D^* = (\Delta/\sigma)^*$ say, and is subject to sampling error. Regions where $|D^*|$ is “very likely to be small” in the sense that $|D^*|$ is very likely to be less than 0.2, that is, where $\Pr(|D^*| < 0.2) > 0.9$, can be identified using the method described in the appendix, where we assume that $(D^* - \Delta/\sigma)/R^2$, where $\Pr(|\Delta^*/\sigma^*| < 0.2) > 0.9$ can be identified using the method described in the appendix, where we assume that $(\Delta^*/\sigma^* - \Delta/\sigma)/R^2$, where $R^2 = [\sum(\Delta_j/\sigma_j - \Delta/\sigma)/(N - 1)]/N^2$, has a t distribution with $N - 1$ degrees of freedom. Here, N is the number of models (24). Note that “0.2” is an arbitrary choice. Others might use different values in their definition of small.

Here we have adopted the formal calibrated language that will be used in the forthcoming IPCC report (Mastrandrea et al. 2011), which assigns the following words to the bracketed probabilities: “likely”

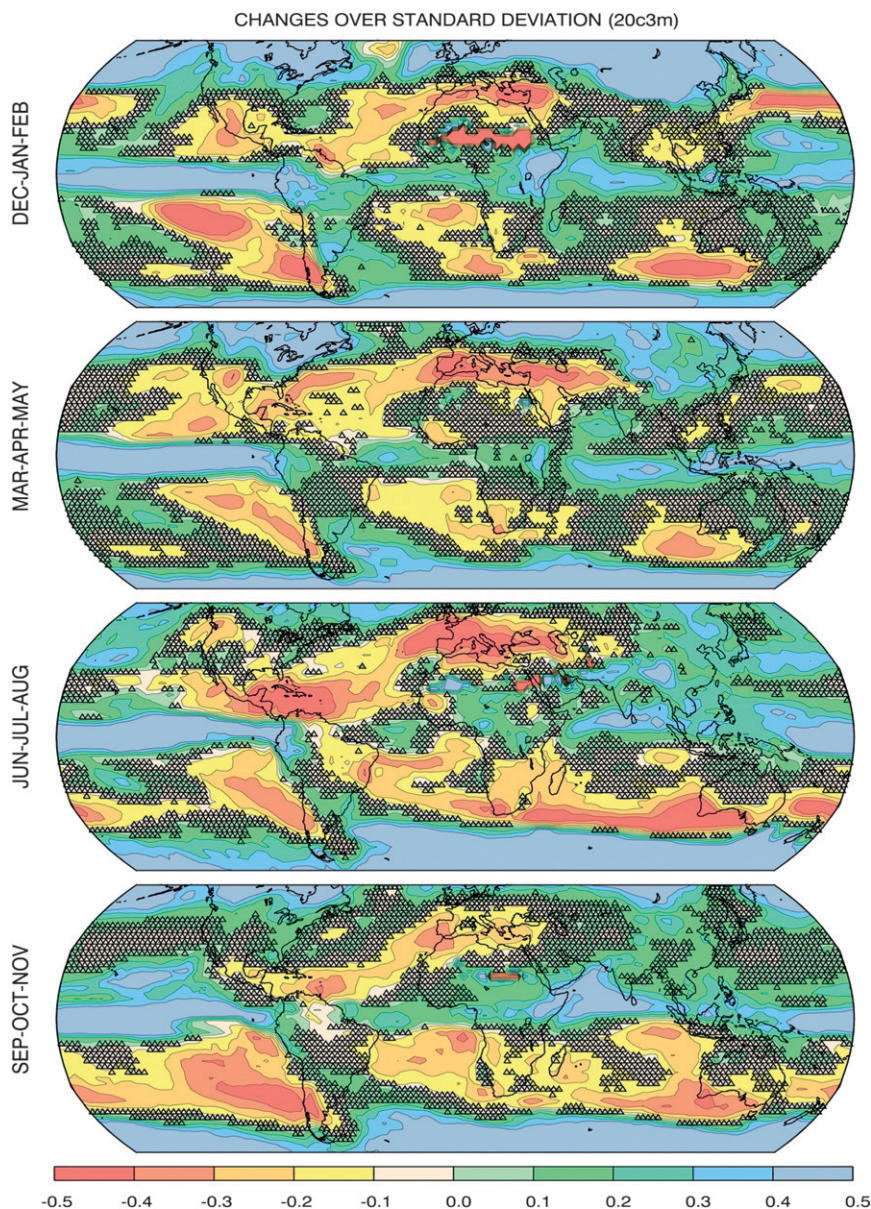


FIG. 2. The MMEM of Δ/σ_j (referred to as Δ/σ in the text) for (top to bottom) each season, where the subscript j refers to model j . Regions where Δ^*/σ^* is “very likely to be small” in the sense that $\Pr(|\Delta^*/\sigma^*| < 0.20) \geq 0.90$ are indicated using small triangles. Likelihood is assessed using the method described in the appendix. Latitudes equatorward of 60°S and 60°N only.

(probability > 0.6), “very likely” (> 0.90), “extremely likely” (> 0.95), and “virtually certain” (> 0.99).

Regions where the normalized change is very likely small (i.e., where $\Pr(|\Delta^*/\sigma^*| < 0.2) > 0.9$) are indicated in Fig. 2 by the small triangles. They are very largely confined to the tropics and subtropics. They cover 26%, 32%, 22%, and 27% of the earth’s surface during December–February (DJF), March–May (MAM), JJA, and September–November (SON), respectively, that is, greatest during the two transition seasons MAM and SON (see

Table 1). The equivalent figures for the Northern Hemisphere only are as follows: 22%, 30%, 18%, and 32%. The corresponding figures for the Southern Hemisphere alone are as follows: 30%, 34%, 24%, and 22%.

During SON, for example, the regions in which the normalized change is very likely small are generally scattered across land and ocean between 50°S and 50°N . During SON, regions where projected normalized changes are very likely small include the following: Japan; much of the United States; and parts of South America, the South

TABLE 1. Percentage area of the globe covered by projections that satisfy the conditions given in the Table. The three groups of four rows gives percentage areas covered using alternative thresholds for small. The values quoted in the paper generally correspond to the first and fourth rows, with the exception noted below. The second group of numbers in rows 5–8 gives the percentage areas covered by projections for which there is 99%, 90%, 80%, and 68% agreement on the sign of change. The third and final group of numbers gives the percentage areas covered by projections, which simultaneously satisfy two conditions: (i) the degree of consensus on the sign of change is <55% and (ii) one of the four conditions stated in the first column.

		Dec–Feb	Mar–May	Jun–Aug	Sep–Nov
	Small signal condition				
1	$\Pr((\Delta/\sigma)^* \leq 0.20) \geq 0.90$	26	32	22	27
2	$\Pr((\Delta/\sigma)^* \leq 0.15) \geq 0.90$	11	17	8	14
3	$\Pr((\Delta/\sigma)^* \leq 0.10) \geq 0.90$	1	4	0	2
4	$\Pr((\Delta/\mu)^* \leq 0.10) \geq 0.90$	27	32	26	29
	Consensus				
5	99%	7	5	3	7
6	90%	19	21	17	20
7	80%	33	35	34	33
8	68%	58	59	60	59
	Small signal condition and no consensus				
9	(a) $\Pr((\Delta/\sigma)^* \leq 0.20) \geq 0.90$	10	12	9	10
10	(b) $\Pr((\Delta/\mu)^* \leq 0.10) \geq 0.90$	9	10	8	8
11	(a) or (b)	11	13	11	11
12	(a) and (b)	7	9	6	7

Pacific, the South Indian Ocean, the Southern Ocean, Russia, the North Pacific, the North Atlantic, northern Australia, and New Zealand. See Fig. 2 for locations of such regions in other seasons.

Note that the areas identified in Fig. 2 correspond to regions where the normalized change, $|(\Delta/\sigma)^*|$, is very likely to be less than 0.2. The percentage areas covered using smaller thresholds of 0.15 and 0.10 are given in Table 1. The areas covered are reduced. For example, the total area covered in SON is 14% and 2% using 0.15 and 0.10, respectively, compared with 27% obtained earlier with 0.2. Larger samples, for example, more models, would be required to locate areas using smaller thresholds.

