Attribution studies of observed land precipitation changes with nine coupled models

F. Hugo Lambert,1,2 Nathan P. Gillett,3 Dáithí A. Stone,2 and Chris Huntingford1

Received 27 May 2005; revised 25 July 2005; accepted 18 August 2005; published 21 September 2005.

[1] Global-land mean observations of 20th century precipitation are compared to modelled values using an optimal regression technique for nine general circulation models. The combined influence of major anthropogenic and natural forcings is detected in five cases. Comparing the accuracy of precipitation and temperature simulation of each model, we find that low temperature simulation accuracy produces low precipitation simulation accuracy, but temperature accuracy does not determine precipitation accuracy in general. Model formulation appears to be more important for accurate precipitation simulation than inclusion of a more complete set of forcings. The implications for possible constraints on land precipitation are discussed. Citation: Lambert, F. H., N. P. Gillett, D. A. Stone, and C. Huntingford (2005), Attribution studies of observed land precipitation changes with nine coupled models, Geophys. Res. Lett., 32, L18704, doi:10.1029/2005GL023654.

1. Introduction

[2] Changes in climate have been attributed to external forcings using optimal regression techniques for a variety of climate variables. Stott et al. [2000] and Tett et al. [2002] attributed changes in large scale near-surface temperature to a combination of natural and anthropogenic factors. Gillett et al. [2003] found that changes in sea-surface pressure cannot be explained by natural variability alone and Barnett et al. [2001] attributed changes in basin-scale ocean heat content to anthropogenic forcings, for example.

[3] Studies attributing changes in precipitation have lagged behind because observing and modelling precipitation accurately is difficult. A recent study using the HadCM3 model by Lambert et al. [2004], however, found an attributable response for global-land precipitation to natural and anthropogenic forcings that appeared to be largely due to volcanic forcing. This was supported by Gillett et al. [2004] who detected the influence of volcanic forcing, but not greenhouse gases (GHGs), anthropogenic sulphates or changes in solar irradiance, using the PCM model. Both studies found that the model simulations they used possibly underestimated the observed response to forcing. Further, Allen and Ingram [2002] showed that the range of modelled global precipitation responses to global warming varies by a factor of ~3. In this work, we therefore compare observed precipitation to results from nine state-of-the-art coupled models. We then compare the accuracy of modelled temperature and precipitation and discuss the implications for possible constraints on land precipitation change.

2. Observed and Modelled Data

[4] Observed 20th century precipitation data are compared to simulations from nine General Circulation Models (GCMs). The observations are land-based gauge data taken from the Hulme data set for 1900–98 interpolated onto a 3.75° longitude by 2.5° latitude grid [Hulme, 1992; New et al., 2000]. Following Gillett et al. [2004], we restrict ourselves to 1944–98, when the largest changes in climate forcings occur.

[5] The modelled data are from the IPCC model output archive and were provided by modelling centres for the imminent IPCC Fourth Assessment Report. The simulations are driven with estimates of historical forcings for the 20th century. We consider only models forced with at least the following four factors: changes in GHGs, the direct effect of tropospheric sulphate aerosols, stratospheric volcanic sulphate aerosols and solar irradiance. Some of the models are also driven by changes in the indirect effects of tropospheric sulphates (IS), black carbon aerosol concentration (BC), stratospheric ozone (SO) and land use (LU), see Table 1.

[6] A global land mean time series is calculated for the observations and each GCM. Observed data are masked so that in each year, only gridboxes having data for at least 7 of 12 months are considered. (The series produced is insensitive to this criterion, however.) The mean climatology from 1961–90 is then removed at each spatial point, ensuring that we are only considering anomalies about a base period. The GCM data are masked identically to the observations and anomalies taken with respect to their own climatologies. Each GCM has a minimum of three ensemble members with different initial conditions (Table 1). These are averaged together to improve the signal-to-noise ratio for the detection of forced signals, because we expect model-simulated internal variability to be uncorrelated between ensemble members. Figure 1 shows land mean time series for observed and model data. In the plot, we differentiate between those models forced by IS, and those which are not. There do not appear to be significant differences between the time series, however.

3. Attribution of Changes in Precipitation

[7] We now compare the observed and modelled time series for 1944–98 using the optimal total least squares regression technique described by Allen and Stott [2003].
First, 11 5-year means of the observed and modelled time series are taken, because this was the timescale on which Lambert et al. [2004] found forcing to be most dominant. The time series are then projected onto the first 11 Empirical Orthogonal Functions (EOFs) of an identically prepared unforced control run, the minimum number necessary to completely describe the data. Signal-to-noise optimisation is achieved by weighting the resulting components by the inverse of the noise variance of the relevant EOF. The control run used is the same in each case, and consists of data from each model concatenated together, because only a short section of control is available in each case. We do not spatially optimise our data because we are interested in the behavior of and constraints on global land mean precipitation.

