GLACE-2

(An AGCM intercomparison study co-sponsored by GLASS and WGSIP)

Brief description: "GLACE-2", the 2nd phase of the Global Land Atmosphere Coupling Experiment, is aimed at quantifying, across a broad range of state-of-the-art forecast models, the subseasonal forecast skill associated with the initialization of land surface state variables. The design of GLACE-2 is based on a successful pilot experiment (section 1c) and is built around a comprehensive suite of 60-day forecasts (section 2) that are evaluated against observed precipitation and air temperature fields.

1. Introduction

a. *Overview of GLACE-1*

The first phase of the Global Land-Atmosphere Coupling Experiment (GLACE-1) addressed the following question: to what extent are variations in meteorological variables such as precipitation and air temperature guided by variations in land surface prognostic states? Because meteorological variables themselves have a strong impact on land surface state variations, the reverse direction of causality -- the impact of land conditions on atmospheric variables -- cannot be teased out directly from standard AGCM diagnostics. GLACE-1's contribution was to define an experiment that could isolate these land impacts in an objective way.

The experiment can be described briefly as follows. Three separate 16-member ensembles of 3-month (JJA) simulations were performed by each participating modeling group. The first ensemble was a standard "AMIP-style" ensemble. (AMIP stands for the Atmospheric Model Intercomparison Project; an AMIP-style simulation is one in which sea surface temperatures, or SSTs, are prescribed at all time steps to realistic, or observed, values.) In the second ensemble, the land surface prognostic states of all ensemble members were forced to be the same – the time-varying and geographically varying values for the land states generated by a single randomly selected member of the first (AMIP-style) ensemble were imposed in all members of the second ensemble. The third ensemble was similar to the second, except that only deep soil moisture states were imposed across the ensemble members, reflecting the potential importance of these particular states to seasonal and subseasonal prediction. For each atmospheric variable and for each of the three ensembles, the "level of agreement" across ensemble members of the variable's time series was determined. A comparison of the levels of agreement obtained in the different ensembles allowed the land's impact on atmospheric variability to be directly and objectively quantified.

Full details of the experimental design and an overview of the GLACE-1 results are provided by Koster et al. (2006). A key summary result is presented in Figure 1 below. Plotted in the figure is a diagnostic describing, in rough terms, the fraction of the week-to-week precipitation variability that is controlled by variations in the deeper (below 10 cm) soil moisture reservoirs. Twelve panels are shown, one for each of the twelve models participating in the study. Two important findings are revealed by the figure: (i) within a given model, the locations for which precipitation variations are controlled by variations in soil moisture state show strong geographical variations, and (ii) the coupling strength between soil moisture and precipitation varies widely between models. To a certain degree, the models' behavior in certain regions (e.g., central North America and the Sahel) is similar. Averaging the results of all models produces the coupling diagnostic distribution shown in Figure 2. As expected, the "hotspots" of relatively strong coupling are, for the most part, found in the transition zones between humid and dry areas – zones for which evaporation is both large and responsive to variations in soil moisture.