4. An alternative definition of small

In the previous section we associated small with the size of the projected change relative to the internal variability. It is also of interest to know if the projected change is very likely to be small relative to average precipitation. Regions where $\Pr(|(\Delta/\mu)^*| < 0.1) > 0.9$, that is, where the projected change is very likely to be small relative to the mean rate of precipitation—again using the method outlined in the appendix—are indicated for MAM by black and blue triangles in Fig. 3a. Here μ is the mean precipitation during 1980–99. We will refer to this approach as “Method 2” (and the method comparing the change to the standard deviation of the variability in section 3 as “Method 1”). Such regions are largely confined to the tropics and the subtropics and occur in many of the same places that were identified using Method 1 (Fig. 2). This is confirmed in Fig. 3a, which also shows where the projected “scaled”

change, that is, $|(\Delta/\mu)^*|$ is very likely small. The black triangles indicate where the change is very likely small using both Methods 1 and 2. Of course the degree of this similarity is a function of the threshold values used in the criteria used in Methods 1 and 2 (i.e., 0.2 and 0.1, respectively).

The criteria used in Methods 1 and 2 are similar if the MEM of $\sigma_j/\mu_j = \alpha = \sigma/\mu \approx 0.5$. This is illustrated in Fig. 3b, which shows the value of σ/μ for MAM: regions with black triangles (i.e., criteria in Methods 1 and 2 both satisfied) tend to occur in some of the regions where $\alpha \approx 0.5$. Red triangles (criterion in Method 1 satisfied but criterion in Method 2 is not) tend to occur in some of the regions where $\alpha \gg 0.5$, while blue triangles (criterion in Method 2 satisfied, but criterion in Method 1 is not) tend to occur in some of the regions where $\alpha \ll 0.5$.

Red triangles tend not to occur in regions where $\alpha \ll 0.5$, and blue triangles tend not to occur where $\alpha \gg 0.5$.

The presence of regions where change is seasonal precipitation is small is consistent with our understanding of the drivers of large-scale precipitation change under global warming. Greater water vapor concentrations with warming due to the Clausius–Clapeyron equation act to enhance the moisture convergence and divergence zones leading to amplified precipitation–evaporation patterns (Mitchell et al. 1987; Held and Soden 2006; Chou et al. 2009). Changes reminiscent of this have been identified in the observational record (Allan et al. 2010). If regions where precipitation is increased are adjacent to regions where precipitation is reduced then the existence of boundary regions where precipitation change is small is not unexpected.

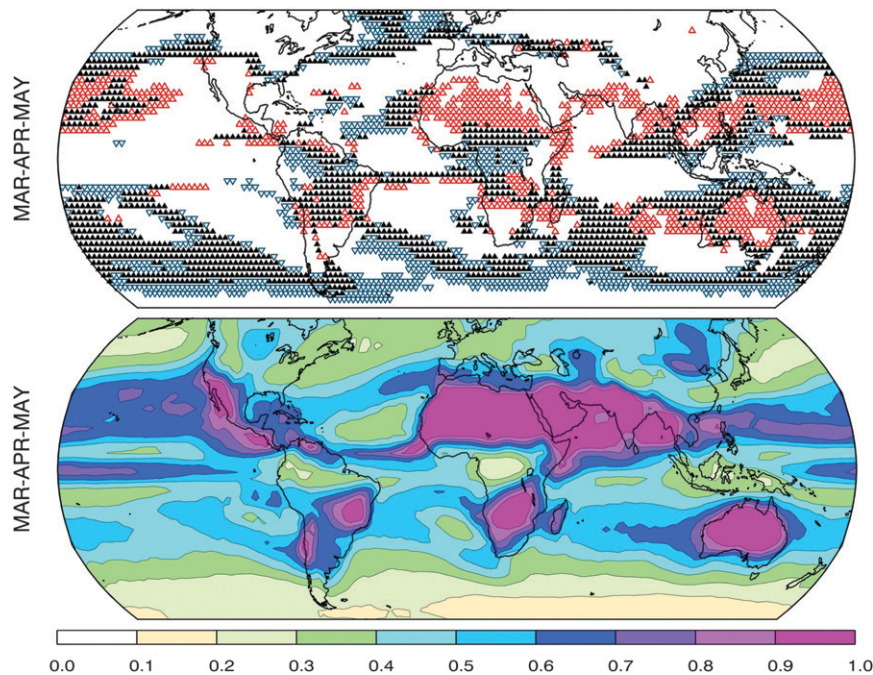


FIG. 3. (a) Areas where (i) $\Pr(|\Delta^*/\sigma^*| < 0.20) \geq 0.90$ and (ii) $\Pr(|(\Delta/\mu)^*| < 0.10) \geq 0.90$ (solid triangles), condition (i) alone (red triangles) and condition (ii) alone (blue triangles); (b) $\alpha = \sigma/\mu$, the MMEM of σ_j/μ_j , where σ_j is the standard deviation of temporal variability in model j and μ_j is the late twentieth-century (1980–2000) average precipitation in model j —for MAM only.

5. Relationship between the different approaches used to identify robust change

As noted in the introduction, the 2007 IPCC report identified regions where (i) there was a consensus on the sign of change among the models or (ii) the MMEM change exceeded the intermodel standard deviation. It is of interest to know how these two alternative indicators of a robust signal discussed in the previous section compare. This is addressed in Fig. 4, which shows the degree of consensus on the sign of the projected change (color shading) and the areas (blue–gray dots) where the MMEM change exceeds the intermodel standard deviation of the change for MAM. There is a very strong relationship between the two statistics: the dotting only occurs where there is a high degree of consensus on sign—typically where 20 or more of the 24 models agree on the sign of change.

Figure 4 also shows that the regions where the normalized change is very likely to be small (black triangles) are all located in regions where the consensus on sign is either weak or absent and where the MMEM change is less than the intermodel standard deviation (i.e., away from the blue–gray dots).

The relationship between the degree of consensus on the sign of projected changes and the MMEM of Δ_j/σ_j (i.e., Δ/σ), is depicted in Fig. 5. This figure shows a scatterplot

between these two variables using all available grid points over the globe during MAM. The plot illustrates that $|\Delta/\sigma|$ increases as the consensus on sign increases. For example, at locations where there is no consensus at all, $|\Delta/\sigma|$ is almost always less than 0.2. At the other extreme, that is, at locations where all models agree on the sign of change, $|\Delta/\sigma|$ exceeds 0.2 everywhere. So the size of $|\Delta/\sigma|$ is a good guide to the degree of consensus, and the degree of consensus is a good guide to the size of $|\Delta/\sigma|$.

Regions where the projected change is very likely small or absent were identified above. It is also of interest to know where there is a consensus among the models that the change is small or absent. The degree of consensus among models on whether the projected precipitation change is unremarkable in the sense that $|\Delta_j/\sigma_j| < 0.4$ is indicated by color shading in Fig. 6 (MAM). Triangles indicate that the normalized change is very likely small. The triangles only tend to appear where there is a strong consensus among the models that $|\Delta_j/\sigma_j| < 0$. There is in fact a strong relationship between these two approaches: if there is a strong consensus among the models that the change is unremarkable in the sense that $|\Delta_j/\sigma_j| < 0.4$ then there is a strong tendency that the normalized change is very likely small or absent [i.e., $\Pr(|D^*| < 0.2) > 0.9$].

