

Regional probabilistic climate forecasts from a multithousand, multimodel ensemble of simulations

C. Piani,¹ B. Sanderson,² F. Giorgi,¹ D. J. Frame,³ C. Christensen,^{1,4} and M. R. Allen²

Received 28 March 2007; revised 28 June 2007; accepted 19 September 2007; published 27 December 2007.

[1] A methodology for constraining climate forecasts, developed for application to the multithousand member perturbed physics ensemble of simulations completed by the distributed computing project *ClimatePrediction.net*, is here presented in detail. The methodology is extended to produce constrained forecasts of mean surface temperature and precipitation within 21 land-based regions and is validated with climate simulations from other models available from the IPCC (AR4) data set. The mean forecasted values of temperature and precipitation largely confirm prior results for the same regions. In particular, precipitation in the Mediterranean basin is shown to decrease and temperature over northern Europe is shown to increase with comparatively little uncertainty in the forecast (i.e., with tight constraints). However, in some cases the forecasts show large uncertainty, and there are a few cases where the forecasts cannot be constrained at all. These results illustrate the effectiveness of the methodology and its applicability to regional climate variables.

Citation: Piani, C., B. Sanderson, F. Giorgi, D. J. Frame, C. Christensen, and M. R. Allen (2007), Regional probabilistic climate forecasts from a multithousand, multimodel ensemble of simulations, *J. Geophys. Res.*, 112, D24108, doi:10.1029/2007JD008712.

1. Introduction

[2] The nature of climate forecasts has evolved since the Third Assessment Report [*Intergovernmental Panel on Climate Change*, 2001]. Future climate change was once discussed in terms of “best guess” scenarios while different research groups debated the relative strengths and weaknesses of their model or forecast methodology. Single “best guess” scenarios from individual models were eventually replaced by averages over ensembles of opportunity, since one can reasonably expect that the mean climate forecast over an ensemble of available independent models will show higher skill than the individual ensemble members. This is in fact the case for seasonal forecasts [*Palmer et al.*, 2004]. In this framework, the spread of the forecasts across the ensemble was taken as a measure of the associated uncertainty and it was suggested that adaptation and mitigation strategists would reduce the economic burden associated with climate change by using forecasts couched in probabilistic terms rather than a single (albeit averaged) deterministic forecast [*Räisänen and Palmer*, 2001]. This paradigm of climate forecasting is now firmly established [*Schneider*, 2001; *Allen and Stainforth*, 2002]. For example, engineers deciding on the height of prospective land-sea defense barriers are little interested in the most probable value of sea level change over the next 100 years being 0.1 m

or 0.5 m. Instead their decisions would be most influenced by the height of the 100 year return event, be it storm surge or tide, given a future average sea level rise equal to the 95th upper percentile of the forecast probability density function (PDF) [*Forest et al.*, 2002, 2006; *Smith*, 2002; *Stott and Kettleborough*, 2002; *Palmer and Räisänen*, 2002].

[3] Many PDFs of future climate change are based on climate simulations from an ensemble of available independent models, i.e., an “ensemble of opportunity.” Recent model intercomparison projects have produced vast data sets that allow for this kind of climate forecast [*Covey et al.*, 2003]. Methods for deriving the PDFs of future climate change from ensembles of opportunity generally involve some form of weighted averaging. One obvious choice is to weigh ensemble members according to their ability to reproduce observable climate [*Giorgi and Mearns*, 2002].

[4] In perturbed physics ensembles (PPE) the model phase space can be explored more objectively than in ensembles of opportunity by systematically perturbing key parameters [*Murphy et al.*, 2004]. The choice of parameters and their range of acceptable values are obtained initially through expert elicitation but can be subsequently altered to span more relevant areas of model space. *Stainforth et al.* [2005] presented a frequency distribution of climate sensitivity from the first 2000+ simulations from the *ClimatePrediction.net* (CPDN) [*Allen*, 1999; *Hansen et al.*, 2001] distributed computing experiment and found results to be too sensitive to the sampling strategy to objectively assign probabilities. Recently *Harris et al.* [2006] presented results from a 129 member PPE in which 29 chosen parameters of HadSM3 (the UK Met Office Slab model) are perturbed both individually and in combinations. In their study, they developed a linear emulator of the response of a coupled

¹Abdus Salam International Center for Theoretical Physics, Trieste, Italy.

²Department of Physics, University of Oxford, Oxford, UK.

³Environmental Change Institute, University of Oxford, Oxford, UK.

⁴Now at Department of Physics, University of Oxford, Oxford, UK.

AOGCM to transient forcing, based on a combination of the large slab model PPE with a smaller selected number of perturbed physics simulations using the fully coupled version of the UK Met Office model, HadCM3. The emulator allows for a cost effective way of exploring a larger area of climate model space. *Harris et al.* [2006] produced frequency distributions of regional transient response of annual temperature and precipitation across model space but do not, for the moment, associate probabilities to their results.

[5] *Piani et al.* [2005, hereinafter referred to as P05] developed a methodology for producing constrained climate forecasts based on the search for a transfer function between observables and forecast variables. In their study, the methodology was applied to the *ClimatePrediction.net* slab model data set to produce constrained forecasts of climate sensitivity. Although climate sensitivity is not a technically forecast variable but a property of the climate system, the methodology still applies. Here we use the methodology developed by P05 to constrain regional climate change projections. In particular, the methodology is applied to an updated and expanded CPDN data set of perturbed physics slab model simulations with the aim to produce regional constrained estimates of precipitation and temperature response to doubling of CO₂ preindustrial concentrations.

[6] In section 2 we give a detailed presentation of the methodology developed by P05 and its application to the CPDN data set, and we describe the experimental setup. In section 3 we present and discuss our results. These are followed by our summary and conclusions in section 4. Appendix A follows the main text where details of the methodology are given that are not essential, in first instance, to the understanding of the study.

2. Methodology

[7] In this study we use the same, albeit expanded, data set from the slab model experiment within the *ClimatePrediction.net* (CPDN) distributed computing project. The CPDN project is presented in detail by, among others, *Stainforth et al.* [2005] and *Allen* [1999]. The database is composed of a total of over 2500 single slab model experiments described in P05. Each experiment is made up of three 15 year phases of which the last two have preindustrial and twice preindustrial concentrations of CO₂ respectively. Results from these two phases are used to extrapolate the equilibrium values of the response of climate variables to CO₂ doubling. Extrapolation is necessary because in some cases the equilibrium is not reached by the model within the 15 years of the 2xCO₂ simulation. Given 15 yearly averages of a climate variable $\alpha_{n=1\dots 15}$ from the 2xCO₂ simulation, the extrapolation technique consists in fitting the exponential function: “ $\alpha_n = \Delta\alpha_{\text{eq}}(1 - e^{n/\tau}) + \alpha_o$ ” to the data to obtain $\Delta\alpha_{\text{eq}}$ which is the response at equilibrium of the variable α while α_o is its preindustrial value and n is year index. This is the same technique as used in P05 and described by *Stainforth et al.* [2005]. An average of the last 8 years of the control simulation, with preindustrial values of CO₂, is taken as the observable climate within each ensemble simulation. The initial 7 years are discarded to eliminate the rarely occurring effects of initial adjustments in the control runs. Finally the model used is the slab version of the United

Kingdom Meteorological Office Unified Model (referred to as HadSM3) described in detail by *Pope et al.* [2000].

[8] The methodology presented here was used by P05 to obtain constrained equilibrium responses of global variables, such as climate sensitivity and the feedback parameter but can be applied to any future climate variable α . The methodology can be generally described as a multiple regression of the forecast variable α against observable climate variables across the ensemble of simulations. Obviously, regressing against all available climate observables would leave us with a covariance matrix ranked $>10^6$, hence we must limit our analysis to conveniently defined subspaces of observable climate. Past climate change detection studies have assumed that the correlation structure of model-data discrepancies can be adequately modeled by internal climate variability [*Hegerl and Allen*, 2002; *Gillett et al.*, 2002]. Consequently we reduce the dimensions of observable space by focusing on the principal EOFs of internal climate variability. An introductory treaty on EOFs and PCAs (Empirical orthogonal functions and principal component analysis) and their application to climate data is given by *Jolliffe* [2002]. Here internal climate variability is modeled by a 496 year control climate simulation of the coupled version of the Unified Model (referred to as HadCM3). A simulation has to be used here because a 500 year observational data set of the real climate with constant preindustrial boundary conditions and external forcings, such as CO₂ concentrations, volcanic eruptions, etc., is a physical impossibility.