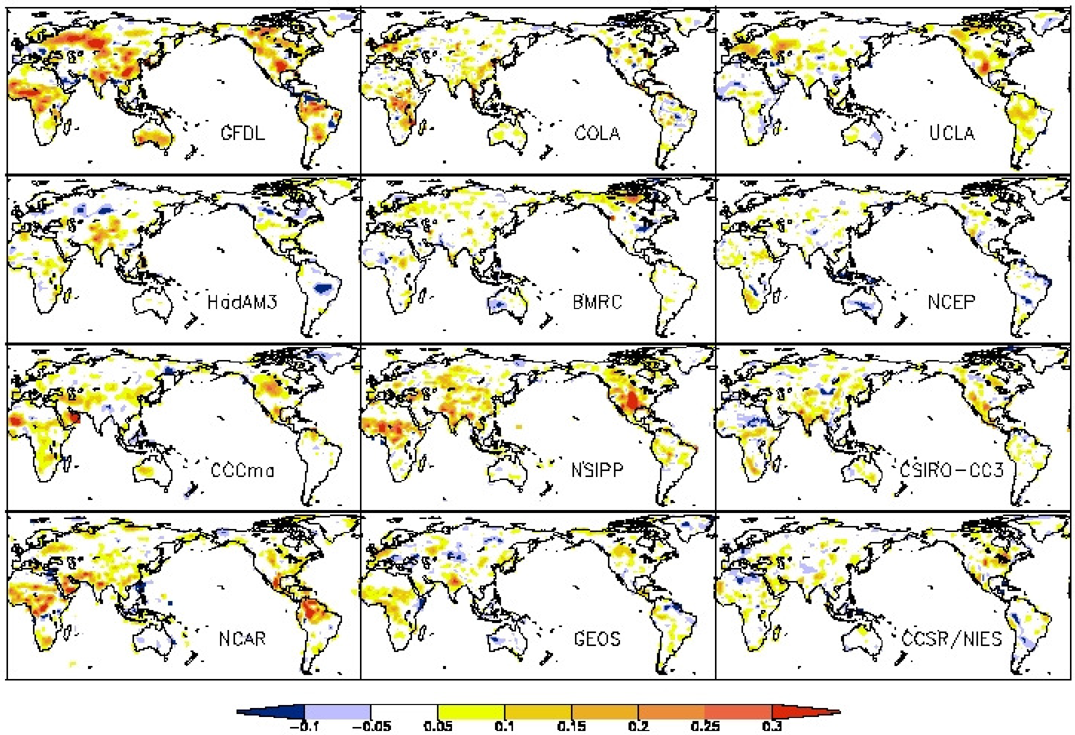

Figure 1. The "Ω(S) – Ω(W)" diagnostic from GLACE-1, showing the degree to which precipitation variations respond to soil moisture variations. (From Koster et al., 2006.)

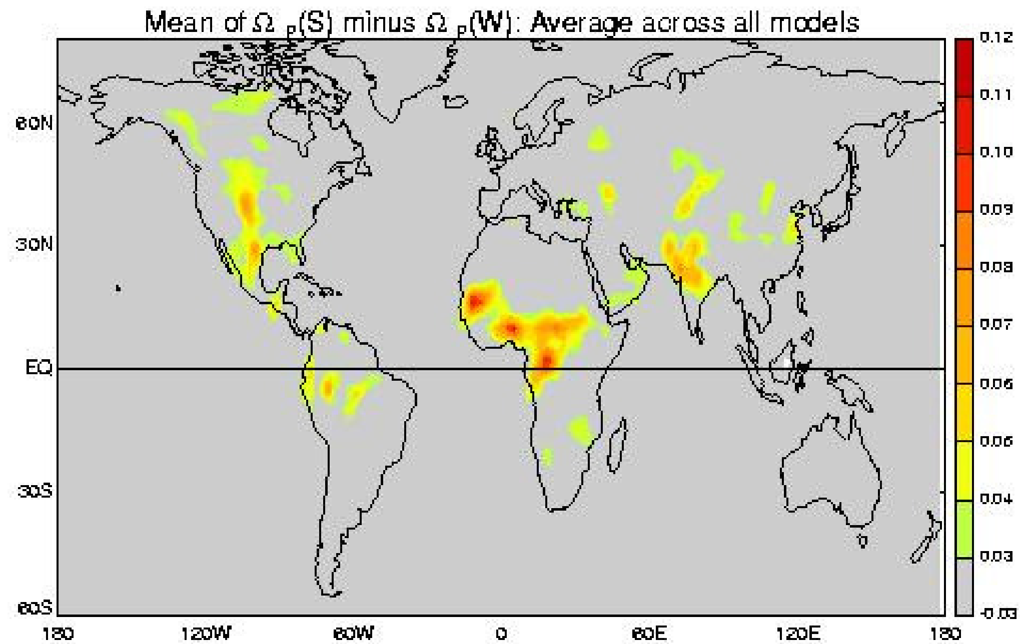

Figure 2. Average of the "Ω(S) – Ω(W)" diagnostic across the models. (From Koster et al., 2006; adapted from Koster et al., 2004b.)