Assuming that model response amplitude may be scaled up or down linearly, the regression coefficients, β, that we calculate, inform us whether or not the model is under or overestimating the observed precipitation response to forcing. If the confidence interval that contains β is entirely greater than zero, then we say that the influence of the forcings applied to the model is detected. If the confidence limit includes 1, and the residual observed variance is consistent with an estimate of natural variability derived from the control run, then we say that we have attributed observed changes in precipitation to the forcings applied to the model.

The regression coefficients we obtain for the period 1944–98 are summarised in Table 1. For GFDL0, HadCM3, and PCM we detect the influence of the forcings applied to the model at the 5% level, and find that residual variance is not inconsistent with control variability at the 10% level using the F-test, despite these models not being forced with the full range of forcings. We also detect the applied forcings in the case of GISS-ER and MIROC, but find that residual and control variability are marginally inconsistent. The observed residual variance is unrealistically large, suggesting that either the control variance underestimates natural variability, or that significant forced responses are not simulated by the models.

We cannot detect the influence of applied forcings with the other models at the 5% confidence level for 1944–98. GFDL1, and MRI, in particular, fail to capture the observed response (Figure 1c). For CCSM3, MRI and GISS-EH the confidence limits that we calculate are unbounded, indicating that the forced signal in these runs is not significantly different from zero. We can attribute forced precipitation response to forcing in 3 of 9 cases - more than 1 in 20 that we would expect from chance at the 5% level. Still, we conclude that Lambert et al. [2004] and Gillett et al. [2004] were fortunate to employ GCMs that adequately simulate forced changes in precipitation.

Although PCM formally underestimates the size of the response (the confidence limits enclosing β are entirely greater than 1), all the models appear to show precipitation changes smaller than those observed. This is confirmed by the ratio of the average variance of individual forced ensemble members to observed variance shown in Table 1, and could be caused by model errors or missing forcings, or measurement errors we have not considered. There is also a small (and statistically insignificant) downward trend in the observations, replicated by MIROC and GISS-ER but not apparent in the other models. We speculate that this may be due to the indirect effects of sulphate aerosols, which are expected to

Table 1. The Size of the Initial Condition Ensemble for Each Model, and Whether or Not it is Forced With IS, BC, SO or LU a

<table>
<thead>
<tr>
<th>Model</th>
<th>Runs</th>
<th>IS</th>
<th>BC</th>
<th>SO</th>
<th>LU</th>
<th>Vol</th>
<th>( \beta_{\text{max}} ), ( \beta_{\text{min}} )</th>
<th>Variance Ratio</th>
<th>( \rho_{\text{prec}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>GFDL CM2 0</td>
<td>3</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>R</td>
<td>0.79, 1.7, 3.4</td>
<td>0.70, 0.70</td>
<td>1.00, 1.00</td>
</tr>
<tr>
<td>GFDL CM2 1</td>
<td>3</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>R</td>
<td>0.79, 1.7, 3.4</td>
<td>0.70, 0.70</td>
<td>1.00, 1.00</td>
</tr>
<tr>
<td>GISS Model E H</td>
<td>5</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>R</td>
<td>0.79, 1.7, 3.4</td>
<td>0.70, 0.70</td>
<td>1.00, 1.00</td>
</tr>
<tr>
<td>GISS Model E R</td>
<td>5</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>R</td>
<td>0.79, 1.7, 3.4</td>
<td>0.70, 0.70</td>
<td>1.00, 1.00</td>
</tr>
<tr>
<td>MIROC 2 medres</td>
<td>3</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>R</td>
<td>0.79, 1.7, 3.4</td>
<td>0.70, 0.70</td>
<td>1.00, 1.00</td>
</tr>
<tr>
<td>MRI CGCM2 3 2a</td>
<td>5</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>S93</td>
<td>0.35, 0.35</td>
<td>0.35, 0.35</td>
<td>0.35, 0.35</td>
</tr>
<tr>
<td>NCAR CCSM3</td>
<td>8</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>S93</td>
<td>0.35, 0.35</td>
<td>0.35, 0.35</td>
<td>0.35, 0.35</td>
</tr>
<tr>
<td>NCAR PCM1</td>
<td>4</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>A</td>
<td>0.35, 0.35</td>
<td>0.35, 0.35</td>
<td>0.35, 0.35</td>
</tr>
<tr>
<td>UKMO HadCM3</td>
<td>4</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>S93</td>
<td>0.35, 0.35</td>
<td>0.35, 0.35</td>
<td>0.35, 0.35</td>
</tr>
</tbody>
</table>

aAll nine models are also forced with GHGs, the direct effect of tropospheric sulphate aerosols, volcanic aerosols and changes in solar irradiance. The Vol column shows whether the volcanic forcing comes from C. M. Ammann et al. (Coupled climate simulations of the 20th-century including external forcing, submitted to Journal of Climate, 2005), A; Hansen and Sato [2001], HS; Sato et al. [1993], S93; or Ramachandran et al. [2000], R, where known. The \( \beta \)s are the 5–95% range of scaling factors derived from regression. The variance ratio is the average ratio of individual model ensemble member variance to observed variance for 5-year precipitation means from 1944–98. \( \rho_{\text{prec}} \) are the correlations between land and global precipitation for the same averaging periods.
cause reductions in precipitation through their effects on cloud water.

