Geophysica (2005), 41(1-2), 19-30

Probability Distributions of CO_2-induced Global Warming as Inferred Directly from Multimodel Ensemble Simulations

Jouni Räisänen

Department of Physical Sciences, Division of Atmospheric Sciences
P.O. Box 64, FIN-00014 University of Helsinki, Finland

(Received: May 2005; Accepted: August 2005)

Abstract

Analytical, purely model-based probability distributions are derived for the instantaneous global mean warming resulting from a gradual doubling of CO_2 (TCR = transient climate response) and for the equilibrium global mean warming caused by a doubling of CO_2 (CS = climate sensitivity). For TCR, the estimated 5-95% uncertainty range based on the results of 20 models is 1.0-2.4 °C when assuming a normal, and 1.1-2.5 °C when assuming a lognormal form of the distribution. The corresponding numbers for CS, based on 15 models, are 2.0-5.0 °C and 2.1-5.3 °C. The limited sample size makes it difficult to estimate the form of the distributions reliably. For TCR, however, the lognormal distribution fits the data better than the normal distribution, although this conclusion is critically dependent on one extreme model. The parameters that define the location (mean or median) and width (standard deviation) of the underlying distribution are also potentially sensitive to sampling variability. For estimating the 5-95% uncertainty range of warming, this aspect of sampling uncertainty dominates over the differences between the normal and the lognormal distributions. The derived probability distribution of CS is generally consistent with estimates based on other methods, although some recent studies have placed the upper bound of the uncertainty range substantially higher than that found in the present analysis.

Key words: climate change, global warming, probability distribution

1. Introduction

In assessment of anthropogenic climate change, the single most widely used parameter is the change in global mean surface air temperature. *Cubasch et al*. (2001) estimated a global warming of 1.4-5.8 °C from 1990 to 2100, taking into account differences between seven climate models and uncertainty about future greenhouse gas and aerosol emissions. To characterize that part of the uncertainty that is directly related to models (rather than to emissions), two numbers are commonly used: climate sensitivity (CS) and transient climate response (TCR). These are both determined from idealized model experiments in which the CO_2 concentration is doubled. CS is the equilibrium global mean warming resulting from a doubling of CO_2. If the CO_2 concentration in the real world would double, and then stabilize at this level, the global mean temperature would approach its new equilibrium in the course of several centuries. TCR measures the warming that would already be realized by the time of the doubling of CO_2, assuming a compound 1% per year increase in CO_2 that leads to doubling in 70 years. Thus, although these two measures of climate change are both idealized, TCR is likely to be the more practically relevant of them in the short run but CS in the long run.

Cubasch et al. (2001) reported CS for 15 and TCR for 19 climate models. Among these models, CS varied from 2.0 °C to 5.1 °C and TCR from 1.1 °C to 3.1 °C. However, the simulated CS and TCR values have apparently not yet been converted to objective, analytical probability distributions. *Wigley and Raper* (2001) derived probability distributions for the global warming from 1990 to 2100, but using in their calculations somewhat arbitrarily specified rather than directly model-based distributions for CS.

In this study, we derive analytical probability distributions for CS and TCR by using the distribution of CO_2-induced global warming in different climate models. We also discuss the question how well these probability distributions can be estimated using a relatively limited sample of model results, as well as the sensitivity of the conclusions to the assumed form of the distributions.

The present study is basically a curve-fitting exercise. As such, it has a very important caveat: there is no *a priori* guarantee that CS and TCR in the real world are within the probability distributions suggested by models. In particular for CS, there have been several attempts to estimate this parameter with more elaborate methods. To put the present results in a perspective, some of these earlier studies are also briefly discussed in this paper.

2. *Data set*

The TCR is evaluated for 20 coupled atmosphere-ocean general circulation models (GCMs) participating in CMIP2, the second phase of the Coupled Model Intercomparison Project (*Meehl et al*., 2000). An 80-year control run with constant ("present-day") CO_2 and an 80-year greenhouse run with a gradual (1% per year compound) increase in CO_2 have been conducted with each model. The TCR is calculated as the difference in 20-year average global mean surface air temperature between the years 61-80 in the greenhouse runs and the same period in the control runs. Except for the Institute for Numerical Mathematics model (INMCM; *Diansky and Volodin* (2002)) that joined CMIP2 later than the others, all 20 models are detailed in Table 9.1 of *Cubasch et al.* (2001). The 15 values of CS used in this study are directly from the same table. In contrast to TCR, the CS values come from experiments in which the atmospheric part of the coupled GCM was connected to a shallow mixed-layer ocean. Thus, these experiments do not account for eventual changes in ocean circulation that might also be relevant on centennial time scales. Feedbacks between climate and changing vegetation distribution are also neglected in the models used for this study.