b. *Land variables and meteorological prediction*

Numerical weather forecasts rely on atmospheric initialization – the accurate specification of atmospheric pressures, temperatures, winds, and humidities at the beginning of the forecast. Such initialization may contribute to forecast skill at leads of up to about ten days. Forecasts at longer leads, however, require a different strategy. They must take advantage of slower modes of the climate system, modes with states that are not so quickly dissipated by chaos. To this end, operational centers now supply seasonal atmospheric forecasts based largely on forecasts of ocean behavior. The idea is simple – if sea surface temperatures (SSTs) can be predicted months in advance, and if the atmosphere responds in predictable ways to the predicted SSTs, then aspects of the atmosphere's behavior can be predicted months in advance.

Soil moisture, another slow variable of the climate system, is beginning to garner attention in the forecast community (e.g., Dirmeyer et al., 2003). The timescales of soil moisture memory are typically 1 or 2 months (Vinnikov and Yeserkepova 1991; Entin et al. 2000), significantly less than those of the ocean. Nevertheless, soil moisture has a special importance. Some atmospheric general circulation model (AGCM) studies (Kumar and Hoerling 1995; Trenberth et al. 1998' Shukla 1998; Koster et al. 2000) note a strong tropical-extratropical contrast in the ocean's impact on climate, with the ocean showing a relatively small impact during summer in midlatitudes. For the prediction of summer midlatitude precipitation over continents at subseasonal and longer leads, soil moisture initialization may be more important than ocean initialization (Koster et al., 2000).

Note that for soil moisture initialization to affect a forecast, two things must happen: (i) the initialized soil moisture anomaly must be remembered into the forecast period, and (ii) the atmosphere must respond in a predictable way to the soil moisture anomaly. GLACE-1 is, in essence, a thorough analysis of the second aspect, the response of a modeled atmosphere to soil moisture anomalies. GLACE-1 thereby filled a critical void, since a broad, multi-model analysis of this important element of the climate system had never before been performed. The first aspect, associated with soil moisture memory, is addressed only indirectly in GLACE-1 through a side analysis by Seneviratne et al. (2006), an analysis that does not, in any case, examine the joint impact of memory and atmospheric response in the context of initialization and forecast skill.

Thus, the full initialization question – how would an accurate land surface initialization affect a meteorological forecast? – is only partially addressed by GLACE-1. Arguably, the full initialization question underlies much of today's research into land-atmosphere interaction, given the societal benefits that could be achieved through improved forecasts. To address the full question, an additional study is needed – one in which ensemble forecasts are performed with and without land surface initialization, so that the impact of the land initialization on the forecast can be directly quantified.

Two approaches could be taken for such a study. First, following the lead of GLACE-1, modelers could run idealized ensembles of forecast simulations that use the full suite of land model states from a single time step of an AMIP-style simulation as the forecast's initial condition. Because all ensemble members use the same initialization, the degree to which the initialization affects the forecast could be directly quantified as a function of space, lead, averaging time, and model used. Such a study would indeed prove invaluable for characterizing the role of land initialization in forecast systems.

The second approach is like the first except that it utilizes land surface initial conditions that reflect the conditions present in the real world at the start of the forecast period. This approach allows (for the most part – see section 6b) the idealized analyses allowed by the first approach. In addition, though, it allows a quantification of forecast skill, given that the forecasted meteorological variables (e.g., precipitation, air temperature) could be compared directly to observations. Thus, with the second approach, a little more effort in the specification of the initialization could lead to valuable additional information – an assessment of the skill now attainable with forecast systems through the realistic initialization of land states.

c. *Impact of land-surface initialization on forecast skill: A pilot study*

The second approach was used by Koster et al. (2004a) (hereafter, K04) in a study using the NASA Seasonal-to-Interannual Prediction Project (NSIPP) seasonal forecast system. Because this study serves as a prototype for the experiment proposed herein, it is described in some detail here.

