



UCRL-JRNL-201134

**PCMDI Report No. 70**

**THE CCPP-ARM PARAMETERIZATION TESTBED (CAPT):  
WHERE CLIMATE SIMULATION MEETS WEATHER PREDICTION**

**Thomas J. Phillips<sup>1</sup>, Gerald L. Potter<sup>1</sup>, David L. Williamson<sup>3</sup>, Richard T. Cederwall<sup>2</sup>, James S. Boyle<sup>1</sup>, Michael Fiorino<sup>1</sup>, Justin J. Hnilo<sup>1</sup>, Jerry G. Olson<sup>3</sup>, Shaocheng Xie<sup>2</sup>, and J. John Yio<sup>2</sup>**

<sup>1</sup> Program for Climate Model Diagnosis and Intercomparison, Lawrence Livermore National Laboratory, Livermore, California

<sup>2</sup> Atmospheric Science Division, Lawrence Livermore National Laboratory, Livermore, California

<sup>3</sup> National Center for Atmospheric Research, Boulder, Colorado

November, 2003

**PROGRAM FOR CLIMATE MODEL DIAGNOSIS AND INTERCOMPARISON  
LAWRENCE LIVERMORE NATIONAL LABORATORY  
UNIVERSITY OF CALIFORNIA  
LIVERMORE, CA 94550**

## DISCLAIMER

**This document was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor the University of California nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or the University of California, and shall not be used for advertising or product endorsement purposes.**

**This report has been reproduced  
directly from the best available copy.**

**Available to DOE and DOE contractors from the  
Office of Scientific and Technical Information  
P.O. Box 62, Oak Ridge, TN 37831  
Prices available from (615) 576-8401, FTS 626-8401**

**Available to the public from the  
National Technical Information Service  
U.S. Department of Commerce  
5285 Port Royal Rd.,  
Springfield, VA 22161**

## ABSTRACT

To significantly improve the simulation of climate by general circulation models (GCMs), systematic errors in representations of relevant processes must first be identified, and then reduced. This effort demands that the GCM parameterizations of unresolved processes, in particular, should be tested over a wide range of frequencies, not just at climate scales. A methodology for developing parameterizations in numerical weather prediction (NWP) models therefore may be applicable, provided this approach is appropriately adapted for climate GCMs. Such an NWP-inspired methodology entails the generation of short-range weather forecasts by a realistically initialized climate model, and the use of high-frequency NWP analyses and observations of parameterized variables to evaluate these forecasts. The efficacy of modified GCM parameterizations also can be tested in such a weather-forecasting framework.

In order to further this approach for improvement of parameterizations in climate GCMs, the U.S. Department of Energy is funding a joint venture of its Climate Change Prediction Program (CCPP) and Atmospheric Radiation Measurement (ARM) Program: the CCPP-ARM Parameterization Testbed (CAPT). In this article, we elaborate the scientific rationale for CAPT, discuss technical aspects of its implementation in a representative climate GCM, and present results that illustrate the CAPT methodology.

## 1. Introduction

Climate simulations performed with general circulation models (GCMs) are widely viewed as the principal scientific basis for developing policies to address potential future global climate change (Houghton et al. 2001). In order to reduce uncertainties in these GCM projections of future climate, there is a compelling need to improve the simulation of processes that produce the *present* climate. This undertaking demands close attention to systematic errors in GCM simulations.

Systematic errors are persistent (average) departures of the model solution from an appropriate observational standard. For example, the GCM systematic climate error is defined by the departure of the simulated climate statistics from the observed (e.g. by a difference in monthly means). For GCMs used in numerical weather prediction (NWP), the relevant systematic error is defined instead by the mean departure of the model forecast from observations at short time scales. As the length of the forecast increases, the systematic forecast error approaches the systematic climate error (Palmer 1999).

Thus, in order to enhance GCM performance ongoing interdependent efforts are needed:

1) to diagnose the details of model systematic errors by comparing GCM simulations with available observations over a range of time scales;

and

2) to reduce these systematic errors by improving the representation of key processes, and thereby increase the accuracy of GCM simulations relative to available observations.

For many years, the Working Group on Numerical Experimentation (WGNE) of the World Climate Research Programme (WCRP) has coordinated efforts to diagnose systematic errors in atmospheric GCMs (AGCMs). WGNE, for example, has sponsored conferences on the characteristics and causes of systematic errors in GCMs designed both for NWP and climate applications (WGNE 1988, Jasper and Meighen 2000). WGNE also has promoted collaborative modelling initiatives such as the Atmospheric Model Intercomparison Project (AMIP) (Gates 1992, Gates et al. 1999) to analyze systematic climate errors in AGCM simulations made with prescribed ocean boundary conditions. The WCRP Working Group on Coupled Modelling

(WGCM) is now extending the diagnosis of systematic errors to climate simulations of coupled ocean-atmosphere GCMs (OAGCMs) through initiatives such as the Coupled Model Inter-comparison Project (CMIP) (Meehl et al. 2000).

Ongoing efforts to reduce GCM systematic errors entail both enhancements of the resolution at which the model state variables (e.g. for AGCMs, the pressure, temperature, moisture, and wind fields) are predicted, and fundamental improvements in the parameterizations of unresolved subgrid-scale processes (e.g. radiation, clouds, convection, precipitation microphysics, turbulent fluxes and diffusion). In climate simulations, the parameterizations are crucially important for correct representation of relevant processes, while the computational costs of increasing resolution are very high. Hence, parameterization development usually is emphasized over resolution enhancement as the chief means of reducing systematic errors in GCMs designed for climate simulation (hereafter, "climate GCMs")

The deciding factor in choosing a new parameterization for a climate GCM is whether its inclusion brings the simulated climate into closer agreement with the observed statistics. However, there are inherent limitations in evaluating GCM parameterizations only in climate-simulation mode. First, the observed climate statistics are only roughly known on the global scale--to greater or lesser degree, depending on the process of interest (e.g. Kistler et al. 2001). Moreover, because the GCM climate state reflects compensating errors in the simulation of many nonlinear processes, it is very difficult to attribute these errors to particular parameterization deficiencies. In such a context also, the parameterizations are driven by an unrealistic large-scale state, so that it is also difficult to evaluate their performance objectively (Schubert and Chang 1996).

For these reasons, climate GCM developers have adopted process-oriented approaches that employ high-frequency observations for evaluating parameterizations. Some effects of introducing a new GCM parameterization can be assessed, for example, within the framework of single-column models (SCMs) or cloud-resolving models (CRMs) (Randall et al. 1996, Xie et al. 2002, Xu et al. 2002, Randall et al. 2003). The strength of this approach is that the column parameterizations are driven by an evolving large-scale dynamical state that is specified from observations. However, because all relevant high-frequency dynamical forcings for the atmospheric column must be specified, there are only a limited number of observational cases

that can be studied with an SCM or CRM. Feedbacks from the column parameterizations to the large-scale dynamics also cannot be represented by these models.

In order to include such feedbacks, some model developers (e.g. Jeuken et al. 1996, Kaas et al. 1999) retain the full GCM, but continuously relax the simulated mass and momentum fields toward a high-frequency global weather analysis. The objective is to constrain the GCM large-scale dynamics close to reality so that the resulting mean high-frequency tendency errors are attributable mainly to parameterized physical processes. Diagnosis of these high-frequency errors then can suggest possible ways to improve the parameterizations, and thereby reduce the low-frequency (climate) systematic errors.

Operational NWP centers follow yet another methodology in developing parameterizations for fine-resolution weather forecasting GCMs: the model dynamics are allowed to evolve freely and to interact fully with the parameterizations, so that all the forcings and feedbacks are generated by the GCM (e.g. Jakob 2003). In this approach, the state variables of the forecast GCM are first initialized by a data assimilation system (DAS) which usually is built around the GCM itself. After ingestion of all available observations (e.g. surface, radiosonde, and satellite measurements), the DAS applies variational methods to produce an optimal analysis of the global weather that defines the initial conditions for the forecast GCM (Daley 1991, Kalnay 2003). In addition, the DAS provides departures of the model from observations that can be used to guide further parameterization development.

Given an accurate analysis, it is reasonable to assume that the model state remains close to "truth" in the early period of the forecasts, so that the systematic forecast error can be attributed largely to parameterization deficiencies. This systematic error is estimated from differences between the mean of a sequence of short-range (~ five-day) forecasts of state variables and corresponding NWP analyses of weather observations. In addition, errors in parameterized model variables (e.g. radiative and turbulent fluxes, cloud properties, precipitation, etc.) are estimated from field observations or other data that are not ordinarily assimilated by the DAS (e.g. Mace et al. 1998, Miller et al. 1999). Selected model parameterizations then are modified so as to ameliorate the perceived deficiencies, and these scheme changes also are evaluated in short-range GCM forecasts to determine whether they reduce the model's high-frequency systematic errors. If that is the case, the new parameterizations are usually also evaluated in model

integrations beyond the deterministic forecast range of ~ 15 days to determine whether they reduce low-frequency systematic errors.