[9] Here the timescale of interest in climate variability is 8 years to match the averaging period in CPDN control runs. Hence the 496 year HadCM3 control climate simulation is divided into 62 nonoverlapping 8 year periods over which the mean is taken. Principal component analysis (PCA) is then conducted over the resulting data set and 62 EOFs of internal climate variability are derived (see Appendix A for details on how the EOF vectors are constructed). Prior to PCA these 62 data segments must be appropriately normalized. Thus we begin by subtracting the mean over the entire 496 year data set from each of the 8 year segments and weighting each of the variables within the segments by the inverse of the globally averaged standard deviation of HadCM3 8 year climate variability. Each segment is composed of the variables listed in Table 1, along with the grids they are defined on. These are the same set of variables used by P05. Different variables may be defined on different grids. For example, in our case, we have surface temperature defined on a latitude by longitude grid, relative humidity defined on a latitude by pressure grid and top of the atmosphere fluxes defined on a latitude only grid. Conducting PCA on such a data set would greatly favor those variables which are defined on a larger number of grid points, hence all grid point variables are weighted by the portion, in mass, of atmosphere in the grid cell. As a result, the sum of the weights for each grid is a constant for all (see Appendix A).

[10] The space of all observable climate variables can now be reduced to the portion of space spanned by the amplitudes of the EOFs in each CPDN control simulation. As anticipated, we can write the following step of the methodology as a simple multilinear regression:

$$\alpha_i = \sum_{j=1}^K x_{ij}\beta_j + \text{noise}, \quad (1)$$

Table 1. Climate Variables Chosen for This Analysis (First Column), Corresponding Observational Data Set Used for Each Climate Variable (Second Column), and Choice of Grid and Spatial Averaging (Third Column)^a

Climate Variable	Source	Space Subset
1.5 m temperature	CRU ²⁴	85°N to 85°S, land only, latitude-longitude grid
Mean sea level pressure	ERA 40 ²⁵	85°N to 85°S, latitude-longitude grid
Precipitation	Xie-Arkin ²⁶	85°N to 85°S over land, 30°N to 30°S over ocean, latitude-longitude grid
Surface sensible and latent heat fluxes	SOC ²⁷	85°N to 40°S ocean only, latitude-longitude grid
Relative humidity, temperature, zonal and meridional winds	ERA 40	85°N to 40°S, zonal mean height-latitude grid
Outgoing longwave and shortwave radiation	ERBE ²⁸	85°N to 40°S, 1-D zonal mean latitude grid

^aThe observational and reanalysis data used are multiyear averages representing near present-day climate. References for the CRU, ERA 40, Xie-Arkin, SOC, and ERBE data sets are given by Jones *et al.* [1999], Gibson *et al.* [1997], Xie and Arkin [1997], New *et al.* [1999], and Harrison *et al.* [1990], respectively.

where α_i is the forecast variable of the i th member of the CPDN ensemble, x_{ij} is the amplitude of the j th EOF in the i th ensemble member (see Appendix A), expressed as anomalies about the ensemble mean, and K is the number of EOFs retained. The values of β_j that minimize the squared prediction error across the CPDN ensemble are given by the standard regression formula:

$$\hat{\beta}_j = \sum_{i=1}^K \mu_{ij}^{-2} x_{ij}^T \alpha_i, \quad (2)$$

where μ_{ij} is the covariance in the j th and i th EOF across the CPDN ensemble.

[11] In this form the observables we are regressing against are not independent across ensemble models (covariance matrix is not diagonal). Hence it would be difficult at best, impossible at worst, to find distinct underlying physical processes that can account for the emergent correlations between the forecast variable and the observables. Consequently we take an intermediate step prior to the regression which does not alter the resulting emerging constraints but does simplify the interpretation. The 62 EOFs, resulting from the initial PCA are rotated so that the resulting new set of vectors, we will refer to these as rotated EOFs or REOFs, are independent across ensemble members. To obtain the REOFs we start by projecting the climates from the ensemble members onto the EOFs. This is best represented in matrix notation:

$$\mathbf{M}^T \cdot \mathbf{E} = \mathbf{P}, \quad (3)$$

where \mathbf{M}^T is the transposed matrix of ensemble climate observables, \mathbf{E} is the matrix of EOF column vectors and the center dot denotes matrix multiplication. Each of the column vectors of \mathbf{M} is obtained in very much the same way as the 62 HadCM3 climate vectors (see Appendix A). Each column is composed of the same set of variables and grids listed in Table 1. The anomaly from the CPDN ensemble mean is then taken and the variables are normalized in the same way, that is, they are weighted by the inverse of the globally averaged standard deviation of HadCM3 8 year climate variability and by the atmospheric mass associated with the grid point. The resulting \mathbf{P} matrix has dimensions $n \times m$, where n (~ 2500) is the number of CPDN ensemble members and m (62) is the number of EOFs. We then perform PCA on the matrix \mathbf{P} :

$$\mathbf{P} = \mathbf{E}_P \cdot \mathbf{\Lambda}_P \cdot \mathbf{C}_P^T, \quad (4)$$

where \mathbf{E}_P is the matrix of empirical orthogonal functions, $\mathbf{\Lambda}_P$ is the diagonal eigenvalue matrix and \mathbf{C}_P^T is the transposed matrix of principal components. We now substitute equations (4) into (3) and left multiply by \mathbf{C}_P :

$$\begin{aligned} \mathbf{M}^T \cdot \mathbf{E} &= \mathbf{E}_P \cdot \mathbf{\Lambda}_P \cdot \mathbf{C}_P^T \Rightarrow \mathbf{M}^T \cdot \mathbf{E} \cdot \mathbf{C}_P \\ &= \mathbf{E}_P \cdot \mathbf{\Lambda}_P \cdot \mathbf{C}_P^T \cdot \mathbf{C}_P \Rightarrow \mathbf{M}^T \cdot (\mathbf{E} \cdot \mathbf{C}_P) = \mathbf{E}_P \cdot \mathbf{\Lambda}_P, \end{aligned} \quad (5)$$

since \mathbf{C}_P is orthonormal and $\mathbf{C}_P^T \cdot \mathbf{C}_P = \mathbf{I}$, where \mathbf{I} is the identity matrix. Left multiplication by \mathbf{C}_P constitutes a rotation and we shall refer to $\mathbf{E} \cdot \mathbf{C}_P$ as the rotated EOFs or REOFs. The REOFs span the same space as the EOFs, which is the space of internal climate variability. Moreover, the REOFs have the added property of being uncorrelated, by construction, across ensemble models as shown by equation (5), since the columns of $\mathbf{E}_P \cdot \mathbf{\Lambda}_P$ are orthogonal (\mathbf{E}_P is orthogonal and $\mathbf{\Lambda}_P$ is diagonal). Substituting EOFs with REOFs simplifies equation (2):

$$\hat{\beta}_j = \sum_{i=1}^K \mu_j^{-2} x_{ij}^T \alpha_i, \quad (6)$$

where μ_j^2 is the variance in the j th REOF across the CPDN ensemble. In this case K is the number of REOFs retained and we should point out that all initial EOFs are retained when identifying the rotation in equations (4) and (5). For consistency with P05, and reasons given therein, we retain the first 10 REOFs ($K = 10$). Assuming that the linear transfer function identified within the CPDN ensemble holds in real world, we can estimate α from observations by using the same regression formula:

$$\alpha_o = \sum_{j=1}^K x_{oj} \hat{\beta}_j, \quad (7)$$

where the x_{oj} is the projection of the observations, also expressed as anomalies about the CPDN ensemble mean and normalized in the same manner, onto the j th REOF; and the variance associated with it is given by:

$$\sigma_o^2 = \frac{1}{M-1} \sum_{m=1}^M \left(\sum z_{mj} \hat{\beta}_j \right)^2, \quad (8)$$

where z_{mj} is a segment of the HadCM3 control climate, independent of the one used to derive the REOFs, also

