Supporting Information for "A Climate of Belief" *2008* Skeptic *14(1)*, 22-30

by Patrick Frank

Except as noted, all the calculations and graphics for the work described here were done using the application Kaleidagraph (Synergy Software, Reading, PA). The data for article Figure 1, Figure 2, and Figure 3, and for SI Figure 14 and Figure 15 were digitized from the published graphics using the program Digitizeit (http://www.digitizeit.de/).

The passive warming model is derived in SI Sections 1 and 2. The goal was to test outputs of GCMs, and no claim is intended about the actual physical behavior of climate. In SI Section 1, an estimate is made of the fractional contribution to Earth climate of the enhanced CO_2 greenhouse effect, *as modeled in GCMs*. The calculation of the enhanced greenhouse effect of CO_2 follows the same assumption made in GCMs, namely that climate globally responds with constant relative humidity to the warming induced by CO_2 forcing. Likewise, no claim is made that this assumption is physically correct. Rather, the condition is included so as to follow the norms of GCM climate models. In SI Section 2, it is shown that the linear propagation of this assumption is virtually all that is necessary to reproduce the temperature projections of complex climate models.

SI Section 3 describes the calculation of GCM cloudiness retrodiction error, and SI Section 4 concerns the description and validation of the propagation of physical uncertainties in cloudiness and GHG forcing, as applied to time-wise GCM global average temperature predictions.

SI Sections 5 and 6 discuss the validity of a GCM projection of the 20th century temperature trend as publicized by the US National Academy of Sciences, and the greening of the Sahel, respectively.

1. Derivation of the Passive Temperature Response Model.

The fractional contribution of the enhanced CO_2 greenhouse effect to the total greenhouse effect, as portrayed in GCM models, was derived from asymptotic intercepts of natural logarithmic fits to the data in Table 4, "Equilibrium temperature of the earth's surface (°K) and the CO_2 content of the atmosphere," of Manabe.[1]

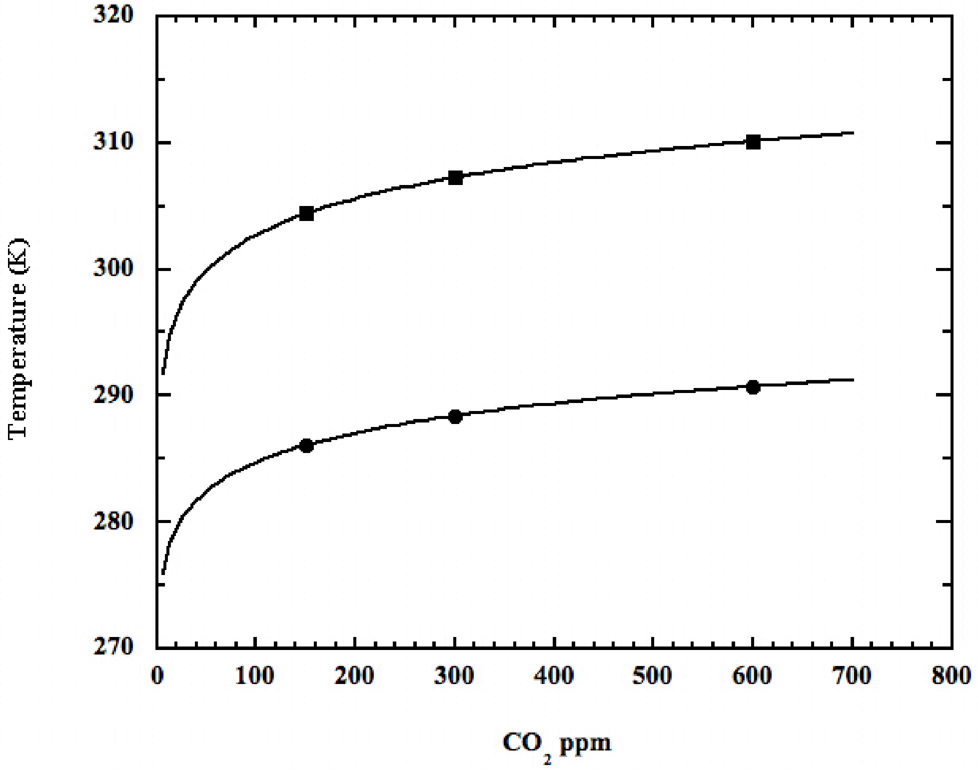


Figure S1: The equilibrium surface temperature under varying concentrations of atmospheric CO_2*, under (—), clear skies; fit:* $T=4.13\times\ln(CO_2)+283.7$*;* $r^2=0.9999$*, and (□), cloudy skies, fit:* $T=3.35\times\ln(CO_2)+269.3$*;* $r^2=0.9999$*. Data from Manabe.*[1]

The fits plus the fit equations are shown in Figure S1. The functional form of the fit is justified by the natural log relation between greenhouse forcing and the concentration of atmospheric CO_2 as given in Myhre, et al.[2]

The derived greenhouse effect on Earth climate at zero ppm CO_2 was the sum of the two asymptotic intercepts (Figure S1) weighted by the ~58.3 % average cloud cover of Earth as determined by integrating the published average 1983-1990 cloudiness (see article Figure 3 and text). Thus, the zero CO_2 greenhouse fraction = 0.583×(269.3 K)+0.417×(283.7 K) = 275.3 K. The surface temperature at pure radiation equilibrium = 254.6 K, so that the zero CO_2 greenhouse contribution = 275.3-254.6=20.7 K. This 21 K represents the greenhouse contribution to Earth surface temperature from water vapor alone. The unperturbed total greenhouse temperature (water vapor alone + water-vapor-enhanced CO_2) is 33 K. The enhanced CO_2 greenhouse contribution is then 33 K-21 K = 12 K.

The fractional enhanced greenhouse contribution to Earth surface temperature, as represented in GCMs, is thus 12/33=0.36. This method of weighting by percent cloudy and clear sky forcing has precedent in the derivation of the current average 2.7 W/m^2 greenhouse gas forcing in Myhre.[2]

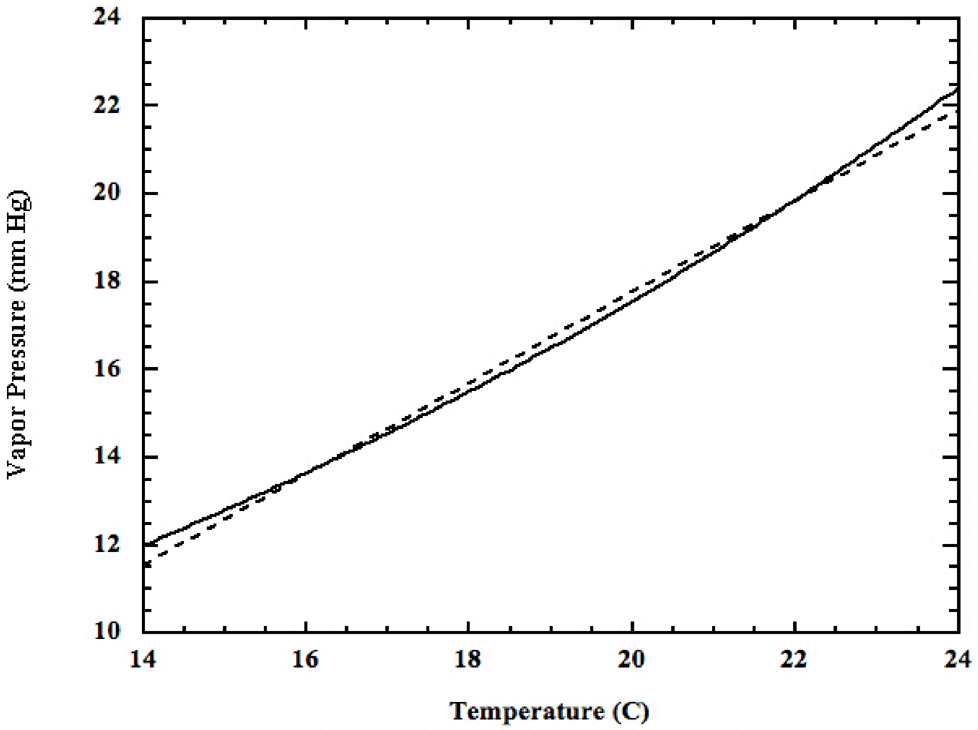


Figure S2. (—), The exponential increase in water vapor pressure with temperature, and; (---), the linear fit From 14-24 C: Vapor Pressure = 1.038×T - 2.98; r^2= 0.995. The covered range includes the purported rise in surface temperature from excess greenhouse warming.

The 0.36×33 C fractional weighting of additional greenhouse forcing in the model automatically and linearly scales the water-vapor enhanced effect, already present in the initial 33 C, into the calculated temperature increases. Therefore, like all current GCMs, the model automatically includes constant relative humidity. The model approximates the increase in water vapor pressure with temperature as a linear function. Although the true relationship is exponential, the linear increase is a reasonable approximation over small temperature ranges, Figure S2.

The historical concentrations of CO_2 for 1880-1957 were obtained from Etheridge, et al.,[3] and for 1958-2004 were obtained from Keeling and Whorf.[4] The historical concentrations of methane were from Etheridge, et al.,[5] and of nitrous oxide were from Table 1 in Khalil, et al.[6]

Figure S3 shows the full fits to the historical greenhouse gas concentrations that were used to calculate greenhouse gas forcings.

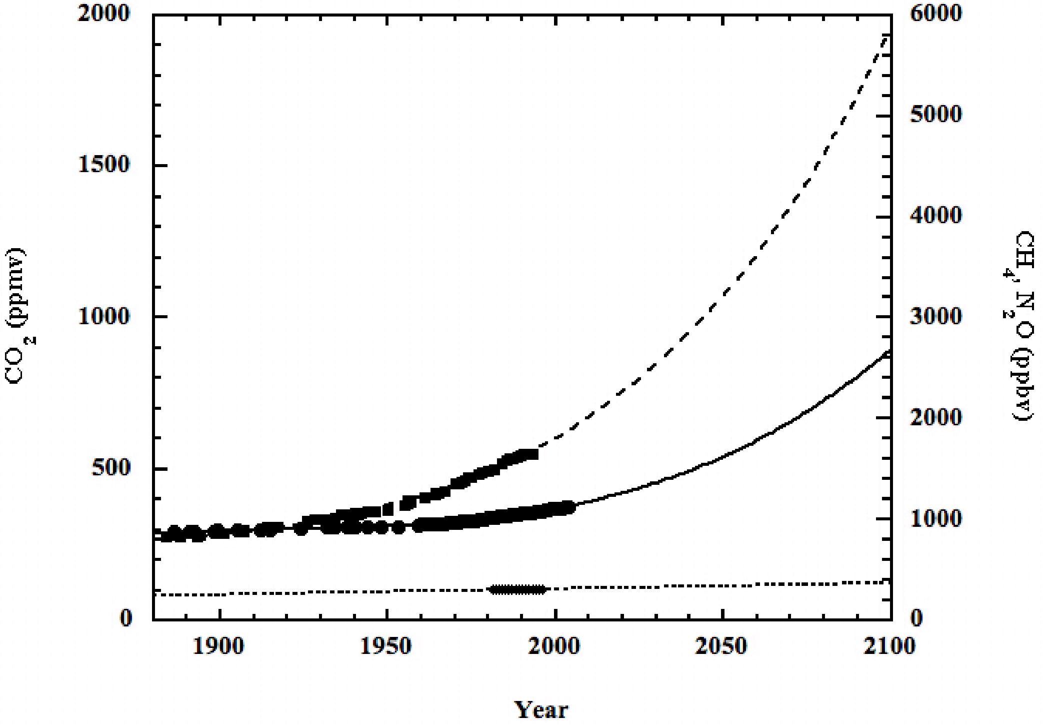


Figure S3. Double-y projection of extrapolated atmospheric concentrations of: (—), CO_2 (left), and; (---), methane and (····), nitrous oxide (right). The points are historical measurements.

The fitted equations extrapolating each of the historical greenhouse gas concentrations in Figure S3 were: For carbon dioxide, CO_2 (ppmv)= $-6.49\times10^5 + Y\times1015.4 - Y^2\times0.529 + Y^3\times9.19\times10^{-5}$, $r^2=0.996$; for methane: CH_4 (ppbv) = $M_0 + M_1\times Y + \ldots + M_6\times Y^6 + M_7\times Y^7$, where $M_0=-2.4434\times10^6$, $M_1=12489$, $M_2=-27.119$, $M_3=0.032436$, $M_4=-2.3081\times10^{-5}$, $M_5=9.7725\times10^{-9}$, $M_6=-2.2801\times10^{-12}$, $M_7=2.262\times10^{-16}$, $r^2=0.997$; for nitrous oxide: N_2O (ppbv) = $0.58662\times Y - 857.04$, $r^2= 0.977$. In all cases, Y is the 4-digit year.