Blue–gray dots again indicate that the MMEM change exceeds the intermodel standard deviation. Regions

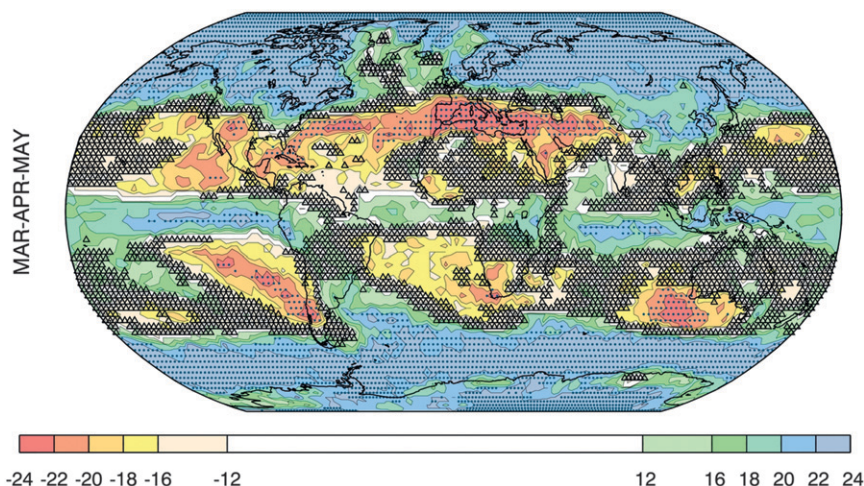


FIG. 4. The degree of consensus on the sign of the projected change (colored shading), the areas where the MMEM change exceeds the intermodel standard deviation (dots), and regions where the projected change is very likely small relative to variability (small triangles). In virtually all cases triangles only appear where the degree of consensus is very low or absent—for MAM only.

with dots do not have triangles, and dots only occur in regions where there is a high level of consensus among the models that $|\Delta/\sigma| > 0.4$. This plot again highlights the tendency toward consistency across the various approaches used to examine robust change.

6. The probability of consensus

As noted in the introduction and as depicted in Fig. 1a (the shaded regions) the degree of consensus in stippled

areas in the IPCC report and in Fig. 1 is 90% or higher, that is, 90% or more of the models have projected changes that have the same sign. Suppose that the signal, that is, the externally forced response in seasonal rainfall, is zero in all models. In this case the sign of the projected change in each model is entirely a function of the internally generated, essentially random variability. How likely is it in this situation to fluke (i.e., obtain by chance) consensus on the sign of change among 90% or more of the models? This question is addressed in Fig. 7, which shows

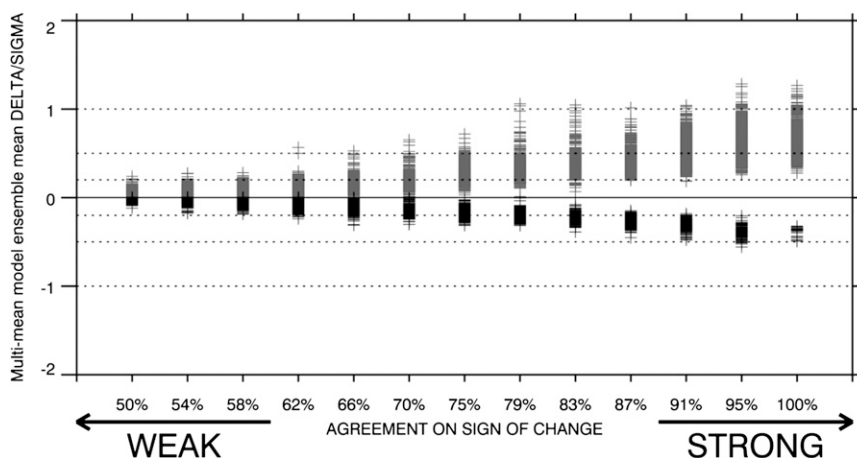


FIG. 5. A scatterplot between degree of consensus on the sign of projected changes and the magnitude of the MMEM of Δ/σ (which is referred to as $|\Delta/\sigma|$ in the text), using all available grid points over the globe. The $|\Delta/\sigma|$ increases as the consensus on sign increases. At locations where there is no consensus at all $|\Delta/\sigma|$ tends to be less than 0.2; ± 0.2 , ± 0.5 , and ± 1.0 are indicated by the dotted lines. At locations where all models agree on the sign of change, $|\Delta/\sigma|$ exceeds 0.2 in the vast majority of cases—for MAM only.

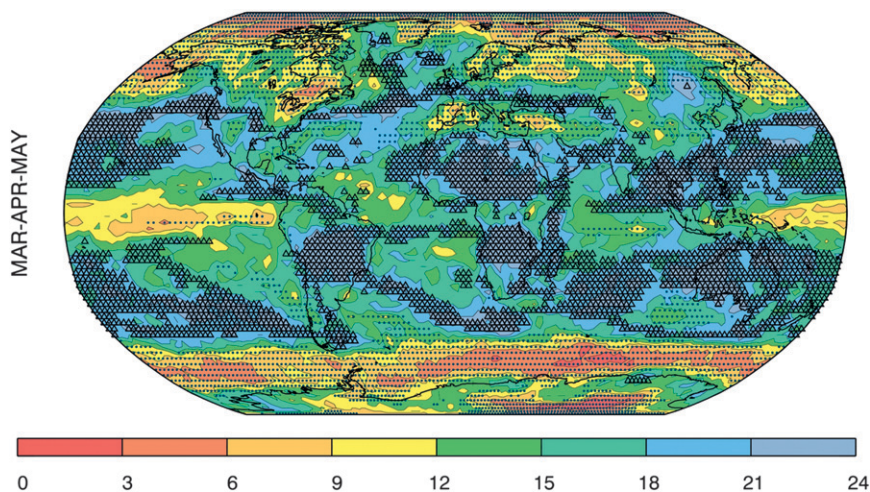


FIG. 6. Degree of consensus among models on whether the projected precipitation change is unremarkable in the sense that $|\Delta j/\sigma j| < 0.4$ (shading). Shading gives number of models for which this is true. Triangles indicate that the change is very likely to be small relative to the variability, using the test outlined in the appendix. The triangles only occur where there is a strong consensus among the models that $|\Delta j/\sigma j| < 0.4$. Dots indicate that the MMEM change exceeds the intermodel standard deviation—in MAM only.

the probability density function (PDF) and the cumulative DF for a binomial distribution with $p = q = 0.5$ and $N = 24$, which is given by the expression $(p + q)^N$. Thus we assume that the probability of the sign of change in each model being positive (p) is 0.5 and the probability that the sign is negative (q) is also 0.5, that is, equally likely. This approach was used in a similar context by Whetton et al. (1996). The PDF shows that with $N = 24$ (i.e., 24 models) the probability of having 90% agreement or more is extraordinarily unlikely (significance level = 99.998%) as is 80% consensus (both 80% and 90% are used in the 2007 IPCC report. 90% is used in the Summary for Policymakers (Alley et al. 2007) and the Technical Summary (Solomon et al. 2007b), while 80% is used in, for example, Figure 10.12 (Meehl et al. 2007a). In fact the 99% significance level corresponds to 74% agreement, while the 95% significance level corresponds to only 68% agreement. So one could use a 68% degree of consensus and still be confident, under our assumption of independence, that such a degree of consensus is very unlikely by chance alone. If we use this smaller value (i.e., of 68% agreement rather than 80% agreement) then an additional 27%–33% of the earth's surface has projections that are robust in this sense during all seasons.

As noted in the introduction, the statistical significance discussed here is a measure of the probability of reaching a given level of consensus on the sign of change among the models through sampling error, under the null hypothesis that there is in fact no change. This is different to the likelihood that there is in fact a change—a difference that was recently highlighted by Ambaum (2010).

To examine the relationship between these two quantities (i.e., the likelihood of change and the statistical significance of apparent change), we identified regions where the projected normalized change is extremely unlikely to be small in the sense that $\Pr(|(\Delta/\sigma)^*| < 0.2) < 0.05$. These areas are identified using dots in Fig. 8. Also shown (shaded) is the degree of consensus on the sign of the projected change. Notice that the dots are only located in regions where 18 or more of the models agree on the sign of the change. This corresponds to a statistical significance level of 99% (Fig. 7). So while statistical significance is not strictly a measure of the likelihood of change, there is a very robust relationship in this context between the two quantities.

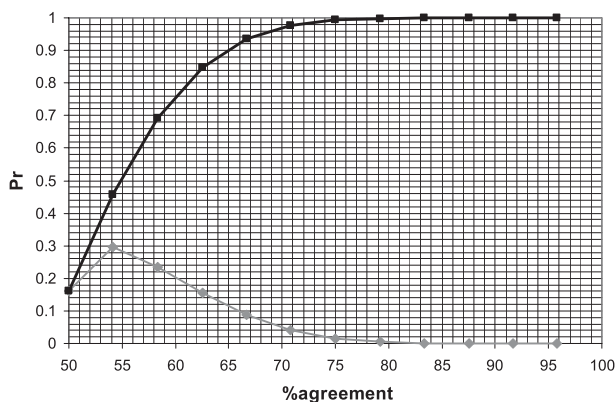


FIG. 7. The PDF and CDF for a binomial distribution with $p = q = 0.5$ and $N = 24$ (i.e., 24 models).

