Exploring basic issues in climate model parameter dependence in a fast general circulation model

J. David Neelin[1,2], Annalisa Bracco[3], Joyce E. Meyerson[1], Jim McWilliams[1,2], Michael Ghil[1,2,4], and Ed Huckle[1]

[1]Department of Atmospheric and Oceanic Sciences, University of California, Los Angeles, CA
[2]Institute of Geophysics and Planetary Physics, University of California, Los Angeles, CA
[3]School of Earth and Atmospheric Sciences, Georgia Institute of Technology, Atlanta, GA
[4]Département Terre-Atmosphère-Océan, Ecole Normale Supérieure, Paris, France

ABSTRACT

A general circulation model with simplified parameterizations and course vertical resolution is used to perform systematic slices in parameter space to explore basic issues in the dependence on model parameters. Even for climatological variables, a fairly large ensemble simulations is required to define the parameter space dependence with reasonable precision. For each of several selected parameters, an admissible range of variation is chosen a priory, and then a series of simulations is conducted over this range. For global measures such as root mean square error relative to reanalysis data sets, or correlation of simulated fields to reanalysis, the parameter dependence tends to be relatively smooth over the admissible range. Defining local optima along each parameter slice, it is found that these often occur at the end of the range. Furthermore, optima for different variables tend to occur at substantially different parameter values. This is a symptom of the well-known problem of improving one aspect of a simulation while causing degeneration in another aspect, and implies that a global optimization procedure would have a strong dependence on the weighting given to each variable when defining the cost function.

INTRODUCTION

A number of basic issues in the dependence of climate model simulations on model parameters are poorly understood. A fast climate model with simple parameterizations, known as Speedy (Molteni 2003; Bracco et al. 2004), is being used to do parameter space exploration with the aim of the elucidating these issues. These include: How typical is it to encounter sudden changes in response for a small change in a parameter, as opposed to smooth evolution in parameter space? There has been interest in use of formal or informal optimization methods to tune parameters for climate simulations (Severijns and Hazeleger 2005; Jones et al. 2005; Kunz et al. 2008) --- can one lay a more thorough groundwork for such applications? The optimization problem for climate has considerable ambiguity in terms of how to weight individual variables in the cost function --- is this likely to be a crucial issue?

Figure 1 shows Speedy precipitation climatology for reference, since precipitation is often one of the more challenging variables to simulate accurately. Compared to observations from the Climate Prediction Center Merged Analysis of Precipitation (CMAP) data set, may be seen that many major features are well simulated at large scales, although as with many climate models, departures are easily found at regional scales.