The increasing forcings due to added greenhouse gases calculated according to the equations in Myhre, et al.,[2] are shown in Figure S4.

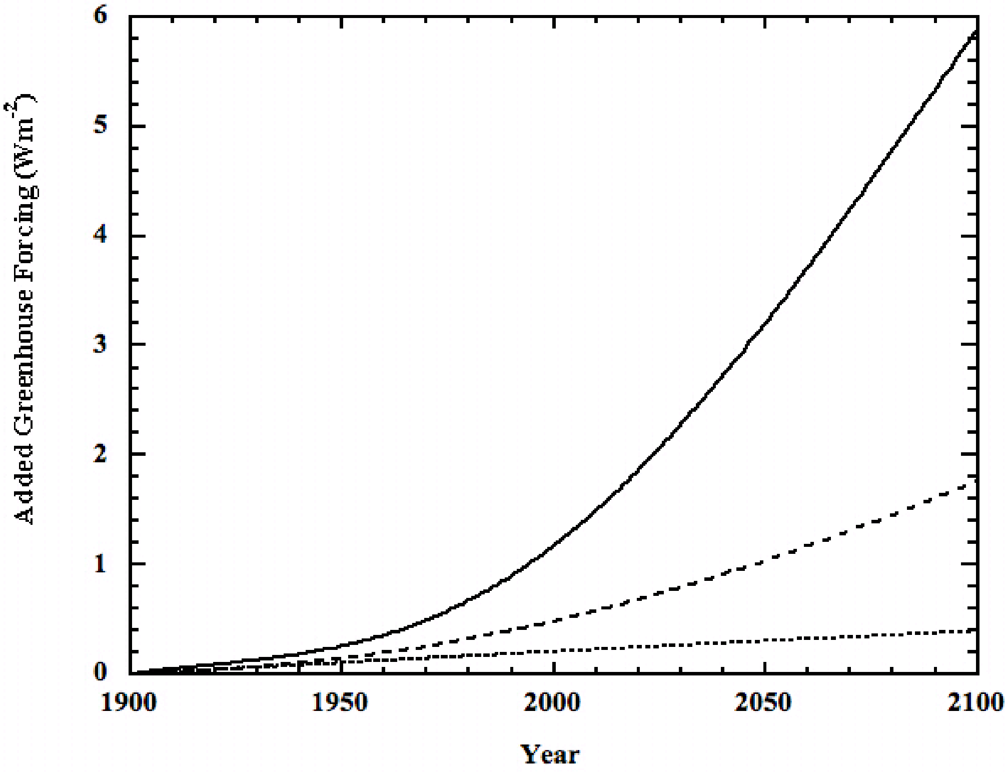


Figure S4. Additional greenhouse gas forcing from the year 1900, for: (—), CO_2; (–––), methane, and; (····), nitrous oxide, as calculated from the fits to the greenhouse gas concentrations of Figure S3, using the equations in Myhre.[2]

The forcings in Figure S4 were summed in order to calculate the increased greenhouse temperature in equation S1, as given in ACoB:

Global Warming=0.36x33Cx[(Total Forcing)÷(Base Forcing)] **S1**,

where "Base Forcing" is the forcing from all three greenhouse gasses as calculated from their historical or extrapolated concentrations for either the year 1900 or 2000, depending on which year was chosen as the zeroth year. "Total Forcing" is "Base Forcing" + the increase in forcing.

When the temperature increase due to a yearly 1 % CO_2 increase was calculated, the increasing CO_2 forcing was adjusted to include the higher atmospheric concentration of this gas, but the increasing forcings due to methane and nitrous oxide were left unchanged at their Figure S4 values.

1.2 On Origins. An anonymous reviewer of an earlier form of the manuscript suggested that equation S1 is "*the standard linearized* *relation between radiative forcing and global-mean temperature response,*" and further stated that the author of ACoB and ACoB SI made a dishonest claim of originality. It is worth reviewing that charge here to forestall any recapitulation of the accusation.

Nowhere in the SI or in the article is any originality claimed. Nor did any such claim appear in the prior version of these documents evaluated by that reviewer. Copies of the prior documents are available.

The reviewer went on to name "*the IPCC report*," as one source for the standard linearized relation. In the 4AR a linear model is discussed in Chapter 2, Section 2.8.4, "Linearity of the Forcing-Response Relationship,"[7] which presumably exemplifies what the reviewer had in mind.

The linear temperature response model given by the IPCC is:

$$\Delta T_s=\lambda_{CO_2}\times E_i\times RF_i, \qquad R1$$

where T_s is the surface temperature, λ_{CO_2} is "the climate sensitivity parameter" for CO_2, E_i is the efficacy of some forcing agent "i" defined in relation to that for CO_2, and RF_i is the radiative forcing for the ith forcing agent.

The following analysis will show that equation S1 is different from equation R1.

When the forcing agents are CO_2 plus methane plus nitrous oxide -- the gases included in the present analysis -- equation R1 becomes:

$$\Delta T_s=(\lambda_{CO_2}\times RF_{CO2})+(\lambda_{CO_2}\times E_{CH4}\times RF_{CH4})+(\lambda_{CO_2}\times E_{N2O}\times RF_{N2O}). \qquad R2$$

Compare R2 to the expanded form of equation S1:

$$\Delta T_s=T_0\times f_{EGE}\times[(\Delta F_{CO_2}+\Delta F_{CH_4}+\Delta F_{N_2O})\div F_0], \qquad S1'$$

where T_s is again the surface temperature, T_0 is the empirically determined base greenhouse contribution to the surface temperature (i.e., 33 C), f_{EGE} estimates the fractional contribution of the enhanced CO_2 greenhouse effect to the surface greenhouse temperature, but *in the magnitude that fraction has within GCMs as deduced from the results reported by Manabe*,[1] as shown in Figure S1 and discussion, ΔF_{GAS} is the canonical change in greenhouse gas forcing[2] due to the increased concentration of each relevant gas after the zeroth year, and F_0 is the reference (i.e., the zeroth year total) greenhouse gas forcing.

Equations S1 and S1' are different from equations R1 and R2 in terms of origination logic, of physical constituents, and of dimensional units. The approach to, and constituents of, equation S1 completely reveal its independent origin. Equation S1 is not the "*the standard linearized relation between radiative forcing and global-mean temperature response.*" It is not, as is equation R1, a derivation of surface temperature from the physical theory of radiative forcing. Equation S1 is a linearized, empirical, and physically reasonable estimator *of what GCMs might produce* when used to calculate average surface temperature.

Equations S1 and R1 are equivalent merely in linearity and calculational focus (i.e., they equate to ΔT_s).

Because S1 and R1 each equate to ΔT_s they must equate to one another, in principle at least. Thus, $\lambda_{CO_2} \times E_i \times RF_i = T_0 \times f_{EGE} \times (\Delta F_{GAS} \div F_0)$. In this case, there should exist a set of transformations that converts equation S1 into equation R1.

Equational parity does not translate into derivational identity, however. To exemplify this distinction by citing a famous equivalence: in order to suppose that equations S1 and R1 are derivationally identical because they can be somehow interconverted through ΔT_s, one must also accept that Erwin Schrödinger's *a posteriori* wave mechanics are derivationally identical to Werner Heisenberg's matrix mechanics because the two have equational parity through $E\Psi$ and can be rigorously interconverted. By the reviewer's logic, it would have been dishonest for Erwin Schrödinger to claim originality for his master work. Clearly that logic fails, both here and above.

Emphatically, no comparison is claimed or implied between the work presented in the SI and anything Schrödinger or Heisenberg attained. Obviously, there can be no such comparison whatever. What is claimed in the above example is a powerful illustration of an impoverished and tendentious logic.

No originality is claimed here. Derivational independence is claimed.

2. Passive Model Validation by Comparison with GCM Projections

Here, the passive warming model as detailed above is shown to be closer to the ensemble average of the 10 GCM projections from the CMIP than any of the individual GCM projections themselves. The test model is a line, while the GCM projections contain excursions from their respective means. In order to remove any bias due to this difference, linear least square fits were made to all 10 GCM projections, so as to determine each mean projection. The comparison to the ensemble average was then made using the passive warming model and the linear means of the 10 GCM projections.

Figure S5 reproduces ACoB Figure 2, except that it shows the linear means of the 10 GCM projections. In order to make the comparison quantitative, the Student t-test for paired data sets was used to compare each of the projections with the 10-GCM ensemble average. Table S1 shows the results of this test.

Table S1: Paired Data Student t-Test: 10 GCMs and Passive Model vs. GCM Ensemble Average.

GCM[a]	Mean Difference	t-value	t-probability
NCAR	0.171	17.39	<0.0001
LMD ISPL	-0.111	-14.76	<0.0001
HadCM3	-0.223	-25.38	<0.0001
HadCM2	-0.046	-11.45	<0.0001
GISS	0.125	11.68	<0.0001
GFDL	-0.175	-11.84	<0.0001
ECHAM3	0.106	24.65	<0.0001
DOE PCM	0.248	18.74	<0.0001
CSIRO	-0.144	-18.90	<0.0001
CFERACS	0.049	11.12	<0.0001
1% GHG[b]	0.014	2.02	0.04496

a. Identification of the GCMs represented by the letter codes is in Covey, et al.[8] b. Passive greenhouse gas warming model.

In Table S1, “Mean Difference” is the difference between the means of the model and the ensemble average. A larger absolute value is a greater difference. The 1% GHG model is far closer to the ensemble average mean than any of the GCMs. The “t-value” is the mean difference divided by the standard deviation of the covariance of the data set – ensemble average pair. A smaller t-value is a closer correlation. For the “t-probability,” a value greater than 0.05 means there is no difference between the data sets. Only the passive warming model comes close to meeting this test.

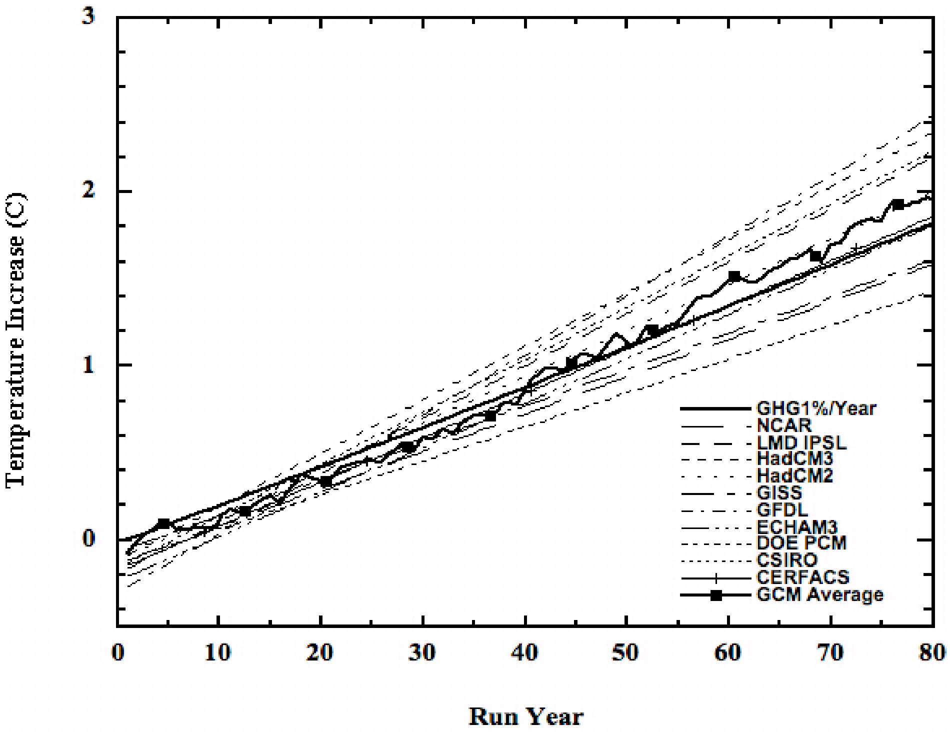


Figure S5: As for ACoB Figure 2, except that the GCM projections are represented as their linear least square means.

The good correlation between the projections of the passive GHG model and the projections of complex GCMs (ACoB Figure 2, Figure S5, and Table S1) validates the calculated forcings and the derived enhanced CO_2 greenhouse fraction of the passive GHG model in terms of the GCMs. The implied validity of the inputs in terms of correspondence among the outputs in turn justifies the SRES error analysis that follows in Section 4.1.