Fig. 1 shows the simulated TCR and CS in the various models, ordered in ascending order of TCR. Although CS appears for 16 models in the figure, only 15 values are used below. Two of the coupled models in CMIP2, CSM and DOE PCM, share the same atmospheric component and thus the same CS (however, CSM and DOE PCM are treated separately in the TCR analysis, although they are close relatives). Also note that in one case (GISS), different versions of the atmospheric model were used when evaluating CS and TCR.

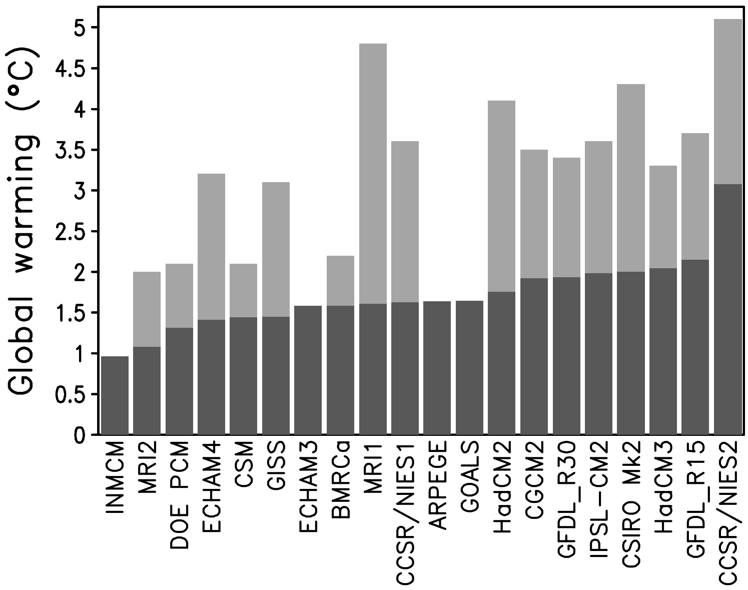

Fig. 1. The values of TCR (darker shading) and CS (lighter shading) in the models used in this study. The model names follow Table 9.1 of *Cubasch et al.* (2001).

Among the models shown in Fig. 1, TCR varies from 1.0 °C in INMCM to 3.1 °C in CCSR/NIES2. The range for CS is 2.0-5.1 °C. Although the two quantities are correlated ($r = 0.72$), the ratio between them also varies substantially. At least a part of this variation is associated with differences in ocean heat uptake between the models (*Raper et al.* 2002). In addition, the balance between the positive and negative feedback processes that regulate the magnitude of warming may change with climate state. Thus, the net feedback may be either more positive or more negative in the new equilibrium climate than during the transient phase of the warming, but models disagree on the direction of the change (*Boer and Yu* 2003). Finally, as already noted, the atmosphere – mixed-layer ocean models used to derive CS cannot account for changes in ocean circulation.

3. *Analytical probability distributions and their sampling uncertainty*

Let us assume that the 20 TCR values and 15 CS values are both random samples of some underlying probability distributions, neglecting the fact that the models are not necessarily independent from each other (because, for example, they often share common parameterization packages for the description of sub-grid scale processes). Let us also assume that the underlying distributions can be expressed in analytical form. To define the distributions, one needs to (i) *choose* their assumed form and (ii) estimate the numeric parameters that determine their location (e.g., the mean or the median) and width (e.g., standard deviation). The first step is largely guessing. Although one can try to eliminate bad candidates of the form by statistical testing, this is in practice difficult when the sample is small (see below).

We consider two alternative forms for the distributions: the normal distribution and the two-parameter (lower bound zero) lognormal distribution, the latter of which was used by *Wigley and Raper* (2001). Fig. 2 shows the resulting normal and lognormal fits to the TCR and CS data (see the Appendix for equations), together with histograms of the original values. An inspection of Fig. 2a suggests that the TCR data are not described well by a normal distribution, although this impression is essentially caused by CCSR/NIES2 that simulates a much larger TCR than any of the other models. The lognormal distribution appears to give a somewhat better fit. For CS (Fig. 2b), discrepancies between the data and each of the two analytical distributions seem more modest.