4. Comparison of Precipitation Simulation to Temperature Simulation

What are key features of the climate system that must be modelled if observed responses to forcing are to be captured? In the case of temperature, a simple model simulating bulk climate sensitivity and the rate of ocean uptake can reproduce the observed global response [e.g., Sokolov and Stone, 1998]. For precipitation, it has been suggested that a linear model of perturbations to the tropospheric energy budget dependent only on temperature and forcings may be able to predict changes [Allen and Ingram, 2002]. This model can explain why precipitation appears to respond strongly to volcanic forcing, but not to anthropogenic GHGs [see also Mitchell et al., 1987; Yang et al., 2003]. However, it fails to account for large differences in land and global precipitation in some models (see GCM correlations in Table 1), despite very high correlations between land and global temperature ($>0.99$ for all 9 models) and quite similar changes in forcing [e.g., Ramaswamy et al., 2001]. This was confirmed by Lambert et al. [2004], who found the energy budget model successful globally, but less applicable when the data were restricted to land in a GCM simulating only a shallow ocean with no dynamics. From our precipitation correlations it appears that different constraints may be important in different models.

Here we compare the accuracy of global precipitation and temperature simulation in the nine coupled models. Temperature changes are important to both the energy budget argument and soil moisture levels, which were cited as a possible alternative constraint on land precipitation changes by Koster et al. [2004]. Figure 2 shows the root mean squared difference between modelled and observed global mean temperature and precipitation for individual ensemble members. (The observed temperature data are taken from CRUTEMP2 [Jones and Moberg, 2003].) The bottom right corner of the figure is devoid of points, indicating that a poor simulation of temperature guarantees a poor simulation of precipitation. A good simulation of temperature, however, does not guarantee a good simulation of precipitation. Examining individual models, we see that temperature and precipitation accuracy in CCSM3 may be linearly related, GISS-ER shows large variations in precipitation accuracy unrelated to temperature and MRI shows large variations in temperature accuracy unrelated to precipitation. It is more difficult to draw conclusions for the models with fewer ensemble members, but we note that HadCM3 and GISS-EH consistently provide accurate simulations and PCM shows little variation in accuracy, consistent with the small natural precipitation variability shown by the model. The differences between individual ensemble members are comparable to the differences between models. As in Section 3, we conclude that the models driven with all eight forcings do not provide significantly better simulations. Differences in model formulation appear to be more important.

5. Conclusion

We detect the influence of applied forcings on global-land mean precipitation in five of nine coupled GCMs for 1944–98. While the forcings we consider are sufficient to explain the bulk of observed precipitation changes during this period using five models, we see that model formulation is important for successful detection. This is confirmed where we incorporate modelling...
uncertainty into a single multi-model detection experiment: we do not detect the applied forcings and find a large range for the possible signal amplitude. This differs from the surface temperature case where detection results are fairly consistent for the same models (D. A. Stone et al., A multi-model update on the detection and attribution of global surface warming, submitted to Journal of Climate, 2005). Although realism of temperature response is important to two precipitation constraints we discuss, we find that it does not fully determine realism of precipitation response. Given the wide range of responses of the nine models, if a simple constraint on land precipitation changes does exist, it does not appear to be a ubiquitous feature of GCMS. More study is needed if we are to find how observed land precipitation is controlled.

Acknowledgments. We thank Mike Hulme at the Tyndall Centre for Climate Change Research at the University of East Anglia (UEA) for the observed precipitation data set. We acknowledge the international modeling groups for providing their data for analysis, the Program for Climate Model Diagnosis and Intercomparison (PCMDI) for collecting and archiving the model data, the JSC/CLIVAR Working Group on Coupled Modelling (WGC) and their Coupled Model Intercomparison Project (CMIP) and Climate Simulation Panel for organizing the model data analysis activity, and the IPCC WG1 TSU for technical support. The IPCC Data Archive at Lawrence Livermore National Laboratory is supported by the Office of Science, U.S. Department of Energy. We also thank Myles Allen, who originally developed the optimal detection code used here, and Peter Stott for a useful discussion. We thank two referees for thorough reviews. FHL acknowledges financial support from CEH.

References