The ensemble average of GCM projections has lower intensity excursions than any of the GCM projections themselves, as shown in ACoB Figure 2. This means the features in the GCM projections are nearly random, and tend to cancel out when the GCM projections are averaged together. Random features have no physical meaning. If the GCMs were able to predict, e.g., El Niños, it might be expected that at least one intense event might show up in 80 years of greenhouse warming oceans, perhaps similar to the 1998 event.[9] GCMs as valid climate models, would all similarly predict such an event after a requisite time of heat build-up. This relatively uniform prediction would produce an atmospheric temperature spike in a similar (but perhaps not identical) year in each 80-year GCM temperature projection. An El Niño temperature signal would not cancel like random noise, but would approximately retain its original intensity with the averaging of GCMs and produce a spike in the ensemble average trend. There are no such features in the ensemble average, and none of the low intensity features remaining are confidently assessable as physically real.

3. Cloudiness Retrodiction Error

Table S2 shows the integrated average observed cloudiness of 1983-1990, and the integrated average retrodicted 1979-1988 cloudiness of 10 revised GCMs, as taken from Figure 18b of Gates, et al.[10]

Table S2: Observed and GCM Retrodicted Global Average Cloudiness Integrated Over the Same Pair-Wise Latitudinal Range.

GCM[a]	GCM Average Cloudiness	Obs'd. Avg. Cloudiness	Absolute Fraction	Lag-1 Error Autocorrelation [R]
LMD	10629	9647.9	0.10169	0.9631
DERF	10389	10291	0.0095164	0.9595
BMRC	9050.7	10346	0.12522	0.9881
CNRM	10659	10306	0.034217	0.9766
NRL	11710	10329	0.13368	0.9850
MPI	11353	10313	0.10084	0.9767
MRI	11709	10435	0.12206	0.9639
DNM	10389	10291	0.0095164	0.9595
SUNGEN	10322	10268	0.0052318	0.9411
YONU	11972	10436	0.14714	0.9704

"Absolute Fraction" = the absolute value of the error fraction [(Observed - GCM)/Observed]. a. Identification of the GCMs represented by the letter codes is in Gates, et al.[10]

To obtain the values in Table S2, the observed cloudiness was interpolated onto the x-grid of each of the GCM projections, so that in every case the covered area was identical. The Lag-1 column is discussed in Section 4.2 below.

In each case, the predicted global average cloudiness was the integrated area of each latitudinal average cloudiness curve for each GCM. The comparative integrated area of the observed cloudiness was the observed cloudiness integrated over the same set of latitudes as covered by each GCM model.

$$\text{Average GCM error} = \sqrt{(\sum_{1}^{10}[(Obs'd - GCM/Obs'd)^2)/9} \qquad \textbf{S2.}$$

Thus, $error=[\surd(0.091989/9)]=\pm 0.1011=\pm 10.1\%$.

The very small absolute errors attending SUNGEN and DERF/DNM arise for the reason mentioned in the text: the residuals from these projections have almost equal positive and negative areas of excursion. A better estimate of the total error remaining after the 10-averages, calculated as the integrated $\surd[(\text{residual})^2]$, yields 16.6 % for DERF/DNM and 15.6 % for SUNGEN. For CNRM and YONU, the total error likewise becomes 31.5 % and 29.4 %, respectively.

Each GCM projection considered here is the average of ten yearly cloudiness realizations (1979-1988). The observed cloudiness is an 8-year average. Averaging across the years tends to smooth out the year-to-year variations. The resulting smoothness removes latitudinal eccentricities from both the observed cloudiness and the retrodicted cloudiness. This homogenization of both retrodiction and observation tends to suppress the detection of errors.

4. SRES A2 Error Propagation

4.1 The temperature uncertainty due to cloud error: Cloud error as present in the A2 scenario was calculated by first extracting greenhouse gas forcings equivalent to those used by the IPCC to calculate the original A2 scenario. The compatibility of the passive warming model projection with real GCM projections (ACoB Figure 2, Figure S5, and Table S1) validates applying the same greenhouse gas forcings as used in the passive GHG model to the A2 projection.

Therefore, the A2 SRES Total Forcing was estimated as,

$$(\text{Year 2000 Forcing})\times[(\text{A2 Temp}+\text{Year 2000 Temp})/(\text{Year 2000 Temp})] \qquad \textbf{S3},$$

where the *Year 2000 Forcing* was the total forcing calculated from the concentrations of CO_2, methane, and nitrous oxide at their year 2000 levels, using the equations in Myhre, et al.[2] The *Year 2000 Temp* was the total increased greenhouse temperature due to the same gasses at their year 2000 levels, relative to their known or extrapolated year 1900 levels (see Figures S3 and S4), calculated according to equation S1. This was added to the SRES A2 year 2000 temperature, because at year 2000, the A2 projection begins at zero (0) degrees.

So,

$$\text{"A2 Total Forcing"} = (35.163)\times(\text{A2 Temp}+13.938)/(13.938) \qquad \textbf{S4},$$

where *35.163* is the total greenhouse forcing in the year 2000, in Wm^{-2}, derived according to Figures S3 and S4, and *13.938* is the water-vapor-enhanced greenhouse gas temperature contribution in Celsius at the year 2000 as calculated using equation S1. In the A2 projection, the net temperature increase starts at zero C in the year 2000.

Per Year Cloud Error for SRES A2 for each year is then,

Yearly Error=(A2 Temp+Year 2000 Temp)×[(Cloud Forcing Error/(Total Forcing)] **S5**,

where *A2 Temp* is the temperature of the A2 projection in each year starting at year 2000, *Year 2000 Temp* is the surface temperature of the year 2000 calculated using equation S1, *Cloud Forcing Error* is ±2.8 Wm^{-2} as obtained by applying the equation S2 result (see the discussion in ACoB), and *Total Forcing* in Wm^{-2} is the extracted A2 forcing, as calculated by equation S4.

4.2 ***The structure of GCM cloudiness retrodiction residuals***: The average cloudiness per latitude as retrodicted by the GCMs (ACoB Figure 3) are one-dimensional (univariate) spatial data, which makes them structurally identical to a linear time series. That allows the unfit residuals to be tested for autocorrelation using simple one-dimensional statistical models. As justified below, the final error propagation model is equivalent to a non-stationary random-walk-like accumulation model in part justified by an assessment that the statistical structure of the unfit average cloudiness residuals are all consistent with unit root autoregressive models.

Figure S6 shows the structure of a Lag-1 plot from a random number series ranging between zero and 100, used here to illustrate an uncorrelated one-dimensional error series.

Figure S7 shows the same series as a bar-graph normalized to the standard deviation (SD) of the series. No pattern emerges. These Figures demonstrate how uncorrelated errors should appear to visual inspection. Uncorrelated Gaussian-normal errors do not accumulate in a long set of iterated outputs from a model.

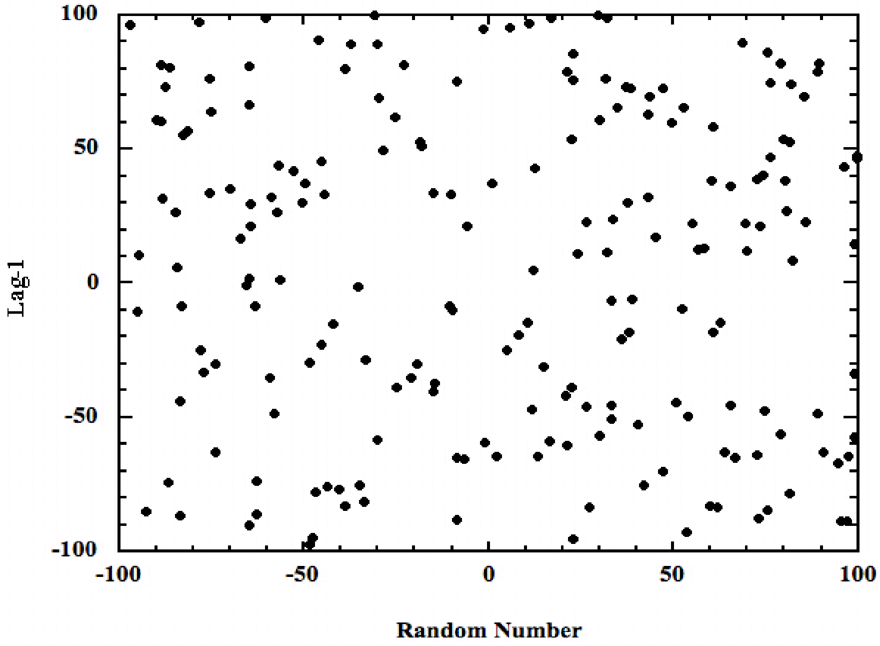


Figure S6: Lag-1 plot of "n" versus "n-1" of a ±100 random number series, where n=1,100. The scatter of points shows the non-correlation that would also appear in a Lag-1 plot of errors in any data set for which the measurement or calculational errors are random and uncorrelated.

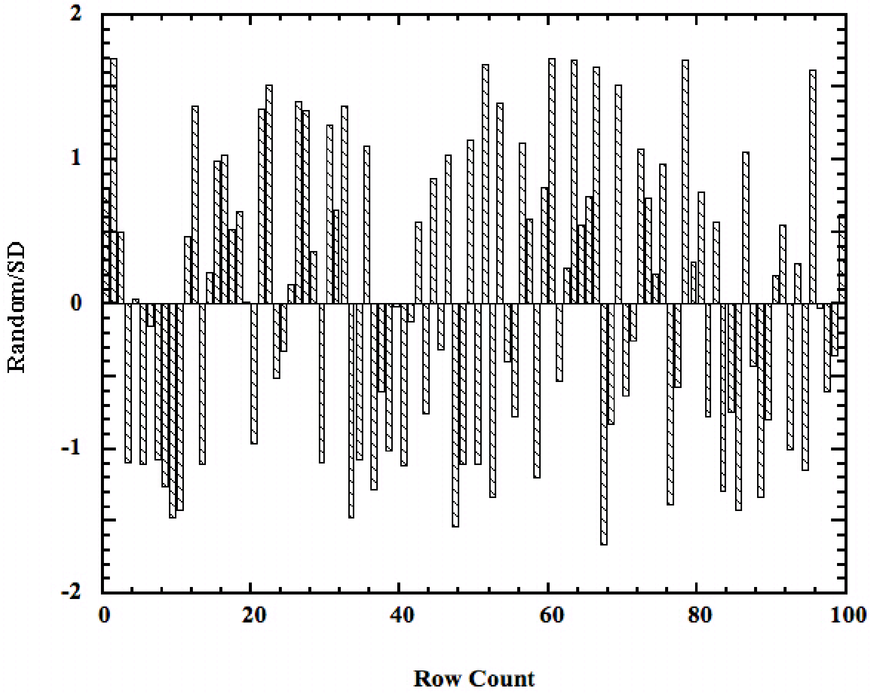


Figure S7: A bar graph of the same random number series as shown in Figure S6, plotted as units of standard deviation. There are no trends or patterns, as expected from an entirely uncorrelated series. Uncorrelated measurement or calculational errors should show a similar lack of pattern.

Figure S8 shows the Lag-1 plot of the unfit residual of the CNRM GCM cloudiness retrodiction shown in ACoB Figure 3. The unfit residual is what is left when the observed cloudiness is subtracted from the GCM retrodicted cloudiness. In a perfectly accurate retrodiction, the unfit residual would be a series of zeros. For the CNRM residual, there is evidence of high autocorrelation in the error values; that is, the value of every Y_n is linearly dependant in some way from the value of the preceding Y_{n-1}. Autocorrelation means that the CNRM unfit residual includes at least one uncompensated trend that extends across the entire data set (there are no outliers). All of the GCM cloudiness error residuals showed high autocorrelation (Table S2).

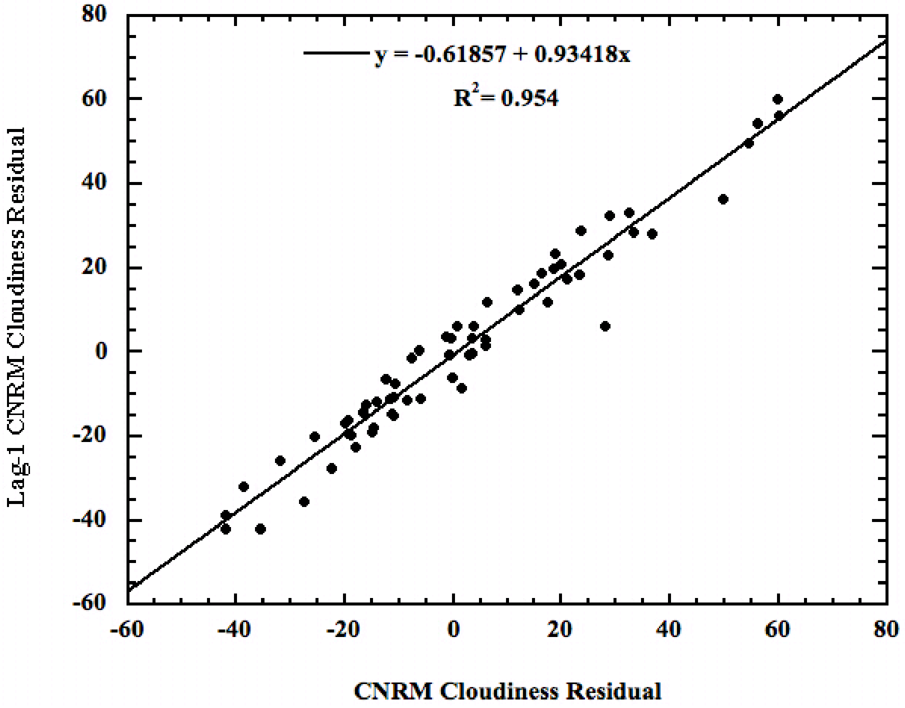


Figure S8: (●), Lag-1 plot of the unfit residual of the CNRM prediction of 1979-1988 average cloudiness, and; (—), a linear regression fit to the data. The negative percents arise when retrodicted cloudiness exceeds observed cloudiness. Compare with Figure S6.

