THE POTENTIAL TO NARROW UNCERTAINTY IN REGIONAL CLIMATE PREDICTIONS

BY ED HAWKINS AND ROWAN SUTTON

Predictions of regional climate change for the next few decades are characterized by high uncertainty, but this uncertainty is potentially reducible through investments in climate science.

Faced with the realities of a changing climate, decision makers in a wide variety of organizations are increasingly seeking quantitative climate predictions. Specifically, they require predictions of the regional and local changes in climate that will impact people, economies, and ecosystems. Such predictions are available (e.g., Solomon et al. 2007) but are subject to considerable uncertainty. Thus, an important issue for these decision makers, and for organizations that fund climate research, is as follows: what is the scope for narrowing the uncertainty through future investments in climate science? Here, we address this question through analysis of twenty-first-century surface air temperature predictions (shown in Fig. 1) in the World Climate Research Programme's (WCRP's) Coupled Model Intercomparison Project phase 3 (CMIP3) multimodel dataset, as used in the Intergovernmental Panel on Climate Change (IPCC) Fourth Assessment Report (AR4; Solomon et al. 2007). This analysis is subject to some caveats, which we acknowledge and discuss.

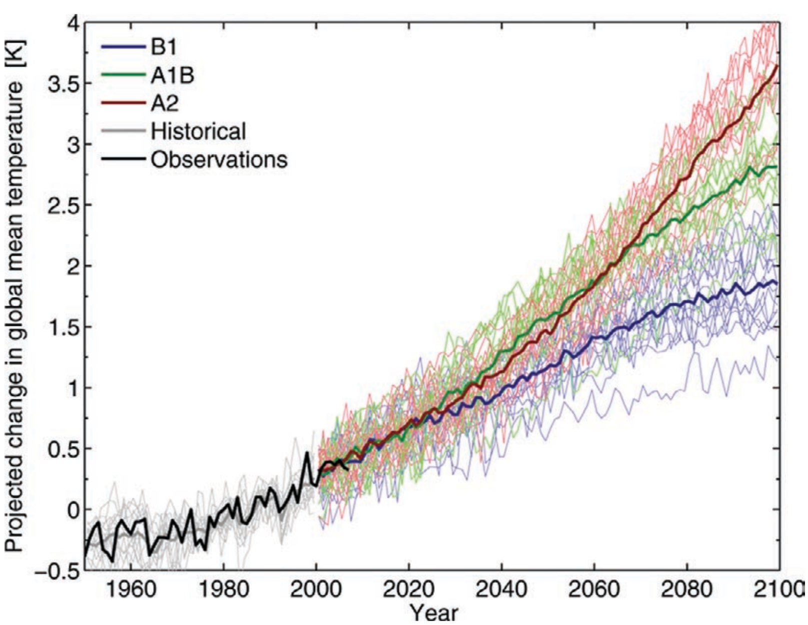


FIG. 1. Global mean, annual mean, surface air temperature predictions from 15 different global climate models under three different emission scenarios from 2000 to 2100 (thin lines): SRES A2 (red), A1B (green), and B1 (blue), designated as high-, medium-, and low-emissions paths, respectively. The same models forced with historical forcings are shown as the thin gray lines, and the observed global mean temperatures from 1950 to 2007 (Brohan et al. 2006) are shown as the thick black line. The multimodel mean for each emissions scenario is shown with thick colored lines demonstrating how uncertainty in future emissions gives rise to uncertainty in climate predictions. The different scenarios give nearly identical predictions until around 2025, demonstrating the delayed effect of future emissions. Each model has a different response to climate forcings, as seen by the spread in results for one particular scenario (or color). The internal (interannual) variability can be seen superimposed on the trend for any one individual prediction. All temperatures are shown as anomalies from the 1971–2000 mean.

PARTITIONING UNCERTAINTY. Uncertainty in climate predictions arises from three distinct sources. The first is the internal variability of the climate system, that is, the natural fluctuations that arise in the absence of any radiative forcing of the planet. Appreciation of these fluctuations is an important matter for decision makers because they have the potential to reverse—for a decade or so—the longer-term trends that are associated with anthropogenic climate change. The second is model uncertainty (also known as response uncertainty): in response to the same radiative forcing, different models simulate somewhat different changes in climate. The third is scenario uncertainty: uncertainty in future emissions of greenhouse gases, for example, causes uncertainty in future radiative forcing and hence climate. The method we use to separate these different sources of uncertainty, using 15 global climate models and three emissions scenarios, is described in appendix A.

The relative importance of the three sources of uncertainty varies with prediction lead time and with spatial and temporal averaging scale (Fig. 2; see also Räisänen 2001). The figure shows that for time horizons of many decades or longer, the dominant sources of uncertainty at regional or larger spatial scales are model uncertainty and scenario uncertainty. However, for time horizons of a decade or two, the dominant sources of uncertainty on regional scales are model uncertainty and internal variability. In general, the importance of internal variability increases at smaller spatial scales and shorter time scales.

There has been recent interest in estimating the fractional uncertainty (i.e., the prediction uncertainty divided by the expected mean change) in predictions of global mean temperature. Cox and

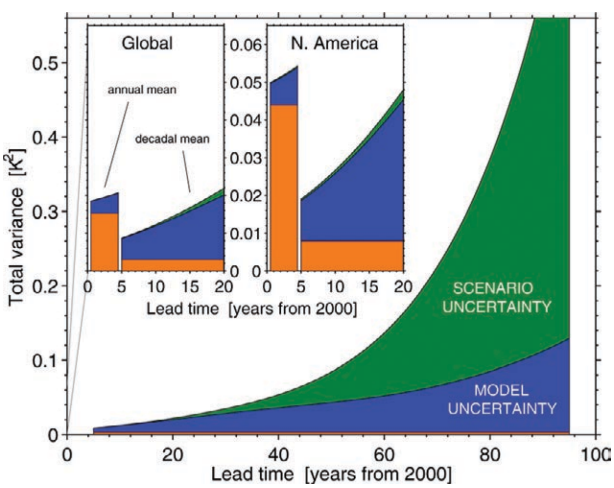


FIG. 2. The relative importance of the three sources of uncertainty changes significantly with region, forecast lead time, and the amount of any temporal meaning applied. Main panel: Total variance for the global-mean, decadal mean surface air temperature predictions, split into the three sources of uncertainty. Insets: As in the main panel, but only for lead times less than 20 yr for (left) the global mean and (right) a North American mean. The orange regions represent the internal variability component. For lead times shorter than 5 yr we plot the results using annual mean data to highlight how the internal variability component is vastly reduced when considering decadal mean data. The uncertainty in the regional prediction is larger than for a global mean.