Precipitation climatology 1978 - 2002

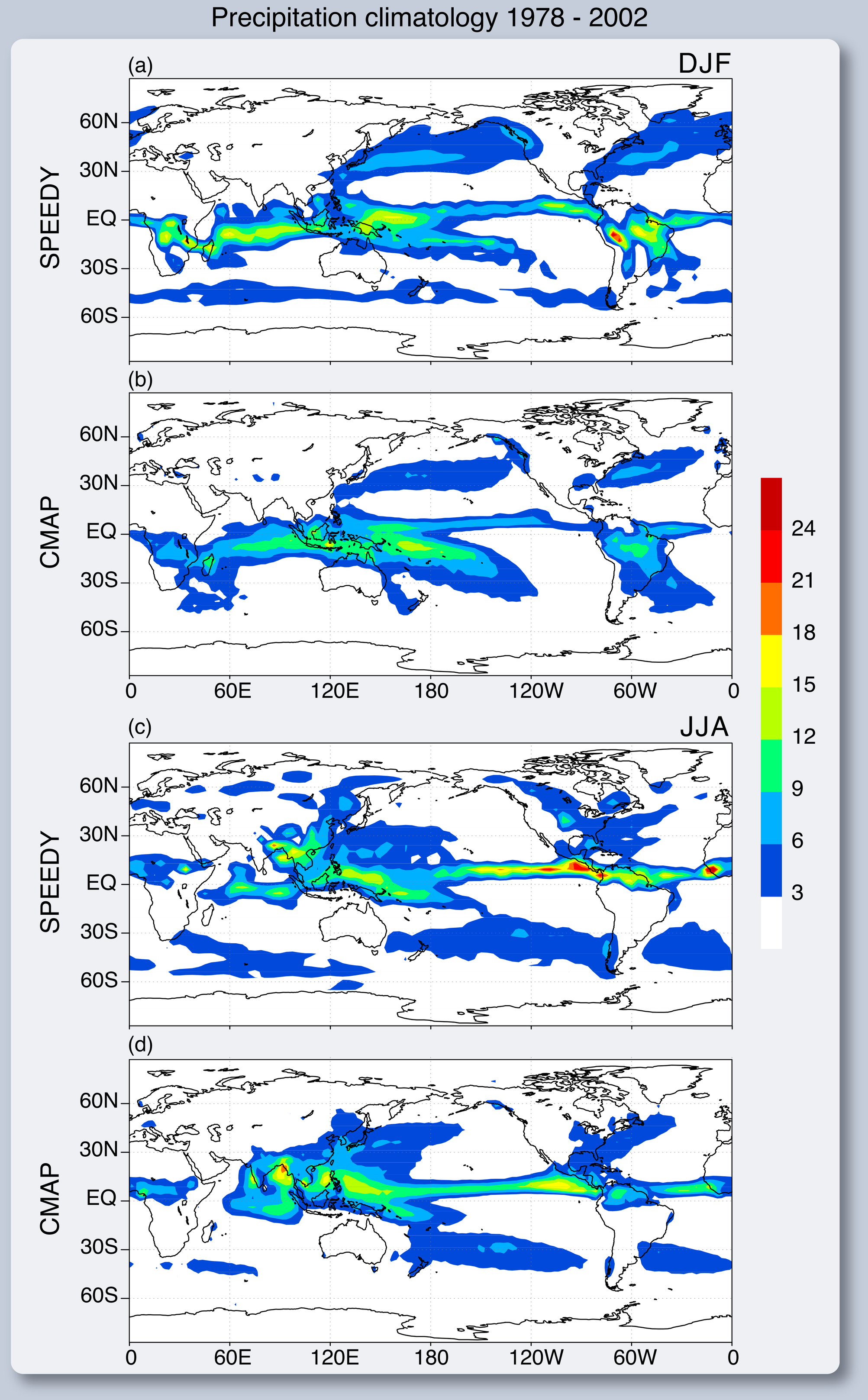

Figure 1 Climatology of precipitation from Speedy simulations with observed sea surface temperature for (a) December-February; (c). June-August compared to observed precipitation from the CMAP data set for (b) December-February; (d) June-August.

PARAMETER SPACE SLICES

To illustrate the parameter space dependence, four parameters are chosen that are known to have significant impacts on different aspects of the model solution. These are: a sub-gridscale wind speed gustiness parameter that creates a minimum wind speed in the bulk formula for surface fluxes; a relative humidity parameter from the deep convective parameterization that controls the moisture towards which convection adjusts the column; a cloud albedo parameter; and a viscosity parameter, here measured by the damping time implied for the shortest spatial scale. For each of these parameters, an "admissible range" is chosen. In most climate model parameterizations, the modeler has external information regarding the range through which the parameters may reasonably be varied. In some cases, there is an absolute limit, such as 100% relative humidity, in some cases information from past runs regarding a viscosity below which numerical instability tends to be encountered, and in some cases a sense from observations that led to the parameterization that values beyond these limits become increasingly implausible. Here, the limits represent the judgment of the values beyond which we would be uncomfortable tuning the parameterization based on a combination of such grounds, combined with the desire to have a reasonable span on either side of the standard value that has been established by past tuning and evaluation of the model. Four values are chosen equally spaced on either side of the standard parameter value, providing a slice with nine parameter values in each parameter direction spanning this admissible range. An ensemble of 10 simulations of 25 years each, forced with observed sea surface temperature, is carried out for each parameter value. National Center for Environmental Prediction (NCEP) reanalysis (Kalnay et al 1996) is used for observational comparison.