Figure S9 shows the CNRM unfit residual as a bar-plot. The residual series of each of the ten GCMs similarly showed a structured profile. These residuals clearly do not have a constant variance (SD^2) across the entire data set, and are consistent with non-stationarity.

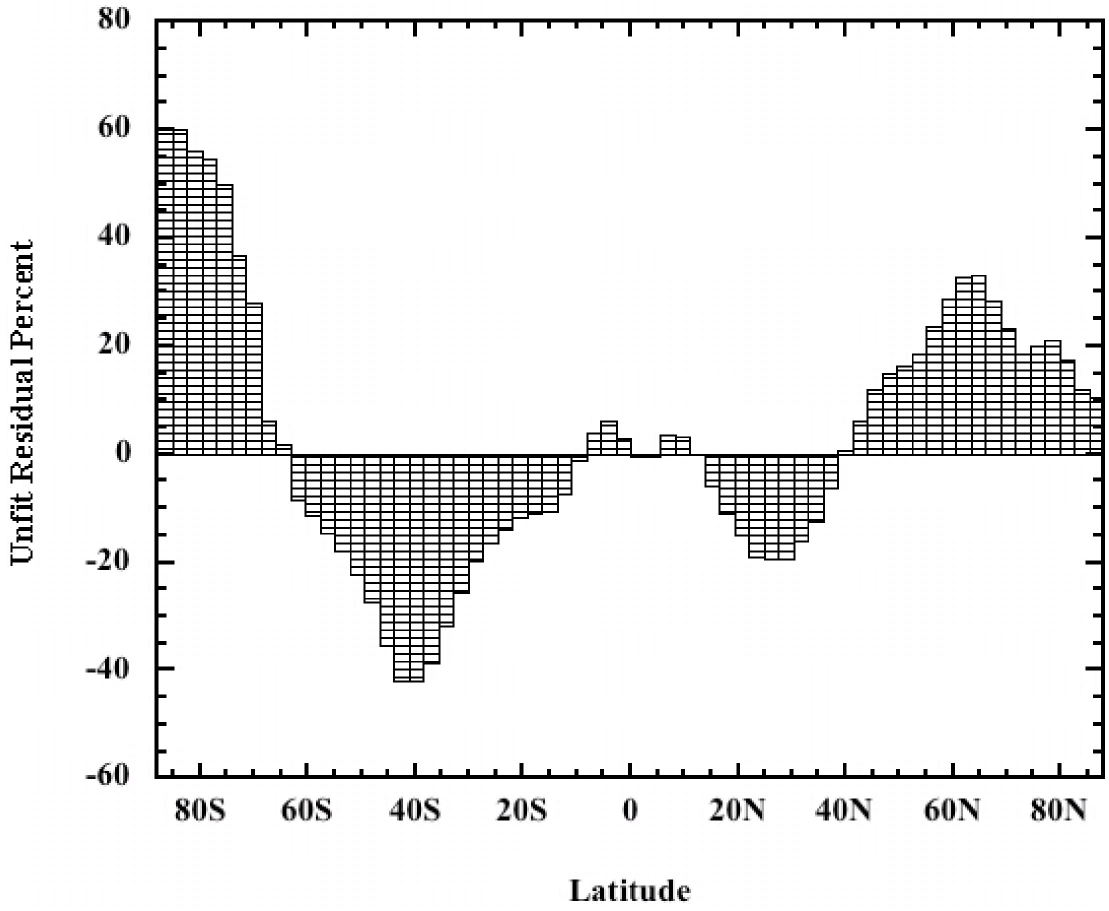


Figure S9: CNRM cloudiness unfit residuals, plotted as (retrodicted minus observed) cloudiness percent difference. Compare with Figure S7.

Figure S6 through Figure S9 show that the GCM cloudiness retrodiction errors do not behave as Gaussian random noise. Non-symmetric error will not self-cancel across an iteratively calculated series.

Random error cancels as √N as N series are averaged. The random noise component in each 10-year GCM average cloudiness projection should already have been reduced by a factor of about √10=3.16-fold, relative to any Gaussian random noise component in each of the ten 1-year realizations of each of the GCM average cloudiness projections. Figure S10 illustrates this effect. If the ±10.1% average cloudiness error was due to residual random noise that survived the averaging process in each of the 10-averages, then the error due to random noise in each 1-year GCM realization necessarily averages about 3.16 times larger. That is, the average cloudiness error due to random excursions alone would be about 32%. That means a about ±8.8 Wm^{-2} uncertainty would be present in every temperature calculation by a GCM. That uncertainty is 3.3× larger than the effect of all the human-produced greenhouse gasses presently in the atmosphere. However, as will be shown below, the ±10.1 % is unlikely to represent random error.

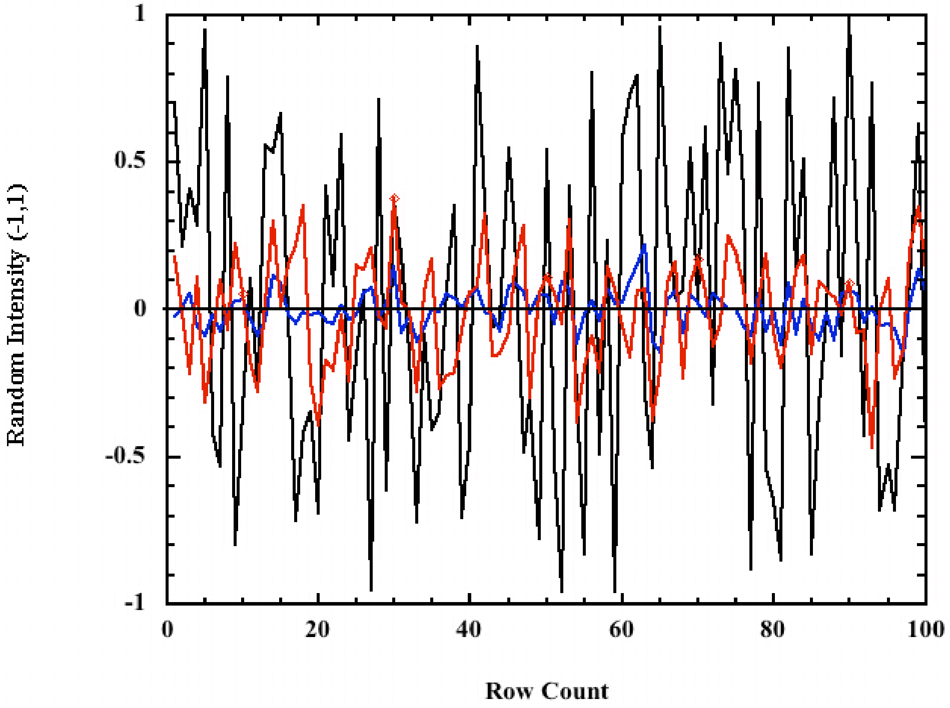


Figure S10: The √N effect of averaging random noise error: (—), random number series between the limits of 1 and -1; (—), the average of 10 such series, and; (—), the average of 80 such series. The standard deviations of these specific series have the ratios of 1, 0.328 (1/√10=0.316), and 0.120 (1/√80=0.112), respectively.

To further test the structure of the GCM cloudiness error, the correlation matrix for all ten residual series was calculated using the Student t-Test for paired data with equal variance. This choice of conditions follows from the fact that all the GCMs targeted the same observed cloudiness, all began with similar (if not identical) empirical inputs, and all employed similar numerical methods. If each residual series represented random errors, their pair-wise correlations should be approximately zero (0). On the contrary, the correlation matrix in Table S3 shows that high to very high pair-wise correlations occur very often between unfit residual series.

Again, the correlation between paired series of random errors should be zero (0). However, happenstantial correlation can occur in small independent data sets. Tests for happenstantial correlation across some computer generated random series of 200 points showed small pair-wise correlations averaging 0.05±0.05. In comparison, of the 45 GCM cross correlations only six are |0.10| or less (BMRC with MPI or YONU; MRI with CNRM; LMD, MPI, and NRL with SUNGEN).

Table S3: Pair-wise Correlation Matrix of the Residuals From Ten GCM Cloudiness Retrodictions[10]

	BMRC	CNRM	DERF	LMD	MPI	MRI	DNM[a]	NRL	SUNGEN	YONU
BMRC	1.0000	0.1959	-0.1141	0.3138	0.08270	0.1664	-0.1141	0.4483	-0.1866	-0.03853
CNRM		1.0000	0.8656	0.4413	0.9748	-0.1021	0.8656	0.6461	-0.3379	-0.4131
DERF[a]			1.0000	0.1712	0.9067	-0.5032	1.000	0.3186	-0.5024	-0.5113
LMD				1.0000	0.3979	0.05247	0.1712	0.5594	0.1875	-0.6244
MPI					1.0000	-0.09002	0.9067	0.5776	-0.2837	0.3163
MRI						1.0000	-0.5032	0.1819	0.8120	0.4598
DNM							1.0000	0.3186	-0.5024	-0.5113
NRL								1.0000	-0.08430	-0.1731
SUNGEN									1.0000	0.2716
YONU										1.0000

a. The revised DERF and DNM projections overlaid exactly in Figure 18b of Gates, et al.[10] It must be presumed here that this surprising coincidence nevertheless reflects valid realizations.

For these GCMs, non-correlation of the residuals might be empirically inferred. However, four correlations are >|0.90| (CNRM with MPI; DERF with MPI and DNM, and; MPI with DNM), thirteen are |0.90|>corr.>|0.50|, and twenty-two are |0.50|>corr.>|0.10|.

Taking a relatively generous ~±0.1 to be the standard deviation for accidental correlation among small data sets of random errors, then of the 45 cross-correlations 31 should be within ±0.1, and a further 12 should be within ±(0.1-0.2). Alternatively, in a normal distribution with SD=±0.1, four random error series with cross correlations >0.9 would appear after sampling about 10^{17} error series. Even discounting the surprising identity of the DERF and DNM average cloudiness projections (see Table S3, footnote 'a.') and thus dropping the number of >0.90 correlations to two, the enormous improbability of the result would not significantly change. Even correlations ≥|0.5| (i.e., 5 SD's) should constitute no more than about 5.6×10^{-6} % of a sampled set of random error series.

Although Lag-1 autocorrelation as in Figure S8 can also arise in net trendless red noise, two lines of evidence argue against this source. First, the GCM cloudiness retrodictions are each the average of at least ten individual model realizations. The intensity of any random redness possibly present in each realization should have been reduced by a factor of about 3.16 across a 10-average. Because the structure of the redness (i.e., the positions and durations of happenstantial trends) in GCM outputs should vary randomly from realization to realization, a ten-average should fragment any redness into small segments that no longer dominate an entire series (if red noise ever did dominate any of the individual GCM cloudiness projections). This result is not consistent with persistent Lag-1 behavior through every 10-year average, nor with the dominance of pair-wise residual correlations in Table 3. Second, the domination of pair-wise correlations across the GCMs residual series (Table S3), in and of itself, also argues against trendless red noise as the source of the correlations unless it can be shown that GCMs produce random errors with consistent inter-GCM reddening trends in the same physical output ranges. In small data sets such uncompensated correlated trends could dominate the residuals and appear as cross-correlations. The necessity of consistently producing correlated random trends both inter-GCM and in the same data ranges savages the usual definition of “random,” however. These considerations argue against uncompensated red noise as the source of the observed correlation in the GCM cloudiness residuals.

Instead, the high Lag-1 and GCM pair-wise correlations of residuals imply that the unfit cloudiness residuals derive from systematic errors in theory that are often similarly expressed among GCMs. From Table 3, where they are wrong, GCMs are often wrong in the same way (or inversely, e.g., LMD vs. YONU).