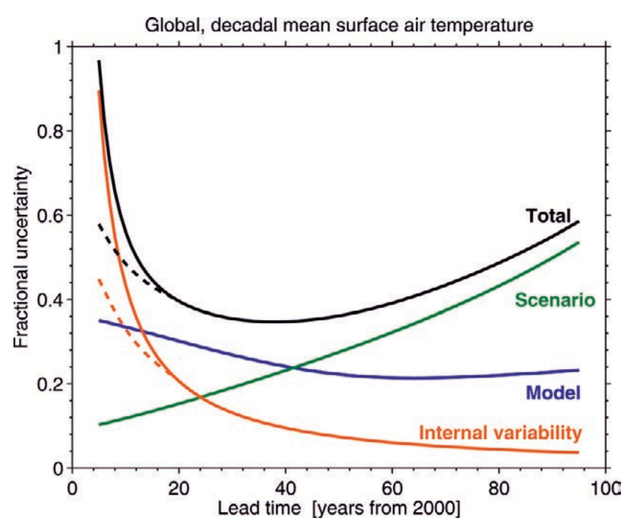


FIG. 3. The relative importance of each source of uncertainty in decadal mean surface air temperature predictions is shown by the fractional uncertainty (the 90% confidence level divided by the mean prediction), for the global mean, relative to the warming since the year 2000 (i.e., a lead of zero years). The dashed lines indicate reductions in internal variability, and hence total uncertainty, that may be possible through proper initialization of the predictions through assimilation of ocean observations (Smith et al. 2007).

AFFILIATIONS: Hawkins and Sutton—NCAS-Climate, Department of Meteorology, University of Reading, Reading, United Kingdom
CORRESPONDING AUTHOR: Rowan Sutton, Department of Meteorology, University of Reading, Reading RG6 6BB, UK
E-mail: r.sutton@reading.ac.uk

The abstract for this article can be found in this issue, following the table of contents.

DOI:10.1175/2009BAMS2607.1

In final form 17 February 2009

Stephenson (2007, hereafter CS07) used a simple climate model to estimate the three different contributions to fractional uncertainty. Knutti et al. (2008) used data from CMIP3 and from simpler climate models in a similar analysis but only quantified the model uncertainty component. Here, we have used the CMIP3 data to estimate the fractional uncertainty associated with all three contributions (Figs. 3, 4a), and extended the analysis to regional scales (Fig. 4b), which are of much greater relevance for adaptation planning. Our results for global mean temperature are consistent with those of Knutti et al. (2008). They also show important similarities to the findings of CS07, but there are also some crucial differences.

Following CS07, Figs. 3 and 4a both show how the contributions to fractional uncertainty vary as a function of prediction lead time. In Fig. 3 the

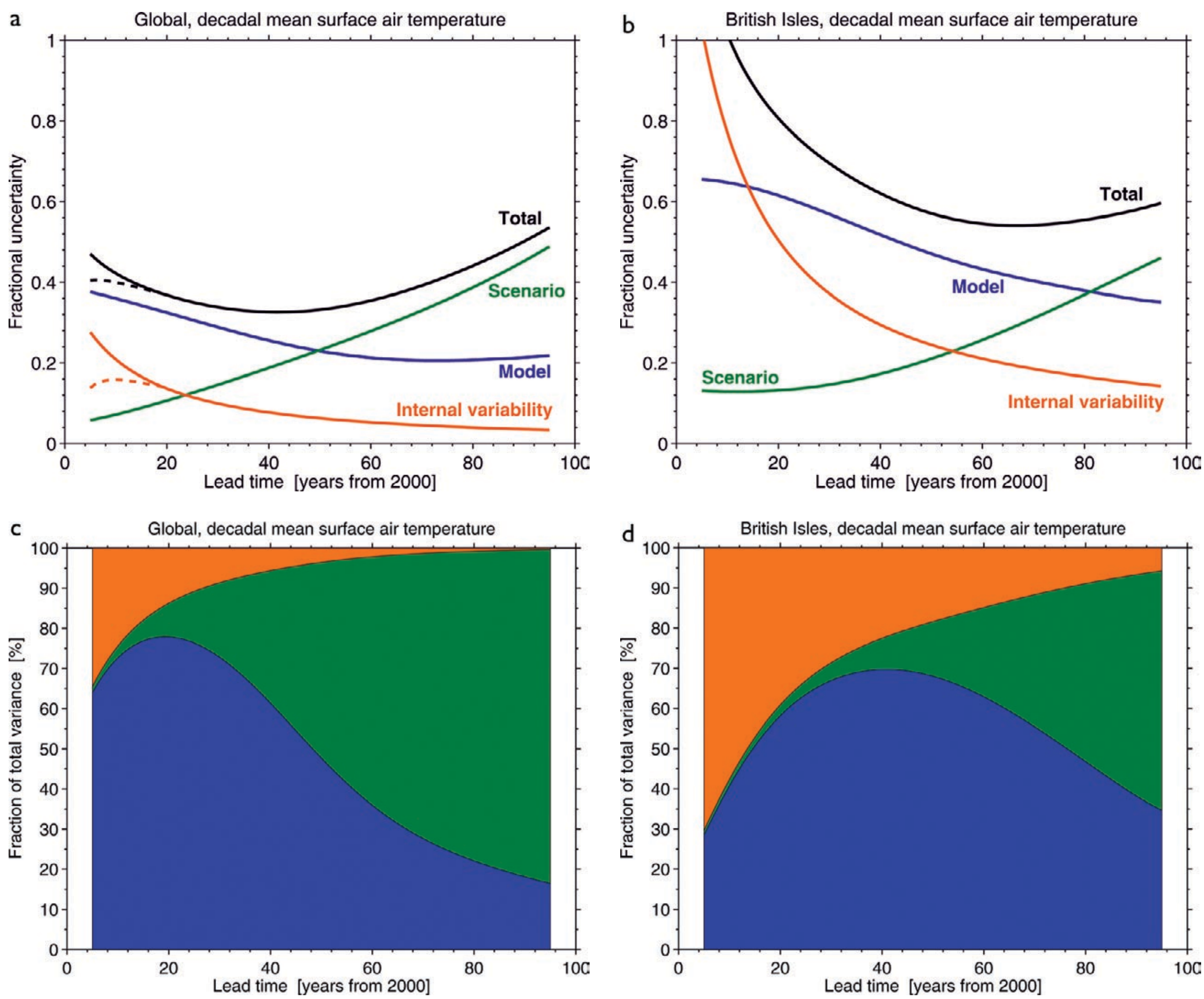


Fig. 4. The relative importance of each source of uncertainty in decadal mean surface temperature projections is shown by the fractional uncertainty (the 90% confidence level divided by the mean prediction) for (a) the global mean, relative to the warming from the 1971–2000 mean, and (b) the British Isles mean, relative to the warming from the 1971–2000 mean. The importance of model uncertainty is clearly visible for all policy-relevant timescales. Internal variability grows in importance for the smaller region. Scenario uncertainty only becomes important at multidecadal lead times. The dashed lines in (a) indicate reductions in internal variability, and hence total uncertainty, that may be possible through proper initialization of the predictions through assimilation of ocean observations (Smith et al. 2007). The fraction of total variance in decadal mean surface air temperature predictions explained by the three components of total uncertainty is shown for (c) a global mean and (d) a British Isles mean. Green regions represent scenario uncertainty, blue regions represent model uncertainty, and orange regions represent the internal variability component. As the size of the region is reduced, the relative importance of internal variability increases.

mean change is computed relative to the present day[1] (hence, it is zero at a lead time of zero); in Fig. 4a the warming is computed relative to a 1971–2000 reference period, assuming a subsequent (up to the year 2000) warming due to changing radiative forcing of 0.22 K (the multimodel mean). The major features of Fig. 3 are the following: 1) consistent with Fig. 2, the dominant contributions to prediction uncertainty are model uncertainty and scenario uncertainty; 2) the contribution from internal variability falls very rapidly with lead time as the signal of climate change strengthens, while the amplitude of internal variability remains constant; 3) the contribution from model uncertainty falls much more slowly with lead time because the spread between models increases with lead time; and 4) the total fractional uncertainty exhibits a minimum at a lead time of around 40 yr associated with the changing dominance of the contributions from model and scenario uncertainty. The major difference in Fig. 4a is that this minimum becomes less pronounced because the climate change signal is larger and the relative contribution from internal variability is smaller.