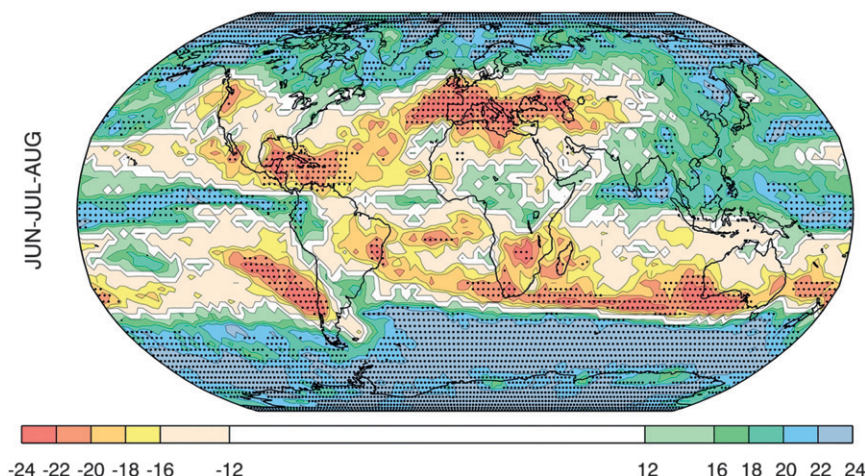


FIG. 8. Level of agreement on the sign of the projected change in precipitation (color shading) and regions where the projected change is extremely unlikely to be small in the sense that $\Pr(|\Delta^*/\sigma^*| < 0.2) < 0.05$ (dots). Notice that the dots only appear in regions where 18 or more of the models agree on the sign of the change. This corresponds to a statistical significance level of 99% (Fig. 7)—for JJA only.

7. More widespread consensus

We can now use Fig. 6 and the results obtained in the previous sections to reassess consensus among model projections making use of the fact that there are regions where the normalized change is very likely to be small, and consensus among 68% or more of models is statistically significant at the 95% level. This new information is presented in Fig. 9 for all seasons. It shows the percentage change in precipitation at all locations. Areas with 66% or more agreement have color shading. Dots indicate that the degree of agreement among the models on the sign of change is statistically significant at the 95% level (i.e., 68% agreement, see Fig. 7). Triangles in the unshaded region are located where there is little or no consensus on the sign of change (<55% agreement). Black solid triangles indicate that the magnitude of the projected change is very likely to be smaller than both (i) 0.2σ (where σ is the MEM of model interannual variability during 1980–99) and (ii) 0.1μ , where μ is the MEM value of the precipitation averaged over the period 1980–99. Red triangles indicate that (i) is true but (ii) is not, while blue triangles indicate that (ii) is true but (i) is not. Likelihood is again assessed using the method outlined in the appendix.

The regions where genuine uncertainty remains toward the end of the twenty-first century can also be identified in, for example, Fig. 9. These are the unshaded regions that do not have triangles. These regions tend to be located in between (i) regions where there is consensus on negative changes and (ii) regions where there is a consensus on positive changes. They also tend to surround regions where

the signal is very likely small. There is very little uncertainty at high latitudes.

Approximately 69% of the globe is now dotted or covered with red or black triangles during JJA. In Fig. 1a we followed the approach used in the 2007 IPCC report and stippled regions where 90% or more of the models agree on the sign of the change. This area covers only 17% of the globe, which might give the impression to some readers that the area covered by robust projections is very limited. Note that the IPCC report (Fig. SPM.7) also shaded regions where 66% or more of the models agreed on the sign of the change. However, no information on the statistical significance of projections in unstippled areas in Fig. SPM.7 or similar plots was provided, and so some readers might have been left with the impression that changes were not robust in all regions that were not stippled.

An alternative simpler approach to presenting this sort of information is given in Fig. 10a. Shading again indicates the % change in precipitation (MAM only). All areas are shaded for transparency. Dots indicate that the degree of agreement among the models on the sign of change is statistically significant at the 95% level (i.e., 68% agreement). To keep things simple the circles indicate where (i) 68% of models agree that the projected normalized change < 0.35 and (ii) where there is little or no agreement on the sign of change. Additional analysis (not shown) indicates that condition (i) corresponds approximately to where the normalized precipitation change is very likely small for 24 independent models.

A similar plot for a “near-term” period of 2016–35 is presented in Fig. 10b. Notice that the area covered by

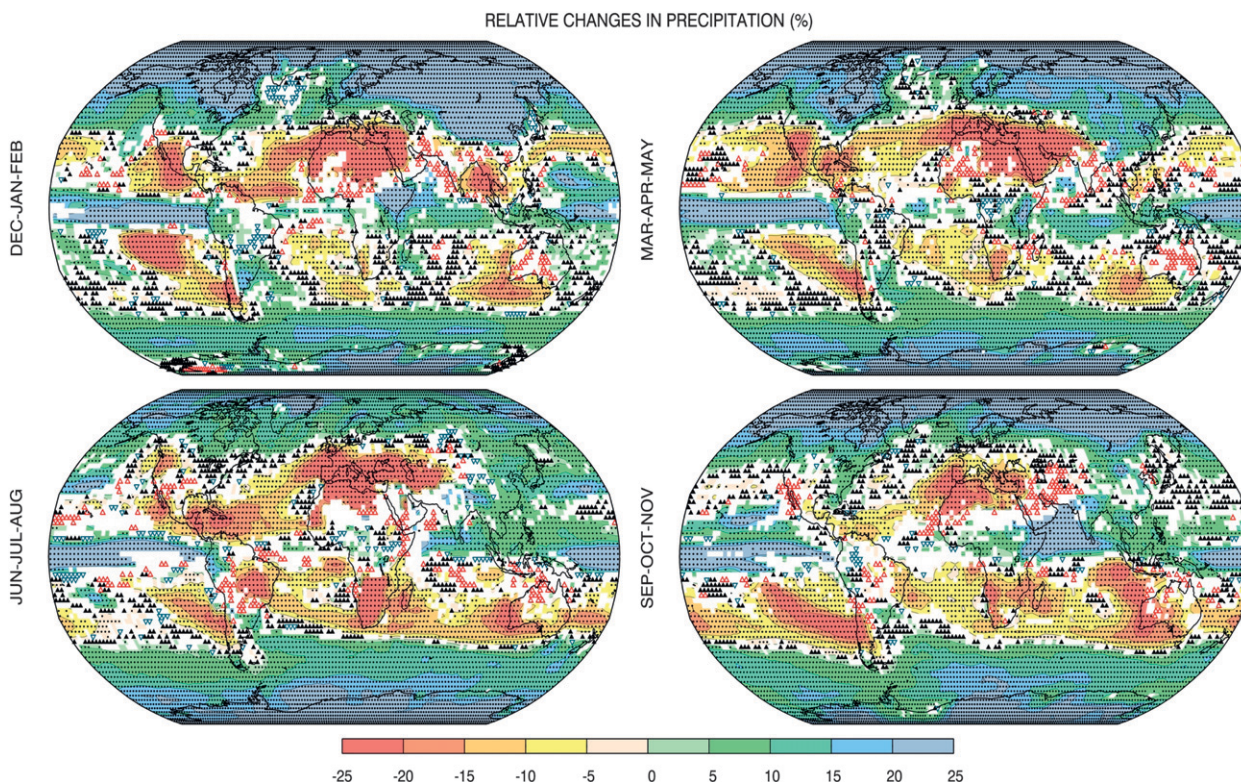


FIG. 9. Percentage (%) change in precipitation averaged over the period 2090–99, relative to precipitation averaged over the period 1980–99 (color shading) for (a) DJF, (b) MAM, (c) JJA, and (d) SON. The figures represent the MMEM of 24 different models forced using the SRES A1B scenario (see section 2 for further details). Only regions where 68% or more of models agree on the sign of change are shaded. All of the triangles in the unshaded region are located where there is little or no consensus on the sign of change ($<55\%$). Black solid triangles indicate that the magnitude of the projected change is very likely to be smaller than both (i) 0.20σ (where σ is the MMEM of model interannual variability during 1980–99) and (ii) 0.10μ , where μ is the MMEM value of the precipitation averaged over the period 1980–99. Red unfilled triangles indicate that (i) is true but (ii) is not, while blue unfilled triangles indicate that (ii) is true but (i) is not. Likelihood is assessed using the method outlined in the appendix.

open circles is much more extensive than for the late twenty-first century (Fig. 10a). Open circles even appear at high southern latitudes. Numerous variants on this approach are of course possible.