Finally, every residual series was subjected to the Phillips-Perron test for unit roots in the AR(1) autoregressive model, $Y_l = \phi Y_{l-1} + \varepsilon_l + u_l$, where Y_l is the unfit cloudiness residual at a given latitude, Y_{l-1} is the latitudinally preceding datum, ϕ is the autocorrelation coefficient, ε_l is white noise, and u_l represents a linear latitude-dependent trend in the residuals. If the coefficient $\phi=1$ (unit root), the series is a non-stationary random walk. In random walks, the error accumulates additively across the series.

Autoregressive models without a trend ($u_l=0$) or with a trend were tested. These tests were carried out by Prof. Ross McKitrick, Department of Economics, University of Guelph, whose generosity of time and attention is acknowledged with gratitude.

The results are presented in Tables S4 and S5.

Table S4: Phillips-Perron Unit Root Test without Trend

GCM Error Series P-P Statistic	Interpolated Dickey-Fuller % Critical Value			MacKinnon p-value for Z(t)
	1%	5%	10%	
BMRC Z(ρ)=0.792	-19.422	-13.532	-10.874	---
BMRC Z(t)=0.302	-3.539	-2.907	-2.588	0.9763
CNRM Z(ρ)=-1.473	-19.134	-13.404	-10.778	---
CNRM Z(t)=-0.505	-3.562	-2.920	-2.595	0.8296
DERF Z(ρ)=-2.911	-19.134	-13.404	-10.778	---
DERF Z(t)-0.848	-3.562	-2.920	-2,595	0.8058
LMD Z(ρ)=-12.454	-18.832	-13.268	-10.680	---
LMD Z(t)=-1.846	-3.587	-2.933	-2.601	0.3577
MPI Z(ρ)=-1.544	-19.134	-13.404	-10.778	---
MPI Z(t)=-0.526	-3.562	-2.920	-2.595	0.8883
MRI Z(ρ)=0.763	-18.900	-13.300	-10.700	---
MRI Z(t)=0.259	-3.500	-2.930	-2.600	0.9744
DNM Z(ρ)=2.911	-19.134	-13.404	-10.778	---
DNM Z(t)=-0.848	-3.562	-2.920	-2,595	0.8058
NRL Z(ρ)=-0.012	-19.242	-13.452	-10.810	---
NRL Z(t)=-0.017	-3.553	-2.915	-2.592	0.9570
SUNGEN Z(ρ)=-3.965	-18.900	-13.300	-10.700	---
SUNGEN Z(t)=-1.056	-3.500	-2.930	-2.600	0.7327
YONU Z(ρ)=0.715	-18.628	-13.172	-10.620	---
YONU Z(t)=0.242	-3.607	-2.941	-2.605	0.9736

Newey-West lags=3 in all cases.

Table S5: Phillips-Perron Test with Trend

GCM Error Series P-P Statistic	Interpolated Dickey-Fuller Percent Critical Value			MacKinnon p-value for Z(t)
	1%	5%	10%	
BMRC Z(ρ)=1.659	-26.686	-20.322	-17.206	---
BMRC Z(t)=0.729	-4.086	-3.471	-3.163	0.9987
CNRM Z(ρ)=0.039	-26.142	-20.034	-16.982	---
CNRM Z(t)=-0.005	-4.121	-3.487	-3.172	0.9955
DERF Z(ρ)=-3.112	-26.142	-20.034	-16.982	---
DERF Z(t)=-0.917	-4.121	-3.487	-3.172	0.9545
LMD Z(ρ)=-8.113	-25.572	-19.724	-16.752	---
LMD Z(t)=-0.971	-4.159	-3.504	-3.182	0.9479
MPI Z(ρ)=-0.550	-26.142	-20.034	-16.982	---
MPI Z(t)=-0.218	-4.121	-3.487	-3.172	0.9924
MRI Z(ρ)=2.580	-25.700	-19.800	-16.800	---
MRI Z(t)=1.196	-4.150	-3.500	-3.180	1.0000
DNM Z(ρ)=-0.005	-26.142	-20.034	-16.982	---
DNM Z(t) =-0.917	-4.121	-3.487	-3.172	0.9545
NRL Z(ρ)=1.678	-26.346	-20.142	-17.066	---
NRL Z(t)=0.761	-4.108	-3.481	-3.169	0.9988
SUNGEN Z(ρ)=-3.960	-25.700	-19.800	-16.800	---
SUNGEN Z(t)=-1.093	-4.150	-3.500	-3.100	0.9300
YONU Z(ρ)=-0.580	-25.188	-19.496	-16.608	---
YONU Z(t)=-0.218	-4.187	-3.516	-3.198	0.9924

Newey-West lags=3 in all cases.

In the Phillips-Perron test, the null hypothesis is that $\phi=1$. Verification of the null hypothesis is met when the calculated $Z(\rho)$ and $Z(t)$ are larger than the Dickey-Fuller Critical Value. In every case but one, the magnitudes of both $Z(t)$ and $Z(\rho)$ are indeed much larger than all the respective Dickey-Fuller Critical Values. The exception is $Z(\rho)$ for LMD with no trend (Table S4), which falls between the 5% and 10% Dickey-Fuller limits, marking a slightly weaker verification (cf. Figure S11). However, $Z(t)$ for this entry is much larger than the Dickey-Fuller values. The same LMD residual series easily passed the unit root test for trended data (Table S5, see also Figure S11), as do the rest of the cloud unfit residual series.

The $\phi=1$ null hypothesis is also verified if $p>0.05$ for $Z(t)$. In every case but one (LMD, Table S4), $p>0.7$ at least and in many cases $p>0.9$. Overall, then, concluding an autoregressive unit root is strongly validated for every series of cloudiness residuals.

An explanation for the LMD result in the Phillips-Perron no-trend test can be found in Figure S11, which compares the LMD and MPI residual series. All the residual series resembled the MPI residual, except the LMD series. The relatively truncated extension of the latter artificially produced a net linear slope across the series that behaves as a latitude-dependent trend in the Phillips-Perron test. Hence the somewhat less powerful unit root result under the no-trend assumption (Table S4), but the very strong unit root result when a trend is included in the test (Table S5).

From a statistical perspective, the GCM cloudiness residuals can thus be "considered as a random walk, which is not covariance-stationary,"[11] across latitude. This result vitiates a statistical argument that cloudiness errors can not accumulate across GCM-derived predictions.

These findings can be applied to time-wise temperature projection errors. The Phillips-Perron tests show the latitudinal residuals are not consistent with random error, but with a systematic accumulating error. Systematic error can be deposited into a time-wise 10-year average only by a systematic error that is propagated time-wise forward from each 1-year GCM realization.

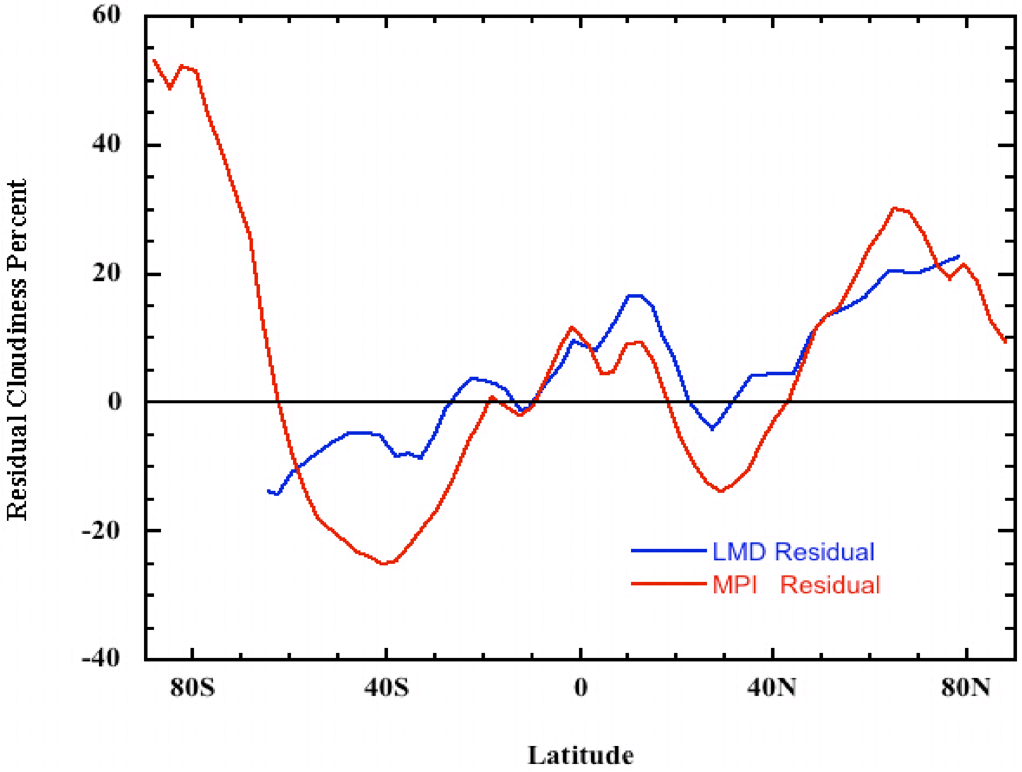


Figure S11: The unfit cloudiness residual for: (—), the LMD retrodiction, and; (—), the MPI retrodiction. All the other GCM residual series resembled the MPI series.

If any autocorrelated latitudinal cloudiness residuals in 1-year GCM realizations were time-wise uncorrelated in multi-year projections, the 1-year latitudinal autocorrelation would be progressively lost during time-wise averaging. Cloud error that is latitudinally autocorrelated but time-forward uncorrelated and Gaussian random would, when time-wise averaged, produce relatively smaller residual error excursions in a 10-year average (as in Figure S10), with weakened or absent latitudinal Lag-1 correlations, and with residuals oscillating in some relatively low-intensity fashion about the cloudiness zero difference line when plotted as in Figure S11. But that is not observed in these residuals, neither visually nor by statistical test. The remaining latitudinal errors in the 10-year averages are in fact strongly autocorrelated (Figure S8, Table S2).

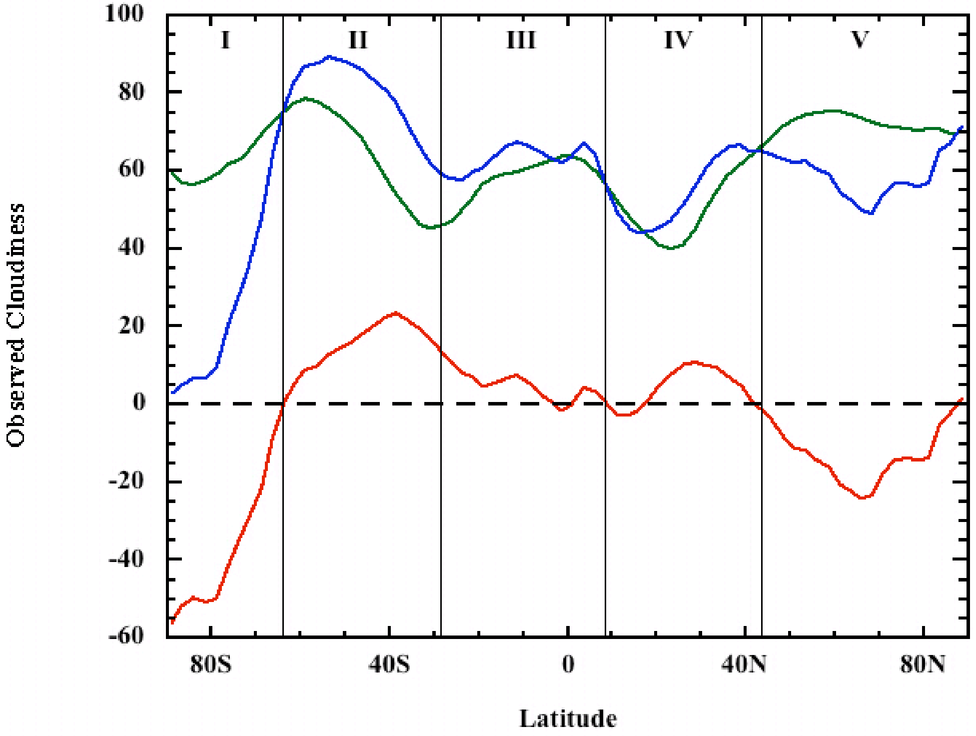


Figure S12: (—), Observed global average cloudiness (1983-1990); (—), 100-Average GCM retrodiction, and; (—), Residual (observed minus retrodicted). The vertical lines mark the regions (Roman numerals) over which integrations were carried out for the retrodiction error metric. See Section 4.4 for this discussion.