The figure shown by CS07 is most comparable to our Fig. 3. An important similarity between our results and those of CS07 is that in both cases the total fractional uncertainty exhibits a minimum at a lead time of 30–50 yr; CS07 note that this feature could be important for adaptation planning. However, in their analysis the location of this feature is determined by the decaying contribution of internal variability and is not significantly influenced by model uncertainty. Indeed, the most important difference between our results and theirs is that in ours the fractional contribution from internal variability is much lower. For example, at a lead time of 40 yr they show a value around 0.4, whereas our value is around 0.1. Uncertainties in estimating this quantity are discussed in appendix A. Further discussion of the comparison between our results and those of CS07 and Knutti et al. (2008) is provided in appendix B. Because planners are likely to be interested in comparisons to the recent past, rather than to a single year such as 2000, our view is that Fig. 4a provides the most appropriate (i.e., relevant to decision making) depiction of the variation with lead time of the various contributions to prediction uncertainty. Consequently, our remaining results are presented relative to a reference period of 1971–2000.

As we have emphasized, predictions of regional change are in many cases of greater relevance to decision making than predictions of global mean temperature. Figure 4b shows the fractional contributions to uncertainty in predictions of decadal mean temperature for the British Isles. As expected, on regional scales, the importance of internal variability is considerably enhanced. A minimum in the total fractional uncertainty is again present, in this case at a lead time of around 65 yr. Our results suggest that this feature is likely to be a robust property of *regional* temperature predictions, albeit that the lead time at which the minimum is found varies between regions (see Fig. 5b).

Figure 4 also shows the contributions to prediction variance plotted as a fraction of the total prediction variance at each lead time, for the global mean (Fig. 4c) and British Isles mean (Fig. 4d) decadal mean temperature. This representation again highlights the dominance of internal variability and model uncertainty at lead times up to a few decades, as well as the declining importance with increasing lead time of internal variability. The fact that scenario uncertainty makes a very small contribution for lead times less than about 30 yr was first shown by Stott and Kettleborough (2002) for predictions of global mean temperature. However, it is important to note that the same result holds for predictions of regional temperature, as illustrated by Fig. 4d. For longer lead times, the relative importance of model and scenario uncertainty varies between regions (see Fig. 6).

Figure 5a shows the signal-to-noise ratio[2] (S/N, the reciprocal of the total fractional uncertainty) for predictions of decadal mean temperature for the global mean and for various regions. All regions show a maximum in S/N (corresponding to a minimum in the fractional uncertainty) at some lead time, primarily as a consequence of the declining importance of internal variability and the increasing importance of scenario uncertainty, but the lead time at which the maximum is found varies between ~30 and ~80 yr. The position of this maximum tends to be at shorter lead times for larger regions because the internal variability component is smaller. For nearly all regions and lead times S/N significantly exceeds 1, indicating that these predictions should have great value for planning purposes. The value of S/N for the next few decades is particularly high for Africa (as has been noted in previous studies; e.g., Stott 2003).

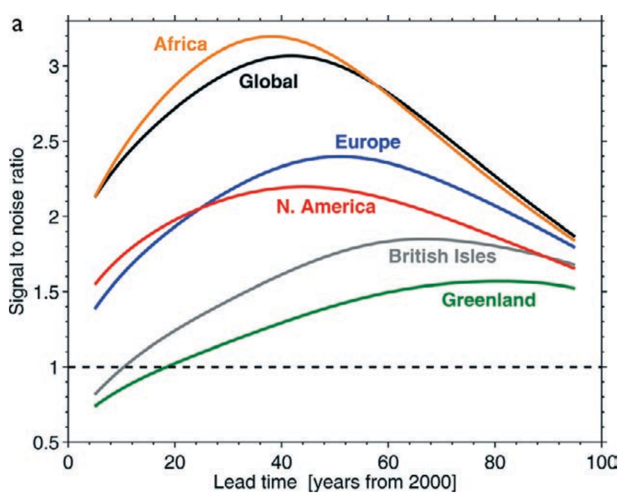
[1] The present day indicates the year 2000 (i.e., the starting year for the IPCC scenarios), and hence the first decade is the mean of 2000–09.

[2] The ratio is the climate change signal divided by the total uncertainty. The precise definition is given in appendix A.

For the mid- and high-latitude regions shown, S/N is lower and varies less with lead time. As would be expected, S/N for the next few decades is generally lower for smaller regions because of the increased importance of internal variability. Maps of S/N (e.g., Fig. 5b) show the highest values in the tropics and considerably lower values at mid-to high latitudes [i.e., predictions of temperature change have the lowest fractional uncertainty in the tropics; this point was illustrated in the IPCC Third Assessment Report (Houghton et al. 2001)]. To allow users to explore the sensitivity of our analyses to choice of region, lead time, and temporal filtering, we have created an interactive web site (see http://ncas-climate.nerc.ac.uk/research/uncertainty/).

THE POTENTIAL TO NARROW UNCERTAINTY. The potential for climate science to deliver reductions in total prediction uncertainty is associated entirely with the contributions from internal variability and model uncertainty. The contribution from internal variability is not reducible far ahead but, as recently reported by Smith et al. (2007), proper initialization of climate predictions with observational data should enable some reduction of this contribution for forecasts of the next decade or so (and can also contribute to reducing model uncertainty, as discussed later). This potential is illustrated by the dashed lines in Figs. 3 and 4a (details in appendix A). Such reductions could prove valuable to a wide range of decision makers. By contrast, there is potential to reduce the model contribution to prediction uncertainty for all lead times. New observations, advances in theory, improved modeling, and—importantly—new methods for bringing together models and observations (e.g., Stott and Forest 2007; Rodwell and Palmer 2007) can be expected to deliver significant progress in this area. These possibilities will improve as we gather more observations of the climate's response to increasing greenhouse gas concentrations (Stott and Kettleborough 2002).