8. Model dependence

In previous sections we assumed that the projection from each model is an independent piece of information. However, as noted in the introduction, models exhibit dependence. Masson and Knutti (2011), for example, examined contrasts in the CMIP3 model simulations of past climate, concluding that the effective number of independent models is 10 only, far less than the total number of models (24). This conclusion is consistent with Pennell and Reichler (2011), who examined contrasts in the skill of CMIP3 models to simulate past climate. They concluded that the effective number of independent CMIP3 models is only 7.5–10.

If we assume that similar figures apply to the degree of independence among projections then the statistical significance of the results presented previously are erroneous. For example, if there are only 10 effectively independent models then the binomial distribution (with $n = 10$) indicates that 8 or more families (i.e., 80% or more) must agree on the sign of change to attain the 95% significance level. This compares to the lower value of 68% obtained in section 6 where we assumed that we had 24 independent models.

If the interdependence among subsets or families of models is strong then rather than presenting multimodel means (MMMs) and providing associated statistics assuming independence it might be better to present multifamily means (MFMs) and statistics based on the assumption of M independent families, where $M < N =$ the total number of models. Such a plot is presented in Fig. 11. We identified 10 families of models using the model genealogy results of Masson and Knutti (2011). The MFM

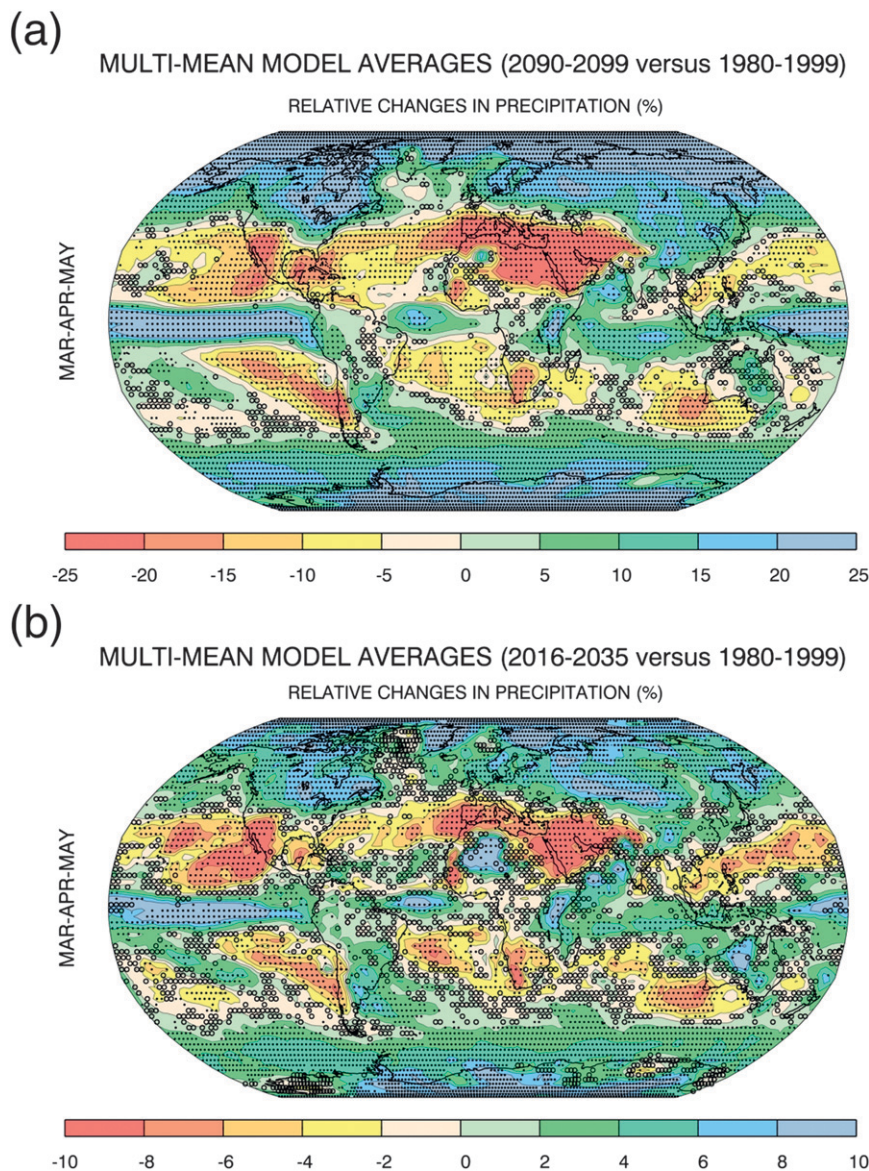


FIG. 10. Percentage (%) change in precipitation averaged over the period (a) 2090–99, relative to precipitation averaged over the period 1980–99 (color shading) for MAM only. Shading is shown at all locations for transparency. The dots indicate where 68% of models agree on the sign of change. The small open circles indicate where (i) 68% of the models agree that the normalized change is less than 0.35, and (ii) where 55% or fewer models agree on the sign of change. (b) As in (a), but for 2016–35.

% change is very similar though not exactly the same as the MMM % change (Fig. 11). The filled dots again indicate that agreement on the sign of the change is significant at the 95% level, but this time for 10 independent families rather than 24 models. The unfilled circles indicate that the normalized change is very likely small using the method outlined in the appendix, but this time using ensemble family mean values for both σ and Δ rather than individual model values and setting $N = 10$.

It is important to note, however, that Masson and Knutti (2011) assessed contrasts in simulations of past climate (relative to the MMEM) while Pennell and Reichler (2011) assessed contrasts in the skill with which the models simulate past climate (relative to observations). Neither study assessed dependence in projections. To infer that the degree of independence they obtained applies to projections, it is necessary to assume that there is close relationship between (i) contrasts in, for example, model simulations of past climate and (ii) contrasts in

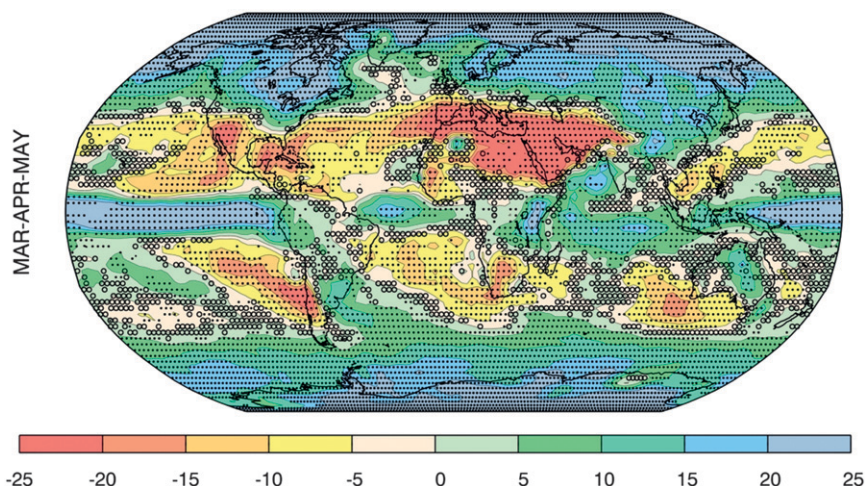


FIG. 11. Multifamily mean change. Percentage (%) change in precipitation averaged over 10 “families” of models (see text), for the period 2090–99, relative to precipitation averaged over the period 1980–99 (color shading) for MAM only. The dots indicate where 80% or more of families agree on the sign of change. This degree of consistency approximates the 95% statistical significance level for 10 independent families. The small open circles indicate where normalized change is very likely to be less than 0.2 and where 60% or fewer families agree on the sign of change.

twenty-first-century projections. The relationship between (i) and (ii) has been investigated previously (Whetton et al. 2007). Whetton et al. (2007), for example, showed that while there are detectable relationships between (i) and (ii) the relationships are generally weak, especially for regional precipitation projections. The degree of similarity in simulated twentieth-century climate typically accounted for only 10% of the contrasts in the projections. This is consistent with the results of Jun et al. (2008), who showed that while climate model biases are correlated with each other, contrasts in the skill of models at simulating mean climate is not strongly related to the ability of the same models to simulate warming. This work is also consistent with studies that have investigated climate feedbacks in models. For example, John and Soden (2007) found that lapse rate and water vapor feedbacks in CMIP3 models were uncorrelated with base-state biases in tropospheric temperature and water vapor.