To obtain an adequate sample size for characterizing the impacts of land initialization on forecast skill, K04 performed 75 independent 1-month forecasts, each forecast consisting of an ensemble of 9 simulations. The 75 forecasts covered the boreal warm season months (May through September) of the years 1979-1993. All ensemble members within a given forecast used the same observations-based land initialization, and the average forecasted precipitation and air temperatures were directly compared to observed values. As a “control”, K04 repeated the full exercise without using a specific land initialization for the ensemble members, that is, by drawing the land surface initial conditions used for an individual ensemble member from a distribution of potential values consistent with the forecast start date’s SST values.

The observations-based land initialization was derived through an offline analysis using data from the Global Land Data Assimilation System (GLDAS) project at NASA Goddard Space Flight Center. In essence, global arrays of observational precipitation, radiation, and other atmospheric forcing covering the period 1979-1993 were used to drive a land-surface model (the same model used in the NSIPP seasonal forecast system) to generate global time series of land surface states. The global arrays of land surface states at the beginning of a forecast period were identified and scaled to account for biases between the climatologies of the forecast model and the real world; these arrays were then used as initial conditions for the ensemble forecasts. Atmospheric conditions were not specified from observations (reanalysis); rather, they were selected randomly from a distribution of atmospheric conditions that were consistent with the prevailing SST distribution. (GLACE-2 participants will, in contrast, initialize atmospheric conditions from observations.)

An idealized analysis was first performed by K04 to determine the potential predictability within the system – the forecast skill that would be obtained with the model if the initialization data, the validation data, and the internal model physical parameterizations and discretized dynamics were all perfect. One ensemble member in each forecast was assumed to be “nature” and the average of the remaining ensemble members made up the “forecast”. The square of the correlation coefficient (r^2) between the multiple (75) independent nature/forecast pairs was then determined. The process was repeated several times, each time using a different ensemble member as “nature”. The resulting average r^2, shown in Figure 3 for precipitation, can be considered a measure of the potential predictability, i.e., of the model’s ability to predict itself in the face of unavoidable chaos in atmospheric dynamics. In fact, what is shown is the difference between the r^2 computed for the forecasts utilizing land initialization and those of the control; the plot thus isolates the precipitation predictability associated with land initialization alone and, due to sampling error, features an occasional negative value. The increase in potential predictability associated with land initialization is essentially negligible in most dry or wet regions; it is large only in the transition zones between wet and dry climates, where evaporation is both significant and responsive to variations in soil moisture. The potential predictability for air temperature (not shown) is much larger.

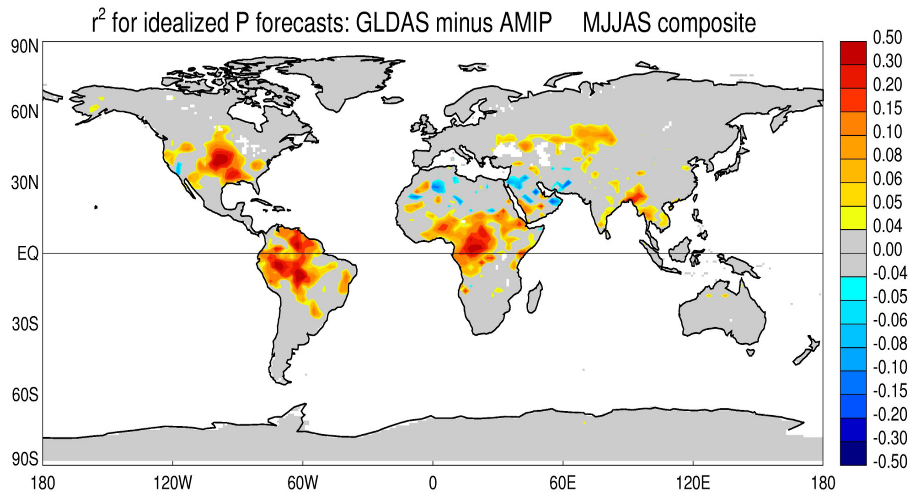

Figure 3. Diagnosed measure of potential predictability associated with land initialization in a specific seasonal forecast system, as determined by K04.