To quantify the potential for reducing uncertainty in predictions of regional climate change, we consider the fraction of the total uncertainty in predictions of decadal mean surface air temperature that is attributable to internal variability, model uncertainty, and scenario uncertainty (Fig. 6). For predictions of the next decade, internal variability accounts for 40%–60% of the total uncertainty in most regions, with higher values in North Africa and northern Europe (Fig. 6, top left). Thus, a 20% reduction in this component would (if spatially uniform) lead to a reduction of typically 8%–12% in the total prediction uncertainty, with—apparently—the largest benefit for Europe. For predictions of the fourth decade ahead, model uncertainty is the dominant contribution over almost all of the globe and accounts for more than 70% of the total variance at high latitudes (Fig. 6,



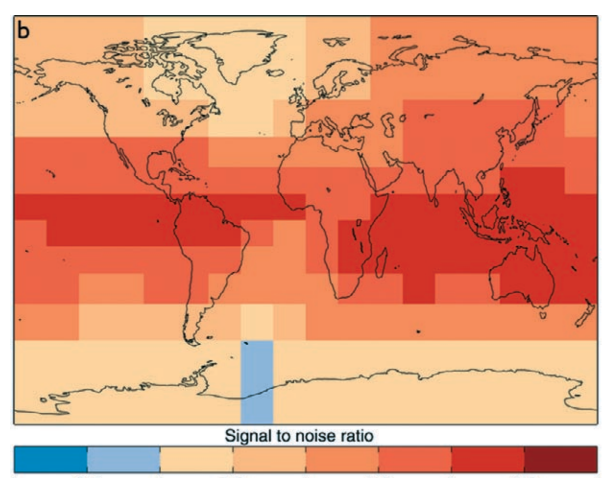


Fig. 5. (a) Signal-to-noise ratio for decadal mean surface air temperature predictions for different regions as labeled (90% confidence levels). The time of the highest S/N is when climate forecasts give most "added value," and this varies with the region as shown. Smaller regions generally have lower signal-to-noise ratios, but Africa does better than a global mean due to its location in the tropics where model uncertainty and internal variability are smaller than average. Greenland has a particularly low signal-to-noise ratio due to uncertainty in high-latitude climate feedbacks. (b) Maps of S/N indicate which regions have more confident predictions. This example shows this ratio for predictions of the fourth decade ahead (90% confidence levels). The tropical regions stand out as having high S/N, whereas Atlantic longitudes have reduced S/N values, perhaps due to uncertainty in the response of the Atlantic Ocean thermohaline circulation to radiative forcings.

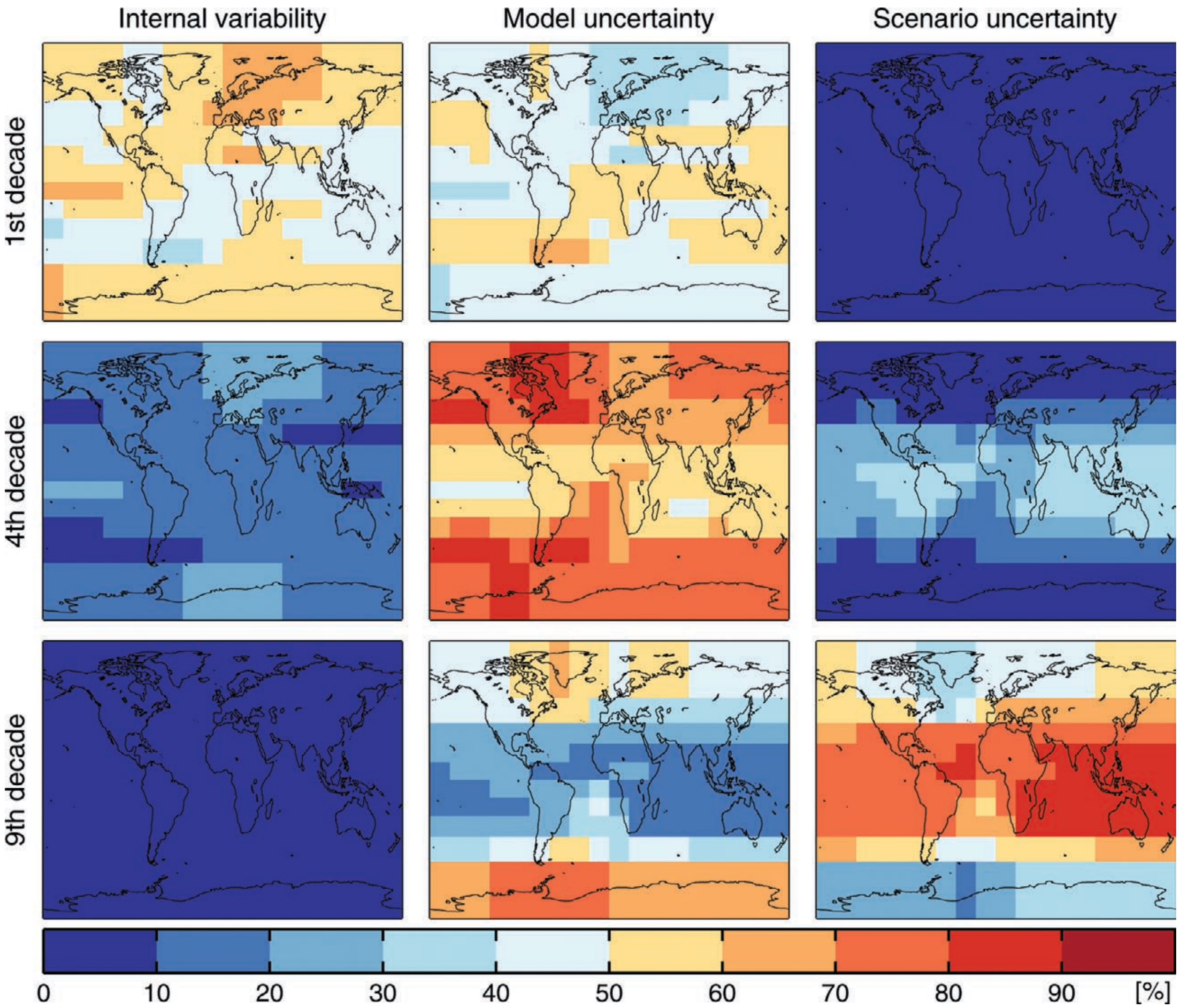


Fig. 6. Maps of the sources of uncertainty for decadal mean surface temperature for various lead times give information on where any reduction in uncertainty will have the most benefit. The columns show the total variance explained by (left) internal variability, (middle) model uncertainty, and (right) scenario uncertainty for predictions of the (top) first, (middle) fourth, and (bottom) ninth decade. It should be noted that (i) even on regional scales, the uncertainty due to internal variability is only a significant component for lead times up to a decade or two, (ii) the largest differences between models occur at high latitudes where climate feedbacks are particularly important, and (iii) even by the end of the century, the emissions scenario is less important than model uncertainty for the high latitudes but dominates in the tropics.

middle). The variation with latitude reflects the fact that model uncertainty has a clear maximum at high latitudes (likely to be a consequence of differences in the representation of the climate feedbacks that lead to high-latitude amplification of the climate change signal), whereas the spatial variation of scenario uncertainty is more complex (not shown, but it can be seen on the previously cited web site). A 20% reduction in model uncertainty would, if spatially uniform, lead to a reduction of significantly more than 10% in total prediction uncertainty in many regions. For predictions of the ninth decade ahead, scenario uncertainty, as expected, is the dominant contribution over much of the globe (Fig. 6, bottom right), but at high latitudes it is still model uncertainty that accounts for the largest fraction of variance.