An indication that projections are not as interdependent as simulations of climatology is given in Fig. 12. It shows Taylor (2001) diagrams for σ (left), Δ (middle), and Δ/σ (right) for SON. Each model is compared with the appropriate MMM field. Each family is assigned its own color. One gets the impression that some of the families do indeed cluster together. It seems that for this season at least, clustering in Δ (middle) and Δ/σ (right) is smaller than it is for σ (left).

These issues are clarified in Fig. 13, which shows the median distance between family members relative to the median distance between all models. Distance is defined as $D = (\Delta\text{NSD}^2 + \Delta\text{Correl}^2)^{1/2}$, where ΔNSD is the difference between the normalized standard deviation between the two models chosen and ΔCorrel is the difference between the spatial correlation of the same two models. We first calculated the distance between all possible pairs of models, regardless of their family, and

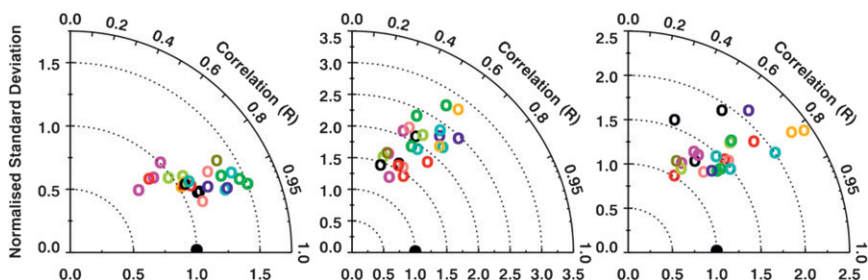


FIG. 12. Modified Taylor (2001) diagrams for (left) σ , (middle) Δ , and (right) Δ/σ for SON. The modification is that instead of comparing each model against the observational mean they are compared with the MMM. Colors denote model “families.”

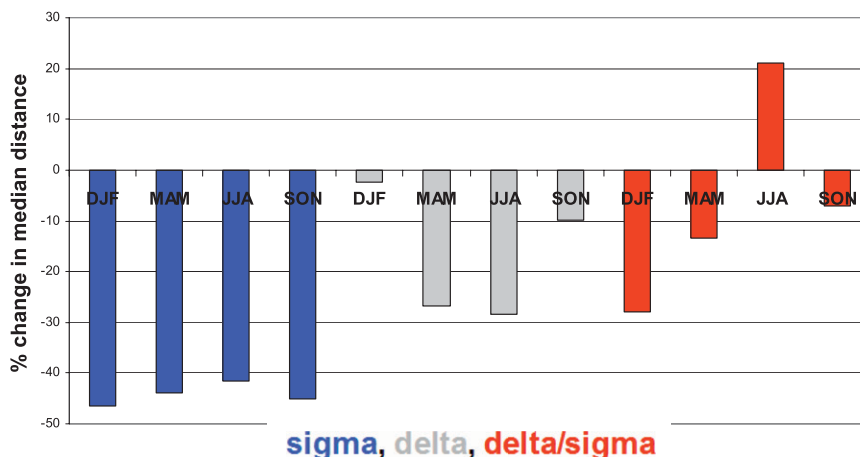


FIG. 13. The median distance between family members relative to the median distance between all models, that is, $\% \text{ Diff} = [M(D_F) - M(D_{ALL})]/M(D_{ALL})$ for each season and for σ (blue), (middle) Δ (gray), and red) Δ/σ . Distance is defined as $D = (\Delta \text{NSD}^2 + \Delta \text{Correl}^2)^{1/2}$, where ΔNSD is the difference between the normalized standard deviation between the two models chosen, and ΔCorrel is the difference between the spatial correlation between the same two models.

then calculated the median of these distances $[M(D_{ALL})]$. We then chose a family and calculated the distances between each member of that family, repeating this for every family with more than one member. We then calculated the median of the distances between family members $[M(D_F)]$, and then the percentage change in the median distance, defined by $\% \text{ Diff} = [M(D_F) - M(D_{ALL})]/M(D_{ALL})$. Values of $\% \text{ Diff}$ were calculated for each season and for σ , Δ , and Δ/σ .

Figure 13 indicates that family ties reduce the median distance in all seasons for σ , Δ , and Δ/σ . The only exception is for Δ/σ in JJA, in which the median distance between family members is greater than the median distance between all models. The averages of $\% \text{ Diff}$ across all seasons are -44% , -17% , and -7% for σ , Δ , and Δ/σ , respectively, suggesting that the impact of family ties is much greater for σ than it is for either Δ or Δ/σ . Additional analysis (not shown) shows that clustering for σ remains much stronger than clustering for Δ if we use means instead of medians. Clustering is, on average, about the same for Δ/σ as it is for σ if we use means. However, we regard the median as a better measure than the mean because it is not as heavily influenced by outliers.

In summary, clustering among projections (Δ) is much weaker than clustering among simulations of climatology (σ). Hence estimates based on contrasts in simulations of the climatology will tend to overestimate dependence in projections. The approach taken in Fig. 11 is therefore very conservative and likely to be suboptimal.

Clearly issues concerning model dependence in projections need further research and there remain considerable

uncertainties. Further investigation is beyond the scope of the present paper.

9. Summary

We analyzed projected precipitation changes for the late twenty-first century under the SRES A1B scenario (Nakicenovic et al. 2000) derived from numerous different coupled general circulation models (CGCMs) from the WCRP/CLIVAR/CMIP3 database (Meehl et al. 2007b). Projected changes were calculated using data averaged over the two periods 2090–99 and 1980–99.

We began by assuming that all models are independent. We identified regions where the projected change in precipitation is very likely to be small in terms of year-to-year variability or in terms of the twentieth-century-average precipitation. We showed that there is a high degree of correspondence between various methods used to assess the robustness of projected changes—including two methods used extensively in the 2007 IPCC report. For example, if there is a high degree of consensus on the sign of change then (i) the multimodel ensemble mean (MMEM) tends to be larger than the intermodel standard deviation of the projected changes in individual models change (Fig. 4) and (ii) the magnitude of the MMEM of normalized change (i.e., Δ_j/σ_j , where Δ_j is the projected change between the late twenty-first century and the late twentieth century, and σ_j is the standard deviation of the variability evident in model j over the late twentieth century) exceeds 0.2 (Fig. 5). Conversely if there is little or no consensus on the sign of the projected change

among models then (i) the MEM is less than the intermodel standard deviation of the projected change and (ii) the magnitude of the MEM of Δ/σ_j is almost always less than 0.2.

We also found a very high level of consistency between regions where (i) small change is extremely unlikely (i.e., $\Pr(|(\Delta/\sigma)^*| < 0.2) < 0.05$) and where (ii) change is statistically significant at or above the 99% level (Fig. 8).

The level of consensus used in the 2007 IPCC report to identify projections that are robust at least in terms of model agreement was shown to be exceedingly unlikely to occur by chance under the standard null hypothesis of no change and model independence. A much lower level of consensus can be used to identify a statistically significant level of agreement on the sign of change under these assumptions. Here $(\Delta/\sigma)^*$ ($=D^*$ in text) is the unknown but estimated externally forced change expressed as a fraction of the late twentieth-century standard deviation of year-to-year variability in precipitation.

By taking this last finding and the identification of areas very likely to have no change or change that is small relative to year-to-year variability in precipitation (or twentieth-century-average precipitation) and no consensus on the sign of change, we showed that the area over which climate change projections among the CMIP3 climate models are consistent is much more widespread than might have previously been appreciated by many climate scientists and readers of the 2007 IPCC report. The total proportion of the entire globe covered by late twenty-first-century projections on which the models display a statistically significant degree of consensus is 67%, 71%, 69%, and 69% during DJF, MAM, JJA, and SON, respectively. If the effects identified here are overlooked, as they have been previously, the corresponding figures are only 19%, 22%, 17%, and 20%, respectively (defining robust as agreement among 90% or more of the models on the sign of change).