Speedy JJA precipitation root-mean-square error relative to NCEP reanalysis

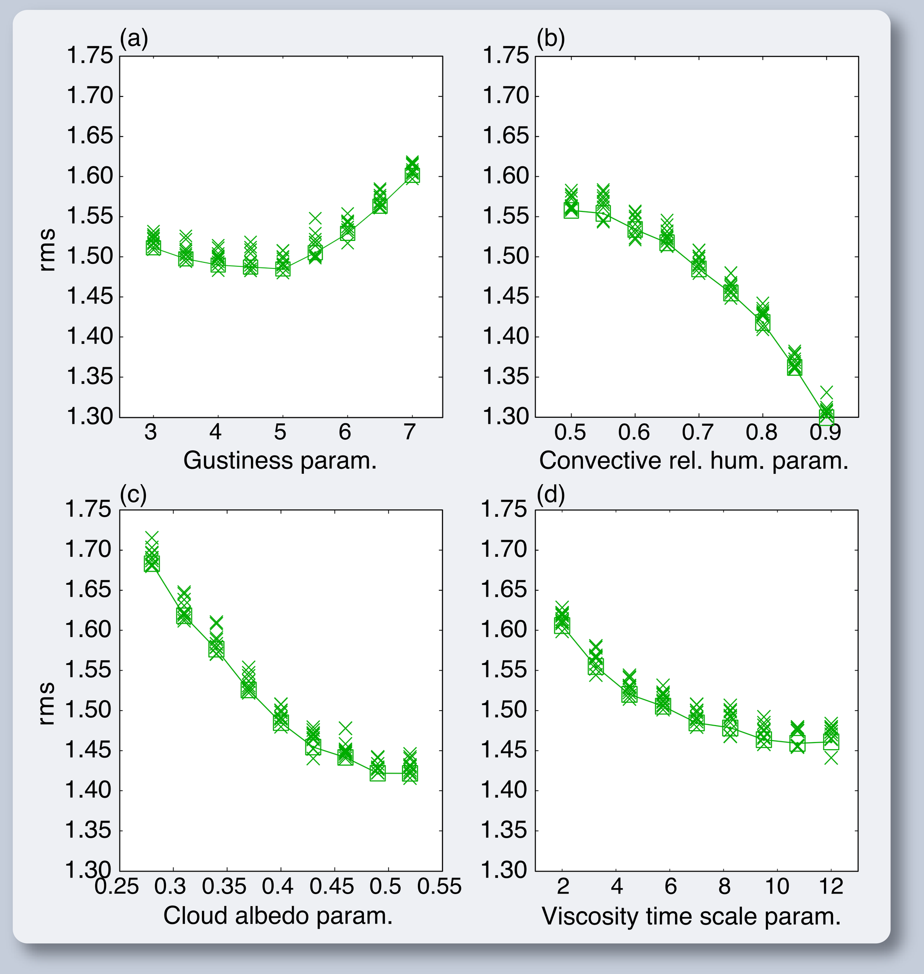

Figure 2 Root-mean-square error of Speedy simulated June-August precipitation relative to NCEP reanalysis, as a function of model parameter. (a) sub grid scale wind speed gustiness parameter; (b) relative humidity parameter from the deep convective parameterization; (c) cloud albedo parameter; and (d) viscosity time scale parameter. Results from each of an ensemble of 10 simulations are shown for each parameter value. Squares connected by lines show the ensemble mean values.