A key result from Fig. 6 is that for predictions of the next few decades, model uncertainty and internal variability are the dominant contributions. An important question is how robust this finding is likely to be for other variables (e.g., changes in precipitation or in the statistics of high-impact events). For any given variable (and spatial and temporal scale), the relative importance of internal variability and model uncertainty will differ, but we see no obvious reason to suppose—at least for predictions of the next few decades—that scenario uncertainty should be more important than has been found for decadal mean temperature.[3] Therefore, we argue that it is likely that *the uncertainty in regional climate predictions for the next few decades is dominated by sources (model uncertainty and internal variability) that are potentially reducible through progress in climate science*. This inference has important implications for managing adaptation to a changing climate, which we discuss in the final section of this paper.

DISCUSSION. Our analysis and arguments are potentially subject to some criticisms, which we now discuss. One potential area of criticism concerns the methods we have used to quantify and partition uncertainty. We acknowledge that the methods we have used (detailed in appendix A) are simple and rely heavily on the IPCC "ensemble of opportunity." The estimate of internal variability is based entirely on data from the climate models, which exhibit a large range in this quantity, although the multimodel mean is in reasonable agreement with an observational estimate (see appendix A). The estimate of model uncertainty incorporates only a very simple observational constraint and neglects many processes, notably carbon cycle feedbacks, that are likely to be important at longer lead times (Knutti et al. 2008). The IPCC AR4 projections do not appear to sample the full range of model uncertainty that is consistent with the observed record (Murphy et al. 2004; Stainforth et al. 2005; Stott and Forest 2007). The estimate of scenario uncertainty is based on only three IPCC scenarios, and future changes in land use, which may be significant on regional scales, are neglected (Feddema et al. 2005). Given these limitations, there is no doubt that our estimates could be improved. The key point, however, is that although the results shown in this paper would change quantitatively, we see no reason to expect them to change qualitatively. In particular, uncertainty in predictions of regional climate change for the next few decades will remain dominated by internal variability and model uncertainty. We accept that for longer lead-time predictions, assessments of the relative importance of model uncertainty and scenario uncertainty may change as a result of including new processes in climate models (e.g., better representation of biogeochemical and ice-sheet feedbacks) and improved understanding of scenarios.

A second potential criticism addresses our arguments concerning the potential to narrow uncertainty. It might be argued that there are fundamental limits that will restrict this potential severely. In the case of internal variability, this is undoubtedly the case: although initialization of climate predictions should permit some reduction of the contribution from internal variability to prediction uncertainty at short lead times and large spatial scales, at longer lead times and smaller spatial scales any such reduction is precluded by chaos. In the case of model uncertainty, it is well known that the uncertainty in equilibrium climate sensitivity has reduced little since the first IPCC assessment report in 1990 (Räisänen 2007), and Roe and Baker (2007) recently argued that there are fundamental reasons why so little progress has been made. However, from the practical perspective of predicting climate over the next few decades, it is transient rather than equilibrium climate change that is of the greatest importance, and Allen and Frame (2007) have argued that the limits on narrowing uncertainty in predictions of transient change may be less serious. Furthermore, narrowing uncertainty in regional climate predictions is not just about climate sensitivity. A much wider range of processes is relevant, and improving the representation of these processes in models is both a major challenge and a real opportunity. Therefore, our response to the proposition that there may be limits to the potential to narrow uncertainty is that such limits may exist, but if so there is no evidence that these limits have yet been reached. Thus, the potential to narrow uncertainty is real, and the need is urgent. An additional point is that acknowledging that progress in climate science may sometimes broaden rather than narrow uncertainty (e.g., as in understanding of carbon cycle feedbacks) is not a reason to avoid focused attention on the requirement and opportunities to narrow uncertainty. Indeed, this requirement should be a major focus of climate science because of its importance to society.

A final important point to emphasize is that the discussion of prediction uncertainty in this study is based on the variance of model predictions ("spread") rather than the variance of prediction errors ("skill," i.e., the difference between predictions and observations). Research in weather forecasting has shown

[3] The one exception is the contribution to scenario uncertainty associated with anthropogenic emissions of aerosol precursors. Aerosol forcing is characterized by a short atmospheric lifetime and large spatial variations. Detailed consideration of this contribution is beyond the scope of this study.

that the relationship between spread and skill is rarely simple. As yet there has been very little published work on the relationship between spread and skill for climate predictions, and this is clearly an area that requires greater attention.

CONCLUSIONS AND IMPLICATIONS. Using data from a suite of climate models, we have carried out a quantitative assessment of the contributions to uncertainty in predictions of regional temperature change. We have estimated the contributions to the total prediction uncertainty from internal variability, model uncertainty, and scenario uncertainty. For lead times of the next few decades, the dominant contributions are internal variability and model uncertainty. It is well known that the importance of internal variability increases at shorter time and space scales, but our analyses suggest that for decadal time scales and regional spatial scales (~2,000 km), model uncertainty is of greater importance than internal variability.

Another important finding from our analyses is that although the total uncertainty increases with lead time, the signal-to-noise ratio for predictions of decadal mean temperature exceeds 1 for almost all regions and lead times and typically shows a maximum at a lead time of some decades. Knowledge of this maximum, which corresponds to a minimum in the fractional uncertainty of the prediction, may by useful for some planning purposes (as noted by CS07). Note, however, that the signal-to-noise ratio is likely to be lower for almost any other climate variable.

The contributions to prediction uncertainty from internal variability and especially from model uncertainty are potentially reducible through progress in climate science. As noted earlier, this conclusion has important implications for managing adaptation to a changing climate. The key point is that greater uncertainty about future climate is likely to be associated with more expensive adaptation: for example, for a given adaptation (e.g., building a sea wall), greater uncertainty, particularly at the high end, implies a need to include tolerance for more extreme events. Because the costs of adaptation are expected to be very large,[4] the clear implication is that *reducing uncertainty in climate predictions is potentially of enormous economic value.* Furthermore, this recognition invites a comparison between a) the cost of various degrees of adaptation, given current levels of uncertainty, and b) the cost of new investments in climate science to reduce current levels of uncertainty. Clearly there is a need for a great deal of further work to assess more fully these costs and benefits, both for specific applications and more generally. In our view, this work should be a high priority. In the face of changing patterns of risk, quantifying the economic value of progress in climate predictions is an urgent issue for society and for scientists.

Finally, our work highlights the importance of targeting climate science investments on the most promising opportunities to reduce prediction uncertainty. In this context, we would highlight the importance of developing predictions that are initialized with observations of the current climate state (e.g., Smith et al. 2007; Meehl et al. 2009). The importance of this approach has three major dimensions. First, as already discussed, it can contribute to reducing the internal variability component of prediction uncertainty. Second, by enabling direct comparison between a climate prediction and specific observations, it provides a powerful new strategy to identify model errors and thereby reduce model-related uncertainty. Third, it is the natural approach to address the key issue that was highlighted at the end of the discussion: understanding the relationship between spread and skill.