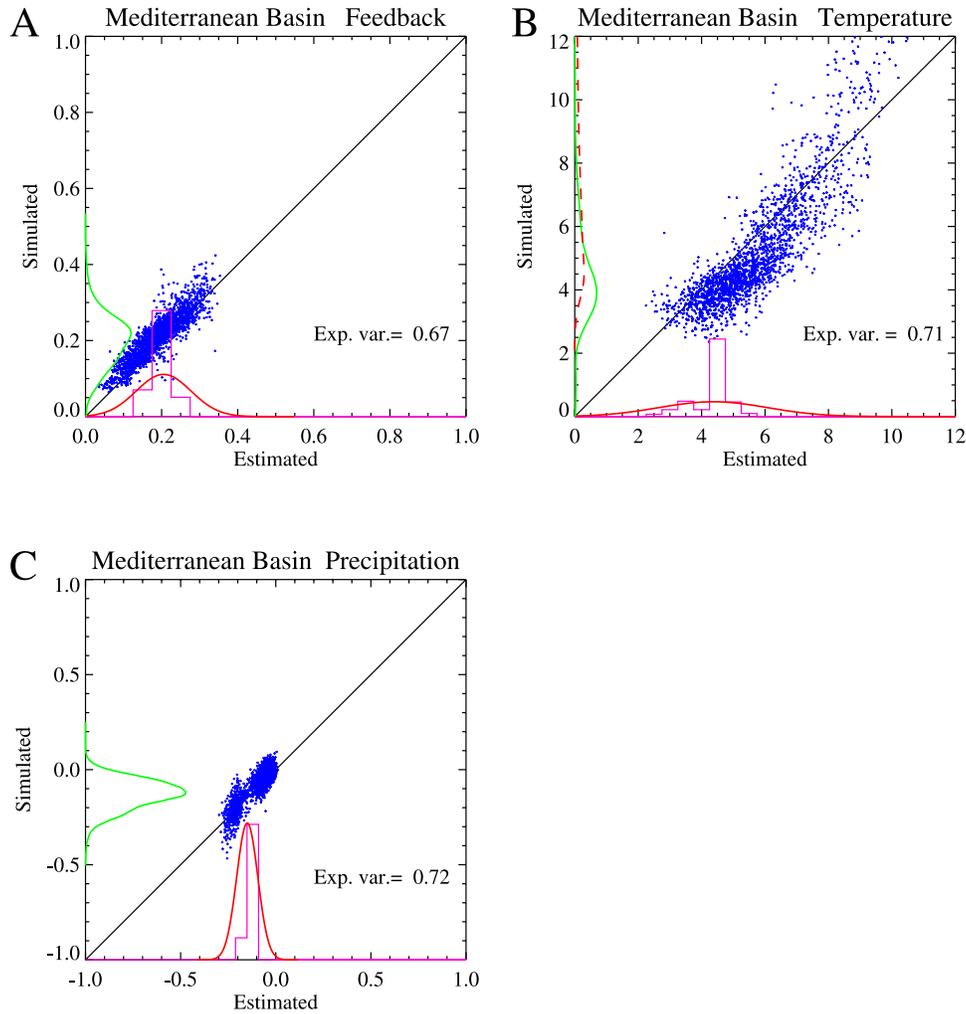


Figure 1. (a) Scatterplot of the simulated feedback parameter over the Mediterranean Basin from the ClimatePrediction.net ensemble (each dot represents one model) versus feedback parameter estimated using equation (1). The thick line on the x axis is the linearly estimated PDF. The thick line on the y axis is the PDF obtained using the scatterplot itself as a transfer function (see methodology section). The histogram on the x axis (thin line) is obtained when the linear predictor described in equation (1) is applied to the GCMs participating in the IPCC AR4. (b) Similar to Figure 1a but for the mean surface temperature. The added PDF (thick dashed line) on the y axis is obtained by replotting the PDF on the x axis in Figure 1a versus temperature. (c) Similar to Figure 1a but for the mean precipitation.

expressed as an anomaly from the CPDN ensemble mean and $M = 64$ is the number of segments of the HadCM3 control climate used to derive the variance. We can now define a Gaussian distribution $G(\alpha_o, \sigma_o)$, with mean and variance defined by equations (7) and (8). This is the result of a linear estimate and concludes the linear portion of the methodology.

[12] The standard deviation σ_o only accounts for the explained variance in the original regression (equation (1)). The unexplained variance, that is the variance associated to the *noise* term in equation (1), is added back to $G(\alpha_o, \sigma_o)$ using a Monte Carlo Methodology which we now describe. Given a climate forecast variable α , we start by plotting α_i versus $\sum_{j=1}^K x_{ij}\beta_j$, that is the simulated forecast variable versus the linearly estimated value for that variable, where “ i ” is the simulation index across the CPDN ensemble. Examples

of such plots are shown in Figures 1 and 2, where the chosen dynamical variable α is the regional feedback parameter (that is, the inverse of the mean surface temperature change at equilibrium upon doubling of CO₂) (Figure 1a), the mean surface temperature change (Figure 1b), and mean precipitation change (Figure 1c), all defined over the Mediterranean Basin. Let us refer to $\sum_{j=1}^K x_{ij}\beta_j$ (the linear prediction of α_i) as α'_i . Each pair (α'_i, α_i) corresponds to a CPDN ensemble member and a dot on the scatterplots in Figures 1a–1c. The PDFs on the x axes (red thick line) is the Gaussian distribution $G(\alpha_o, \sigma_o)$, obtained by applying observations to the linear predictor. We now define a bin size for the forecast variable “ α ” small enough to resolve $G(\alpha_o, \sigma_o)$, but still wide enough so that a significant number of α'_i fall within each bin. The j th bin, defined on the x axis, identifies the set $[\alpha'_i]_j$ of α'_i that fall within that bin. Also, it identifies the set $[\alpha_i]_j$ of α_i ,

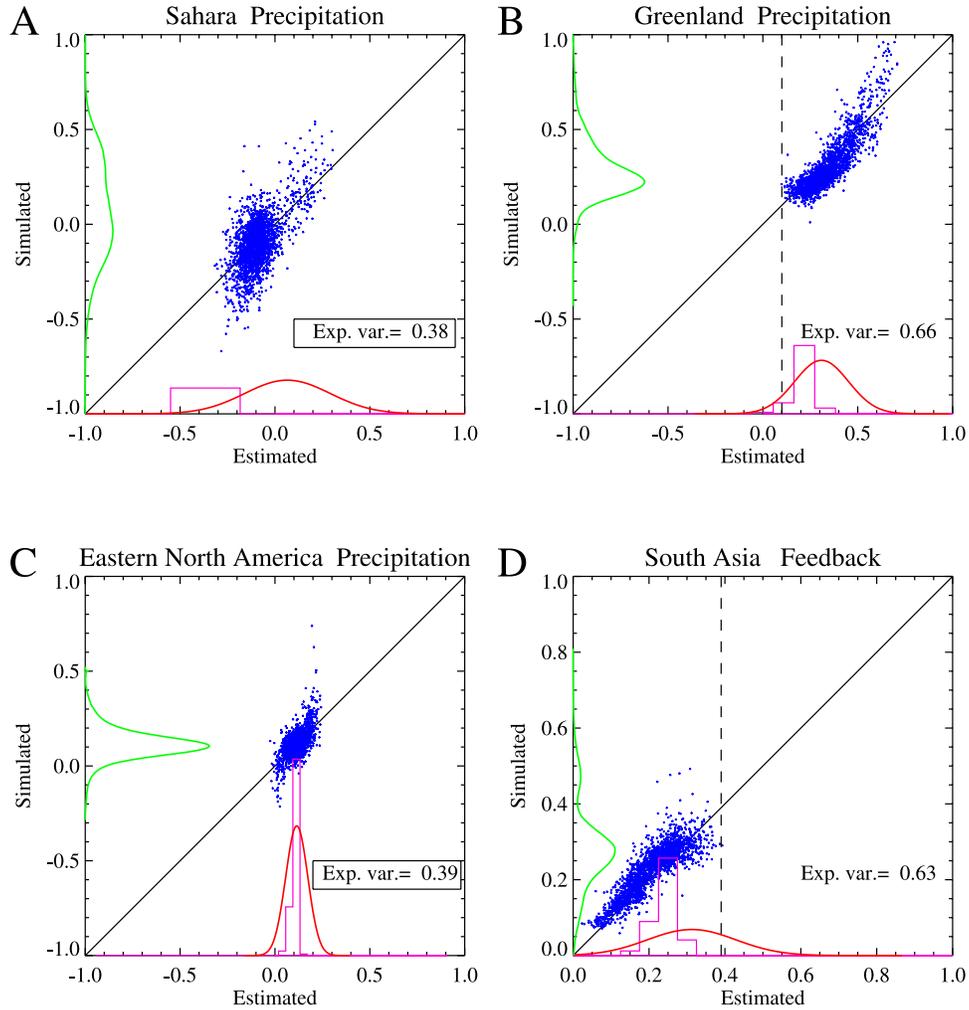


Figure 2. Similar to Figures 1a–1c but for precipitation over (a) the Sahara region, (b) Greenland, (c) eastern North America and (d) for the feedback parameter over south Asia. In Figures 2a and 2c the regression yields a comparatively small correlation. Here the amount of explained variance is highlighted with a text box, also note the scatter of blue points is more isotropic than for Figures 2b and 2d. In Figures 2b and 2d, there are no simulation points (blue dots) to fill in either the low (Figure 2b) or the high (Figure 2d) 5% end of the observed distribution (solid line on x axis).

defined on the y axis, associated with each α'_i in that bin. We now generate a large population of random values $\beta_{m=1,2,\dots,10^6}$ that follow the distribution $G(\alpha_o, \sigma_o)$. Each β_m from this population will fall within one of the J bins defined above. Hence, to each β_m we can associate a set of values $[\alpha]_j$ that are defined on the y axis. Finally to each β_m we can randomly associate one of the members of the set $[\alpha]_j$. If we refer to this value as α_m , then we have defined a population whose members α_m are distributed according to $G(\alpha_o, \sigma_o)$ with an added noise term determined by the scatter of α_i points across the $y = x$ line in the plots. The scatter may or may not be Gaussian and generally depends on the position of the j th bin on the x axis. The large population of points α_m is used to derive a final PDF (green thick line on the y axis, Figures 1a–1c).