Speedy JJA precipitation correlation to NCEP reanalysis

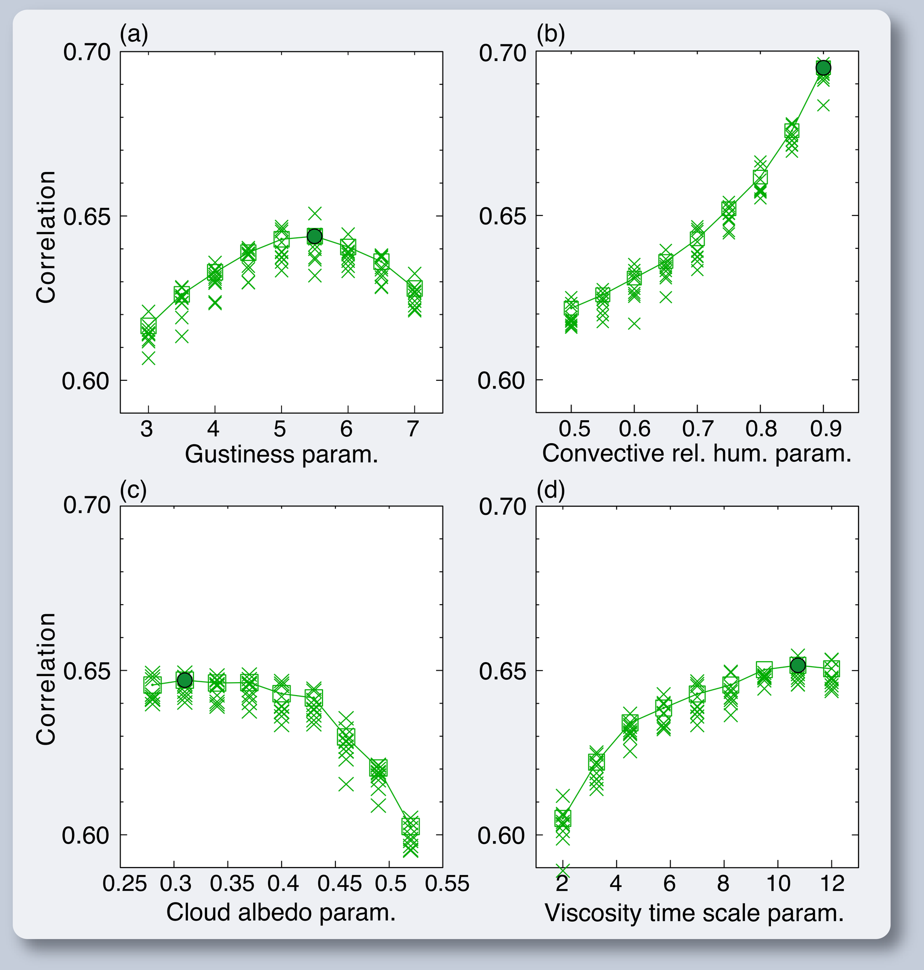

Figure 3 As in Figure 2, but for correlation of Speedy simulated June-August precipitation to NCEP reanalysis, as a function of the same model parameters. The optimum within each parameter slice, here as measured by maximum correlation, is indicated by a circle.

Figures 2 and 3 provide examples of model parameter dependence. Figure 2 shows root-mean-square error of the simulation relative to NCEP, and Figure 3 shows spatial correlation to NCEP, using the example of June-August precipitation climatology, with both values evaluated over the whole globe. One issue facing climate model evaluation can immediately be seen in the spread among the ensemble members, even for global correlations on a climatological quantity. This limits the precision to which any cost function can be evaluated. It is worth noting that this cannot be entirely overcome by increasing the ensemble size, because the observations themselves have comparable error.

The ensemble mean values (squares) are seen to typically evolve smoothly as a function of parameter, for these global measures, and to this precision. This is apparent good news for potential optimization procedures, since it implies that there may be substantial ranges for which a reasonable cost function constructed from climatological variables is not too rough. However, there are a number of other features which signal caution. In Figures 2a and 3a, the optimum solution, i.e., the minimum rms error and the maximum correlation to observations, respectively, are achieved within the admissible parameter range. A roughly quadratic parameter dependence occurs about this optimum, again a factor that might favor optimization procedures. However, the optimum determined by minimum rms and that determined by maximum correlation do not occur at exactly the same parameter value, an indication that there will be some dependence on how the cost function is constructed.