In view of the benefits of applying this NWP development methodology to fine-resolution forecast models, some scientists (e.g. Hollingsworth 1998, Miller 1999, Jakob 2003) have advocated adoption of analogous procedures for developing parameterizations in coarse-resolution climate GCMs. To realize this goal, alternatives to NWP procedures that rely on a GCM-compatible DAS must be developed, since many climate models lack such a resource. Working relationships also must be forged between GCM developers and parameterization specialists who do not always share the same institutional affiliations.

The U.S. Department of Energy (USDOE) is well-positioned to foster these scientific collaborations because of the support it provides to GCM developers through the Climate Change Prediction Program (CCPP) and to parameterization specialists through the Atmospheric Radiation Measurement (ARM) Program. Moreover, the extensive high-frequency ARM field data that have been collected over the last decade (Stokes and Schwartz 1994, Ackerman and Stokes 2003) are potentially very useful for evaluating GCM parameterizations (e.g. Morcrette 2002). Hence, the USDOE has established a new joint initiative, the CCPP-ARM Parameterization Testbed (CAPT), in order to resolve technical problems in implementing the NWP development methodology in climate GCMs, and to facilitate the needed scientific collaborations as well.

The remainder of this article reports on the progress of CAPT to date. In Section 2, we discuss the scientific premise of the project and outline the steps in the CAPT diagnostic protocol. In Section 3, we elaborate on technical aspects of implementing this protocol for version 2.0 of the Community Atmosphere Model (CAM2), an AGCM which was developed under the auspices of the National Center for Atmospheric Research (NCAR) (Collins et al. 2003). We present preliminary results of applying the CAPT methodology to this climate GCM in Section 4, and close with a brief summary in Section 5.

## 2. Premise and Protocol

CAPT is promoting a diagnostic approach which is new for climate models that are not associated with operational forecast centers: the use of six-hourly (6-h) global NWP analyses and high-frequency unassimilated observations of parameterized variables (such as provided by ARM) to evaluate short-range weather forecasts made with climate GCMs that are initialized realistically.

The CAPT premise is that, as long as the dynamical state of the forecast remains close to that of the verifying analyses, the systematic forecast errors are predominantly due to deficiencies in the model parameterizations. In that event, it is appropriate to compare parameterized variables with available high-frequency observations collected under the same dynamical conditions, and to modify the parameterizations so as to better match these evaluation data. Finally, if the modified parameterizations are also able to reduce the systematic forecast errors, it is probable that the GCM climate simulation will improve as well.

The basic elements of the CAPT protocol are illustrated in Figure 1. First the climate GCM is initialized without recourse to a DAS, but with its atmospheric state specified from actual synoptic conditions, while also being in approximate dynamical balance (see details in Section 3b). Next, the climate model is run in a short-range forecast mode, and these predictions are compared against the actual evolving atmospheric state, as determined both from NWP analyses and unassimilated observations of parameterized variables. Differences between the model predictions and these evaluation data are diagnosed in order to learn more about the performance of the model parameterizations, and to suggest needed changes. The efficacy of modifying the parameterizations then can also be evaluated in a short-range forecasting framework

However, the overriding goal is not that the climate GCM produce the "best" weather forecast, but only a good approximation thereof, so that the parameterizations respond to a realistic large-scale state. Thus, even though the weather forecasts of a coarse-resolution climate GCM may be inferior to those of a fine-resolution NWP model, *relative* decreases in systematic error are still indicative of improved parameterizations in the climate GCM. Moreover, the rich variety of weather phenomena allows the model parameterizations to be evaluated over a wide range of conditions, and at much less computational expense than is required in climate-

simulation mode. In CAPT, therefore, weather forecasting is viewed as a context for learning more about climate GCM parameterizations, and not as an end in itself.

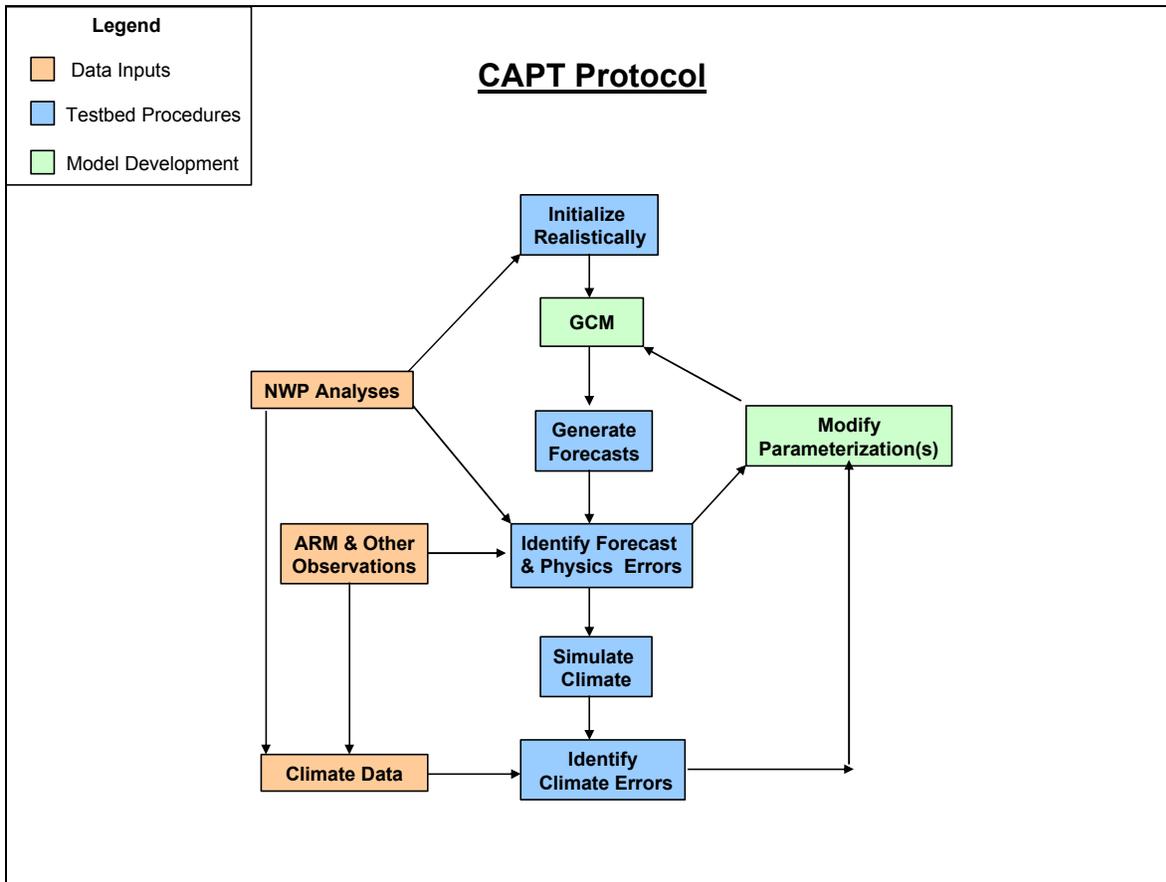


Figure 1: Flow diagram of the CAPT protocol.

But will the CAPT methodology enhance the performance of the GCM in climate simulations? In principle, yes: modified parameterizations that reduce systematic forecast errors should also improve the simulation of climate statistics, which are just aggregates of the detailed evolution of the model. In fact, connections between forecast errors and climate errors are often observed in practice. An example of this in CAM2 is an anomalous split in the Inter-Tropical Convergence Zone (ITCZ) which manifests itself very early in the model forecasts, and then grows more pronounced with time (Figure 2).