To compute forecast skill, the square of the correlation coefficient (r^2) between the 75 independent nature/forecasts pairs was again computed, this time taking observations – i.e., what really happened, to the extent that it can be measured – as "nature" and the average of the nine ensemble simulations as the forecast. Results are shown in Figure 4 over North America, the only large-scale region having both a high potential predictability in the model and adequate rain gauge coverage. Skill scores are small but significant. (Again, what is actually shown is the difference between the skill obtained with land initialization and the skill inherent in the control.) Land initialization is seen to have an impact precisely where expected from the earlier estimates of potential predictability. The skill scores in Figure 4 fall well below the potential values in Figure 3 presumably due to deficiencies in the forecast model and in the initialization and validation data. In other words, as models and data improve, the values shown in Figure 4 should increase toward those in Figure 3. In essence, Figure 4 illustrates the current real value of land state initialization for seasonal forecasts within a single, state-of-the-art seasonal prediction system.

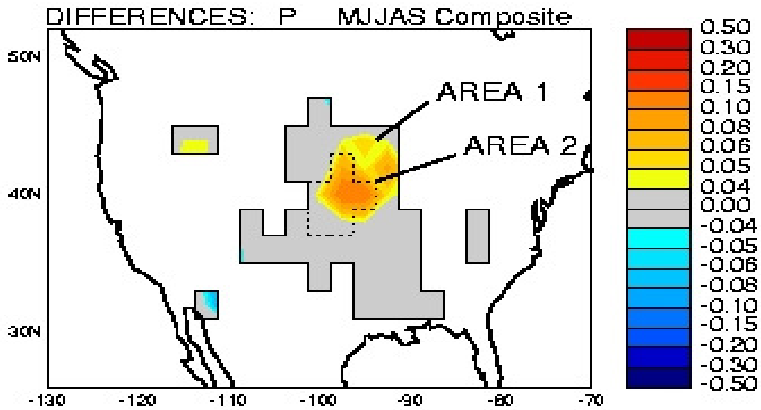

Figure 4. Increase in forecast skill associated with land initialization, from K04. For clarity, results are shown only where both the potential predictability is adequate (from Figure 3) and the rain gauge density over the forecast suite period is adequate, according to independent uncertainty measures (see K04 for details). "Area 1" and "Area 2" refer to potential predictability levels above 0.1 and 0.3, respectively. Differences of 0.05 and 0.08 are significant at the 95% and 98% significance levels, respectively.

d. *Motivation for GLACE-2*

Much of today's research into land-atmosphere interaction has the underlying, eventual goal of determining the extent to which land surface state initialization can improve a forecast. Again, GLACE-1 addresses one element of this goal, namely, quantifying land-atmosphere coupling across a broad range of AGCM systems. The K04 study does address all elements of the goal, but the results are relevant only to a single prediction system. Using the K04 study as a prototype, the community is now poised to embark on GLACE-2, an international model intercomparison experiment that will provide a multi-model, comprehensive view of the impact of land state initialization on forecast skill. Participants in GLACE-2 will learn, probably for the first time ever, the quantitative benefits that can stem from the incorporation of realistic land surface state initialization into their forecast algorithms.

e. *GLACE-2 Sponsors*

GLACE-2, like GLACE-1 before it, is jointly sponsored by the Global Land-Atmosphere System Study (GLASS) panel of the Global Energy and Water Cycle Experiment (GEWEX) and the Working Group on Seasonal-to-Interannual Prediction of the Climate Variability Experiment (CLIVAR). The joint emphasis on land surface initialization and subseasonal prediction clearly spans GEWEX and CLIVAR, making GLACE-2 a strong contribution to the COPES (Coordinated Observation and Prediction of the Earth System) program through the Task Force for Seasonal Prediction (TFSP).