Furthermore, another typical behavior is seen in Figures 2b and 3b as a function of the convective relative humidity parameter. The optimum occurs at one end of the permissible parameter range. Pragmatically, this may be taken to indicate that blind optimization may often lead to the model operating near the limits of validity of some of the parameterizations. Figures 2c,d and Figure 3c,d reinforce both of these points. The optima occur very close to the limit of the admissible range. Furthermore, in the case of cloud albedo, the optimum determined by correlation occurs near the low end, while the optimum determined by minimum rms occurs near the high end.

How do these parameter dependences compare for different climatological variables? Figure 4 shows correlations for several variables as a function of the four parameters. The situation ones hopes for is typified by the dependence on the viscosity, in which multiple parameters tend to have the correlation improve as one moves towards lower viscosity (longer damping time). However, the situation for cloud albedo appears to be more typical. The optima for different variables occur at different parameter values.

Speedy correlation to NCEP reanalysis for several variables

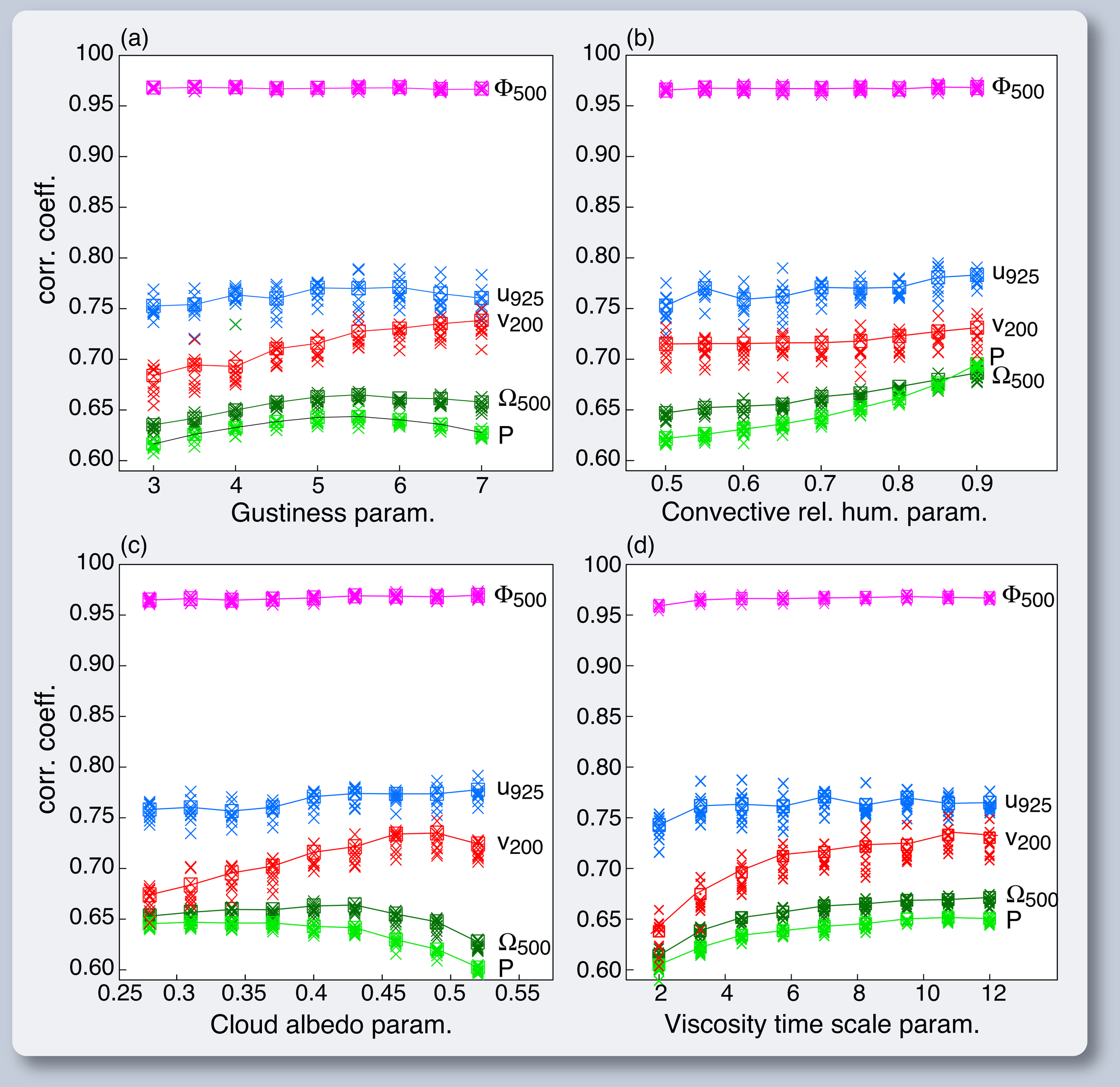

Figure 4 Correlation of Speedy simulated variables to NCEP reanalysis, as a function of the same model parameters as in Figures 2 and 3, but showing values for several variables. Variable names follow the convention: u_{925} zonal wind at a near surface level, u_{200} indicates the zonal wind upper troposphere, and similarly for v meridional wind, T temperature, Φ geopotential and Ω vertical pressure velocity.

DISTRIBUTION OF OPTIMA