The regions where genuine uncertainty remains in the late twenty-first century tend to be located in between regions where there is consensus on negative-positive changes. They also tend to surround regions where the projected change is very likely small. There is very little uncertainty at high latitudes.

As noted in the introduction, the statistical significance discussed above is a measure of the probability of reaching a given level of consensus on the sign of change amongst the models through sampling error, under the null hypothesis that there is in fact no change. This is different to the likelihood that there is in fact a change—a difference that was recently highlighted by Ambaum (2010). We showed, however, that while statistical significance is not strictly a measure of the likelihood of

change, there is a very robust relationship in this context between the two.

In the results discussed above we assumed that we have 24 independent models. The issue of model dependence [through biases arising from the sharing of (imperfect) parameterizations amongst models] was investigated. Previous studies have identified model dependence among models in their ability to simulate climatology (Masson and Knutti 2011; Pennell and Reichler 2011). We used this earlier work to identify 10 “families” of models and produced a plot showing the multifamily mean (MFM) change in precipitation, identifying regions where there is agreement among the families on the sign of change and where the MFM change is very likely small.

We pointed out, however, that links between model biases and contrasts in projections (Whetton et al. 2007; John and Soden 2007) or simulations of observed trends (Jun et al. 2008) are weak and that the links to regional rainfall are especially weak (Whetton et al. 2007).

We also showed that, while model dependence is evident in both simulations of σ and in Δ , the interdependence in Δ seems much smaller than it is for σ . Hence estimates of interdependence based on contrasts in simulations of the climatology very likely overestimate interdependence among projections.

10. Discussion

The 2007 IPCC report gave very great emphasis to identifying regions, processes and phenomena for which there appeared to be robust change or where change exceeded the intermodel standard deviation. For example Chapter 11 of the IPCC report neatly summarized robust findings on regional precipitation changes that were considered likely or very likely. However, very little information was given on regions where change is very likely to be small or absent. We have shown that in some locations where there is no consensus on the sign of change, the projected change is very likely to be small or absent, and there is little or no consensus on the sign of change. This is a subtle but very important difference. Discriminating between these two different situations (i.e., where projections are genuinely uncertain and where the projections are very likely small or absent), helps us to more accurately represent projections information and will help us convey more useful information to the wider community.

Consensus among models is an important factor to consider when we assign confidence to projections. Of course even complete model consensus on a particular change is not a guarantee of high confidence because all the models might share a common systematic bias that impacts upon the particular projection one is examining.

So we cannot assume that a consensus among the models that change is small or likelihoods based entirely on model output will necessarily be described as a robust projection. The degree of statistical significance assuming model independence is a single line of evidence. To form a judgment on our confidence in the projections we need to also examine other lines of evidence. This includes consistency with projections from earlier generations of models or from simplified physically based modeling systems, the degree to which the physics underpinning the projected change is understood and accepted, the degree of confidence we might have in the treatment of physical processes in our climate models relevant to the projection being considered, and the degree to which the projected changes are already evident in the observational record (Allan et al. 2010). It also includes the ability of models to simulate climatological features and observed trends relevant to what is being projected, and, of course, the degree to which model projections are independent.

It would be very useful if the scientific community could accurately quantify the impact of model interdependence on the statistical significance of projected changes. While research in this area is being conducted (Masson and Knutti 2011; Pennell and Reichler 2011; Jun et al. 2008; section 8), we are not currently able to provide such accurate quantification. Nor is there any guarantee that we will ever be able to. One current worthwhile approach is to suppose that the impact of interdependency on the statistical significance of a projected change can be accomplished by assuming that interdependency reduces the number of models to a smaller “effective” number of models. In fact we exploited this in section 8 in our grouping of models into families. However, there is no guarantee that the level of interdependence will be the same for all variables, for all seasons, or for all locations. It could be that in some locations and some seasons the model-dependent deficiencies in parameterizations underpinning model interdependence might not have as strong an impact as they do in other regions or other seasons. In an extreme though very plausible case perhaps the consequences of a particular physical parameterization are unimportant for a particular region at a particular time of the year. Hence it might not be appropriate to assume that the degree of model interdependence can be characterized in terms of a single unchanging effective model number for all variables, locations, and seasons.

We showed that clustering in projections is weaker than clustering in simulations of past climate. It is also important to realize that any clustering that does exist might have a relatively small impact on the magnitude and sign of projections in certain regions or seasons. Clustering per se is therefore a necessary though insufficient condition for having appreciably reduced independence among

model projections. The magnitude of the expression or impact of the clustering on projections is also crucial in determining the extent to which independence is reduced by that clustering.

In this paper we have focused entirely on seasonal precipitation. While the need to identify regions where changes are likely small is still important for extreme (e.g., daily) precipitation, specific conclusions on the location of the regions described here for seasonal rainfall are not necessarily applicable to extreme precipitation. Conclusions drawn here are not valid at all for some variables, for example, surface temperature, because models agree that surface temperature increases in all locations (Meehl et al. 2007a). On the other hand, many of our high-level conclusions (e.g., that in many regions the projected change is very likely small in some sense) might be valid for some other variables, for example, mean sea level pressure (MSLP) and wind speed (see, e.g., McInnes et al. 2011).

We also, by necessity, made choices for thresholds used to define a “small change.” Others might choose different values. In addition we know that even small changes in precipitation can lead to larger changes in dependent variables including streamflow (e.g., Power et al. 2005) and crop yield (Allan et al. 2010). So a small change in seasonal precipitation does not necessarily correspond to a small impact.

11. Recommendations

It is recommended that approaches similar to those outlined above (e.g., Figure 10) are adopted in the analysis of CMIP5 (Taylor et al. 2011) model output and in the next IPCC report. More broadly, we recommend that the next IPCC report and climate projections research generally give greater attention to identifying quantities, regions, processes, phenomena, and features where projected changes are either small or absent and distinguishing such projections from situations where there is genuine uncertainty in projected change. This is especially important for near-term (e.g., 2016–35) climate (cf. Fig. 10b with Fig. 10a). We also recommend that more attention is given to quantifying the degree to which models agree on the sign of projected changes that corresponds to commonly used levels of statistical significance, that is, that choices made for thresholds used in the statistical tests are linked to the “calibrated” uncertainty language (i.e., the formal assignment of terms like “very likely” to specific and agreed probabilities) that will be used in the next IPCC report (Mastrandrea et al. 2011).

Clearly the quantitative treatment of model interdependence, especially in projections, will continue to be an important area of research. However, it is a complex area, and it may be many years before the impact of

interdependence on the statistical significance of projected changes can be quantified accurately (assuming it can be achieved). We therefore take the pragmatic view that the statistical significance assuming independence is a useful, well-quantified single line of evidence. We therefore recommend that significance information based on the assumption of independence continues to be provided (e.g., in the forthcoming IPCC report), provided that it is given in association with caveats relating the presence of some uncertain degree of interdependence. We also suggest that conservative choices are made for statistical significance levels, for example, that we use 99% levels (assuming independence) to compensate for interdependence.

Acknowledgments. We wish to thank the Climate Variability and Predictability (CLIVAR) Working Group on Coupled Models (WGCM) Climate Simulation Panel for devising, initiating, and managing the WCRP/CLIVAR/WGCM Coupled Model Intercomparison Project 3 (CMIP3) experiment, agencies participating in CMIP3, the Program for Climate Model Diagnostics and Intercomparison (PCMDI) in the United States for collecting, collating, and making CMIP3 output available, and Lawson Hanson for providing ready access to the data here in the Bureau of Meteorology. We wish to thank Thomas Stocker, Ed Hawkins, Rowan Sutton, Matt Collins, Richard Seager, anonymous reviewers, Julie Arblaster, and Greg Kociuba for reviewing earlier drafts. This project was supported by the Australian Climate Change Science Program, the Pacific Climate Change Science Program, and the Pacific Australia Climate Change Science and Adaptation Planning Program (PACCSAP).