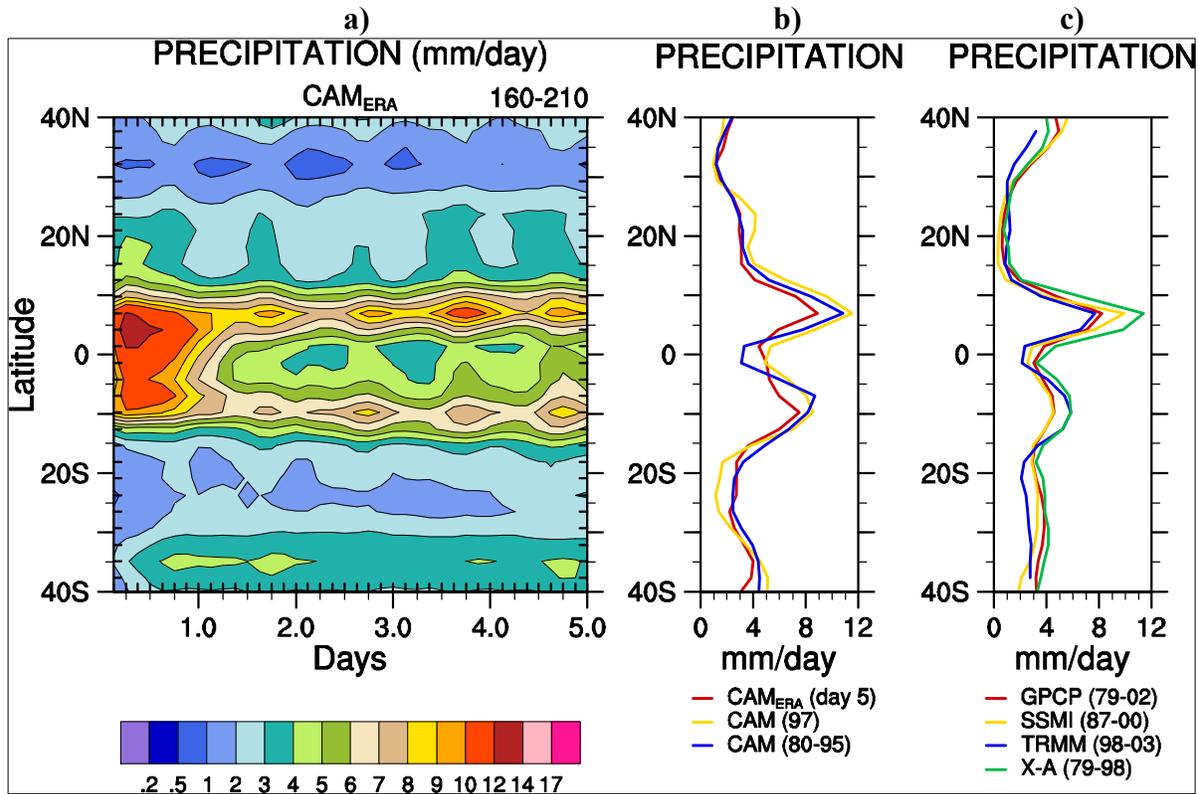


Figure 2: Depiction of Tropical Pacific precipitation anomalies in the CAM2 model. In a) a contour plot of the CAM2 mean five-day forecast of 3-h accumulations of longitudinally averaged (between 180 E and 210 E) Tropical Pacific precipitation is shown for June 1997. (Mean is based on 30 five-forecasts, initiated at 00Z each day.) Here, the precipitation pattern indicates that a split ITCZ develops early in the mean forecast. In b) it is seen that the CAM2 split ITCZ grows more pronounced with time, as evinced by the 180-210 E zonal average of 24-h precipitation in the mean five-day forecast and in the monthly mean for June 1997, as well as in a 1980-1995 June climatology, both from AMIP simulations. Note that the southern branch of the zonal-average split ITCZ in CAM2 precipitation in b) is anomalously intense compared to various estimates of observed zonal-average precipitation climatologies, as shown in c) for the Global Precipitation Climatology Project (GPCP, for the period 1979-2002), the Special Sensor Microwave Imager (SSMI, for 1987-2000), the Tropical Rainfall Measuring Mission (TRMM, for 1998-2003), and the Xie-Arkin (X-A, for 1979-1998) data sets. In all cases, precipitation is given in units of  $\text{mm day}^{-1}$ .

However, some systematic climate errors develop more slowly. An example in CAM2 is a cold bias in the tropical tropopause temperature that sets up gradually, presumably because the controlling processes have long natural time scales. We acknowledge that slow climate errors such as these are not as readily amenable to reduction by a forecast-based approach.

Once parameterization improvements are provisionally indicated by better short-range forecasts, enhancements in model performance also must be demonstrated in progressively longer (extended-range, seasonal, inter-annual, decadal, etc.) simulations. GCM parameterizations that are improved at short time scales also may require some further "tuning"

of free parameters in order to achieve radiative balance in climate mode. Parameterization evaluation in climate simulations is therefore a necessary part of the CAPT protocol (Figure 1).

### 3. Technical Details

Here we elaborate on several aspects of the CAPT protocol, as it has been applied thus far to the CAM2 model.

#### *a. Evaluation data*

The efficacy of the CAPT methodology depends crucially on the accuracy of current NWP analyses. For instance, earlier attempts (e.g. Williamson and Daley 1986) to diagnose parameterizations by using analyses to evaluate the weather forecasts of climate GCMs were thwarted by the strong influence of the NWP model on the analysis, which had especially negative impacts on the accuracy of the analyzed atmospheric moisture and parameterized variables.

NWP analyses now are much better approximations of the actual atmospheric state, as shown by recent findings (Hollingsworth et al. 2002) that representative operational short-range weather forecasts can track atmospheric observations with an accuracy that lies within current measurement uncertainties. Hence, in observation-rich regions (e.g. continental U.S. and Europe), the analyses from modern NWP operational DAS's (and, by extension, multi-decadal reanalyses) can be regarded as reliable references for identifying errors in GCM short-range forecasts. We therefore are using the latest high-frequency (6-h) reanalyses of the European Centre for Medium-Range Weather Forecasts (ECMWF ERA-40, ECMWF 2002) and of the National Centers for Environmental Prediction (NCEP/DOE R2, Kanamitsu et al. 2002) as the main data for global evaluation of the CAM2 short-range weather forecasts.

NWP reanalyses are not sufficient, however, to evaluate all aspects of a GCM forecast, since they cannot furnish precise checks on physical forcings. (Estimates of these forcings from current reanalyses strongly depend on the parameterizations of the analysis GCM.) Thus, ancillary high-frequency observations such as the ARM field data are indispensable for *independent* evaluation of GCM parameterizations. Moreover, field observations of state variables can corroborate the NWP analyses in identifying local forecast errors.

ARM field data at 6-h and higher frequencies (in some cases, at frequencies comparable to a GCM time step of 30 minutes) are available at sites in the U.S. Southern Great Plains (SGP), the North Slope of Alaska (NSA), and the Tropical West Pacific (TWP) (ARM 2002). The most comprehensive high-frequency observations are supplied during sporadic intensive observation periods (IOPs) at the ARM SGP site, such as April 1997 and June/July 1997 (Table 1).

Table 1. Available observations during ARM intensive observation periods (IOPs) at the Southern Great Plains (SGP) site. Note: NOAA denotes the National Oceanic and Atmospheric Administration, GOES the Geostationary Operational Environment Satellite, SIRS the Solar Infra Radiation Station, EBBR the Energy Budget Bowen Ratio, MWR the Microwave Radiometer. U, V, T, and RH represent horizontal winds, temperature, and relative humidity, respectively. LWT and SWT are the top-of-atmosphere longwave and shortwave radiative fluxes. CLDTOT, CLDHGH, CLDMED, and CLDLOW are the total, high, middle, and low level cloud amounts, respectively. LWS and SWS are the surface longwave and shortwave radiative fluxes. LH is the surface latent heat flux, SH the surface sensible heat flux, PW the column precipitable water, CLW the column cloud liquid water, and PREC the surface precipitation. Us, Vs, Ps, Ts and RHs are the surface u and v wind components, pressure, air temperature, and relative humidity, respectively.

<b>Instruments</b>	<b>Sampling frequency</b>	<b>Measured fields</b>
ARM Radiosondes	3 hours	U, V, T, RH
NOAA Wind Profilers	1 hour	U, V
GOES	30 minutes	LWT, SWT, CLDTOT, CLDHGH, CLDMED, CLDLOW
SIRS	1 minute	LWS, SWS
EBBR	30 minutes	Surface LH, SH
MWR	5 minutes	Column PW and CLW
Surface Mesonet Stations	5-30 minutes	Surface PREC, Ts, Us, Vs, Ps, RHs
Cloud Radar, Micropulse Lidar	10 seconds -1 minute	Cloud frequency and properties

In contrast to methods relying on SCMs or CRMs, CAPT can utilize other data sets that are not as comprehensive as those of ARM, such as coordinated satellite, aircraft, and surface measurements that have been collected during one-time field campaigns. For example, data sets of this type have been centralized for investigations conducted by participants in the Global Energy and Water Cycle Experiment (GEWEX) Cloud System Study (GCSS 2002, Randall et al. 2003). Similar field data at some 30 other sites are being collected during the 2003-2004 GEWEX Coordinated Enhanced Observing Period (CEOP) (Bosilovich and Lawford 2002).