To provide a sense of how the typical these behavior are, we can plot parameter space location of the optima for various variables. Figure 5 shows a three-dimensional visualization of these optima as a function of three of the parameters. Note that these are only determined along one-dimensional parameter slices, but they provide a sense of what might happen in the process of optimizing in three dimensions simultaneously. Optima such as those seen in Figures 4 and 5 are each indicated by a dot, color-coded according to the climate variable for which the optimum was obtained. Optima occurring at the ends of the permissible range are very common. Furthermore, the optima often occur at substantially different parameter values for different climate variables. Fortunately, there is considerable tendency for a given variable to optimize a similar location in different seasons, although this is not guaranteed.

The tendency for optima of different variables to occur at different parameter space locations may be described as a "tension" among different metrics for the accuracy of the model simulation. It implies that determination of a global optimum would depend strongly on the weighting of each variable in the cost function. For instance, a cost function that gave heavy weight to the quality of the precipitation simulation would yield optimization to a different set of parameters than one that gave more weight to low-level wind simulation. In essence, this helps to quantify a frustration long known to climate modelers of improving one aspect of simulation while making another aspect worse. Together with the very high cost of evaluation of the simulation at each point in parameter space, this points toward developing optimization strategies that retain a large amount of model information at selected parameter values and consider a range of potential cost functions that reflect the requirements of different climate model users.