With respect to the error intensity reduction following averaging, Figure S12 shows the cloudiness residuals remaining from the ensemble average of all 100 GCM cloudiness projections. Any random or red noise component in the individual 1-year GCM realizations should have been reduced by a factor of 1/√(100)=1/10th. Alternatively, any remaining Gaussian or red noise components in each of the 10-average GCM cloudiness projections should have been reduced by 1/√(10)=0.316. However, the 100-average residuals are only modestly diminished relative to the 10-average residuals of Figure S11. This result is further discussed in Section 4.4.

Figure S13 shows a Lag-1 plot of the 100-average residual. The strong autocorrelation of the residual of the 100-avreage is completely comparable with those of the 10-year GCM averages, Figure S8 and Table S2. There is no reduction in latitudinal autocorrelation of error. This behavior is not consistent with time-wise uncorrelated Gaussian errors when they are propagated through a further ten-average. The result is consistent, however, with systematic theory-bias errors that remain correlated in a time-wise projection and that can be opposingly expressed in different GCMs; a trait already indicated by correlations of residuals that can be either strongly positive or strongly negative (Table S3).

This reduces or removes the possibility that time-wise (as opposed to latitudinal) GCM cloudiness errors are uncorrelated or reflect only red-noise trends. The remaining possibility is that the 10-average and 100-average Lag-1 autocorrelations are due to a systematic bias in the cloudiness retrodictions that can propagate forward in a time series projection. Therefore, time-series GCM projections do not produce only uncorrelated non-accumulating cloud errors (also see the discussion of numerical uncertainty propagation in Section 4.3, the lowest limit theory-bias uncertainty estimate in Section 4.4, and the discussion of time-wise propagation of Gaussian error in Section 4.5). Cloud-error propagates time-wise forward, as well as latitudinally, and can accumulate in a futures temperature projection producing an increasing uncertainty.

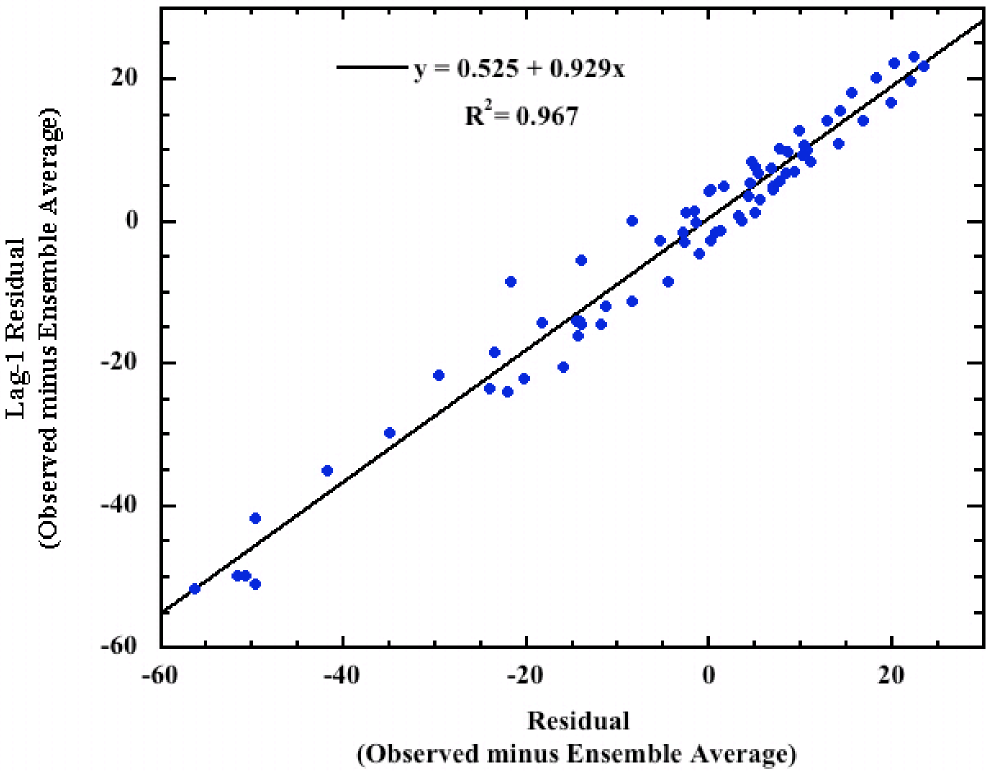


Figure S13: Lag-1 plot of the residuals remaining from the difference between the 1983-1990 observed average cloudiness and the further 10-fold ensemble average retrodiction of the original 10-year averaged retrodiction of the 10 GCMs in article Figure 3.

4.3 The theory-bias physical uncertainty propagation model:
The coincidence of the line produced by Equation 1 with the lines produced by the climate models, as displayed in article Figure 2, shows that increasing forcing propagates linearly through the climate models to produce increasing temperature. As does the mean value of forcing, the mean±(uncertainty) in forcing, i.e., the high/low values of forcing represented by the mean±1SD, likewise linearly propagate through the models to produce ±1SD in calculated temperature.

The statistical analysis of errors in SI Section 4.2 showed that the cloud error was unit root, and so a latitudinal accumulation (Brownian) error model was justifiable. The consistent autocorrelation of errors in the 10- and 100-averages showed these errors were very <u>un</u>likely to be time-wise Gaussian. Persistent red noise is unlikely to survive 10- and 100-averages. The error correlation matrix, Table S3, showed that the various GCMs produced highly correlated errors, which are statistically very unlikely to be the result of independent random processes. This result again implies systematic rather than Gaussian cloud errors. The same inference turned up in the analysis described in SI Section 4.4 (see below), where cloud error does not cancel as expected from a purely Gaussian model, i.e., as $1/\sqrt{N}$. The conclusory inference is that the cloud error represents theory-bias, rather than Gaussian noise, and so does not cancel in a time-wise projection.

Proceeding from Section 4.2, assessment of the SRES A2 uncertainty limits followed a physical model wherein each GCM calculated future temperatures as $t_n=f(t)+e_1(t_n)+e_2(t_n)$, where t=temperature, $f(t)$ is the temperature generator (e.g., equation S1), $e_1(t)$=statistical error (i.e., the measurement and numerical errors, each assumed to be Gaussian), and $e_2(t)$=systematic error (i.e., theory bias). This analysis is concerned with the physical theory-bias errors in GCM projections and so the statistical errors, $e_1(t)$, were ignored (but see Section 4.5).

In theory-bias error, the modeled value of t_n and the physically true value of t_n diverge systematically, continuously, and linearly (at best) by some unknown amount within $\pm e_2(t_n)$ for every iterated new calculation of t_n. This error arises because of the inaccuracies, approximations, or incompleteness of the theoretical description.

The justification for the error model described below is that for each prior year, Y_{n-1}, in a time-wise projection, the climate with its forcings f_{n-1}, represents the initial conditions entering the calculation of the climate of each subsequent year Y_n, along with the nth year forcings f_n. This has been described as "the time-marching method."[12]

For each year Y_n, the limiting values of $f_n+e_2(f_n)$ and $f_n-e_2(f_n)$, empirically calculated from the mismatch between prediction and measurement, capture the uncertainty from theory-bias. The resultant discrete predictive divergence of each calculated t_n away from the true t_n due to theory bias is unknown, but is present in every calculated t_n. Because of the ambiguity in both the magnitude and the sign of the theory-bias error, each of which may vary from GCM to GCM, every prediction of a new t_n from theory can only be represented by the value range of its theory-bias uncertainty limits estimated from the difference between projection and observation.

The two limiting uncertainty values of forcing, $f_n \pm e_2(f_n)$, are themselves necessarily valid physical estimates of forcing because they represent the average range of the possible magnitude of each f_n. That is, the mean magnitude, f_n, does not exhaust the magnitude of the forcing for year Y_n. Each of the two limiting values of the uncertainty in calculated temperature, i.e., the $t_n \pm e_2(t_n)$'s calculated from the $f_n \pm e_2(f_n)$'s, are thus as valid an estimate of the future true t_n as the calculated mean t_n value itself. The limiting uncertainty values, $f_n \pm e_2(f_n)$, therefore cannot be excluded from a serially iterative projection. Every calculated yearly forcing, f_n (n=1,2,3,...), must thus include the mean value of f_n and the high and low values of f_n representing the average systematic theory-bias uncertainty $\pm e_2(f_n)$, i.e., $f_n \pm e_2(f_n)$, in each iterative step. The uncertainty in forcing due to theory bias, $\pm e_2(f_n)$, in each Y_n then produces an uncertainty in the corresponding calculated temperature t_n,.

The yearly modeled increased temperature t_n (n=1,2,3,...,n) then equals the starting temperature, t_0, plus the sum of all the calculated fractional increases in temperature, $\Delta t_{1,2,\ldots,n}$. But every $\Delta t_{1,2,\ldots,n}$ includes the high and low values, $\pm e_2(t_{1,2,\ldots,n})$, that capture the uncertainty in forcing due to theory bias. These values of $\Delta t_{1,2\ldots n} \pm e_2(t_{1,2,\ldots,n})$ are temperature estimates that have a physical validity equal to that of the mean value, $\Delta t_{1,2,\ldots,n}$, itself. All three values, t_n, $t_n+e_2(t_n)$ and $t_n-e_2(t_n)$, enter the representation of each new t_n and are propagated forward in a time-wise projection, as in S6:

$$t_n=t_0+\sum_{1}^{n}\ [\Delta t_1 \pm e_2(t_1)]+[\Delta t_2 \pm e_2(t_2)]+[\Delta t_3 \pm e_2(t_3)]+\ldots+[\Delta t_n \pm e_2(t_n)] \qquad \textbf{S6},$$

where the $\Delta t_{1,2\ldots,n}$ is the fractional increase in temperature in Celsius for each given year n due to increased greenhouse gas forcing Δf_n, and $\pm e_2(t_n)$ contains the high and low physical estimates for Δt_n, based on the $\pm e_2(f_n)$ cloud error or GHG forcing error. This systematic error propagates additively with every iteration of a progressive calculation and produces an accumulating uncertainty.

The physically modeled temperature projected over time therefore spreads into a vertex of values, with the diameter across the vertex increasing as the magnitude of $\pm e_2(t_n)$. This vertex captures the predictive uncertainty in future global temperature.

The cloud forcing error contribution to temperature for each $\pm e_2(t_n)$ was estimated by equation S5. Estimation of the temperature uncertainty due to direct uncertainties in GHG forcing proceeded similarly, using the ranges given in Myhre.[2] Following from the analysis in Section 4.2, the cloudiness error estimate can reasonably be taken to capture a lower-limit average range of theory-bias uncertainty inherent in the 10 GCMs evaluated, which in turn arises from the way cloudiness is physically represented in these GCMs. Likewise the uncertainty in the GHG forcing will produce a systematic divergence in predicted and true temperatures, of unknown magnitude but of linear accumulation, which divergence is in turn captured within the forcing uncertainty range.

Thus for SRES scenario A2, the error $\pm e_2(t_n)$ for each year after, e.g., year 2000, was calculated as the simple sum of the antecedent errors. For example, the total systematic error at, e.g., projection year 100, $\pm e_2(t_{100})$, was calculated according to equation S7.

$$\text{Propagated Error} = \sum_{n=1}^{100}(peryearclouderror) \qquad \textbf{S7}$$

The total cloud uncertainty for each year includes the error for that year plus the sum of the errors of the antecedent years, finally summed across 100 years for a centennial projection.

4.4 A lowest-limit estimate of cloud theory-bias error: The error analysis was extended to the ensemble average of all ten GCM cloudiness retrodictions, shown in Figure S12. This ensemble amounts to a 100-average, i.e., (10 GCMs)×(10 years per GCM).

A standard error of the 100-average was calculated by first dividing the latitudinal percent cloudiness in Figure S12 into five regions. The observed or projected average cloudiness was then integrated in each region. The percent error represented by each of the five latitudinal regions was calculated from the integrated residual, as [(observed-retrodicted)/observed]×100, yielding a Δ % error for each region. The overall standard error for the 100-average was calculated by the usual formula,

$$s = \sqrt{[\sum_{n=1}^{5}(\Delta\%)_n{}^2]/4}.$$

This error was ±6.0 percent.

For comparison, the direct error calculated as in equation S2 was very comparable: [(total integrated ensemble cloudiness residual)÷(integrated observed cloudiness)]×100=6.2%.

This result allows extraction of a lowest-limit estimate for accumulating theory-bias error. Suppose, speculatively, that the minimal ±10.1 % error calculated in Section 3 was entirely due to uncorrelated noise, rather than due to the theory-bias that the analysis of Section 4.2 implies. Then the further 10-average should have reduced the error magnitude by √10-fold=1/3.16. That is, the 100-average standard error should have become (±10.1/√10) %=±3.2 %, if all of the ±10.1 % was due to uncorrelated Gaussian noise.