To make such observations fully relevant for model evaluation, the data should be aggregated to the scale of a GCM grid box. Hence, observations of forcings in different locations at the ARM SGP site have been spatially averaged, and atmospheric state variables also have been subjected to objective variational analysis to ensure overall conservation of heat, moisture, and momentum (Zhang and Lin 1997, Zhang, et al. 2000). On the other hand, some GCM parameterized variables (e.g. cloud properties) need to be translated into quantities that can be compared more readily with observations. For example, we are currently exploring use of the International Satellite Cloud Climatology Project (ISCCP) simulator (Klein and Jakob 1999, Webb et al. 2001) for this purpose in CAPT.

#### *b. Initialization procedures*

Like many climate models, the CAM2 lacks a compatible DAS (although a community data assimilation testbed is under development--see online information at <http://www.cgd.ucar.edu/DART>), and so it is necessary to devise simple alternatives to standard NWP initialization procedures. Because of their high accuracy, it is feasible to use NWP reanalyses to initialize the model. It is also desirable to use atmospheric reanalyses from diverse NWP analysis models, so as to estimate the sensitivity of the CAM2 parameterizations to this difference. Hence, we are using both the ECMWF ERA-40 and NCEP/DOE R2 reanalyses to initialize the CAM2 model. This entails a three-dimensional mapping of finer-resolution reanalysis data to the coarser (spectral T42/L26) CAM2 resolution.

Operational NWP centers routinely map state variables from high-resolution operational analyses to lower resolutions in order to provide initial conditions for their ensemble forecast systems, and for developmental research forecasts. These mapping procedures also have been applied when exchanging analyses between operational centers, for example to investigate the effect of the initial conditions on specific forecasts (e.g. Harrison et al. 1999). We have successfully adopted the relevant NWP algorithms (White 2001), even though these normally are not applied for resolutions as coarse as those of typical climate GCMs.

As noted previously, our objective is to obtain a good estimate of the atmosphere/land initial conditions, rather than a state which gives the best forecast. Although operational centers might still include a mass-momentum balancing phase ("initialization", in NWP parlance) after the resolution change, we have found our forecasts with CAM2 initialized from both the ECMWF

ERA-40 and NCEP/DOE R2 reanalyses to be relatively noise-free. Should initialization noise arise in specific cases, a temporal digital filter could easily be included (Lynch and Huang 1992, Polavarapu et al. 2000).

For our prototype implementation, initial values for the prognostic parameterized variables (e.g. cloud water in the CAM2 model) are obtained via spin-up procedures that are described below in conjunction with the land initialization. These methods are presently adequate because the time scales for the adjustment of the parameterized variables are relatively fast, and the model errors are currently very large. As a future refinement, we will need to develop mapping procedures for these prognostic parameterized variables as well.

Initialization of the land is particularly problematical because it is difficult to map discrete and discontinuous land variables between different resolutions, especially when there may be different definitions of these variables and different land types in the systems that are involved in this mapping. Thus far, we have applied two procedures to spin up land and atmospheric parameterized variables. Both allow the land model (and parameterizations) to interact with and respond to the forcing from the atmospheric model while the latter is constrained to follow the evolution of the observed atmosphere. We refer to these two methods as "forecast/analysis" and "nudging".

The forecast/analysis method periodically updates (e.g. at 6-h intervals) the atmospheric state variables with the interpolated analyses, and lets the coupled land/atmosphere system evolve until the next update time. This is akin to the current 6-h update cycle of an NWP DAS (such as that used for the NCEP R1 reanalysis--see Kalnay et al. 1996), except that here the atmospheric assimilation phase is replaced by the mapping of a high resolution reanalysis to the climate model grid.

The nudging method involves the addition of terms to the atmospheric equations to relax predicted state variables toward the reanalysis at a specified (e.g. 6-h) time scale. This procedure has been used, for example, to generate a smooth start for NWP model forecasts (Hoke and Anthes 1976).

In order to indicate the minimum period needed for either of these spin-up methods to converge to the best state it is able to provide, we have performed "perfect model" experiments (i.e. using the outputs of the CAM2 as input to the forecast/analysis initialization). In these

simulations, the CAM2 soil moisture, for example, spins up to "correct" values in a few months, except when snow is present. Hence, these simple spin-up methods might be generally applicable in tropical regions, as well as in warm-season observational periods at the ARM SGP site and other mid-latitude locations. Spin-up methods appropriate for snow-covered regions will require further development.

In mapping from analyses to the climate GCM, there is some (probably small) risk of obtaining a false response to a parameterization change. However, in order to produce an initial atmospheric state that is model-consistent, it would be necessary to use a DAS that is based on the GCM, thereby excluding from consideration many climate models that lack this capability. Moreover, to produce the best estimate of the initial state, the DAS would need to be run at higher resolution than is typical for climate simulation, since even the large scales in NWP analyses are improved with enhanced resolution.

### *c. Model forecasts*

The current CAPT practice is to generate five-day (0-120 h) GCM forecasts for each day during the time period of interest (e.g. an ARM IOP), and to archive the forecast data at intervals that match the sampling of the field observations (e.g. at 3-h frequencies for comparison with ARM variational analysis data). For each forecast, the model atmosphere is initialized by applying either the nudging or forecast/analysis methods described previously. We also compute the mean, at some elapsed time, of a sequence of forecasts that are initialized on different days. This mean forecast may be calculated from model predictions that are initiated on consecutive days, or alternatively from forecasts that are stratified according to similar initial conditions, so as to assess the sensitivity of the model parameterizations to particular synoptic or seasonal conditions (Jakob 2003).

We also compute the difference between the mean forecast and corresponding evaluation data, as a way of estimating the GCM systematic forecast error. Both the magnitude and pattern of this systematic error are of diagnostic value, and these can be quantified using standard NWP metrics defined by the World Meteorological Organization (WMO 1999). For example, the mean bias and root-mean-square (RMS) statistics provide information on error amplitudes, while anomaly correlations supply error-pattern information.

## 4. Results

In this section, we show selected results of applying the CAPT protocol to the CAM2 model that illustrate the concepts previously discussed.

In order to verify that our simple initialization procedures are able to produce a large-scale dynamical state that is close to that of the verifying analyses, we have evaluated the skill of the CAM2 forecasts of the 500 hPa height field, following established guidelines (WMO 1999). For example, we computed the mean anomaly correlation (AC) of these forecasts (a commonly accepted measure of forecast skill), where the verification anomalies were calculated from a thirty-year monthly mean climatology of the ECMWF ERA-40 reanalysis for the period 1970-1999. (The AC calculations were found to be insensitive to the choice of climatology.)

Figure 3 shows the AC decay (spatially averaged mean AC as a function of forecast day) of the CAM2 model forecasts, initialized from both the ECMWF ERA-40 and the NCEP/DOE R2 reanalyses, during the April and June/July 1997 ARM IOPs. These are compared with the AC decay of analogous forecasts from the models that generated the ECMWF ERA-40 and NCEP/DOE R2 reanalyses, where values less than 0.6 indicate an absence of forecast skill.

. In general, the CAM2 forecasts of 500 hPa heights are seen to be surprisingly "competitive" with those from the two NWP models. In particular, the AC decay of the CAM2 in the first two forecast days is small, implying that its dynamical state remains close to those of the reanalyses during the early part of the forecast. The skill scores of the CAM2 forecasts show greater separation in the Southern Hemisphere (Figures 3b and 3d), indicating a generally stronger dependence on initial conditions, and a somewhat higher accuracy of the ECMWF ERA-40 reanalysis relative to the NCEP/DOE R2. As expected, the AC decay of all forecasts is more rapid during boreal summer (Figure 3c) when mid-latitude synoptic control is weaker and forecast skill is more strongly influenced by physical processes. The comparatively low decay of the ECMWF ERA-40 model skill in this season thus attests to the quality of its physical parameterizations.