2. Experiment overview

Participants in GLACE-2 will be asked to perform the following two series of forecasts:

<u>Series 1</u>

Description: Forecast simulations using realistic land surface state initialization.
Length of each forecast: 2 months (more precisely, 60 days)
Start dates: April 1, April 15, May 1, May 15, June 1, June 15, July 1, July 15, August 1, and August 15 in each of the years 1986-1995.
Total number of start dates: 100
Number of ensemble members per forecast: 10
Equivalent number of simulation months: 2000 (=166.7 years)

<u>Series 2</u>

Description: Forecast simulations not using realistic land surface state initialization.
Length of forecasts:
Start dates: (Same as for
Total number of start dates: Series 1)
Number of ensemble members per forecast:
Equivalent number of simulation months:

These two series represent the base set of simulations to be analyzed in GLACE-2; they are the only "mandatory" simulations. However, modeling groups with the necessary interest and computational resources will be asked to perform some additional forecast simulations, as described in section 6a. The non-mandatory additional forecasts cover a much broader span of years and will thereby improve the statistical quantification of forecast skill.

Optimally, initial land surface states for Series 1 will be established through participation in the Global Soil Wetness Project – Phase 2 (GSWP-2). GSWP is an ongoing environmental modeling research activity of the Global Land-Atmosphere System Study (GLASS) and the International Satellite Land-Surface Climatology Project (ISLSCP), both of which are contributing projects of the Global Energy and Water Cycle Experiment (GEWEX). Through GSWP-2, modelers produce global fields of land surface fluxes, state variables, and related hydrologic quantities by driving their models offline with global arrays of observations-based meteorological forcing. This forcing spans the period 1986-1995 at a resolution of 1 degree. GSWP-2 model states at the forecast start times can be used to initialize the 2-month forecasts in GLACE-2.

For Series 2, the initial land states for a given forecast ensemble are not identical; rather, they are drawn from a distribution of potential states, the distribution determined from long term simulations with the model. Series 2 is identical to Series 1 in every way except for the fact that it does not benefit from realistic land state initialization.

If possible, the raw GSWP-2 states should not be used directly in the forecast model. The states should first be scaled to the forecast model's climatology, to account for possible biases between the model's climate and nature. The scaling procedure is outlined in detail in section 3b.

Outputs, to be provided to the GLACE-2 data center, include 15-day-averaged precipitation and air temperature fields for each ensemble member and daily soil moisture values for a subset of the ensemble members. A detailed output listing is provided in section 4.

3. Technical details

a. *Model resolution*

The choice of the model resolution for both the land and atmosphere components of the prediction system is left up to the participating modeling group. Model outputs (section 4) can be provided to the GLACE-2 Data Center at the model's resolution, as well; to interpret these data, the Data Center will need information regarding the model grid.

b. *Land surface variable initialization*

(i) Step 1: Attain land surface states through offline simulation.

Several options are available for computing the realistic land surface initialization states to be used for the ensemble forecasts. They are listed here in order of preference, with the first being the most desired approach and the third used only if the first two are impractical.

> -- Option 1: Regrid the GSWP-2 atmospheric forcing to the forecast system's atmospheric resolution and "redo" the GSWP exercise at this resolution. The GSWP-2 atmospheric forcing is available at a 1^o x 1^o resolution; this forcing can be aggregated, disaggregated, or interpolated to the grid used by the forecast modeling system. Once regridded, the forcing data can be used to drive the forecast system's land model globally, using the same land resolution and boundary parameters (vegetation type, soil type, etc.) used by the full forecast system. The land states so generated can then be rescaled (if desired) in Step 2 below prior to being used in the forecasts.
>
> -- Option 2: Regrid GSWP-2 land state outputs to the forecast system's resolution. Rather than regridding the atmospheric forcing, the land model can be forced at the 1^o x 1^o resolution, and the resulting land states can be scaled to match the resolution of the atmospheric model. This approach may be tempting to those groups who have already participated in GSWP-2, since for those groups, the first task in this option would already be done – all that would be required is the regridding. Note, however, that boundary conditions (e.g., vegetation type) used in GSWP-2 may differ from that normally used in the operational system, and this may induce some suboptimality. To some extent, the effect of these differences is ameliorated by the scaling in Step 2.