[13] To further clarify this second phase of the methodology, we point out that if the dots in Figure 1a all fell on the line $y = x$, then the green PDF on the y axis would be identical to the red PDF on the x axis. Also, if the dots in

Figure 1 formed a cloud normally distributed about the line $y = x$, then the green PDF would be a Gaussian $G'(\alpha_o, \sigma')$, where σ' is obtained by simply adding on the variance of the distribution about the $y = x$ line.

[14] To validate the linear portion of the methodology, we can estimate α from the preindustrial slab GCM simulations participating in the IPCC AR4 [Meehl *et al.*, 2007] by using the usual regression formula:

$$\alpha_a = \sum_{j=1}^K x_{aj} \hat{\beta}_j, \quad (9)$$

where “ a ” is the index of AR4 models and x_{aj} is the projection of the j th predictor onto the normalized anomaly of the a th simulation. In equation (9) the anomalies are relative to the mean CPDN climate, as for all other regressions in this study, hence equation (9) tells us how the relative model biases project onto the linear estimator

Table 2. Regional Temperature Responses^a

Region Name	Variance Explained	Surface Temperature Response (ΔK)			IPCC AR4
		5% Cutoff	Median	95% Cutoff	
AUS	0.64	1.9 K	3.7 K	6.9 K	3.1 K
AMZ	0.63	1.7 K	4.0 K	7.3 K	4.3 K
SSA	0.69	2.0 K	3.6 K	9.2 K	3.3 K
CAM	0.65	2.6 K	5.0 K	9.4 K	3.9 K
WNA	0.68	2.4 K	4.5 K	7.2 K	4.2 K
CNA	0.71	4.0 K	6.8 K	11.7 K	4.9 K
ENA	0.69	2.9 K	5.1 K	8.4 K	4.9 K
ALA	0.62	2.0 K	4.4 K	7.1 K	4.9 K
GRL	0.67	2.5 K	5.0 K	8.3 K	5.0 K
MED	0.68	3.2 K	5.7 K	10.5 K	4.3 K
NEU	0.68	3.0 K	5.1 K	8.8 K	4.4 K
WAF	0.63	1.3 K	4.2 K	8.7 K	3.6 K
EAF	0.63	2.3 K	4.6 K	9.4 K	3.5 K
SAF	0.65	1.5 K	3.7 K	7.8 K	4.0 K
SAH	0.65	2.7 K	5.0 K	9.3 K	4.2 K
SEA	0.70	1.6 K	3.3 K	5.9 K	2.6 K
EAS	0.62	3.2 K	5.3 K	9.2 K	4.3 K
SAS	0.66	1.9 K	4.1 K	8.0 K	3.7 K
CAS	0.64	3.6 K	6.2 K	11.0 K	4.8 K
TIB	0.66	3.7 K	6.4 K	11.6 K	4.8 K
NAS	0.71	3.5 K	6.2 K	11.0 K	5.1 K

^aThe explained variance column refers to the regression against the feedback parameter (inverse surface temperature). The three “surface temperature response” columns refer to the PDF obtained from the constraints on the feedback parameter then mapped onto the surface temperature. The last column is the mean predicted value from the AR4 models. Italics correspond to cases where the <5% or >95% tails of the PDF on the x axis are not populated by the scatterplots (condition 2 from methodology section).

defined by equation (6). The results for the AR4 models are plotted on Figure 1 as a histogram on the x axes. The extent to which the histograms follow the red PDFs on the x axis, can be considered a validation of the linear estimator. The acronyms of the IPCC SLAB models used are: CCMA (versions T47 and T63), MICRO3-2 (versions MIDRES and HIRES), MPI-ECHAM5, CSIRO-MK3, MRI-CGCM2, GISS-SLABM, NCAR-CCSM3, INMCM3, UKMO-HADGEM1, MICRO3-2HIRES. See the IPCC 4th Assessment Report 2007 for individual model description [Randal *et al.*, 2007, chap. 8].

[15] In some cases the tails of the $G(\alpha_o, \sigma_o)$ distributions extend beyond the scatter of α' points, that is to say that none of the dots that compose the scatterplot populate the <5% or >95% tails of the x axis distribution as is the case in Figures 2b and 2d. When this happens β_m may fall into a J bin for which the set $[\alpha]_j$ is empty. In this case we assume that, if we had ensemble members in that particular bin, they would be distributed normally with variance equal to the unexplained variance from the regression. This is tantamount to simply adding back the unexplained variance as a normally distributed linear noise term. This may work well when the noise term is clearly linear and the tails of the $G(\alpha_o, \sigma_o)$ distribution do not extend far beyond the scatter of α' points, as it is in most cases for the regional feedback parameters (Figure 1a), but it does not work where the noise term is clearly non linear, as in most cases for the mean surface temperature (Figure 1b), or when the tails of the $G(\alpha_o, \sigma_o)$ distribution extend well beyond the scatter of α' points, as it is in some cases for the regional precipitation (Figures 2b and 2d). An arbitrary tolerance threshold, for

the extension of the $G(\alpha_o, \sigma_o)$ distribution beyond the scatter of α' points, was set at 5%. Hence, when none of the α' points fall within the <5% and >95% tails of the $G(\alpha_o, \sigma_o)$ distribution, we will flag the resulting constrain by italicizing the relative row in the final summarizing Tables 2 and 3 and highlight the region name in red in Figures 3 and 4. This tends to happen mostly when constraining precipitation and is mostly, but not exclusively, associated with low values of explained variance.

[16] The final distribution derived from the transfer function methodology can be used, as it was in P05 to derive 95% and 5% cutoff values for the forecast variables which we here take as the constraining limits for the climate variable in question. One of the results obtained by P05 was that observations scaled with the feedback parameter and, consequently, not with climate sensitivity. The authors concluded that it was reasonable to use the emerging constraint on the feedback parameter, adequately transformed, to derive a constraint for climate sensitivity. From a Bayesian point of view, the choice between constraining surface temperature or the feedback parameter is tantamount to a choice of the prior distribution. Choosing to constrain the feedback parameter in lieu of surface temperature corresponds to assuming a uniform prior in the observables, which is arguably the best represents the “zero information” hypothesis. In P05, however, the constraint from the global feedback parameter matched the constraint on the mean global surface temperature (that is, climate sensitivity) remarkably well. This is not the case for the corresponding regional values (for example, Figure 1b) as should be expected since the assumptions underlying the one-dimensional climate model equation, on which the feedback parameter is based, does not necessarily hold for regional values.

Table 3. Regional Precipitation Responses^a

Region Name	Variance Explained	Precipitation Response ($\Delta\%$)			IPCC AR4
		5% Cutoff	Median	95% Cutoff	
AUS	0.41	-36%	-11%	18%	6%
AMZ	0.41	-23%	-9%	6%	-14%
SSA	0.27	-8%	-0%	8%	-1%
CAM	0.40	-45%	0%	48%	-10%
WNA	0.56	-14%	-2%	11%	2%
CNA	0.60	-28%	-5%	17%	-1%
ENA	0.40	0%	12%	27%	10%
ALA	0.72	4%	27%	64%	19%
GRL	0.67	10%	28%	57%	21%
MED	0.72	-26%	-13%	-1%	-13%
NEU	0.39	4%	15%	29%	14%
WAF	0.61	-41%	-2%	25%	-1%
EAF	0.70	2%	15%	33%	16%
SAF	0.50	-6%	0%	7%	-7%
SAH	0.39	-34%	7%	52%	-37%
SEA	0.51	-1%	0%	12%	1%
EAS	0.60	4%	10%	18%	12%
SAS	0.65	2%	13%	27%	17%
CAS	0.31	-26%	-9%	10%	-6%
TIB	0.48	-6%	7%	25%	4%
NAS	0.64	8%	22%	44%	18%

^aThe response is given in terms of precipitation percentage changes. Italics correspond to cases where conditions 1 and/or 2 apply, in which case the significance of the results is impacted.