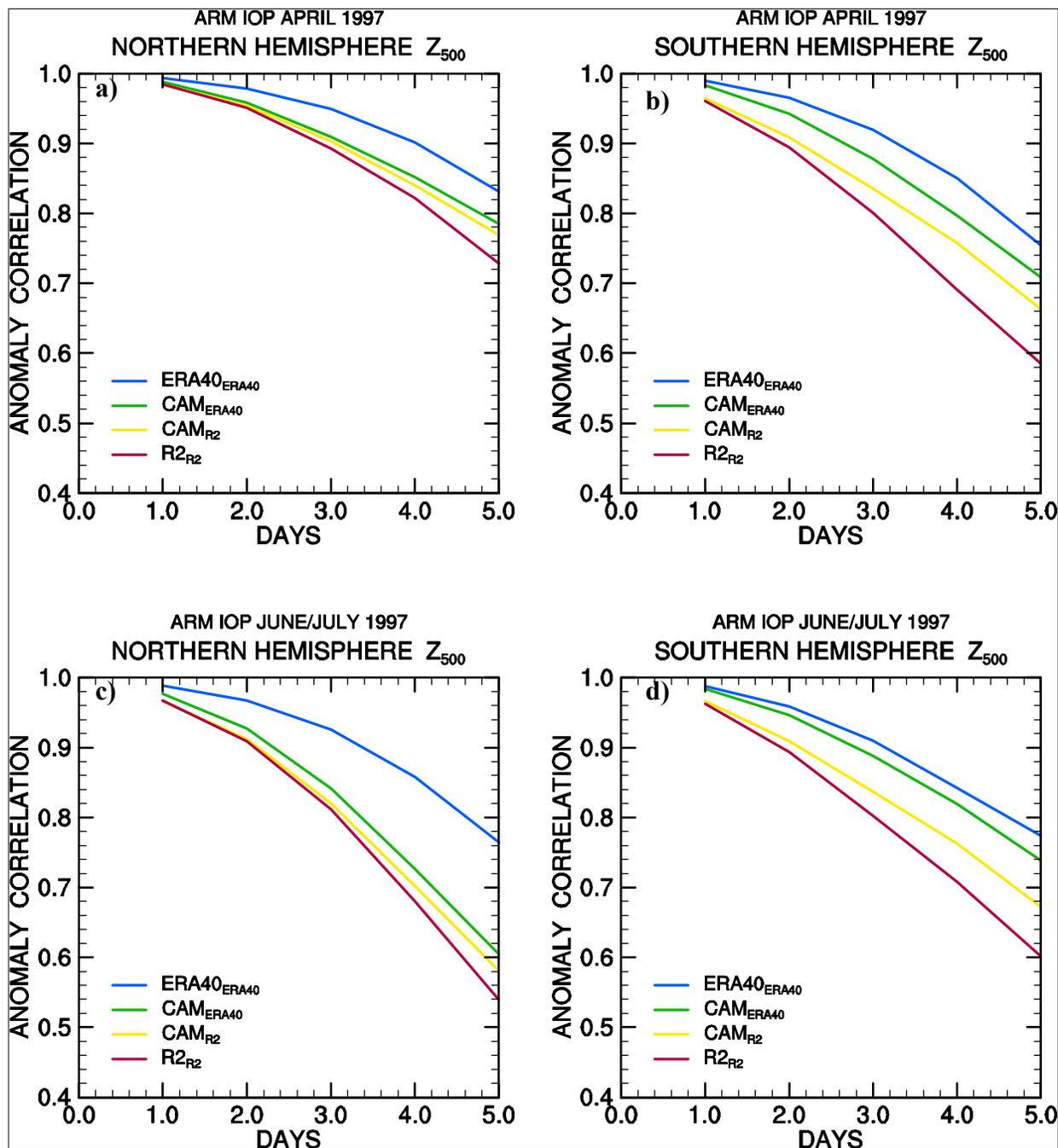


Figure 3: Mean anomaly correlation (AC) for a sequence of forecasts of 500 hPa heights made with three GCMs as a function of forecast day during the April 1997 ARM IOP (panels a and b), and during the June/July 1997 IOP (panels c and d). In all cases, the AC is interpolated to a common 2.5-degree global grid and spatially averaged (with cosine-latitude weighting) over the mid-latitudes of the Northern and Southern Hemispheres (20 N-90N and 20S-90S, respectively). Results are shown for the ECMWF ERA-40 reanalysis model initialized with its own analyses (in blue); the NCEP/DOE R2 reanalysis model initialized with its own analyses (in red); the CAM2 model initialized with ECMWF ERA-40 reanalyses (in green); and the CAM2 model initialized with NCEP/DOE R2 reanalyses (in yellow).

To place these results in perspective, we also compare the CAM2 scores at day 5 with two operational NWP models for the April 1997 IOP period and for subsequent Aprils in the years 2000 and 2003 for the Northern Hemisphere mid-latitudes (Table 2). The fact that the CAM2 five-day AC is higher than representative operational NWP models in 1997 demonstrates that a coarse-resolution climate GCM can make credible forecasts of the large-scale synoptic flow when initialized with an accurate NWP reanalysis (e.g. ECMWF ERA-40). Further, the higher accuracy of the ECMWF ERA-40 reanalysis implied in Figure 3 is in part a result of more recent (circa 2000) NWP technology, as indicated by the skill trend of the operational models from 1997-2003.

Table 2. Anomaly correlation at day 5 of CAM2 forecasts of Northern Hemisphere 500 hPa heights for the April 1997 (1997-04) ARM IOP period compared with that of forecasts from the NCEP and ECMWF operational models. CAM.ERA-40 denotes the CAM2 model initialized with ECMWF ERA-40 reanalysis; NCEP ops and ECMWF ops are the respective operational model scores, which also are shown for comparison in April of 2000 (2000-04) and 2003 (2003-04).

<b>Model</b>	<b>1997-04</b>	<b>2000-04</b>	<b>2003-04</b>
CAM.ERA-40	0.83		
NCEP ops	0.74	0.73	0.84
ECMWF ops	0.78	0.84	0.89

However, the relatively high skill of CAM2 forecasts of the 500 hPa heights does not generally carry over to predictions of large-scale atmospheric moisture, which is tied more directly to the model's physical parameterizations. This model shortcoming is found locally as well, for example in evaluating a sequence of CAM2 daily forecasts of atmospheric relative humidity at the ARM SGP site during the period 19-25 June 1997 (Figure 4). Here it is seen that the temporal variation of the vertical profile of relative humidity obtained from the ARM measurements (Figure 4a) and from the ECMWF ERA-40 reanalysis (Figure 4b) are quite similar. Relative to these evaluation data, the CAM2 forecasts a lower troposphere that is too dry, and an upper troposphere that is too moist in the first few days of this period. (The CAM2 does successfully forecast the moisture plumes associated with convective events on June 23/24, however.)

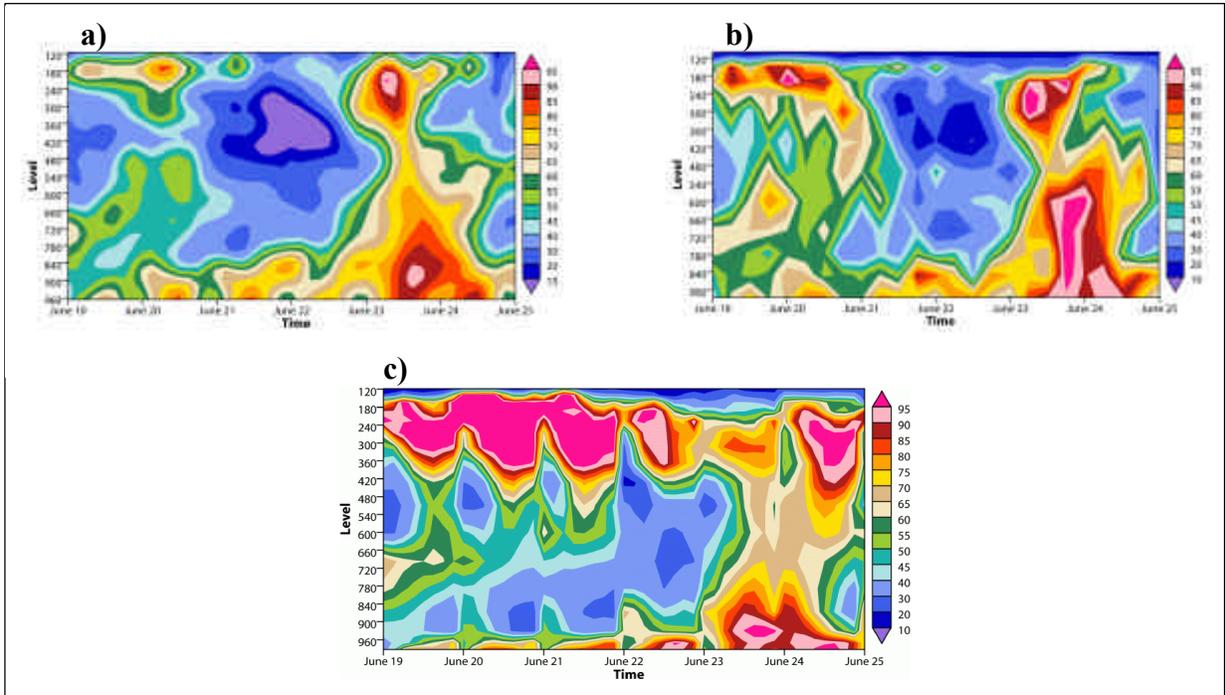


Figure 4: Plots of the vertical profile of atmospheric relative humidity (in %) at the ARM SGP site are shown at 3-h intervals for the period for 19-25 June 1997, as obtained from a) ARM observations, b) the ECMWF ERA-40 reanalysis, and c) a sequence of CAM2 forecasts that are initialized at 00Z each day and valid for the period 03Z-24Z (but with the 00Z value shown for June 19 supplied by the 24Z forecast for June 18). Note, the apparent diurnal cycle in the relative humidity profile in c) is, in actuality, evidence of the rapid departure of the CAM2 forecasts from a realistic humidity profile after their initialization at 00Z each day.