ACKNOWLEDGMENTS. We thank Peter Stott, Jonathan Gregory, and Jason Lowe for valuable discussions and for providing data. We also thank Peter Cox and three anonymous reviewers for their thoughtful comments, which helped improve the paper. EH is funded by the U.K. Natural Environment Research Council under the thematic Rapid Climate Change programme (RAPID). RS is supported by a Royal Society University Research Fellowship. We acknowledge the modeling groups, the Program for Climate Model Diagnosis and Intercomparison, and the WCRP's Working Group on Coupled Modelling for their roles in making available the WCRP CMIP3 multimodel dataset. Support of this dataset is provided by the Office of Science, U.S. Department of Energy.

APPENDIX A: DATA AND METHODS. We use the global mean, annual mean, surface air temperature predictions for the twentieth and twenty-first century (Fig. 1), from 15 different global climate models under historical forcings and three different future Intergovernmental Panel on Climate Change

[4] The Stern Review (Stern 2006) estimated the cost of doing nothing at 5%–10% of global Gross Domestic Product (GDP) by 2100, and the costs of mitigation at 1% of global GDP by 2050. The costs of adaptation are highly uncertain, but if they are as little as 10% of the costs of mitigation, they would still be ~0.1% of global GDP by 2050.

Special Report Emissions Scenarios (SRES A1B, A2, and B1, giving a total of 45 predictions). These future scenarios are summarized in the latest IPCC report (Solomon et al. 2007). Although some of the models have several realizations of these simulations, we use just one ensemble member for each model to treat all of the models equally. The particular models used in the analysis were chosen purely on the basis of data availability (i.e., we only included the models for which all three scenario simulations were available).

The method used for separating the three components of uncertainty was as follows:

- Each individual prediction was fit, using ordinary least squares, with a fourth-order polynomial over the years 1950–2099. The raw predictions X for each model m, scenario s and year t can be written as

$$X_{m,s,t} = x_{m,s,t} + i_{m,s} + \varepsilon_{m,s,t}, \tag{1}$$

  where the reference temperature is denoted by i, the smooth fit is represented by x, and the residual (internal variability) is ε. The reference temperatures used were the year 2000 (Fig. 3) and the mean of the years 1971 to 2000 (for all other analyses), both of which were estimated from the smooth fits.
- We weight the models by their ability to simulate the global mean warming from the mean of 1971–2000, up to the year 2000. Thus, each model is given a weight,

$$w_m = \frac{1}{x_{obs} + \left|x_{m,2000} - x_{obs}\right|}, \tag{2}$$

  where $x_{m,2000}$ is the model global mean warming at the year 2000, relative to 1971–2000, and x_{obs} = 0.25 K is an observational estimate derived from fitting a similar polynomial to the observations. These weightings can also be expressed as normalized quantities:

$$W_m = \frac{w_m}{\sum_m w_m}. \tag{3}$$

  This scheme thus downweights those models that have warmed by too large or too small an amount. More complex weighting schemes exist (e.g., Giorgi and Mearns 2002), but we choose to keep the methodology simple, especially because the weighting does not affect the results greatly.
- The internal variability for each model was defined as the variance of the residuals from the fits, estimated independently of scenario and lead time. The multimodel mean of these variances is taken to be the internal variability component,

$$V = \sum_m W_m \operatorname{var}_{s,t}(\varepsilon_{m,s,t}), \tag{4}$$

  where $\operatorname{var}_{s,t}$ denotes the variance across scenarios and time and V is constant in time. Future changes in internal variability are likely (Solomon et al. 2007), and decision makers also need such information. In this study of surface air temperatures, we assume these changes are negligible, although we note that there is a small downward trend in internal variability that merits future investigation. The individual models have considerable differences in this quantity for the global mean surface air temperature, with a mean $\sqrt{V} \approx 0.12$ K (range 0.06–0.20 K) for the twenty-first-century predictions and a mean $\sqrt{V} \approx 0.13$ K (range 0.08 to 0.21 K) for data from 1950–2007, which includes volcanic forcings. Comparison with the historical record is difficult because of incomplete observations and knowledge of climate forcings in the twentieth century, but a similar polynomial fit to the observed global mean surface air temperature record provides an interannual estimate of $\sqrt{V_{\text{obs}}} \sim 0.10$ K (using years 1950 to 2007). An estimate of the internal variability on decadal time scales is very uncertain, but the models ($\sqrt{V_{\text{dec}}} \sim 0.05$ K) may overestimate natural ($\sqrt{V_{\text{obs,dec}}} \sim 0.02$ K) decadal fluctuations (also see Solomon et al. 2007). However, the observational variance may be underestimated by fitting a polynomial to such a short observed time series.
- The model uncertainty for each scenario is estimated from the weighted variance (var^W) in the different model prediction fits. The multiscenario mean is taken as an estimate of the model uncertainty component,

$$M(t) = \frac{1}{N_s}\sum_s \operatorname{var}_m^W(x_{m,s,t}), \tag{5}$$

  where N_s is the number of scenarios. When the reference temperature is the year 2000, this quantity is zero for a lead time of zero years.
- The scenario uncertainty is simply the variance of the weighted multimodel means for the three scenarios:

$$S(t) = \operatorname{var}_s\left(\sum_m W_m x_{m,s,t}\right). \tag{6}$$

- We assume that there are no interactions between the three sources of uncertainty (i.e., they can be treated independently); thus, the total variance is then

$$T(t) = V + S(t) + M(t), \tag{7}$$

and the mean change of all the predictions,[A1] above the reference temperature, is

$$G(t) = \frac{1}{N_s} \sum_{m,s} W_m x_{m,s,t}. \tag{8}$$

- The fractional uncertainty (90% confidence level) shown in Figs. 3 and 4 is then

$$F(t) = \frac{1.65\sqrt{T(t)}}{G(t)}. \tag{9}$$

The signal-to-noise ratio shown in Fig. 5 is the inverse of this quantity.

For different temporal means of the predictions (e.g., the decadal means generally considered in the text), $\varepsilon_{m,s,t}$ is smoothed before the variance is calculated in Eq. (4). This procedure can also be repeated for any region of interest. In the maps derived from model data in Figs. 5–7 we consider 180 regions of equal area (20° in longitude and 10 latitude bands of varying size from ~11° to 37°). This methodology could also be extended to any climate variable of interest, although it is important to note that intermodel agreement is smaller and the relative importance of internal variability is larger for precipitation and sea level pressure than for surface temperature (Solomon et al. 2007; Räisänen 2007).