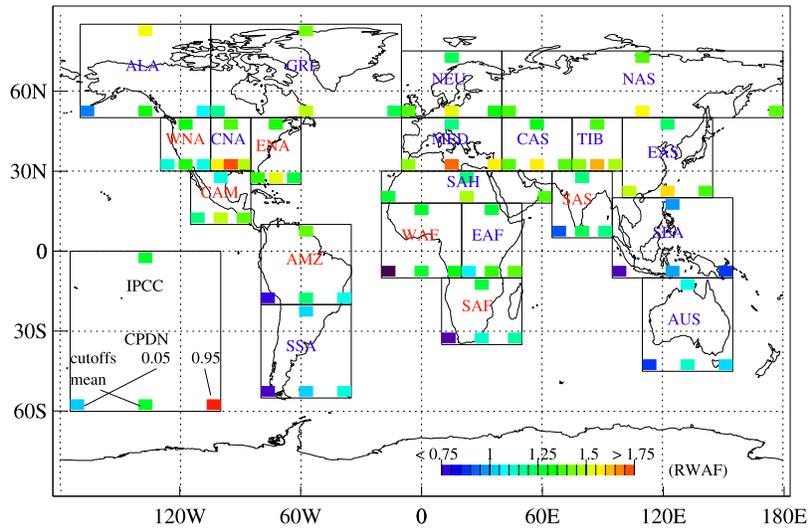


Figure 3. Constrained estimates of the regional temperature response to a doubling of preindustrial CO₂ concentrations. The response is given in terms of Regional Warming Amplification Factor (RWF). For each region the mean from the CPDN ensemble is shown in the bottom middle. 5% (95%) cutoff values of the regional surface temperature response divided by 2.2 K (6.8 K) are shown on the lower left (right) corner. The median linearly predicted RWF from the AR4 ensemble of slab model simulations is shown on the top middle of each region box. Where the region names are given in red, this indicates that conditions 1 or 2 or both, as defined in the results section and which impact the significance of the result, occur.

[17] In this study we will use the same methodology used in P05 to constrain the equilibrium response in regional mean surface temperatures to doubling CO₂ concentrations.

3. Results

[18] The methodology described above was applied to the CPDN data set to constrain the response of precipitation and

surface temperature to a doubling of preindustrial carbon dioxide concentrations in 21 different land regions of the globe. The regions used are similar, but not identical, to the ones defined by *Giorgi et al.* [2001]. Region names and northeast-southwest corners are given in Table 4. The latitude-longitude limits in Table 4 define rectangular regions from which only land points are considered. Natu-

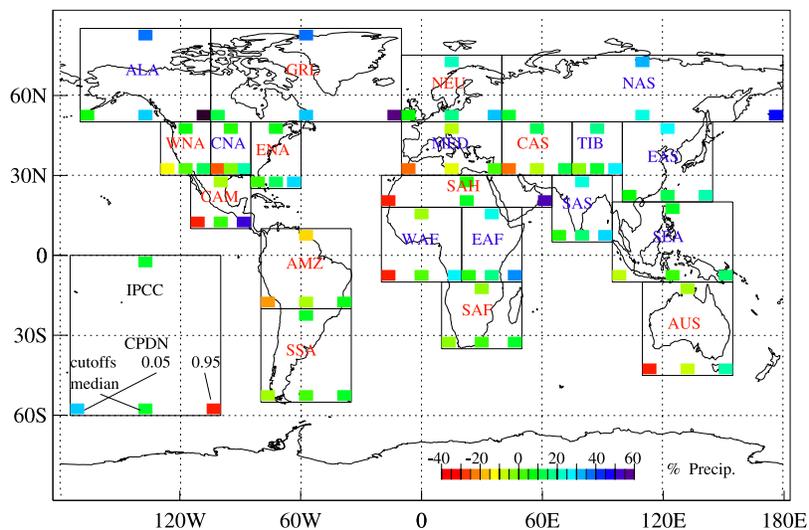


Figure 4. Constrained estimates of the regional precipitation response to a doubling of preindustrial CO₂ concentrations. The response is given in terms of Regional percentage change in annual precipitation. For each region the mean from the CPDN ensemble is shown in the bottom middle. 5% (95%) cutoff values of the regional percentage change in annual precipitation are shown on the lower left (right) corner. The mean linearly predicted values from the AR4 ensemble of slab model simulations is shown on the top of each region box. As in Figure 3, where the region names are given in red, this indicates that conditions 1 or 2 or both, as defined in the results section and which impact the significance of the result, occur.

Table 4. List of Regions and Region Boundaries^a

Region Name	Description	Eastern Limit	Western Limit	Northern Limit	Southern Limit
AUS	Australia	155°E	110°E	10°S	45°S
AMZ	Amazon Basin	35°W	80°W	10°N	20°S
SSA	southern South America	35°W	80°W	20°S	55°S
CAM	Central America	85°W	115°W	30°N	10°N
WNA	western North America	130°W	105°W	50°N	30°N
CAN	central North America	105°W	85°W	50°N	30°N
ENA	eastern North America	85°W	60°W	50°N	25°N
ALA	Alaska	170°W	105°W	85°N	50°N
GRL	Greenland	105°W	10°W	85°N	50°N
MED	Mediterranean Basin	10°E	40°W	50°N	30°N
NEU	northern Europe	10°E	40°W	75°N	50°N
WAF	western Africa	20°E	20°W	18°N	10°S
EAF	eastern Africa	20°W	50°W	18°N	10°S
SAF	southern Africa	10°W	50°W	10°S	35°S
SAH	Sahara	10°E	65°W	30°N	18°N
SEA	Southeast Asia	95°W	155°W	20°N	10°S
EAS	east Asia	100°W	145°W	50°N	20°N
SAS	south Asia	65°W	95°W	30°N	5°N
CAS	central Asia	40°W	75°W	50°N	30°N
TIB	Tibet	75°W	100°W	50°N	30°N
NAS	northern Asia	40°W	180°W	75°N	50°N

^aCoordinate limits describe a rectangular area on the globe though, for each region, only land points are used.

rally, given the limits of the grid resolution, some boundary points may be partially affected by the oceanic values.

[19] As anticipated, the methodology was not equally successful in all cases. Less successful cases are those where either the multilinear regression fails to explain an acceptable fraction of the variance (condition 1), or where the tails of the linearly estimated PDFs extend beyond the limits of the scatterplot (condition 2). In Table 2 we present the results from the methodology applied to the regional mean surface temperature. Each row corresponds to a region, and shows the three letter name, the variance explained by the linear regression, the 5% cutoff, median and 95% cutoff values and, finally, the median value obtained when applying the linear estimator to the AR4 models. The italicized rows correspond to cases where the tails of the linearly estimated PDFs extend beyond the limits of the scatterplot. We must stress that this last condition is diagnosed on the feedback parameter scatterplot, and not directly on the temperature plot, since it is the feedback parameter constrain, albeit replotted, that is being used. In Table 3 we present the results from the methodology applied to the regional precipitation. The columns are similar to those in Table 2. Here the response in precipitation is given as a percentage change rather than in absolute values.

[20] Often conditions 1 and 2, which impact the significance of the resulting constrain, tend to occur together. In Figures 2a–2d we show cases where these conditions occur. In Figure 2a we show the results when the methodology is applied to the Sahara precipitation. Here the explained variance is only 38% (note, the AR4 models as well do not validate the methodology). In the case of the Greenland precipitation (Figure 2b) the explained variance is acceptable but the lower cutoff of the red PDF extends too far left (In Figure 2b the vertical dashed line marks the lower extent of the scatter cluster). The opposite is true when the methodology is applied to precipitation over eastern North America where only 39% of the variance is explained by the linear estimator while the red PDF lies within the scatterplot limits. Figure 2d shows a case where condition 2 occurs and

the >95% tail of the x axis distribution is unpopulated by points in the scatterplot.

[21] Results for all regions are shown in Figure 3 for temperature and in Figure 4 for precipitation. The rectangular regions defined by the latitudinal and longitudinal limits, given in Table 1, are shown as boxes in Figures 3 and 4. The results for the regional temperature responses (Figure 3) are shown in terms of Regional Warming Amplification Factor or RWAf [Giorgi and Bi, 2005, hereinafter referred to as GB05]. The RWAf is defined as the ratio of regional mean surface temperature change to the global mean surface temperature change. Here this definition has been extended to statistically derived quantities. Consequently the values of the mean (or expected value) of temperature response to a doubling of preindustrial carbon dioxide is normalized by 3.3 K, which is the global mean response in surface temperature from P05. Similarly the 5% cutoff values of the regional response are normalized by 2.2 K, which is the 5% cutoff value of the global response and the 95% cutoff values of the regional response are normalized by 6.8 K, which is the 95% cutoff value of the global response. These three values are given, for each region, along the bottom of the corresponding bounding box. The 5% cutoff value is the first on the left, the mean value in the centre and the 95% cutoff value is located to the right. The colors represent the value of the ratios as given in the color table at the bottom of the figure. Along the top of each bounding box, we give the mean value of the linear predictor when applied to the slab models (global atmospheric models with a mixed layer ocean) participating in the IPCC AR4. As for the results from the CPDN simulations, the AR4 climate model results are also normalized by 3.3 K, the mean value of the global temperature response from the CPDN ensemble. Consequently the top and bottom middle values can be compared directly. Upper and lower bounds are not given for the AR4 ensemble. This is because double carbon dioxide experiments were not carried out with slab models. Hence the second part of the methodology presented here, which is necessary to produce constrained forecasts and which

requires a scatterplot of simulated versus linearly predicted response values, could not be carried out. Finally, in our analysis we compare our results with those of GB05. In the GB05 study, the Reliability Ensemble Averaging (REA [Giorgi and Mearns, 2002]) methodology was used to produce deterministic forecasts of precipitation and surface temperature changes for the 21st century over 26 land-based regions which are similar, but not identical, to the ones used here. The REA, in the version used in GB05, is a weighted average of individual deterministic forecasts from different ensemble models, where the weight is a function of the present climate bias in that model [see Giorgi and Mearns, 2002, equations (3) and (4)]. Consequently this method shows the same weaknesses ascribed to any weighted averaging method. First, that, during the development process, climate models are already tuned to reproduce present climate, and therefore there is no assurance that the models with the lowest errors in simulating present climate are also the most believable in simulating climate response to large external forcings. Secondly, results based on Bayesian methods are intrinsically dependent on the prior distribution which, in these cases, would be entirely left to chance and unlikely to be an objective representation of the “space of all possible models” [Allen and Stainforth, 2002]. Ensembles of opportunity are unlikely to be objective representations of model space primarily because of their tendency to cluster around preexisting results. If 100 model simulations were normally distributed we would expect there to be 1 or 2 sitting 2σ away from the consensus (mean), but no group would want to author such a model, hence, certain areas of model space tend to get relatively under sampled.