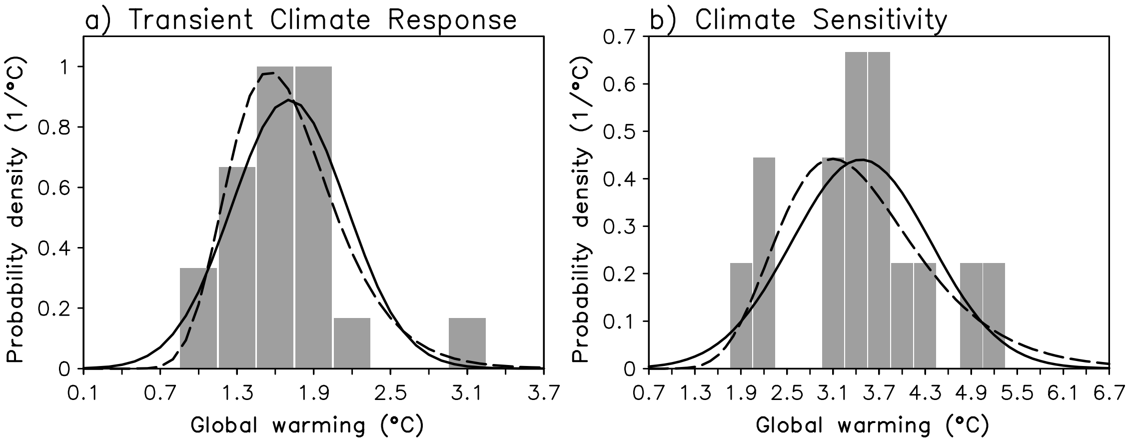

Fig. 2. Normal (solid line) and lognormal (dashed line) probability distributions fitted to (a) TCR in 20 models and (b) CS in 15 models. The bars show histograms of the original data with a bin width of 0.3 °C.

To check the visual impressions statistically, the skewness (A3) and kurtosis (A4) of the TCR and CS data sets were computed, first for the data as such and then after a logarithmic transformation (which converts a lognormal distribution to normal). The resulting values were then compared with the corresponding sampling distributions derived from Monte Carlo simulations with normally distributed random numbers. Both the skewness and the kurtosis of the TCR data exceed the 99th percentiles obtained in the Monte Carlo simulations (Table 1), so that the normal distribution can be rejected at 2% significance level using a two-sided test. Because the logarithmic transformation reduces the skewness and kurtosis substantially, the lognormal distribution cannot be rejected. For CS, the two distributions both provide statistically acceptable fits to the data.

Table 1. Test results for the form of the TCR and CS distributions. The p-values give the fraction of Monte Carlo experiments in which skewness and kurtosis were smaller than the values calculated from the actual data (see text).

	TCR	ln(TCR)	CS	ln(CS)
Skewness	1.16	0.07	-0.02	-0.56
(p-value)	(0.990)	(0.561)	(0.486)	(0.133)
Kurtosis	2.57	0.84	-0.54	-0.49
(p-value)	(0.994)	(0.925)	(0.260)	(0.308)

If TCR and CS are to be presented by the same form of a distribution, the lognormal distribution appears to be a statistically better candidate than the normal distribution. The lognormal distribution also has the physically desirable property of giving zero probability for negative values of TCR and CS. On the other hand, the test results for TCR are radically affected by a single model (CCSR/NIES2), without which both the skewness and the kurtosis of the distribution would be negative. Thus, if there were *a priori* reason to exclude CCSR/NIES2 from the calculations, the conclusion on the relative merits of the two distributions would change. Moreover, it is naturally possible that the actual underlying distributions are neither normal nor lognormal, but rather of some other form.

More generally, it is difficult to distinguish between different forms of the distribution on the basis of small samples. We performed Monte Carlo simulations in which 20-number random samples were drawn from the lognormal distribution estimated from the TCR data. About 85% of these samples were positively skewed, but the skewness exceeded the corresponding 95th percentile for the normal distribution in only 30% of the samples. The same fraction for 15-number CS samples was only 26%. Kurtosis was found to give a less powerful means to distinguish between the two types of distributions than skewness.