The anomalous CAM2 relative humidity profile on June 19-22 is indicative of a model systematic error, as revealed both by the mean five-day model forecast relative to ARM observations during June/July 1997 (Figure 5a) as well as by the CAM2 June/July climatology relative to the ECMWF ERA-40 and NCEP/DOE R2 reanalyses (Figure 5b). Given the relatively skillful model forecast of large-scale dynamics during this period (Figure 3c), these results imply that there are both high- and low-frequency systematic deficiencies in the CAM2 moist physics parameterizations.

It is just such parameterization deficiencies that are appropriate to study further in the CAPT framework. For instance, the overly dry CAM2 lower troposphere during June/July 1997 at the ARM SGP site is consistent with the model's propensity to rain out moisture nearly every day, rather than in the episodic bursts that are observed (Figure 6b). In contrast, the agreement between CAM2 precipitation forecasts and observations is generally much better during the April 1997 IOP (Figure 6a), when large-scale advective forcing is a more significant contributor to the column moisture balance (analysis not shown).

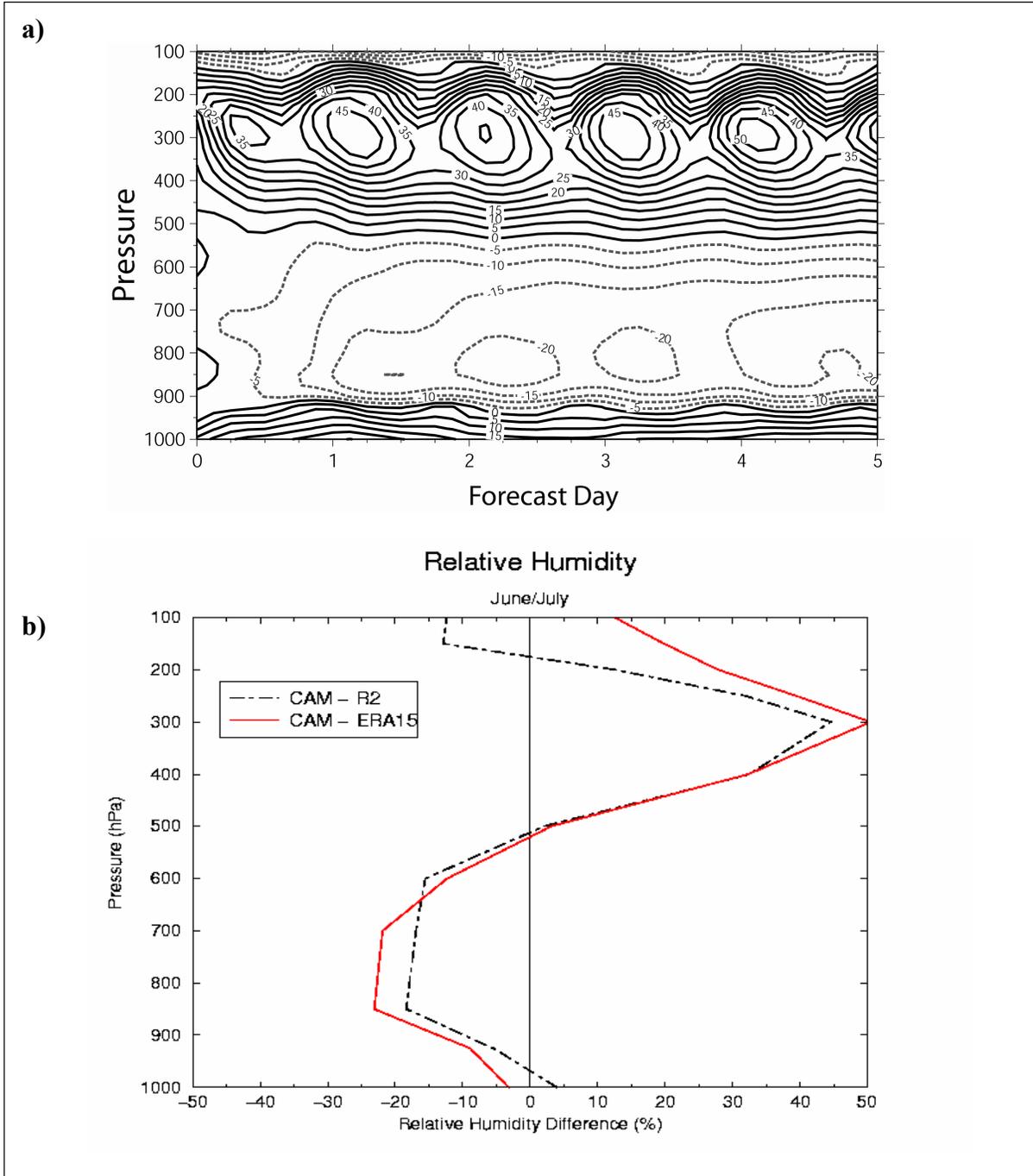


Figure 5: Systematic errors in CAM2 predictions of the vertical profile of atmospheric relative humidity at the ARM SGP site relative to ARM observations (in %) are displayed for different time scales. In a) the evolution of the mean difference between CAM2 five-day forecasts of the relative humidity profile (initialized at 00Z each day of the June/July 1997 IOP) and the corresponding ARM observations is shown. Note the rapid growth of CAM2 relative humidity errors early in the mean five-day forecast and their subsequent diurnal variation, especially in the upper troposphere. In b) differences are shown between a ten-year June/July climatology of the CAM2 relative humidity profile (generated in an AMIP simulation) and corresponding climatologies obtained from the ECMWF ERA-40 reanalysis (solid line) and from the NCEP/DOE R2 reanalysis (dashed line). Note similarities in the vertical distribution of the CAM2 systematic errors (overly dry lower troposphere above the boundary layer, and overly moist upper troposphere) at both forecast and climate time scales.