However, the ±10.1 % error instead was reduced to ±6.0 % by the new 10-average. That means at most only part of the ±10.1 % could have been due to Gaussian noise, while the rest must be due to a systematic error, namely theory-bias. This theory-bias would represent the very minimum lowest-limit estimate of GCM theory-bias that could be in the calculated cloud error, because it is derived from a lower-limit estimate of total cloud error.

If the total error in every case is the sum of an uncorrelated Gaussian noise component, $e_1(t)$, plus a theory-bias component, $e_2(t)$, then two linear equations can be set up:

1. $\pm[e_1(t)+e_2(t)]=\pm 0.101$, from the individual 10-averages in equation S2.
2. $\pm\{[e_1(t)/\sqrt{10}]+e_2(t)\}=\pm 0.060$, from the further all-GCM 10-average of Figures S12 and S13.

Solving these equations for $e_1(t)$ and $e_2(t)$ yields a lowest-limit estimate of the constant average theory-bias error, as well as of the Gaussian noise error remaining in the original GCM average cloudiness retrodictions (i.e., the original 10-averages).

These are: random error = $e_1(t)=\pm 0.060$, and theory-bias error = $e_2(t)=\pm 0.041$, for the respective contributions to the ±10.1 % average error as present in each of the original 10-average GCM cloud predictions.

From that, the average Gaussian noise error, $e_1(t)$, in each single-year GCM cloudiness projection of the ten individual 10-year GCM cloudiness projections is:

GCM per-year projection average $e_1(t)=\pm0.06\times\sqrt{10}=\pm0.19$, or ±19 %.

This error would also accumulate over an iterative calculation as noted in Section 4.5 below, and is a minimum estimated average uncertainty stemming from the average cloudiness error of the ten GCMs that would contribute to the uncertainty of projected future global temperature, equivalent to an instantaneous global average forcing uncertainty of ±5.3 Wm^{-2}.

More interesting is the lowest-limit theory-bias average cloudiness error estimate of ±4.1 %. This error would accumulate as increasing uncertainty in an iterative calculation, and would not be reduced by performing arbitrarily large numbers of GCM realizations at each t_{n-1} projection point.

From equation S5, the cloudiness error calculation is linear. Thus, to transfer this lowest limit estimate of theory-bias error to the temperature uncertainties calculated following Figure 4 in the article, merely multiply the numbers there by (0.041/0.101)=0.41. That is, the cumulated SRES A2 uncertainties from cloudiness error are reduced to 41% of their Figure 4 values.

So, from application of a bare lowest-limit estimate of theory-bias error, the year 2100 SRES A2 projected temperature increase with its associated physical uncertainty would be (3.7±46) C. Likewise, after only 4 years the SRES A2 projected warming plus its physical uncertainty would be (0.25±1.8) C, and after 50 years, (1.3±23) C.

That is, even with the lowest-limit estimate of cloud theory-bias error, the accumulating uncertainties make GCM temperature projections predictively useless. Our knowledge of future climates is left at zero, even over the short term.

4.5 An aside about numerical error in GCM temperature projections.

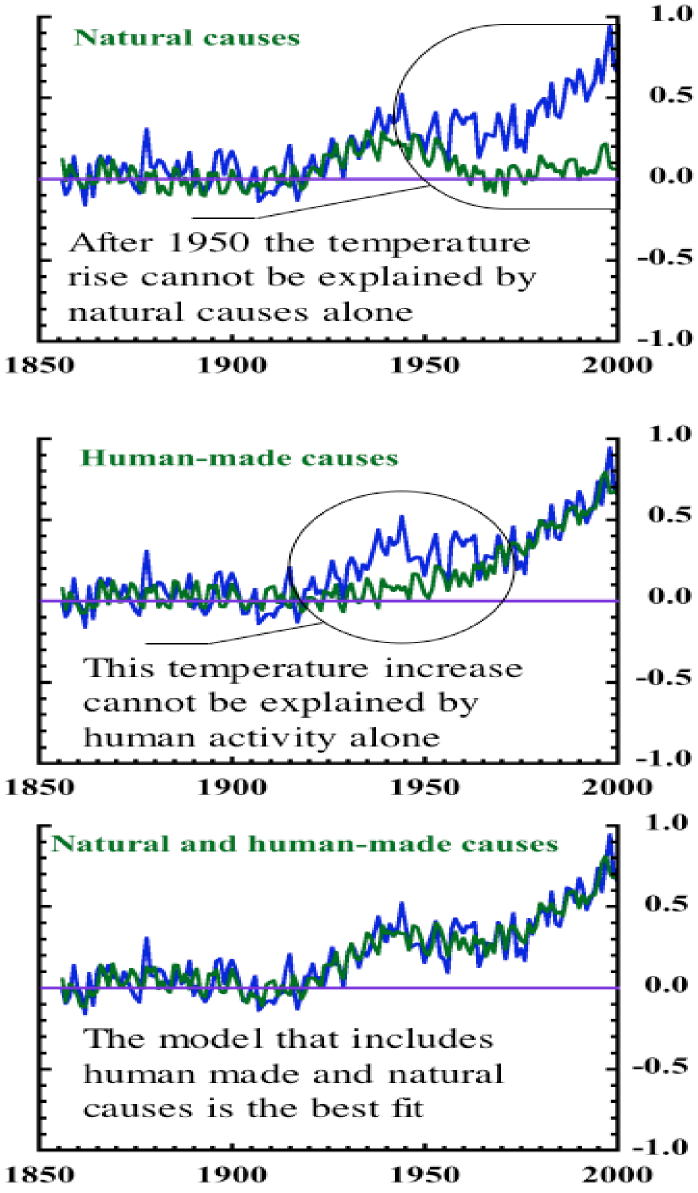
As an important aside, the numerical uncertainty in GCM temperature projections is typically presented as though it were associated with a time-series of independently measured points, each point having Gaussian error. When data points are independently measured or independently calculated, the Gaussian error associated with each point does not accumulate across a time-series. However, time-series GCM global average temperature projections are not at all like empirical independently measured time-series constructions; nor are they independently calculated. In a GCM projection, each point, t_n, in a time-wise temperature series projection is calculated using the t_{n-1} point as its initial condition. This calculational connection is entirely absent in time-series reconstructions with independent empirically measured or independently calculated data points, and entirely changes the meaning of the numerical uncertainty in a GCM time-wise climate projection.

For example when employing GCMs to produce the trends in Figure SPM-5, or , e.g., Figure SPM-4 or Figure TS-22 of the 4AR Technical Summary, WG1 and the IPCC are using the mean of each t_{n-1} point as the initial condition for calculation of each subsequent t_n, but discarding the ±(numerical uncertainty) of each t_{n-1} at each iterated forward step of the GCM temperature projection. Every t_{n-1} is thus iterated forward as though it were perfectly accurate, physically. But this is incorrect. The numerical uncertainty, $\pm e_1(t_{n-1})$, is a true uncertainty in the magnitude of each calculated temperature, t_{n-1}. That same numerical uncertainty becomes a real physical uncertainty when the calculated temperature t_{n-1} is propagated forward to calculate further physical quantities. That is, as soon as each t_{n-1} is propagated into and through the physical theory of a GCM, its uncertainty is transformed from numerical into physical. That means the entire numerical uncertainty around each t_{n-1} becomes a true physical uncertainty in the magnitude of each t_{n-1} and must become part of the physical initial state of each subsequent t_n. Every $\pm t_{n-1}$ must be included in the calculation of each subsequent t_n.

In practice, this means the numerical uncertainty accumulates across the years of a time-wise GCM temperature projection in a manner identical to the propagation of theory-bias error as described by equation S6 above. The accumulating numerical uncertainty then also produces a widening vertex of physical uncertainty about the mean temperature series, and the radius of this vertex increases with the number of projection years.

One could reduce the numerical uncertainty portion of the total physical uncertainty to an arbitrarily small (±) range by performing an arbitrarily large number of GCM realizations at each t_{n-1}. But Gaussian numerical uncertainty only asymptotically approaches zero and would never reach it. The residual numerical uncertainty would always accumulate as physical uncertainty across an iterative time-wise global average temperature projection. Therefore the numerical uncertainty limits about the GCM projected means, as displayed in Figures SPM-4, SPM-5, and TS-22 of the 4AR, are misrepresentations. The part of the total physical uncertainty in the displayed temperature trends, that which arises from the GCM numerical error alone, is far larger than actually presented and rapidly increases across the projected time series.

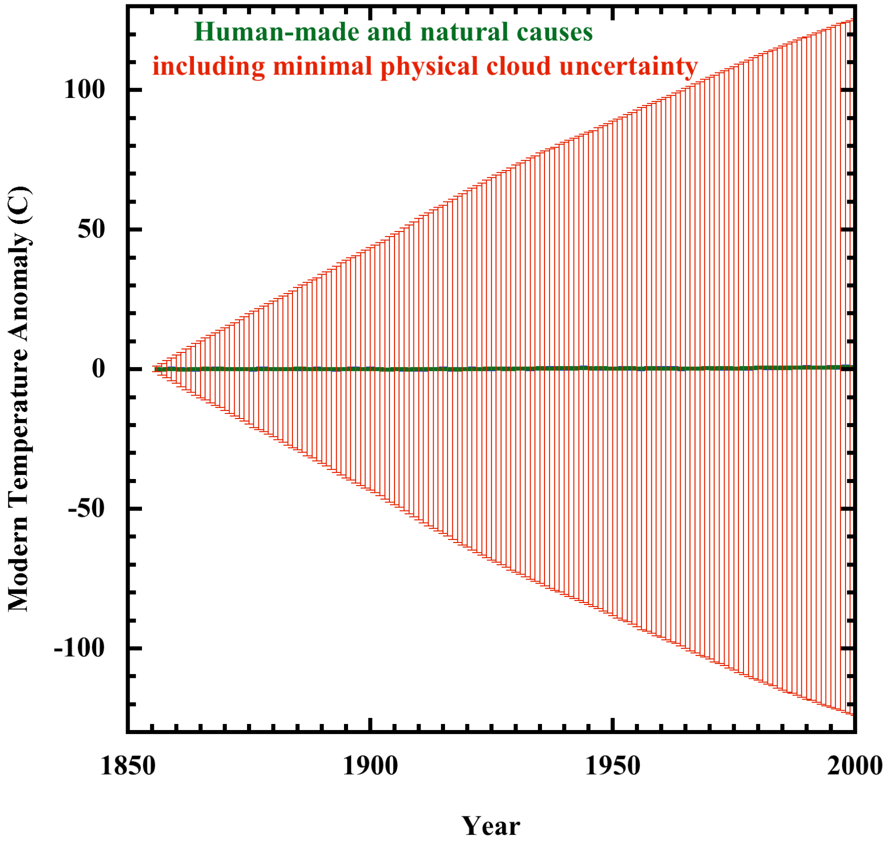
5. Concerning Figure 4 from the U.S. National Academy of Sciences brochure, “Understanding and Responding to Climate Change.”



Cartoon representing the message of Figure 4 from the National Academy brochure, “Understanding and Responding to Climate Change,” *dated March 2006.*[13] *The* **blue** *line is the global surface temperature record. The* ***green*** *lines represent GCM fits under the described forcing assumptions. The inset captions are direct quotes. The validity of this representation is easily checked by visiting the NAS website at* http://dels.nas.edu/basc/.

The original Figure 4 Legend is: "***Figure 4.*** *Simulations of past temperature more closely match observed temperature when both natural and human causes are included in the models. The gray [**green**] lines indicate model results. The red [**blue**] lines indicate observed temperatures. Source: Intergovernmental Panel on Climate Change.*"

The blue line is the canonical global average surface temperature record.[14] The green line is a cartoon representing the message conveyed by the General Circulation Model fit offered by the IPCC. As in NAS figure 4, the cartoon version of the GCM fit displays no error bars. The IPCC/NAS offer the fit as though it has perfect physical accuracy. Only very high physical accuracy can justify the unqualified statements of the inset captions.



How panel 3 of NAS Figure 4 might have looked if it included the physical uncertainty of minimal cloud error.

If the NAS/IPCC had opted to display the projection uncertainty due to minimal cloud error, however, panel 3 from Figure 4 might have appeared as above. The error bars were calculated as described in SI Section 4.

Minimal cloud error alone would make the uncertainty in the GCM projection about ±100 C by 1970 and about ±130 C by 2000 relative to claimed projected temperature anomalies of ~0.4 C and ~0.8 C, respectively.

This unreliability was somehow overlooked by the authors and reviewers of the National Academy of Sciences, as well as by the authors and reviewers of the IPCC.

6. The Greening of the Sahel

The Sahel (the southern borderlands of the Sahara Desert in Africa) has been getting greener.[15] The Normalized Difference Vegetation Index (NDVI) of the Sahel is plotted against the trend in atmospheric CO_2 in Figure S14. The NDVI is monitored by satellite and measures the change in vegetation. It is reckoned by the authors as particularly suited to study vegetation changes in semi-arid regions.