The inferred 5th, 50th (median) and 95th percentiles of the TCR and CS probability distributions are given in Table 2. Although the previous analysis suggests a slight preference for the lognormal form, results for the normal distribution fit are also shown. The 5-95% uncertainty range for TCR is 1.1-2.5 °C, and that for CS 2.1-5.3 °C, when using the lognormal distribution. The corresponding numbers for the normal distribution are 1.0-2.4 °C and 2.0-5.0 °C. Thus, both the 5th and 95th percentiles are slightly lower for the normal than the lognormal distributions, whereas the medians are slightly higher. The median for CS (about 3.4 °C) is twice the value for TCR (1.7 °C), even though it should be noted that the two distributions have been derived using slightly different sets of models. The widely cited range (1.5-4.5 °C) and frequently given best estimate (2.5 °C) for CS are somewhat on the lower side of the model results. Finally, one may note that the normal distribution fit for TCR classifies CCSR/NIES2 as a much more extreme outlier (probability of larger warming 0.1%) than the lognormal fit (0.7%).

Table 2. Percentiles of TCR and CS, as derived from the fitted analytical probability distributions. In each table entry, the first value gives the best estimate and the next two (in parentheses) the 5-95% uncertainty range derived from Monte Carlo simulations.

	Percentile	Normal distribution	Lognormal distr.
TCR	5%	0.97 (0.72-1.23)	1.09 (0.95-1.27)
	Median	1.71 (1.55-1.87)	1.66 (1.51-1.82)
	95%	2.44 (2.18-2.69)	2.51 (2.17-2.89)
CS	5%	1.98 (1.39-2.59)	2.11 (1.76-2.55)
	Median	3.47 (3.08-3.85)	3.35 (2.97-3.76)
	95%	4.96 (4.34-5.54)	5.31 (4.39-6.35)

Being based on a limited number of model results, the derived probability distributions suffer from sampling uncertainty. This source of uncertainty was estimated by first drawing a large number of artificial 15- and 20-model random samples from the best-fit normal or lognormal distributions, and by then repeating the distribution fitting for each of these samples. For the percentiles considered in Table 1, the sampling uncertainty is larger than the differences between the normal and the lognormal distributions. The best-estimate 5th, 50th and 95th percentiles for the normal distribution are always within the 5-95% sampling range for the lognormal distribution, and vice versa. It is only in the extreme tails of the distribution (roughly, $< 1\%$ and $> 99\%$) where the differences between the normal and the lognormal form begin to dominate over the sampling uncertainty (not shown).

In the case of the lognormal fit, the sampling uncertainty is largest in the upper end of the distribution; note in particular the large (4.39-6.35 °C) uncertainty in the 95th percentile of CS. For the normal distribution fit, both the lower and upper ends suffer from larger sampling uncertainty than the middle of the distribution.

4. Comparison with other studies of climate sensitivity

Due to its great importance, climate sensitivity (CS) has been a topic of intense research especially in the last few years. In addition to the simple method applied in this study, there are at least two more sophisticated alternatives for estimating this quantity. One of these is based on so-called perturbed-parameter model simulations and the other on instrumental observations of climate variability during the last century or proxy data of earlier climate variations.

The differences in CS among climate models arise from inter-model differences in the parameterisation of sub-grid scale phenomena such as boundary layer turbulence, radiation transfer and (most importantly) cloud processes. However, it has been argued that the uncertainty associated with the parameterisation problem may not be captured properly by “unplanned” multi-model ensembles such as the one used in this study. To explore the issue in a more systematic manner, *Murphy et al*. (2004) and *Stainforth et al*. (2005) used perturbed-parameter ensembles, in which a large number of uncertain model parameters were varied within what modelling experts consider as reasonable limits. *Murphy et al*. (2004) determined CS for 53 versions of the Hadley Centre model HadAM3, including the standard version of the model and 52 modified versions obtained by changing one model parameter at a time. From the results of their perturbed-parameter simulations they estimated the 5-95% uncertainty range of CS as 2.4-5.4 °C (to be compared with 2.0-5.0 °C or 2.1-5.3 °C in the present study). To obtain this estimate they weighted the CS values from the individual model versions according to the ability of the various versions to simulate the present-day climate; the corresponding uncertainty range without the weighting was 1.9-5.3 °C.