The reliability of these uncertainty estimates needs to be considered carefully. As further discussed in the main text, it is likely that each estimate should be considered as a lower limit on the uncertainty. For the case of the internal variability we use a subset of the models that have more than one ensemble member and fit a smooth polynomial to each ensemble member in turn. By comparing the derived internal variability it is found that for individual models, the internal variability estimates have a spread of ~±10% for decadal means. Assuming that a large part of this spread is random across models, this spread would become insignificant when averaged over the 15 models considered. As a further check, we also examined the internal variability in the preindustrial control simulations of the models considered. It was found that the mean internal variability of global mean surface air temperature is $\sqrt{V} \approx 0.12$ K—the same as that found from the residuals from our polynomial fit. These tests give us confidence that our estimates are robust.

The potential reductions in the internal variability component through initialization of forecasts shown with dashed lines in Figs. 3 and 4a are derived from Smith et al. (2007). Using the Met Office's Decadal Prediction System (DePreSys), Smith et al. (2007) showed that an initialized forecast of decadal mean, global mean surface air temperature reduced forecast error variance by 50% in the first decade when compared to uninitialized projections. We assume here that this reduction in variance shrinks to zero at a lead time of 20 yr; it is an illustrative estimate from just one model.

Using this methodology, we can also consider how the models have performed over the recent past. The multimodel mean temperature response in the year 2000, relative to the mean of 1971–2000, is shown in Fig. A1a, with the spread between models shown in Fig. A1b. A similar polynomial fit to the observations from 1950 to 2007 allows an estimate of the warming at the year 2000 above the 1971–2000 mean to be made (Figs. 7c,d). Here we have used the HadCRUT3v dataset (Brohan et al. 2006) of sea surface temperatures and surface air temperatures and have regridded to a similar scale as the models (Fig. 7c) and just used land-based data for comparison (Fig. 7d, not regridded). Because of the nonuniform nature of the observations, we have required an observation in every month in at least 80% of the years since 1950, and for the regridding, observations for at least three subgrid cells. The multimodel mean appears to underestimate the warming at high northern latitudes, but with a large intermodel spread, possibly due to different albedo feedbacks. The lack of observations in the Arctic makes detecting any statistically significant differences between the observations and the models difficult (Gillett et al. 2008). The multimodel mean is in reasonable agreement with observations over tropical regions, but the observations are naturally more noisy as they are just a single realization of the climate system.

APPENDIX B: COMPARISON WITH PREVIOUS STUDIES. Although our emphasis is on estimating the dominant sources of uncertainty on regional scales, we can compare our results for the global mean with previous studies that also estimated

[A1] Note that it should not be assumed that the ensemble mean is the best prediction, or even an unbiased prediction.

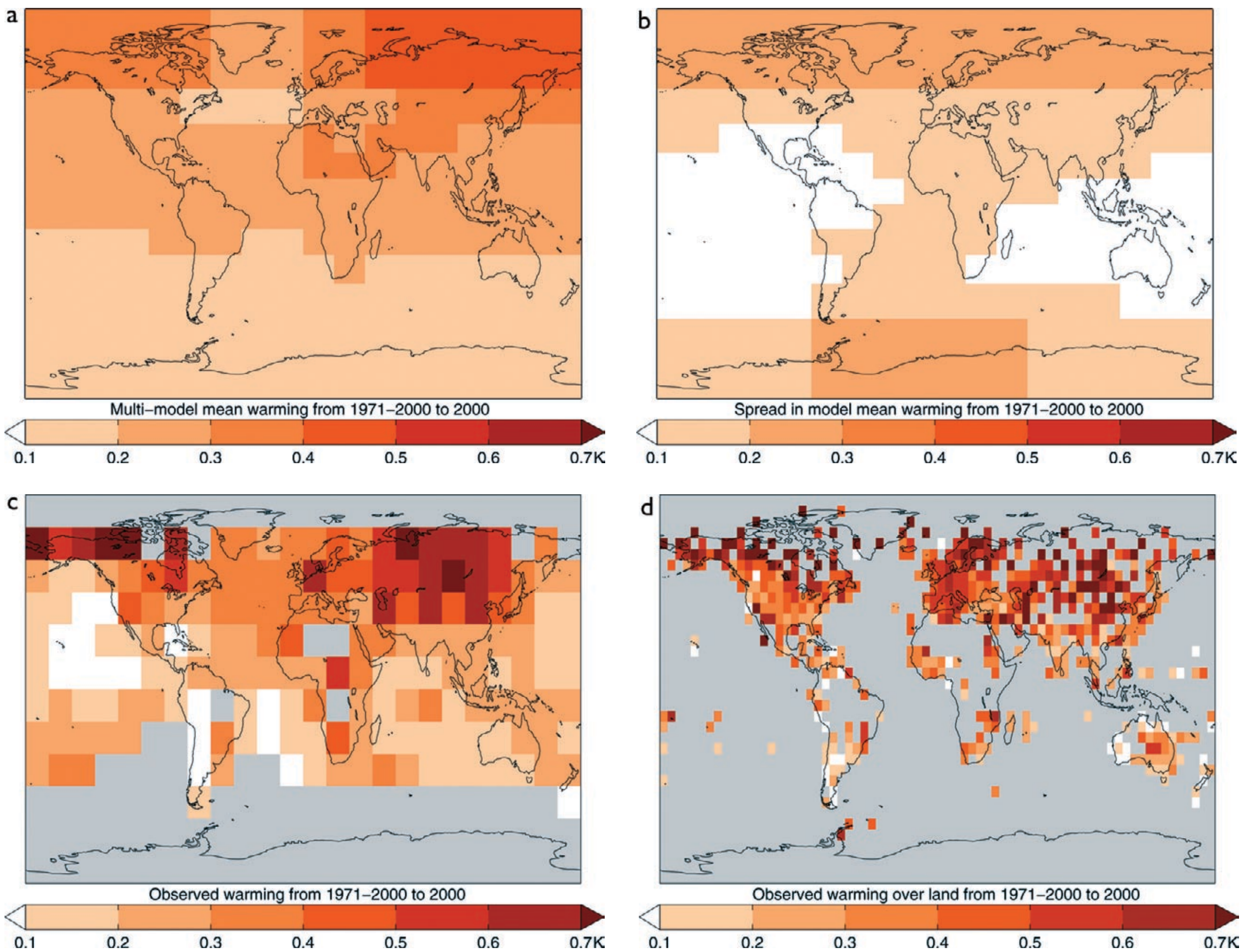


Fig. A1. (a) The multimodel mean surface air temperature warming at the year 2000, relative to the mean of 1971–2000. The global model average is 0.22 ± 0.09 K (90% confidence level), but the warming is spatially nonuniform, with enhanced warming in the Arctic due to climate feedbacks from ice melt. An estimate from observations gives a global average warming of around 0.25 K for the same period. (b) The spread (half of the 5%–95% uncertainty range) in model projections, showing large model differences in the high latitudes, especially the Arctic. (c) Estimated observed warming of surface air temperature over land and sea surface temperature over the ocean from the HadCRUT3v dataset (Brohan et al. 2006), regridded onto a 15° × 15° grid. (d) Just land-based data, without regridding (a 5° × 5° grid). Gray regions indicate insufficient data. The multimodel mean is naturally smoother than the observed estimates, which are based on just one realization of the climate system, but the models do seem to underestimate the climate feedbacks at high northern latitudes, although this is probably not significant given the spread in model projections for the Arctic.

the fractional uncertainty of global mean, decadal mean, surface air temperature predictions.