[22] In comparing our results with GB05, however, two issues should be kept in mind. First that GB05 analyzed transient AOGCM simulations under three IPCC emission scenarios rather than double CO₂ experiment with coupled slab models, and secondly, that GB05 produced deterministic forecasts not PDFs. Therefore the comparison is only indicative.

[23] Initially, it is worth noting that the color table in Figure 3 spans mostly ratios >1 and has a lower limit at 0.75. This is the result of three effects. First, regional cutoff limits tend to be broader rather than narrower than their global counter part. This is expected since the energy fluxes in and out of a limited region are affected by interannual variability to a greater extent than that of the Earth as a whole. Secondly, the lower cutoff value in the temperature response is much less sensitive to observational variability than the upper cutoff value. This is a direct consequence of the fact that observations do not scale directly with the surface temperature response but with its inverse, the feedback parameter, as stated earlier. As for the global case, we can define a regional feedback parameter as the inverse of the regional surface temperature response. We find that regional feedback parameters also scale linearly with observations, hence the resulting PDFs for regional temperature response tend to be skewed to higher values with a sharp lower bound and a fat tail toward higher values. This makes the upper bound more sensitive to observational variability (P05). As a consequence, regional upper cutoff values can be significantly larger than the global counter parts but lower cutoff values can only be slightly lower. Finally, most of the regions have mean responses which are higher than

the global value (3.3 K) because they are composed solely of land grid points while global values are also composed of ocean points which are strongly affected by unresponsive slab model SSTs. This finding is in agreement with the regional climate projections from Christensen [2007, chapter 11, hereinafter referred to as AR4-11].

[24] The mean values shown in Figure 3 generally agree with previous results from GB05. The strongest warmings are in the northern regions, where most of the regional climate “hot spots” occur [Giorgi, 2006, hereinafter referred to as G06]. In our study four regions have RWFs ~ 1 , that is SEA, SSA, AUS and SAF, of these, only SEA has RWF < 1 . In GB05, SEA showed a RWF < 1 as well, while the region named SSA in this study was split into two regions in GB05 the southernmost of which did present RWF < 1 while the northernmost region presented a RWF ~ 1 . Regions AUS and SAF were divided in GB05 as well with both parts presenting a RWF ~ 1 and no other regions in GB05 presented RWF < 1 . Of these 4 regions however, only SEA and AUS showed a comparatively tight constrain in our study, with both cutoff values presenting ratios close to or less than their global counterpart. For SAF, italicized in Table 2, condition 2 occurs while for SSA the upper cutoff value is rather high at 9.2 K. In all 4 regions the linear predictions obtained using AR4 model simulations compares well with the median values (top and bottom middle colors are similar) and we should also note that Southeast Asia, Australia and South America were identified by AR4-11 as the only land regions where the annual mean warming is not greater than the global mean.

[25] The results for the northernmost regions (ALA, GRL, NEU and NAS) compare well with the results from GB05 and AR4-11, displaying all a RWF > 1 . In GRL and NEU the AR4 model’s linear predictions compare well with the median of our constrained forecasts while for ALA and NAS the linear prediction is of stronger and weaker warming, respectively. In ALA, although the mean constrained value has RWF ~ 1.25 , both cutoff values, most significantly the higher limit, show ratios ~ 1 . This gives us a comparatively tight constraint on the climate response in that region. The median constrained values in MED, CAS, TIB and EAS, the Euro-Asian midlatitude regions, have somewhat higher RWF than in GB05 ranging from 1.5 to 1.75. In particular the Mediterranean region presented a high RWF close to 1.75 which supports the identification of this region as a climate hot spot [Giorgi, 2006] and does not contrast the findings in AR4-11. Linearly predicted values from the AR4 ensemble models yielded results that were very much in line with GB05 and with little variation among them (RWF ~ 1.25). Regions WNA, CNA, ENA and CAM, North American midlatitudes, generally agree with GB05 though in 3 out of 4 cases, condition 2 arises.

[26] Figure 4 shows the constrained regional precipitation responses to a doubling of preindustrial CO₂ concentrations. In Figure 4, darker colors represent increases in precipitation. All the values shown in each region box are given as ratios to the mean preindustrial precipitation, unlike in Figure 3 where the mean and cutoff values were given as ratios to their corresponding global counterparts. As an example, in cases such as GRL, where even the lower cutoff is positive, our methodology yields a positive response for precipitation with, at least, a 95% certainty. To

simplify the discussion of these results, we have divided the regions according to their response in three groups: regions where the lower cutoff is positive, hence the response is defined and positive within the accepted uncertainty (group 1), regions where the lower cutoff is negative and the upper cutoff is positive, hence the response is not defined within the accepted uncertainty (group 2), and regions where the upper cutoff is negative, hence the response is defined and negative within the accepted uncertainty (group 3). The northernmost regions (ALA, GRL, NEU, NEE and NAS), along with ENA, EAS, SAS, and EAF fall within group 1. These are the very same regions, other than TIB, that showed a distinctive increase in precipitation in GB05. An increase in precipitation is also predicted by AR4-11 in northern, eastern, southern and southeastern Asia, while a decrease in precipitation is expected for central Asia where, in our study, condition 2 arises. In our study TIB falls into group 2 since the lower cutoff value is slightly negative. In group 3 we find only MED which has a definite drying signal in GB05 and AR4-11. Also in GB05, CAM, AUS and SAH showed a drying response, but in our case the mean response is weak. All other regions fall into group 2. When the associated uncertainty is taken into account, these results compare well with prior studies. In comparing with GB05 we also recall that we analyze here annual precipitation, while GB05 divided the precipitation into 6-month wet and dry seasons, so that, when expressed in terms of percentage, the annual precipitation change would be dominated by the wet season one.

[27] Arguably, the most important result here is that large uncertainties are associated with all the precipitation response estimates and that, out of all the regions where condition 2 does not occur, only in 9 does the explained fractional variance exceed 0.5. Southeast Asia is the region with the tightest constrain, and the Mediterranean basin (one of the most prominent hot spots identified by G06) is the only region with a comparatively tight constraint and a non zero response. Very large upper cutoff values are given for GRL, ALA, NEU and NAS, the last two of which are identified as climate hot spots by G06. Large values of the upper cutoff are also found in SAH and CAM. In these regions the large percentage changes are mostly due to the low values of preindustrial precipitation. By comparison, the large uncertainties in WAF precipitation response reflect a broad uncertainty ranges in the absolute values. In some cases we obtain a comparatively tight constrain for a zero, or very small, response. This is the case for AMZ, SSA, SAF, SEA which appear to be relatively less affected by climate changes in the analysis of G06. This is a very different result from those of other no-response regions such as CAM and WAF where the median is accompanied by comparatively large uncertainties. Finally, we point out that in most regions the linear predictions from the AR4 slab model ensemble agree reasonably well with the median of our estimates. The difference between the two values is largest in CAS and EAS.

4. Summary and Conclusions

[28] In this study we describe a methodology for producing constrained estimates of the climate system response to external forcing and we apply it to the multithousand

member ensemble of slab model simulations from the distributed computing project that is *ClimatePrediction.net*. Our goal is to constrain the response in 21 nonoverlapping land regions of mean surface temperature and precipitation to the doubling of preindustrial concentrations of CO₂ with all available observations.

[29] The regional mean values are in general agreement with previous regional climate change studies and the linear section of the methodology described here yields similar results when applied to the AR4 slab model simulations. Land-based temperature responses are generally higher than the global value. This, together with the fact that regional forecasts tend to have a looser constraint (broader PDF), results in relatively high upper limits. The regions that show a negative or very small response are mostly confined to the southern hemisphere while those that show a very strong warming are confined to northern part of the Euro-Asian landmass.