Stainforth et al. (2005) also perturbed the parameters of HadAM3 but, in contrast to *Murphy et al*. (2004), they changed several parameters at a time. Among the over 400 model versions studied by them, CS varied from 1.9 °C to 11.5 °C, with 4.2% of the versions having sensitivity above 8 °C. Although one might expect the most sensitive model versions to simulate the present climate poorly, *Stainforth et al*. (2005) found no evidence of this when studying the long-term mean climates in the various model versions. However, it is not yet known if this conclusion would hold after a more complete evaluation of the simulated climates, including higher-order climate statistics such as the amplitude of the simulated interannual variability.

Both *Murphy et al*. (2004) and *Stainforth et al*. (2005) used the same model (HadAM3) as their starting point. If a similar exercise were repeated for other models with different standard-version sensitivities, this might result in an even wider range of CS estimates.

A purely model-based assessment cannot tell the final truth of CS (or TCR) in the real world. Several investigators have therefore attempted to estimate CS from observations of temperature variability during the last 100-150 years or from proxy data of earlier variations. Apart from estimates of temperature variability, such studies require quantitative estimates of the external factors (usually in terms of radiative forcing) that have caused the temperature changes. In addition, simulations with simple climate models have been used in many of these studies to help the interpretation of the observations. Some of the derived CS estimates are given in Table 3.

Table 3. Estimates of climate sensitivity (CS, the equilibrium global mean warming resulting from a doubling of CO_2) based on either instrumental observations during the last 100-150 years (I) or on proxy data of earlier temperature variations (P). In studies 1, 2, 3 and 7, range refers to the derived 5-95% uncertainty range. The uncertainty interval covered in the other studies has not been specified in probabilistic terms.

	I or P	Reference	Range	Best estimate
1	I	Andronova and Schlessinger (2000)	1.0-9.3 °C	2.0 °C
2	I	Forest et al. (2002)	1.4-7.7 °C	2.9 °C
3	I	Gregory et al. (2002)	1.6 °C - ∞	6.1 °C
4	I	Harvey and Kaufmann (2002)	1.0– 3.0 °C	2.0 °C
5	P	Hoffert and Covey (1992)	1.4-3.2 °C	2.3 °C
6	P	Barron et al. (1995)	2.5-4.0 °C	not given
7	P	Schneider von Deimling et al. (2004)	1.5-4.7 °C	2.1-3.6 °C

Without going to the details of methodology, which vary substantially across the various studies, we can make a few general observations. Studies based on temperature variations during the instrumental period tend to give highly skewed probability distributions of CS. The lower bound of the derived uncertainty range is at 1.0-1.6 °C and the best estimate (median of the distribution) is with one exception at 2-3 °C. However, the upper bound is in most studies very high, with one exception at least 7.7 °C, and in the most extreme case (*Gregory et al*., 2002) even infinitely high climate sensitivity was found to be consistent with observations. The difficulties in determining an upper bound for CS are associated with the great uncertainty in the magnitude of anthropogenic aerosol forcing, which is thought to have opposed the warming driven by increased greenhouse gas concentrations (*Boucher and Haywood,* 2001; *Andreae et al*., 2005). If the negative aerosol forcing has cancelled a large fraction of the positive greenhouse gas forcing, then large climate sensitivity is required to explain the observed warming. At the limit where the aerosol forcing is strong enough to make the net radiative forcing to approach zero, infinitely high CS is implied, although some of the assumptions in the calculations might break down at this limit.

In *Harvey and Kaufmann* (2002), a much lower upper bound (3.0 °C) for CS was found than in other studies based on instrumental data. CS exceeding 3.0 °C was found inconsistent with the relatively modest cooling that followed the Mount Krakatau eruption in 1883. However, this conclusion may be sensitive to errors in temperature observations in the 1880s and in the radiative forcing resulting from the Mount Krakatau eruption.

Studies based on proxy data of earlier climate variations have generally focused either on the Last Glacial Maximum (about 21 thousand years ago) or on the Cretaceous warm period (about 100 million years ago), or both. There are a number of complications when inferring CS from such distant climate variations. For example, the studies generally treat changes in continental ice sheets and their effect on the planetary energy balance as part of the external forcing, although they are in reality a feedback from changing climate. This is justified by the fact that the ice sheet feedback is too slow to be important in the period (a few hundreds years) usually considered to be of greatest interest in the context of current anthropogenic climate change. Furthermore, many although not all (*Schneider von Deimling et al.* (2004) is an exception) of the pre-instrumental studies have implicitly assumed that CS is independent of the basic state of the global climate. If this is not the case, as has been suggested by some model simulations (e.g., *Stouffer and Manabe,* 2003), then studies based on periods with very different climate conditions may give biased estimates for the sensitivity of the present-day climate system.