CS07 used a very simple model of the climate system,

$$\Delta F = \Delta Q + \lambda \Delta T, \tag{10}$$

where F is the radiative forcing, λ is the climate sensitivity, T is the temperature response, and Q is the ocean heat uptake, to estimate the fractional uncertainty. The model uncertainty was derived by changing λ (i.e., by exploring parameter uncertainty rather than the structural uncertainty we consider), and the scenario uncertainty by changing F.

Their fractional uncertainty was estimated relative to the year 2000, allowing a direct comparison with our Fig. 3. As discussed in the main text, our results show some similarities to and some differences from their findings. The most important difference concerns the relative importance of internal variability and model uncertainty. A further difference between our results is that in CS07 the fractional contribution from model uncertainty is zero at a lead time of zero and initially increases with lead time, whereas our results show no such increase. In fact, the behavior of the fractional model uncertainty as the lead time t approaches zero is by no means obvious. It is necessary to consider uninitialized predictions and initialized predictions separately.

In the case of uninitialized predictions relative to the year 2000, at $t = 0$ (i.e., the year 2000) the fractional uncertainty (uncertainty/signal) is formally not defined because the signal is zero. If one argues that the signal should be computed relative to some earlier reference climate period (e.g., 1971–2000), then the model uncertainty should also be computed relative to the same reference climate. In this case the model uncertainty is equivalent to the uncertainty in the warming attributable to changes in radiative forcing (Hegerl et al. 2007) and is certainly nonzero. This is the approach that we took in calculating Fig. 4a.

In the case of initialized predictions, both the model uncertainty and the signal are zero at $t = 0$, so the fractional uncertainty is again not defined. Whether the fractional uncertainty initially increases or initially decreases depends on which term, numerator or denominator, grows most rapidly. To assess the rate of growth of model uncertainty, one would need large initial condition ensembles for a set of models all initialized (to within observational uncertainty) from the same initial climate state. We are not aware of any research on this question, but a multimodel intercomparison is planned to provide input to the next IPCC report (Meehl et al. 2009). In view of this lack of knowledge, where we attempt schematically to illustrate the potential impact of initialization (Figs. 3 and 4a), we have not attempted to show any impact on fractional model uncertainty, and we do not show results for lead times less than 5 yr.

Knutti et al. (2008) explored different methods to estimate the fractional model uncertainty, and their results for the CMIP3 models are similar to those presented here (compare to their Fig. 3). They also find that the model uncertainty component decreases with lead time, with a fractional uncertainty of around 0.4 at a lead time of 20 yr, decreasing to around 0.25 at a lead time of 100 yr.

REFERENCES

Allen, M. R., and D. J. Frame, 2007: Call off the quest. *Science,* **318,** 582–583, doi:10.1126/science.

Brohan, P., J. J. Kennedy, I. Harris, S. F. B. Tett, and P. Jones, 2006: Uncertainty estimates in regional and global observed temperature changes: A new dataset from 1850. *J. Geophys. Res.,* **111,** D12106, doi:10.1029/2005JD006548.

Cox, P., and D. Stephenson, 2007: A changing climate for prediction. *Science,* **317,** 207–208, doi:10.1126/science.1145956.

Feddema, J. J., K. W. Oleson, G. B. Bonan, L. O. Mearns, L. E. Buia, G. A. Meehl, and W. M. Washington, 2005: The importance of land-cover change in simulating future climates. *Science,* **310,** 1674–1678, doi:10.1126/science.1118160.

Gillett, N. P., D. A. Stone, P. A. Stott, T. Nozawa, A. Y. Karpechko, G. C. Hegerl, M. F. Wehner, and P. D. Jones, 2008: Attribution of polar warming to human influence. *Nature Geosci.,* **1,** 750–754, doi:10.1038/ngeo338.

Giorgi, F., and L. O. Mearns, 2002: Calculation of average, uncertainty range, and reliability of regional climate changes from AOGCM simulations via the Reliability Ensemble Averaging (REA) method. *J. Climate,* **15,** 1141–1158.

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.

Houghton, J. T., Y. Ding, D. J. Griggs, M. Noguer, P. J. van der Linden, X. Dai, K. Maskell, and C. A. Johnson, Eds., 2001: *Climate Change 2001: The Physical Science Basis*. Cambridge University Press, 881 pp.

Knutti, R., and Coauthors, 2008: A review of uncertainties in global temperature projections over the twenty-first century. *J. Climate,* **21,** 2651–2663.

Meehl, G. A., and Coauthors, 2009: Decadal prediction: Can it be skillful? *Bull. Amer. Meteor. Soc.,* in press.

Murphy, J. M., D. M. H. Sexton, D. N. Barnett, G. S. Jones, M. J. Webb, M. Collins, and D. A. Stainforth, 2004: Quantification of modeling uncertainties in a large ensemble of climate change simulations. *Nature,* **430,** 768–772, doi:10.1038/nature02771.

Räisänen, J., 2001: CO_2-induced climate change in CMIP2 experiments: Quantification of agreement and role of internal variability. *J. Climate,* **14,** 2088–2104.

——, 2007: How reliable are climate models? *Tellus,* **59A,** 2–29.

Rodwell, M. J., and T. N. Palmer, 2007: Using numerical weather prediction to assess climate models. *Quart. J. Roy. Meteor. Soc.,* **133,** 129–146.

Roe, G. H. and M. B. Baker, 2007: Why is climate sensitivity so unpredictable? *Science,* **318,** 629–632, doi:10.1126/science.1144735.

Smith, D. M., S. Cusack, A. W. Colman, C. K. Folland, G. R. Harris, and J. M. Murphy, 2007: Improved surface temperature prediction for the coming decade from a global climate model. *Science,* **317,** 796–799, doi:10.1126/science.1139540.

Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K. Averyt, M. M. B. Tignor, and H. L. Miller, Eds., 2007: *Climate Change 2007: The Physical Science Basis*. Cambridge University Press, 996 pp.

Stainforth, D. A., and Coauthors, 2005: Uncertainty in the predictions of the climate response to rising levels of greenhouse gases. *Nature,* **433,** 403–406, doi:10.1038/nature03301.

Stern, N., 2006: *The Economics of Climate Change: The Stern Review*. Cambridge University Press, 712 pp.

Stott, P. A., 2003: Attribution of regional-scale temperature changes to anthropogenic and natural causes. *Geophys. Res. Lett.,* **30,** 1728–1731, doi:10.1029/2003GL017324.

——, and J. A. Kettleborough, 2002: Origins and estimates of uncertainty in predictions of twenty-first century temperature rise. *Nature,* **416,** 723–726, doi:10.1038/416723a.

——, and C. E. Forest, 2007: Ensemble climate predictions using climate models and observational constraints. *Philos. Trans. Roy. Soc. London,* **365A,** 2029–2052, doi:10.1098/rsta.2007.2075.