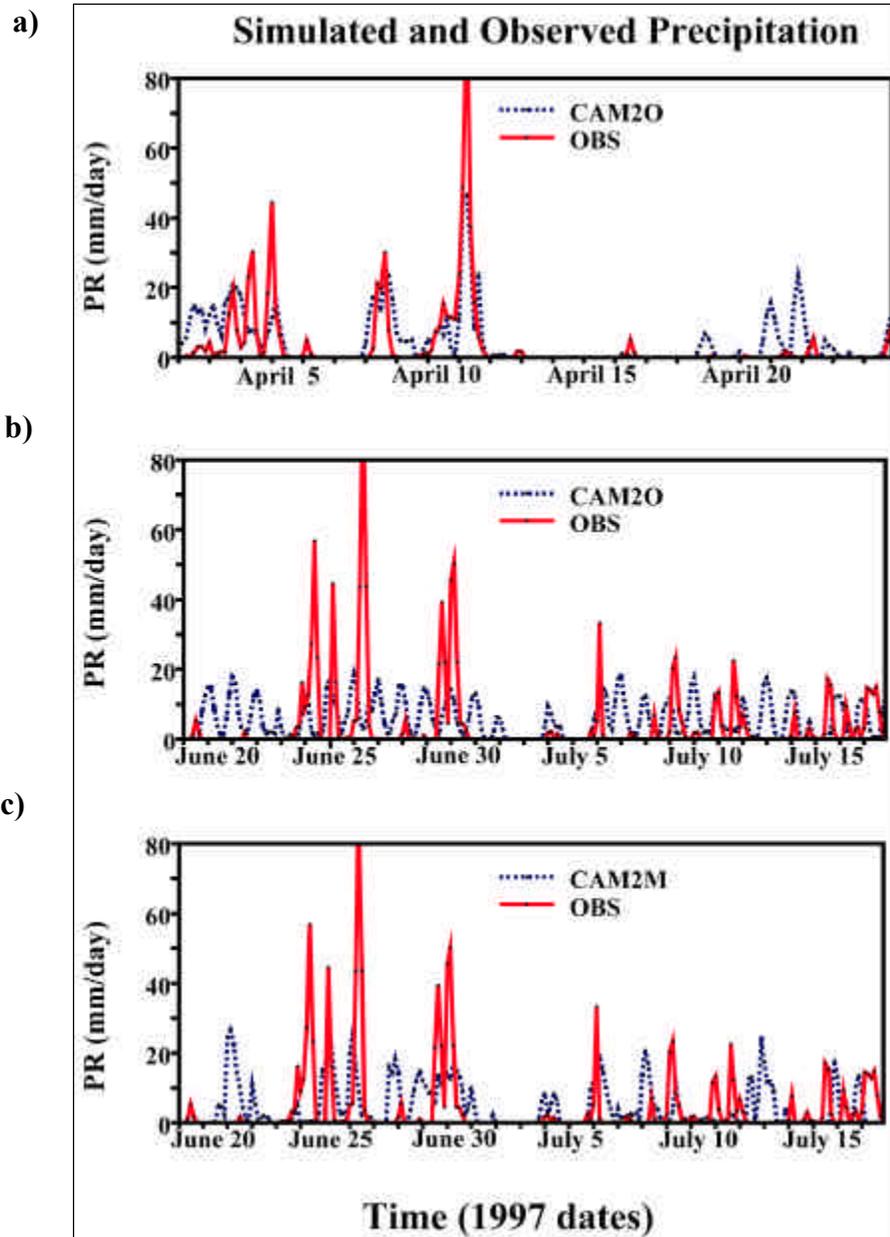


Figure 6: Comparisons of 3-h accumulations of precipitation (in units of mm/day) at the ARM SGP site, as observed (solid line), and as forecast by the CAM2 (dashed line), where the large-scale atmospheric state of the model was reinitialized each day from ECMWF ERA40 reanalysis. In a), the comparison is for the April 1997 IOP, while in b) it is for the June/July 1997 IOP; in both cases, the forecasts were made with the standard version of CAM2 (denoted as CAM2O) that employed the Zhang-McFarlane parameterization of deep convection. In c) are shown forecasts of June/July 1997 precipitation made with a version of the CAM2 (denoted as CAM2M) that included a modified (DCAPE) convective triggering mechanism. See text for further details.

This apparent seasonal sensitivity implies there may be deficiencies in the CAM2 parameterization of convective precipitation (although this is certainly not the only possible source of error). In particular, the characteristics of the CAM2 precipitation displayed in Figure 6b are reminiscent of problems previously identified in the triggering mechanism of the model's Zhang-McFarlane (1995) deep convection scheme when implemented in an SCM (Xie and Zhang 2000). These deficiencies were alleviated by replacing the standard trigger based on positive convective available potential energy (CAPE) with one based on the rate of dynamic CAPE (DCAPE) generation by large-scale advective tendencies of temperature and moisture.

We recently have implemented the DCAPE convective triggering mechanism in the CAM2 and have analyzed the effects on its precipitation forecasts. Because CAPE can accumulate before convection occurs in the modified CAM2, stronger but less frequent precipitation events are produced by the new scheme, yielding generally better agreement with ARM observations (Figure 6c). Further evaluation of the performance of this new convective scheme in an AMIP climate simulation is currently in progress.

## **5. Summary**

CAPT is motivated by the experience of GCM developers that it is very difficult to unravel model parameterization deficiencies solely by diagnosing the simulated climate, which includes systematic errors resulting from nonlinear interactions of many different processes. Our premise is that studying climate GCM parameterizations in a weather-forecasting framework is an effective way to identify their deficiencies and to gain insights on their improvement.

The overarching goal of CAPT is to enhance the performance of GCM parameterizations, as indicated by a reduction of systematic errors that is manifested initially in short-range weather forecasts, but ultimately in climate simulations. We acknowledge, however, that slowly developing systematic climate errors probably will remain resistant to significant reduction by this methodology.

Thus, CAPT is not a panacea for improving climate GCM parameterizations at all time scales, but just one choice from a "toolkit" that may also include, for example, SCMs, CRMs, and simplified GCMs. Nonetheless, we expect that insights obtained from adopting this NWP-inspired methodology will contribute significantly to the general improvement of GCM climate simulation.

## **Acknowledgments**

This work was performed under the auspices of the U.S. Department of Energy (USDOE) Office of Science, Biological and Environmental Research (BER) program by the University of California, Lawrence Livermore National Laboratory under Contract W-7405-Eng-48. This work also was partially supported at the National Center for Atmospheric Research (NCAR) by the Climate Change Prediction Program (CCPP), which is administered by the USDOE Office of Science, BER program. NCAR is sponsored by the National Science Foundation.

## References

- Ackerman, T.P., and G.M. Stokes, 2003: The Atmospheric Radiation Measurement Program. *Physics Today*, **56**, 38-44.
- ARM, 2002: ARM cloud and radiation test bed sites. Accessible online at <http://www.arm.gov/docs/sites.html> .
- Bosilovich, M.G., and R. Lawford, 2002: Coordinated Enhanced Observing Period (CEOP) international workshop. *Bull. Amer. Meteor. Soc.*, **83**, 1495-1500.
- Collins, W.D., J.J. Hack, B.A. Boville, P.J. Rasch, D.L. Williamson, J.T. Kiehl, B. Briegleb, J.R. McCaa, C. Bitz, S-J. Lin, R.D. Rood, M. Zhang, and Y. Dai, 2003: Description of the NCAR Community Atmosphere Model (CAM2). Accessible online at <http://www.cesm.ucar.edu/models/atm-cam/docs/description/index.html> .
- Daley, R., 1991: *Atmospheric Data Analysis*. Cambridge University Press, 457 pp.
- ECMWF, 2002: ECMWF re-analysis ERA. Accessible online at <http://www.ecmwf.int/research/era> .
- Gates, W.L., 1992: AMIP: The Atmospheric Model Intercomparison Project. *Bull. Amer. Meteor. Soc.*, **73**, 1962-1970.
- Gates, W.L., J.S. Boyle, C. Covey, C.G. Dease, C.M. Doutriaux, R.S. Drach, M. Fiorino, P.J. Gleckler, J.J. Hnilo, S.M. Marlais, T.J. Phillips, G.L. Potter, B.D. Santer, K.R. Sperber, K.E. Taylor, and D.N. Williams, 1999: An overview of the results of the Atmospheric Model Intercomparison Project (AMIP). *Bull. Amer. Meteor. Soc.*, **80**, 29-55.
- GCSS, 2002: GEWEX Cloud System Study data integration for model evaluation. Accessible online at <http://gcss-dime.giss.nasa.gov/> .
- Harrison, M., T.N. Palmer, D.S. Richardson, and R. Buizza, 1999: Analysis and model dependencies in medium-range ensembles: Two transplant case studies. *Quart. J. Roy. Meteor. Soc.*, **125**, 2487-2515.
- Hoke, J.E., and R.A. Anthes, 1976: The initialization of numerical models by a dynamic-initialization technique. *Mon. Wea. Rev.*, **103**, 757-763.
- Hollingsworth, A. (ed.), 1998: ECMWF Ten-Year Plan, 1999-2008. European Centre for Medium-Range Weather Forecasts, Shinfield Park, Reading, UK.
- Hollingsworth, A., P. Viterbo, and A.J. Simmons, 2002: The relevance of numerical weather prediction for forecasting natural hazards and for monitoring the global environment. ECMWF Technical Memorandum 361, European Centre for Medium-Range Weather Forecasts, Shinfield Park, Reading, UK. Also accessible online at [http://www.ecmwf.int/publications/library/ecpublications/\\_pdf/tm361.pdf](http://www.ecmwf.int/publications/library/ecpublications/_pdf/tm361.pdf) .
- Houghton, J., Y. Ding, D.J. Griggs, M. Noguer, P.J. Van Der Linden, and D. Aiaosu (eds.), 2001: *Climate Change 2001: The Scientific Basis*, Cambridge University Press, 881 pp.