The CS estimates from the pre-instrumental studies are broadly comparable with the estimates based on instrumental data, but the uncertainty ranges tend to be narrower (lower half of Table 3). In particular, estimates of temperature variability in the pre-instrumental past do not appear to support the very high upper bound of CS found in many of the instrumental studies. However, this conclusion should be treated with some caution because the uncertainty analysis in the pre-instrumental studies has been less complete than that in the instrumental studies (e.g., regarding uncertainties in radiative forcing).

In conclusion, our model-based 5-95% uncertainty interval for CS (approximately 2.1-5.3 °C) is more or less consistent with the estimates obtained from other methods, except for the fact that some of these other methods suggest a substantially higher upper bound for the uncertainty range.

5. Conclusions

Analytical probability distributions of the transient climate response (TCR) and equilibrium climate sensitivity (CS) caused by a doubling of CO_2 have been derived from available model results, using the assumption that all models provide equally likely projections of climate change. The main findings are listed below.

- For the 20 CMIP2 models, the derived median TCR is about 1.7 °C. The best-estimate 5-95% uncertainty range based on a normal distribution is 1.0-2.4 °C and that based on a lognormal distribution 1.1-2.5 °C.
- For the 15 equilibrium doubled CO_2 simulations, the median CS is about 3.4 °C. The best-estimate 5-95% uncertainty range based on a normal distribution is 2.0-5.0 °C and that based on a lognormal distribution 2.1-5.3 °C.
- Limited sample size makes it hard to distinguish between different candidates for the form of the distributions. In the case of TCR, the lognormal distribution fits the data better than the normal distribution, but this conclusion is critically dependent on one extreme model.
- Limited sample size also causes uncertainty in specifying the parameters that determine the location (mean or median) and the width (standard deviation) of the distribution. For estimating the 5-95% ranges of TCR and CS, this is a larger uncertainty than the differences between the normal and the lognormal distribution.
- Our purely model-based 5-95% uncertainty range of CS is in reasonable agreement with estimates based on instrumental observations and proxy data of pre-instrumental climate variability.

Acknowledgments

All CMIP2 modelling groups are acknowledged for conducting and making available the simulations requested by the CMIP Panel. CMIP is supported and the model data are distributed by the Program for Climate Model Diagnosis and Intercomparison (PCMDI) at the Lawrence Livermore National Laboratory (LLNL). This article benefited from helpful comments by Dr. Heikki Tuomenvirta and an anonymous reviewer.

Appendix. Equations

Let us denote the simulated global mean temperature changes in a sample of N different models as x_i, $1 \le i \le N$. The mean (m), standard deviation (s), skewness (*skew*) and kurtosis (*kurt*) of the sample are calculated from

$$m = \frac{1}{N}\sum_{i=1}^{N} x_i \tag{A1}$$

$$s^2 = \frac{1}{N}\sum_{i=1}^{N} (x_i - m)^2 \tag{A2}$$

$$skew = \frac{1}{Ns^3}\sum_{i=1}^{N} (x_i - m)^3 \tag{A3}$$

$$kurt = \frac{1}{Ns^4}\sum_{i=1}^{N} (x_i - m)^4 - 3 \tag{A4}$$

If the N model results are a random sample of some underlying probability distribution, the unbiased estimate for the variance of the distribution is

$$\sigma^2 = \frac{N}{N-1} s^2 \tag{A5}$$

The best-estimate normal distribution fit to the sample gives the probability density function

$$\phi_{Norm}(x) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{(x-m)^2}{2\sigma^2}\right] \tag{A6}$$

The median of this distribution is m, and the 5th and 95th percentiles are $m \pm 1.6449\sigma$. Similarly, the density function for the two-parameter lognormal distribution is obtained from

$$\phi_{Log}(x) = \frac{1}{\sqrt{2\pi}x\sigma_{Log}}\exp\left[-\frac{(\ln(x) - m_{Log})^2}{2\sigma_{Log}^2}\right] \tag{A7}$$

where m_{Log} and σ_{Log} are estimated by applying (A1), (A2) and (A5) to the natural logarithm of the temperature change. The median of the distribution is $\exp(m_{Log})$, and the 5th and 95th percentiles are $\exp(m_{Log} \pm 1.6449\sigma_{Log})$.