[30] Regional constraints on precipitation are generally looser than those we obtain for surface temperature and for a large group of regions we are unable to determine the sign of the response, with a 95% confidence level. Precipitation is expected to increase in the northernmost regions and for Alaska and Greenland the 95% upper cutoff level is extremely high. In arid and semiarid regions percentage changes in precipitation are practically unbounded. In these regions absolute values of precipitation might prove to be a better quantity to constrain.

[31] Of all the 21 regions, only the Mediterranean Basin shows a relatively tight constrain on precipitation showing a drying response to doubling of preindustrial carbon dioxide concentrations. The later, combined with the strong regional surface temperature response, to supports the identification of the Mediterranean Basin as a climate hot spot (see G06).

[32] The loose constraints and small amounts of explained variance that we obtain in some cases could be improved by shifting our focus from annual to seasonal means. This is particularly true for those regions that showed significantly different responses in different seasons in GB05. In particular, the precipitation response in South America, India and South Africa would be expected to yield tighter constrains divided according to the season. Also splitting regions that have opposite seasonal cycles in different areas, like northern and southern Australia, would likely give us a tighter constraint. A further natural development will be to apply this methodology to a multithousand member ensemble of fully coupled atmosphere-ocean general climate models as they become available from the *ClimatePrediction.net* project.

Appendix A: Step by Step Construction of Climate EOFs and Their Projections Onto Ensemble Control Climate Simulations—“ x_{ij} ” Mentioned in Equation (1)

[33] The initial data set from the HadCM3 control climate simulation, that is with preindustrial values of CO₂, is composed of the 11 variables listed in Table 1 along with the grids they are defined on. Temperature at 1.5 m, mean sea level pressure, precipitation, and sensible and latent surface heat fluxes are defined on the following grid: [time, latitude, longitude, season] of size: [$N_t = 62$, $N_{lat} = 73$,

$N_{lon} = 96$, $N_s = 4$]. Zonally averaged relative humidity, temperature, and zonal and meridional winds are defined on the following grid: [time, pressure level, latitude, season] of size: [N_t , $N_p = 14$, N_{lat} , N_s]. Finally Zonally averaged top of atmosphere OLR and OSR the grid is: [time, latitude, season] and size [N_t , N_{lat} , N_s]. Here the “time” dimension represents 8 year means as explained in the main text.

[34] We start by taking the anomaly of each variable, that is, we subtract the mean over the first dimension, for example in the case of surface temperature:

$$T'(n, i, j, s) = T(n, i, j, s) - \frac{1}{N_t} \sum_n T(n, i, j, s),$$

where n, i, j, s are indexes for time, latitude, longitude and season respectively, T is surface temperature and T' is the anomaly field. Then we normalize by the globally averaged standard deviation of 8 year climates, again for the case of surface temperature:

$$T''(n, i, j, s) = T'(n, i, j, s) / \sigma_T(s),$$

where

$$\sigma_T(s) = \frac{1}{N_{lon} \sum_i \cos(lat(i))} \sum_i \cos(lat(i)) \cdot \sum_j \sqrt{\frac{1}{N_t - 1} \sum_n T'(n, i, j, s)^2},$$

and $lat(i)$ is the latitude of the HadCM3 grid i th latitude point. The $\cos(lat(i))$ is a mass weighting term.

[35] We normalize variables defined on other grids in an equivalent way, for example for zonally averaged relative humidity and zonally averaged top of atmosphere OLR:

$$RH'(n, k, i, s) = RH(n, k, i, s) - \frac{1}{N_t} \sum_n RH(n, k, i, s)$$

$$RH''(n, k, i, s) = RH'(n, k, i, s) / \sigma_{RH}(s)$$

$$\sigma_{RH}(s) = \frac{1}{pres(1) \sum_i \cos(lat(i))} \sum_k (pres(k) - pres(k + 1)) \cdot \sum_i \cos(lat(i)) \sqrt{\frac{1}{N_t - 1} \sum_n RH'(n, k, i, s)^2},$$

where $pres(k)$ is the pressure of the k th HadCM3 pressure level and $pres(N_p + 1)$ is assumed to be 0. The $(pres(k) - pres(k + 1))$ term is a mass weighting term.

$$OLR'(n, i, s) = OLR(n, i, s) - \frac{1}{N_t} \sum_n OLR(n, i, s)$$

$$OLR''(n, i, s) = OLR'(n, i, s) / \sigma_{OLR}(s)$$

$$\sigma_{OLR}(s) = \frac{1}{\sum_i \cos(lat(i))} \cdot \sum_i \cos(lat(i)) \sqrt{\frac{1}{N_t - 1} \sum_n OLR'(n, i, s)^2}$$

[36] As mentioned in the text, all variables are now normalized by the portion of atmospheric mass in the grid cells, for example for surface temperature:

$$T'''(n, i, j, s) = T''(n, i, j, s) \times w_{surface}(i),$$

where

$$w_{surface}(i) = \frac{1}{N_{lon}} \times \frac{\cos(lat(i))}{\sum_i \cos(lat(i))}.$$

[37] Again other variables defined on different grids were normalized accordingly, for example for relative humidity and OLR:

$$RH'''(n, k, i, s) = RH''(n, k, i, s) \times w_{zonal}(k, i)$$

where

$$w_{zonal}(k, i) = \frac{pres(k) - pres(k + 1)}{pres(1)} \times \frac{\cos(lat(i))}{\sum_i \cos(lat(i))}.$$

$$OLR'''(n, i, s) = OLR''(n, i, s) \times w_{latitude}(i)$$

where

$$w_{latitude}(i) = \frac{\cos(lat(i))}{\sum_i \cos(lat(i))}.$$

[38] At this point, the variables are still defined on the grids described above which coincide with the entire global model grids. However, in Table 1 we point out that many variables are defined only on a subset of the grid points. For example surface temperature (T) is defined on a [time, latitude, longitude, season] grid but only land points are considered. Another example is precipitation, where only 85°N to 85°S over land and 30°N to 30°S over ocean points are considered. All data corresponding to points outside the areas defined in Table 1, column 3, are tagged as missing data and will later be removed. Also pole grid points, that is latitude points $i = 1$ and $i = N_{lat} - 1$ are tagged as well.

[39] For each variable, all dimensions but the first are mapped onto a one-dimensional array. Using the same example as above for the different grids:

$$\underline{T}(n, ijs) = T'''(n, i, j, s),$$

where “ ijs ” is a new index and

$$ijs = (iN_{lon}N_s) + (jN_s) + s,$$

and similarly:

$$\underline{RH}(n, kis) = RH'''(n, k, i, s),$$

where “ kis ” is a new index and

$$kis = (kN_{lat}N_s) + (iN_s) + s,$$

and

$$\underline{OLR}(n, is) = \underline{OLR}'''(n, i, s),$$

where “is” is a new index and

$$is = (iN_s) + s.$$

The following step is to create a single one-dimensional array for each time (index “n”) by appending sequentially all normalized nondimensional variables, we will refer to the final vector as \underline{X} .

$$\underline{X}(n, i) = \underline{T}(n, i), \quad \text{for } i = 0 \text{ to } (N_{lat}N_{lon}N_s),$$

$$\underline{X}(n, i + (N_{lat}N_{lon}N_s)) = \underline{SLP}(n, i), \quad \text{for } i = 0 \text{ to } (N_{lat}N_{lon}N_s),$$

$$\underline{X}(n, i + 2 \times (N_{lat}N_{lon}N_s)) = \underline{P}(n, i), \quad \text{for } i = 0 \text{ to } (N_{lat}N_{lon}N_s),$$

$$\underline{X}(n, i + 3 \times (N_{lat}N_{lon}N_s)) = \underline{SSHf}(n, i),$$

for $i = 0$ to $(N_{lat}N_{lon}N_s)$,

$$\underline{X}(n, i + 4 \times (N_{lat}N_{lon}N_s)) = \underline{SLHF}(n, i),$$

for $i = 0$ to $(N_{lat}N_{lon}N_s)$,

where T, SLP, P, SSHF and SSHF are surface temperature, sea level pressure, precipitation, surface sensible heat flux and surface latent heat flux respectively.

$$\underline{X}(n, i + 5 \times (N_{lat}N_{lon}N_s)) = \underline{RH}(n, i), \quad \text{for } i = 0 \text{ to } (N_pN_{lon}N_s),$$

$$\underline{X}(n, i + 5 \times (N_{lat}N_{lon}N_s) + (N_pN_{lat}N_s)) = \underline{Tz}(n, i),$$

for $i = 0$ to $(N_pN_{lon}N_s)$,

$$\underline{X}(n, i + 5 \times (N_{lat}N_{lon}N_s) + 2 \times (N_pN_{lat}N_s)) = \underline{Uz}(n, i),$$

for $i = 0$ to $(N_pN_{lon}N_s)$,

$$\underline{X}(n, i + 5 \times (N_{lat}N_{lon}N_s) + 3 \times (N_pN_{lat}N_s)) = \underline{Vz}(n, i),$$

for $i = 0$ to $(N_pN_{lon}N_s)$,

where RH, Tz, Uz and Vz are zonally averaged relative humidity, temperature, zonal wind and meridional wind respectively.

$$\underline{X}(n, i + 5 \times (N_{lat}N_{lon}N_s) + 4 \times (N_pN_{lat}N_s)) = \underline{OLR}(n, i),$$

for $i = 0$ to $(N_{lat}N_s)$,

$$\underline{X}(n, i + 5 \times (N_{lat}N_{lon}N_s) + 4 \times (N_pN_{lat}N_s) + (N_{lat}N_s))$$

$= \underline{OLR}(n, i), \quad \text{for } i = 0 \text{ to } (N_{lat}N_s),$

where OSR and OLR are top of atmosphere outgoing shortwave and longwave radiation respectively. Finally all i th grid points tagged as missing data are removed from $\underline{X}(\mathbf{n}, \mathbf{i})$. Here we will refer to the column vectors of matrix $\underline{X}(\underline{X}(\mathbf{n}, *))$ as the climate vectors or simply climate. In this study we use this same procedure to calculate “climate vectors” from observations, from our ensemble control simulations and from IPCC AR4 slab model control simulations.