The authors looked hard to find an explanation for the greening trend. They wrote: "*To conclude, the strong secular trend of increasing vegetation greenness over the last two decades across the Sahel cannot be explained by a single factor such as climate. Increasing rainfall does explain some of the changes but not conclusively. Another potential explanation could be improved land management, which has been shown to cause similar changes in vegetation response elsewhere. However, the fact that millet yield and increasing vegetation greenness were unrelated does not support this explanation. The third factor, land use changes as a result of migration, is a plausible contributing explanation, but more empirical research is needed to verify this.*"

At the beginning of their paper, Olsson et al., wrote: "*The aim of this paper is therefore also to discuss and hypothesize possible factors other than climate* ***even though we are unable to determine clear links between vegetation change and these factors.***" (bolding added)

On the other hand, the correlation $r^2=0.71$ for the concurrent trends in Sahel NDVI and atmospheric CO_2.

Interestingly, in another article,[16] the same authors collaborated with climate modelers to further study the same question. They showed that most of the "*aggregated simulated vegetation changes*" could be explained by changes in precipitation. "*Aggregated simulated changes*" is a revealing description of what was explained. It means that the model explained its own predictions, which are only presumed to reflect the same causal chain as those responsible for the greening of the Sahel.

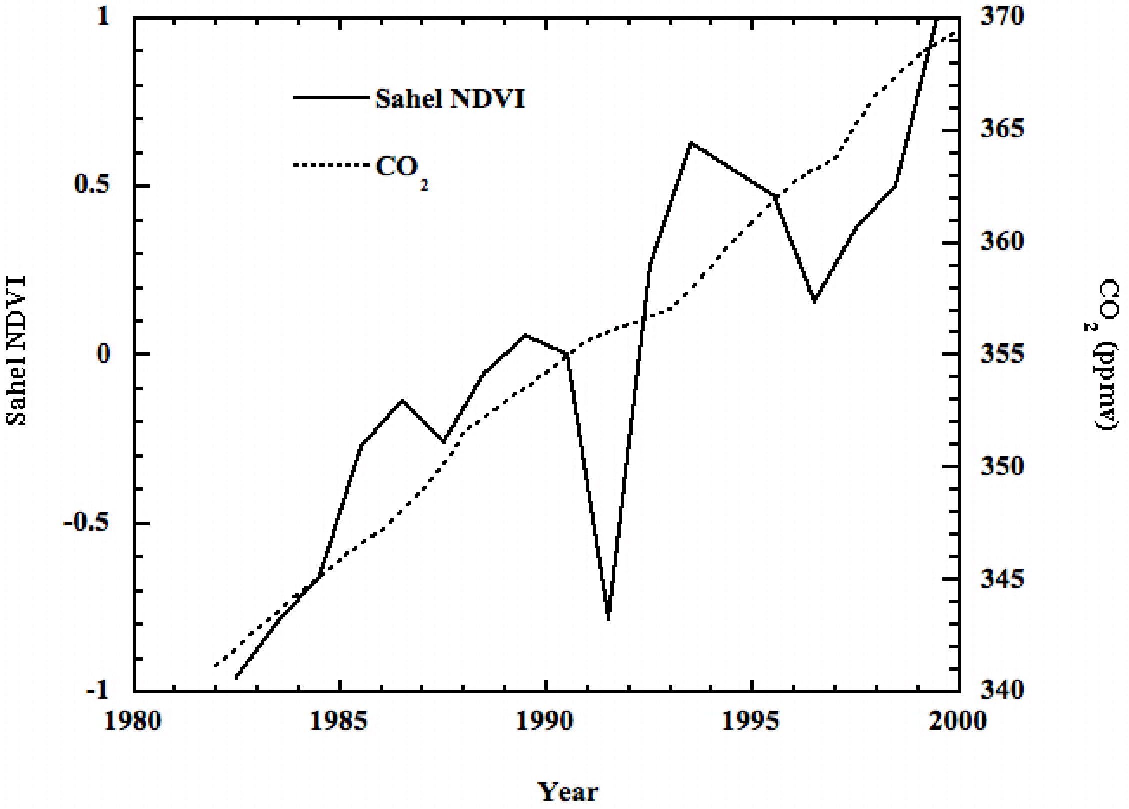


Figure 14: A double-y plot of: (—), the greening of the Sahel represented by the NDVI, and; (····), the trend in atmospheric CO_2, between 1982-1999.

The Sahel NDVI is shown plotted with the Sahel precipitation in Figure S15. The correlation $r^2=0.33$, or less than half that attending the trend in CO_2. Part of this difference is due to the smooth yearly CO_2 trend, relative to the jagged yearly variation in precipitation.

However, the correlation $r^2=0.57$ between the trend in Sahel precipitation and atmospheric CO_2. So, the correlation between Sahel precipitation and Sahel NDVI is partly confounded by the correlation between Sahel precipitation and atmospheric CO_2.

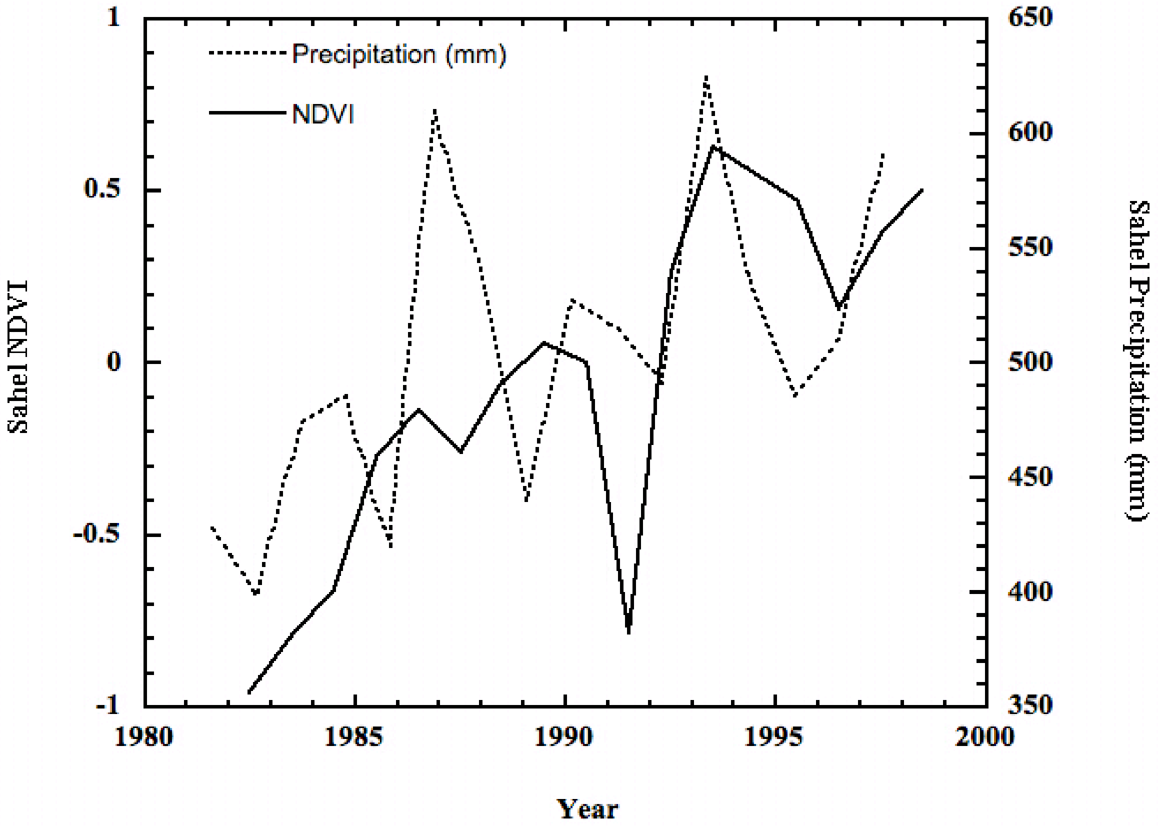


Figure S15: *A double-y plot of: (—), the greening of the Sahel represented by the NDVI, and; (····), the trend in Sahel precipitation between 1982-1998.*

Correlation is indeed not causation. But from a *political* point of view, anyone who claims rising CO_2 is causing Earth climate to warm up must then insist just as fervently that rising CO_2 is causing Earth ecology to green up.

References:

1. S. Manabe and R. T. Wetherald (1967) *Thermal Equilibrium of the Atmosphere with a given Distribution of Relative Humidity* Journal of the Atmospheric Sciences 24, 241-259.

2. G. Myhre, E. J. Highwood, K. P. Shine and F. Stordal (1998) *New estimates of radiative forcing due to well mixed greenhouse gases* Geophysical Research Letters 25, 2715-2718.

3. D. M. Etheridge, L. P. Steele, R. L. Langenfelds, R. J. Francey, J.-M. Barnola, et al. (1996) *Natural and anthropogenic changes in*

atmospheric CO_2 over the last 1000 years from air in Antarctic ice and firn Journal of Geophysical Research 101, 4115-4128.

4. C. D. Keeling and T. P. Whorf *Atmospheric CO_2 concentrations (ppmv) derived from in situ air samples collected at Mauna Loa Observatory, Hawaii* Scripps Institution of Oceanography (SIO), University of California 2005 http://cdiac.ornl.gov/trends/co2/sio-mlo Last accessed on: 14 September 2007.

5. D. M. Etheridge, L. P. Steele, R. J. Francey and R. L. Langenfelds *Historical CH_4 Records Since About 1000 A.D. From Ice Core Data. In Trends: A Compendium of Data on Global Change* Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy 2002 http://cdiac.ornl.gov/trends/atm_meth/lawdome_meth.html Last accessed on: 14 September 2007.

6. M. A. K. Khalil, R. A. Rasmussen and M. J. Shearer (2002) *Atmospheric nitrous oxide: patterns of global change during recent decades and centuries* Chemosphere 47, 807-821.

7. IPCC *Working Group I: The Physical Science Basis of Climate Change* 2007 http://ipcc-wg1.ucar.edu/wg1/wg1-report.html Last accessed on: 14 September 2007.

8. C. Covey, K. M. AchutaRao, S. J. Lambert and K. E. Taylor *Intercomparison of Present and Future Climates Simulated by Coupled Ocean-Atmosphere GCMs PCMDI Report No. 66* Lawrence Livermore National Laboratory 2001 http://www-pcmdi.llnl.gov/publications/pdf/report66/ Last accessed on: 14 September 2007; C. Covey, K. M. AchutaRao, U. Cubasch, P. Jones, S. J. Lambert, et al. (2003) *An overview of results from the Coupled Model Intercomparison Project* Global and Planetary Change 37, 103-133.

9. *Climate of 2006 - October in Historical Perspective* National Climate Data Center, NOAA 2006 http://lwf.ncdc.noaa.gov/img/climate/research/2006/oct/ratpac-ytd-oct-pg.gif Last accessed on: 14 September 2007.

10. W. L. Gates, J. S. Boyle, C. Curt Covey, C. G. Dease, C. M. Doutriaux, et al. (1999) *An Overview of the Results of the Atmospheric Model Intercomparison Project (AMIP I)* Bulletin of the American Meteorological Society 80, 29-55.

11. *Autoregressive moving average model* Wikimedia Foundation 2007 http://en.wikipedia.org/wiki/Autoregressive_moving_average_model Last accessed on: 14 September 2007.

12. T. S. Saitoh and S. Wakashima "An efficient time-space numerical solver for global warming" In: Energy Conversion Engineering Conference and Exhibit (IECEC) 35th Intersociety Ed. IECEC: Las Vegas 2000 pp. 1026-1031.

13. A. Staudt, N. Huddleston and S. Rudenstein *Understanding and Responding to Climate Change* The National Academy of Sciences USA 2006 http://dels.nas.edu/basc/ Last accessed on: 14 September 2007. The "low-res pdf" is a convenient download.

14. P. D. Jones, T. J. Osborn, K. R. Briffa and D. E. Parker *Global Monthly and Annual Temperature Anomalies (degrees C), 1856-2004 (Relative to the 1961-1990 Mean)* CRU, University of East Anglia, and Hadley Centre for Climate Prediction and Research 2005 http://cdiac.ornl.gov/ftp/trends/temp/jonescru/global.dat Last accessed on: 14 September 2007.

15. L. Olsson, L. Eklundh and J. Ardo (2005) *A recent greening of the Sahel-trends, patterns and potential causes* Journal of Arid Environments 63, 556-566.

16. T. Hickler, L. Eklundh, J. W. Seaquist, B. Smith, J. Ardo, et al. (2005) *Precipitation controls Sahel greening trend* Geophysical Research Letters 32, L21415 21411-21414.