- Jakob, C., 2003: An improved strategy for the evaluation of cloud parameterizations in GCMs. *Bull. Amer. Meteor. Soc.*, **84**, 1387-1401.
- Jasper, J.D., and P.J. Meighen (eds.), 2000: 'Model Systematic Errors': Extended abstracts of presentations at the Twelfth Annual BMRC Modelling Workshop (co-sponsored by WCRP/WGNE), 16-20 October 2000, Melbourne, Australia. BMRC Research Report No. 80, Bureau of Meteorology Research Centre, 288 pp.
- Jeuken, A.B.M., P.C. Siegmund, and L.C. Heijboer, 1996: On the potential of assimilating meteorological analyses in a global climate model for the purpose of model validation. *J. Geophys. Res.*, **D12**, 101, 16939-16950.
- Kaas, E., A. Guldborg, W. May, and M. Deque, 1999: Using tendency errors to tune the parameterization of unresolved dynamical scale interactions in atmospheric general circulation models. *Tellus*, **51A**, 612-629.
- Kalnay, E., 2003: *Atmospheric Modeling, Data Assimilation, and Predictability*. Cambridge University Press, 341 pp.
- Kalnay, E., M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha, G. White, J. Woollen, Y. Zhu, M. Chelliah, W. Ebisuzaki, W. Higgins, J. Janowiak, K.C. Mo, C. Ropelewski, J. Wang, A. Leetmaa, R. Reynolds, R. Jenne, and D. Joseph, 1996: The NCEP/NCAR 40-year reanalysis project. *Bull. Amer. Meteor. Soc.*, **77**, 437-471.
- Kanamitsu, M. W. Ebisuzaki, J. Woollen, S-K. Yang, J.J. Hnilo, M. Fiorino, and G.L. Potter, 2002: NCEP/DOE AMIP-II Reanalysis (R-2). *Bull. Amer. Meteor. Soc.*, **83**, 1631-1643.
- Kistler, R., W. Collins, E. Kalnay, R. Reynolds, S. Saha, G. White, J. Woollen, Y. Zhu, M. Chelliah, W. Ebisuzaki, J. Janowiak, M. Kanamitsu, K. Mo, C. Ropelewski, R. Jenne, D. Joseph, and M. Fiorino, 2001: The NCEP/NCAR 50-year reanalysis: Monthly-means, CD ROM and documentation. *Bull. Amer. Meteor. Soc.*, **92**, 247-267.
- Klein, S.A., and C. Jakob, 1999: Validation and sensitivities of frontal clouds simulated by the ECMWF model. *Mon. Wea. Rev.*, **127**, 2514-2531.
- Lynch, P., and X.Y. Huang, 1992: Initialization of the HIRLAM model using a digital filter. *Mon. Wea. Rev.*, **120**, 1019-1034.
- Mace, G.G., C. Jakob, and K.P. Moran, 1998: Validation of hydrometeor occurrence predicted by the ECMWF model using millimeter wave radar data. *Geophys. Res. Lett.*, **25**, 1645-1648.
- Meehl, G. A., G. J. Boer, C. Covey, M. Latif, and R. J. Stouffer, 2000: The Coupled Model Intercomparison Project (CMIP). *Bull. Amer. Meteor. Soc.*, **81**, 313-318.
- Miller, M., 1999: Discussion of the 'Transpose AMIP' Project. In Report of the Fourteenth Session of the CAS/JSC Working Group on Numerical Experimentation, R. Newsom (ed.), CAS/JSC WGNE Report No. 14, pp. 7-8, WMO/TD-No. 964.
- Miller, S.D., G.L. Stephens, and A.C.M. Beljaars, 1999: A validation survey of the ECMWF prognostic cloud scheme using LITE. *Geophys. Res. Lett.*, **26**, 1417-1420.

- Morcrette, J.-J., 2002: Assessment of the ECMWF model cloudiness and surface radiation fields at the ARM SGP site. *Mon. Wea. Rev.*, **130**, 257-277.
- Palmer, T.N., 1999: Predicting uncertainty in forecasts of weather and climate. ECMWF Technical Memorandum 294, European Centre for Medium-Range Weather Forecasts, Shinfield Park, Reading UK.
- Polavarapu, S., M. Tanguay, and L. Fillion, 2000: Four-dimensional variational data assimilation with digital filter initialization. *Mon. Wea. Rev.*, **128**, 2491-2510.
- Randall, D.A., K.-M. Xu, R. J. C. Somerville and S. Iacobellis, 1996: Single-column models and cloud ensemble models as links between observations and climate models. *J. Climate*, **9**, 1683-1697.
- Randall, D.A., S. Krueger, C. Bretherton, J. Curry, P. Duynkerke, M. Moncrieff, B. Ryan, D. Starr, M. Miller, W. Rossow, G. Tselioudis, and B. Wielicki, 2003: Confronting models with data: The GEWEX Cloud Systems Study. *Bull. Amer. Meteor. Soc.*, **84**, 455-469.
- Schubert, S. and Y. Chang, 1996: An objective method for inferring sources of model error. *Mon. Wea. Rev.*, **124**, 325-340.
- Stokes, G.M., and S.E. Schwartz, 1994: The Atmospheric Radiation Measurement (ARM) Program: Programmatic background and design of the cloud and radiation test bed. *Bull. Amer. Meteor. Soc.*, **75**, 1201-1221.
- WGNE, 1988: Workshop on systematic errors in models of the atmosphere, Septemeber 1988, Toronto, Canada, CAS/JSC Working Group Report No. 12, WMO/TD-273, Geneva, Switzerland, 382 pp.
- Webb, M., C. Senior, S. Bony, and J.-J. Morcrette, 2001: Combining ERBE and ISCCP data to assess clouds in the Hadley Centre, ECMWF and LMD atmospheric climate models. *Climate Dyn.*, **17**, 905-922.
- White, P.W. (ed.), 2001: FULL-POS Postprocessing and Interpolation, in *IFS Documentation Part VI: Technical and Computational Procedures (CY23R4)*. European Centre for Medium-Range Weather Forecasts, Reading, UK. Also accessible online at [http://www.ecmwf.int/research/ifsdocs\\_old/TECHNICAL/index.html](http://www.ecmwf.int/research/ifsdocs_old/TECHNICAL/index.html) .
- Williamson, D. L., and R. Daley, 1986: The effect of model climate drift on medium range weather prediction. Extended abstracts, WMO/IUGG International Symposium on Short- and Medium-Range Numerical Weather Prediction, 4-18 August, Tokyo, WMO/TD No. 114, 553-555.
- WMO, 1999: *Commission for Basic Systems Abridged Final Report with Resolutions and Recommendations*, Extraordinary Session Karlsruhe, 30 September-9 October 1998, Report WMO-No. 893, Geneva, Switzerland. Also accessible online at <http://www.wmo.ch/web/www/reports/CBS-Ext98.pdf> .
- Xie, S.C., and M.H. Zhang, 2000: Impact of the convective triggering function on single-column model simulations. *J. Geophys. Res.*, **105**, 14983-14996.

- Xie, S.C., K.-M. Xu, R.T. Cederwall, P. Bechtold, A.D. Del Genio, S.A. Klein, D.G. Cripe, S.J. Ghan, D. Gregory, S.F. Jacobellis, S.K. Krueger, U. Lohmann, J.C. Petch, D.A. Randall, L.D. Rotstajn, R.C. J. Somerville, Y.C. Sud, K. Von Salzen, G.K. Walker, A. Wolf, J. J. Yio, G.J. Zhang, and M.H. Zhang, 2002: Intercomparison and evaluation of cumulus parameterizations under summertime midlatitude continental conditions. *Quart. J. Roy. Meteor. Soc.*, **128**, 1095-1135.
- Xu, K.-M., R.T. Cederwall, L.J. Donner, W.W. Grabowski, F. Guichard, D.E. Johnson, M. Khairoutdinov, S.K. Krueger, J.C. Petch, D.A. Randall, C.J. Seman, W.-K. Tao, D. Wang, S. C. Xie, J.J. Yio, and M.-H. Zhang, 2002: An intercomparison of cloud-resolving models with the ARM measurement summer 1997 IOP data. *Quart. J. Roy. Meteor. Soc.*, **128**, 593-624.
- Zhang, G.J., and N.A. McFarlane, 1995: Sensitivity of climate simulations to the parameterization of cumulus convection in the Canadian Climate Centre general circulation model. *Atmos. Ocean*, **33**, 407-446.
- Zhang, M. H., and J. L. Lin, 1997: Constrained variational analysis of sounding data bases on column-integrated budgets of mass, heat, moisture, and momentum: Approach and application to ARM measurements. *J. Atmos. Sci.*, **54**, 1503-1524.
- Zhang, M.H., J.L. Lin, R.T. Cederwall, J.J. Yio, and S.C. Xie, 2000: Objective analysis of ARM IOP data: Method and sensitivity. *Mon. Weather Rev.*, **129**, 295-311.