[40] To obtain the \mathbf{E} matrix of EOF column vectors mentioned in the methodology sector we simply perform singular value decomposition on the matrix \underline{X} :

$$\underline{X} = \mathbf{E}\mathbf{A}\mathbf{C}^T,$$

where \mathbf{E} is the matrix of empirical orthogonal functions, \mathbf{A} is the diagonal eigenvalue matrix and \mathbf{C}^T is the transposed matrix of principal components.

[41] Finally the amplitudes of the j th EOF in the i th ensemble member “ x_{ij} ” referred to in equation (1) are obtained as a dot product between the j th column of \mathbf{E} and the climate vector obtained, in very much the same way, from the control simulation of the i th model ensemble member.

[42] **Acknowledgments.** The authors thank the research staff of the ClimatePrediction.net group for their support. Special thanks to the international modeling groups for providing their data for analysis and to the Program for climate Model Diagnosis and Intercomparison (PCMDI) for collecting and archiving the model data. This work was supported by the United Kingdom Natural Environmental Research Council’s COAPEC, e-Science and fellowship programmes; by the Commission of the European Communities through the WATCH (Water and Global Change) project (036946); and by the hundreds of thousands of members of the public who participated in the ClimatePrediction.net program.

References

- Allen, M. R. (1999), Do-it-yourself climate prediction, *Nature*, *401*, 627.
- Allen, M. R., and D. A. Stainforth (2002), Towards objective probabilistic climate forecasting, *Nature*, *419*, 228.
- Christensen, H. D. (2007), Regional climate projections, in *Climate Change 2007: The Physical Science Basis—Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by S. Solomon et al., pp. 847–891, Cambridge Univ. Press, New York.
- Covey, C., et al. (2003), An overview of results from the coupled model intercomparison project, *Glob. Planet Change*, *37*(1–2), 103–133.
- Forest, C. E., et al. (2002), Quantifying uncertainties in climate system properties with the use of recent climate observations, *Science*, *295*, 113–117.
- Forest, C. E., P. H. Stone, and A. P. Sokolov (2006), Estimated PDFs of climate system properties including natural and anthropogenic forcings, *Geophys. Res. Lett.*, *33*, L01705, doi:10.1029/2005GL023977.
- Gibson, J. K., et al. (1997), ERA description, *ECMWF Reanal. Proj. Rep. Ser.*, *1*, Eur. Cent. for Med.-Range Weather Forecasts, Reading, U. K.
- Gillett, N. P., F. W. Zwiers, A. J. Weaver, G. C. Hegerl, M. R. Allen, and P. A. Stott (2002), Detecting anthropogenic influence with a multi-model ensemble, *Geophys. Res. Lett.*, *29*(20), 1970, doi:10.1029/2002GL015836.
- Giorgi, F. (2006), Climate change hot-spots, *Geophys. Res. Lett.*, *33*, L08707, doi:10.1029/2006GL025734.
- Giorgi, F., and X. Bi (2005), Updated regional precipitation and temperature changes for the 21st century from ensembles of recent AOGCM simulations, *Geophys. Res. Lett.*, *32*, L21715, doi:10.1029/2005GL024288.
- Giorgi, F., and L. O. Mearns (2002), Calculation of average uncertainty range and reliability of regional climate changes from AOGCM simulations via the “Reliability Ensemble Averaging (REA)” method, *J. Clim.*, *15*, 1141–1158.
- Giorgi, F., et al. (2001), Emerging patterns of simulated of regional climatic changes for the 21st century due to anthropogenic forcings, *Geophys. Res. Lett.*, *28*, 3317–3320.
- Hansen, J. A., et al. (2001), Casino-21: Climate simulation of the 21st century, *World Res. Rev.*, *13*, 187–189.

- Harris, G. H., et al. (2006), Frequency distributions of transient regional climate change from perturbed physics ensembles of general circulation model simulations, *Clim. Dyn.*, 27(4), 357–375.
- Harrison, E. F., et al. (1990), Seasonal variation of cloud radiative forcing derived from the Earth Radiation Budget Experiment, *J. Geophys. Res.*, 95, 18,687–18,703.
- Hegerl, G. C., and M. R. Allen (2002), Origins of model-data discrepancy in optimal fingerprinting, *J. Clim.*, 15, 1348–1356.
- Jolliffe, I. T. (2002), *Principal Component Analysis*, Springer, New York.
- Jones, P. D., M. New, D. E. Parker, S. Martin, and I. G. Rigor (1999), Surface air temperature and its variations over the last 150 years, *Rev. Geophys.*, 37, 173–199.
- Intergovernmental Panel on Climate Change (2001), Third Assessment Report, Geneva, Switzerland. (Available at <http://www.ipcc.ch/pub/online.htm>)
- Meehl, G. A., et al. (2007), THE WCRP CMIP3 multimodel dataset: A new era in climate change research, *Bull. Am. Meteorol. Soc.*, 88, 1383–1394.
- Murphy, J. M., et al. (2004), Quantifying uncertainties in climate change using a large ensemble of general circulation model predictions, *Nature*, 430, 768–772.
- New, M., M. Hulme, and P. Jones (1999), Representing twentieth-century space-time climate variability. Part I: Development of a 1961–90 mean monthly terrestrial climatology, *J. Clim.*, 12, 829–856.
- Palmer, T. N., and J. Räisänen (2002), Quantifying the risk of extreme seasonal precipitation events in a changing climate, *Nature*, 415, 512–517.
- Palmer, T. N., et al. (2004), Development of a European multi-model ensemble system for seasonal-to-interannual prediction (DEMETER), *Bull. Am. Meteorol. Soc.*, 85, 853–872.
- Piani, C., D. J. Frame, D. A. Stainforth, and M. R. Allen (2005), Constraints on climate change from a multi-thousand member ensemble of simulations, *Geophys. Res. Lett.*, 32, L23825, doi:10.1029/2005GL024452.
- Pope, V. D., et al. (2000), The impact of new physical parameterizations in the Hadley Centre climate model, HadAM3, *Clim. Dyn.*, 16, 123–146.
- Räisänen, J., and T. N. Palmer (2001), A probability and decision model analysis of a multimodel ensemble of climate change simulations, *J. Clim.*, 14, 3212–3226.
- Randal, D. A., et al. (2007), Climate models and their evaluation, in *Climate Change 2007: The Physical Science Basis—Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by S. Solomon et al., pp. 589–663, Cambridge Univ. Press, New York.
- Schneider, S. H. (2001), What is dangerous climate change?, *Nature*, 411, 17–19.
- Smith, L. (2002), What might we learn from climate forecasts?, *Proc. Natl. Acad. Sci. U. S. A.*, 99, 2487–2492.
- Stainforth, D., et al. (2002), Uncertainty in predictions of climate response to rising levels of greenhouse gases, *Nature*, 433, 403–406.
- Stott, P. A., and J. A. Kettleborough (2002), Origins and estimates of uncertainty in predictions of twenty-first century temperature rise, *Nature*, 416, 723–726.
- Xie, P., and P. A. Arkin (1997), Global precipitation: A 17-year monthly analysis based on gauge observations, satellite estimates and numerical model outputs, *Bull. Am. Meteorol. Soc.*, 78, 2539–2558.

M. R. Allen, C. Christensen, and B. Sanderson, Department of Physics, University of Oxford, Oxford OX1 3PU, UK.

D. J. Frame, Environmental Change Institute, University of Oxford, Oxford OX1 3QY, UK.

F. Giorgi and C. Piani, Abdus Salam International Center for Theoretical Physics, I-34014 Trieste, Italy. (cpiani@ictp.it)