Locations of Ensemble-Mean Maximum Correlation

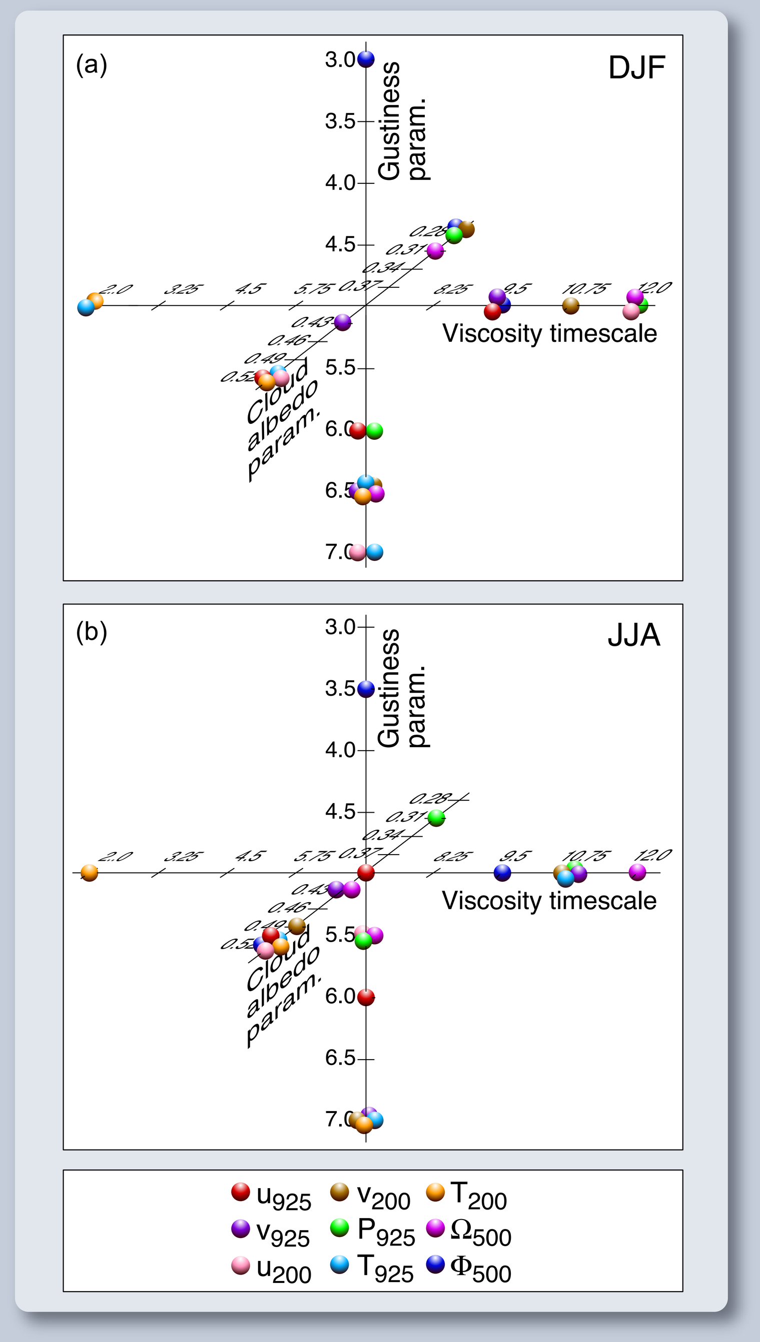

Figure 5 Local optima determined as maxima in correlation of ensemble mean simulated values to NCEP for individual climate variables along three directions in parameter space. These are determined separately for (a) December-February, and (b) June-August. Dots indicate the positions of the optima, colors indicate the climate variable as shown in the legend (variable names as in Figure 4).

REFERENCES

Molteni, F., 2003: Atmospheric simulations using a GCM with simplified physical parameterizations. I. Model climatology and variability in multi-decadal experiments. *Clim. Dyn.*, **20**, 175-191.

Bracco, A., F. Kucharski, R. Kallummal, F. Molteni, 2004: Internal variability, external forcing and climate trends in multi-decadal AGCM ensembles. *Clim. Dyn.*, **23**, 659-678.

Jones, C., J. Gregory, R. Thorpe, P. Cox, J. Murphy, D. Sexton, P. Valdes, 2005: Systematic optimisation and climate simulation of FAMOUS, a fast version of HadCM3. *Clim. Dyn.* **25**, 189-204.

Kalnay, E., and coauthors, 1996: The NCEP/NCAR 40-year reanalysis project. *Bull. Amer. Meteorol. Soc.*, **77**, 437-471.

Kunz, T., K. Fraedrich, and E. Kirk, 2008: Optimisation of simplified GCMs using circulation indices and maximum entropy production. *Clim. Dyn.*, **30**, 803-813.

Severijns, C. A., and W. Hazeleger, 2005: Optimizing Parameters in an Atmospheric General Circulation Model. *J. Climate*, **18**, 3527-3535.