APPENDIX

The Probability that $(\Delta/\sigma)^* < 0.2$ or that $(\Delta/\mu)^* < 0.1$

Here we calculate $\Pr(|D^*| < 0.2)$, where $D^* = (\Delta/\sigma)^*$, so we can identify regions where this probability exceeds 0.9, that is, where $|D^*|$ is very likely to be less than 0.2. Here, as in the text, $D^* = (\Delta/\sigma)^*$ is the unknown but estimated externally forced change expressed as a fraction of the late twentieth-century standard deviation of year-to-year variability in precipitation. Note that D^* and the Greek variables in this Appendix vary spatially.

Each of the N models has its own Δ_j/σ_j , where $j = 1, 2, 3, \dots$, and N is a model reference index. The MMEM of the (Δ_j/σ_j) is defined here and in the text as Δ/σ . The variance of the Δ_j/σ_j is, by definition, equal to $S^2 = \sum(\Delta_j/\sigma_j - \Delta/\sigma)/(N - 1)$. We assume that our estimation of D^*

has a t distribution with $N - 1$ degrees of freedom, with mean Δ/σ and variance $R^2 = S^2/N$.

If $|\Delta/\sigma| \geq 0.2$ then it can be shown that $\Pr(|\Delta^*| < 0.2 \sigma)$ with: $\Pr(|D^*| < 0.2)$ and so we need only consider cases with $|\Delta/\sigma| < 0.2$. This leaves the following three cases to consider: (i) $\Delta/\sigma = 0$, (ii) $-0.2 < \Delta/\sigma < 0$, and (iii) $+0.2 < \Delta/\sigma < 0$. In case (i) $\Pr(|D^*| < 0.2) = 1 - 2\Pr(t \geq 0.2/R)$, where t is the t value of a Student's t test with $N - 1$ degrees of freedom. In cases (ii) and (iii) $\Pr(|D^*| < 0.2) = 1 - \Pr[t \geq (0.2 + \Delta/\sigma)/R] - \Pr[t \geq (0.2 - \Delta/\sigma)/R]$.

Identical formulae can be used to calculate $\Pr(\Delta/\mu < 0.1)$ if we set $S^2 = [\sum(\Delta_j/\mu_j - \Delta/\mu)/(N - 1)]$, where Δ/μ is the MMEM of the Δ_j/μ_j , we again set $R^2 = S^2/N$, and we replace “0.2” with “0.1.”

REFERENCES

- Allan, R. P., B. J. Soden, V. O. John, W. Ingram, and P. Good, 2010: Current changes in tropical precipitation. *Environ. Res. Lett.*, **5**, 025205, doi:10.1088/1748-9326/5/2/025205.
- Alley, R. B., and Coauthors, 2007: Summary for policymakers. *Climate Change 2007: The Physical Science Basis*, S. Solomon et al., Eds., Cambridge University Press, 1–18.
- Ambaum, M. H. P., 2010: Significance tests in climate science. *J. Climate*, **23**, 5927–5932.
- Brown, J., S. B. Power, and F. Delage, 2010: Simulation of the SPCZ in AR4 climate models. *J. Climate*, **24**, 1565–1582.
- Chou, C., J. D. Neelin, C.-A. Chen, and J.-Y. Tu, 2009: Evaluating the “Rich-Get-Richer” mechanism in tropical precipitation change under global warming. *J. Climate*, **22**, 1982–2005.
- Christensen, J. H., and Coauthors, 2007: Regional climate projections. *Climate Change 2007: The Physical Science Basis*, S. Solomon et al., Eds., Cambridge University Press, 847–940.
- Hegerl, G. C., and Coauthors, 2007: Understanding and attributing climate change. *Climate Change 2007: The Physical Science Basis*, S. Solomon et al., Eds., Cambridge University Press, 663–745.
- Held, I. M., and B. J. Soden, 2006: Robust responses of the hydrological cycle to global warming. *J. Climate*, **19**, 5686–5699.
- Irving, D. B., and Coauthors, 2011: Evaluating global climate models for climate change projections in the Pacific region. *Climate Res.*, **49**, 169–187.
- John, V. O., and B. J. Soden, 2007: Temperature and humidity biases in global climate models and their impact on climate feedbacks. *Geophys. Res. Lett.*, **34**, L18704, doi:10.1029/2007GL030429.
- Jun, M., R. Knutti, and D. W. Nychka, 2008: Spatial analysis to quantify numerical model bias and dependence: How many climate models are there? *J. Amer. Stat. Soc.*, **103**, 934–947.
- Masson, D., and R. Knutti, 2011: Climate model genealogy, 2011. *Geophys. Res. Lett.*, **38**, L08703, doi:10.1029/2011GL046864.
- Mastrandrea, M. D., K. J. Mach, G.-K. Plattner, O. Edenhofer, T. F. Stocker, C. B. Field, K. L. Ebi, and P. R. Matschoss, 2011: The IPCC AR5 guidance note on the consistent treatment of uncertainties: A common approach across the working groups. *Climatic Change*, **108**, 675–691, doi:10.1007/s10584-011-0178-6.
- McInnes, K., T. A. Erwin, and J. M. Bathols, 2011: Global Climate Model projected changes in 10 m wind speed and direction due to anthropogenic climate change. *Atmos. Sci. Lett.*, **12**, 325–333, doi:10.1002/asl.341.

- Meehl, G. A., and Coauthors, 2007a: Global climate projections. *Climate Change 2007: The Physical Science Basis*, S. Solomon et al., Eds., Cambridge University Press, 747–845.
- , C. Covey, T. Delworth, M. Latif, B. McAvaney, J. F. B. Mitchell, R. J. Stouffer, and K. E. Taylor, 2007b: The WCRP CMIP3 multimodel dataset: A new era in climate change research. *Bull. Amer. Meteor. Soc.*, **88**, 1383–1394.
- Mitchell, J. F. B., C. A. Wilson, and W. M. Cunningham, 1987: On CO₂ climate sensitivity and model dependence of results. *Quart. J. Roy. Meteor. Soc.*, **113**, 293–322.
- Nakicenovic, N., and Coauthors, 2000: *Special Report on Emissions Scenarios: A Special Report of Working Group III of the Intergovernmental Panel on Climate Change*. Cambridge University Press, 599 pp.
- Pennell, C., and T. Reichler, 2011: On the effective number of climate models. *J. Climate*, **24**, 2358–2367.
- Power, S. B., B. Sadler, and N. Nicholls, 2005: The influence of climate science on water management in western Australia: Lessons for climate scientists. *Bull. Amer. Meteor. Soc.*, **86**, 839–844.
- Randall, D. A., and Coauthors, 2007: Climate models and their evaluation. *Climate Change 2007: The Physical Science Basis*, S. Solomon et al., Eds., Cambridge University Press, 589–662.
- Solomon, S., D. Qin, M. Manning, M. Marquis, K. Averyt, M. M. B. Tignor, H. L. Miller Jr., and Z. Chen, Eds., 2007a: *Climate Change 2007: The Physical Science Basis*. Cambridge University Press, 996 pp.
- , and Coauthors, 2007b: Technical summary. *Climate Change 2007: The Physical Science Basis*, S. Solomon et al., Eds., Cambridge University Press, 19–92.
- Taylor, K. E., 2001: Summarizing multiple aspects of model performance in a single diagram. *J. Geophys. Res.*, **106** (D7), 7183–7192.
- , R. J. Stouffer, and G. A. Meehl, cited 2011: A summary of the CMIP5 experiment design. 33 pp. [Available online at http://www.climar.org/organization/wgcm/references/Taylor_CMIP5.pdf.]
- Trenberth, K. E., and Coauthors, 2007: Observations: Surface and atmospheric climate change. *Climate Change 2007: The Physical Science Basis*, S. Solomon et al., Eds., Cambridge University Press, 235–336.
- Whetton, P. H., M. H. England, S. O'Farrell, I. G. Watterson, and A. B. Pittock, 1996: Global comparison of the regional rainfall results of enhanced greenhouse coupled and mixed layer ocean experiments: Implications for climate change scenario development. *Climatic Change*, **33**, 497–519.
- , I. Macadam, J. Bathols, and J. O'Grady, 2007: Assessment of the use of current climatic patterns to evaluate regional enhanced greenhouse response patterns of climate models. *Geophys. Res. Lett.*, **34**, L14701, doi:10.1029/2007GL030025.