References

Andreae, M.O., C.D. Jones and M.O. Cox, 2005. Strong present-day aerosol cooling implies a hot future, *Nature*, **435**, 1187-1190.

Andronova, N.G. and M.E. Schlessinger, 2001. Objective estimation of the probability density function of climate sensitivity, *J. Geophys. Res*., **106**, 22 605-22 611.

Barron, E.J., P.J. Fawcett, W.H. Peterson, D. Pollard and S.K. Thompson, 1995. A ‘simulation’ of mid-Cretaceous climate, *Paleoceanography*, **10**, 953-962.

Boer, G.J. and B. Yu, 2003. Climate sensitivity and climate state, *Climate Dynamics*, **21**, 167-176.

Boucher, O. and J. Haywood, 2001. On summing the components of radiative forcing of climate change, *Climate Dynamics*, **18**, 297-302.

Cubasch, U., G.A. Meehl, G.J. Boer, R.J. Stouffer, M. Dix, A. Noda, C.A. Senior, S. Raper and K.S. Yap, 2001. Projections of future climate change. Climate change 2001. The Scientific Basis, J.T. Houghton, Y. Ding, D.J. Griggs, M. Noguer, P.J. van der Linden, X. Dai, K. Maskell, and C.A. Johnson, Eds., Cambridge University Press, 525-582.

Diansky, N.A. and E.M. Volodin, 2002. Simulation of present-day climate with a coupled atmosphere-ocean general circulation model, *Izvestia, Atmospheric and Oceanic Physics*, **38**, 732-747.

Forest, C.E., P.H. Stone, A.P. Sokolov, M.R. Allen and M.D. Webster, 2002. Quantifying uncertainties in climate system properties with the use of recent climate observations, *Science*, **295**, 113-117.

Gregory, J.M., R.J. Stouffer, S.C.B. Raper, P.A. Stott and N.A. Rayer, 2002. An observationally based estimate of the climate sensitivity, *J. Climate*, **15**, 3117-3121.

Harvey, L.D.D. and R.K. Kaufmann, 2002. Simultaneously constraining climate sensitivity and aerosol radiative forcing, *J. Climate*, **15**, 2837-2861.

Hoffert, M.I. and C. Covey, 1992. Deriving global climate sensitivity from paleoclimatic reconstructions, *Nature*, **360**, 573-576.

Meehl, G.A., G.J. Boer, C. Covey, M. Latif and R.J. Stouffer, 2000. The Coupled Model Intercomparison Project (CMIP), *Bull. Am. Meteorol. Soc*., **81**, 313-318.

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 modelling uncertainties in a large ensemble of climate change simulations, *Nature*, **430**, 768-772.

Raper, S.C.B., J.M. Gregory and R.J. Stouffer, 2002. The Role of Climate Sensitivity and Ocean Heat Uptake on AOGCM Transient Temperature Response, *J. Climate*, **15**, 124-130.

Schneider von Deimling, T., H. Held, A. Ganopolski and S. Rahmsdorf, 2004. Climate sensitivity range derived from large ensemble simulations of glacial climate constrained by proxy data. IPCC Working Group I Workshop on Climate Sensitivity, Paris, France, 26-29 July 2004, pp. 145-147.

Stainforth, D.A., T. Aina, C. Christensen, M. Collins, N. Faull, D.J. Frame, J.A. Kettleborough, S. Knight, A. Martin, J.M. Murphy, C. Piani, D. Sexton, L.A. Smith, R.A. Spicer, A.J. Thorpe and M.R. Allen, 2005. Uncertainty in predictions of the climate response to rising levels of greenhouse gases, *Nature*, **433**, 403-406.

Stouffer, R.J. and S. Manabe, 2003. Equilibrium response of thermohaline circulation to large changes in atmospheric CO_2 concentration, *Climate Dynamics*, **20**, 759-773.

Wigley, T.M.L. and S.C.B. Raper, 2001. Interpretation of high projections for global-mean warming, *Science*, **293**, 451-455.