-- Option 3: Regrid GSWP-2 multi-model output (which has already been produced by other centers) to the resolution of the participant's forecast model. GSWP-2 provides multi-model estimates of soil moisture content, surface temperature, and other land state variables for every day of the 10-year GSWP-2 period. In lieu of performing the GSWP-2 exercise themselves, a modeling group can utilize these multi-model data directly after regridding them to their forecast model's grid. This option is suboptimal, however, and if chosen, the scaling outlined in Step 2 provides the only hope that the fields used will have any practical benefit.

(ii) Step 2: Scale results to forecast system's climatology.

"Scaling" of the land surface states generated in Step 1 is necessary for the second and third options outlined above, and it is encouraged for the first.

Option 1: The fields generated may be inconsistent with the climatology of the forecast system used, due to biases in the forecast model. A relatively dry state obtained through the GSWP-2 exercise for a given region may be a relatively wet state in the forecast system because the forecast system may be biased dry in the region. By scaling, a relatively dry state in nature (and presumably, then, in the GSWP-2 exercise) can be converted to a correspondingly dry state for the coupled model system. We note, however, the counter-argument that the soil moistures generated offline, while biased relative to the model's climatology, are still more "realistic" and thus may be preferable. This is a philosophical argument, and different groups will have different views on the usefulness of scaling in this situation. Because scaling for option 1 will probably not have a significant impact on the results, and because we want groups to perform the experiment in a way they feel would be most beneficial to their institution's needs, scaling is only *encouraged* for option 1.

Option 2: Scaling is needed because of the climate model biases discussed for Option 1 above *and* because the boundary conditions used in GSWP-2 may differ from those used in the forecast system.

Option 3: Scaling is needed for the reasons outlined for Options 1 and 2 above and (first and foremost) because the soil moisture generated by one model is not directly transferable to another model (Koster and Milly, 1997).

Scaling a soil moisture state X (e.g., soil moisture content in a given layer) at a given point requires four quantities:

$M_{X,\ GSWP\text{-}2}$: The mean of X at the point and for the time of year, as determined from the GSWP-2 exercise (or from the multi-model mean fields if Option 3 is used).

$M_{X,\ forecast}$: The mean of X at the point and for the time of year, as determined from long-term runs with the forecast system.
$\sigma_{X,\ GSWP\text{-}2}$: The standard deviation of X at the point and for the time of year, as determined from the GSWP-2 exercise (or from the multi-model mean fields if Option 3 is used).
$\sigma_{X,\ forecast}$: The standard deviation of X at the point and for the time of year, as determined from long-term runs with the forecast system.

Given an unscaled value of X ($X_{unscaled}$) from Step 1, the scaled value (X_{scaled}) that can be used to initialize a forecast can be computed with:

$$(X_{scaled} - M_{X,\ forecast}) / \sigma_{X,\ forecast} = (X_{unscaled} - M_{X,GSWP\text{-}2}) / \sigma_{X,\ GSWP\text{-}2}$$

This scaling amounts to the matching of standard normal deviate values – forcing the standard normal deviate of an anomaly in the coupled modeling system to match that in the observations. Suitable constraints must, of course, be applied to the scaled fields, ensuring, for example, that soil moistures don't fall below the wilting point or become supersaturated.

For the Series 2 experiments, the land initial conditions are not the same amongst the ensemble members; instead they are drawn from a distribution of potential initialization datasets. Most groups have archived restart files spanning decades or more for their modeling system; the land states from these restart files on a given start date (for ten different years, spaced as far apart as possible) can be used to initialize the different ensemble members. Even better, if the modeling group has archived restart files from parallel AMIP simulations, the multiple restart files for a given start date and year can be used to provide the same number of sets of independent land initial conditions; in this way, each set of land initial conditions would be consistent with that year's SST distribution.

If a modeling group chooses Option 1 above to initialize the Series 1 experiments *and* chooses not to scale the resulting soil moistures to the forecast system's climatology, then another option presents itself for the Series 2 initialization: the ten sets of soil moisture states on a given day-of-year from the ten years of the GSWP experiment could be used to initialize the ten ensemble members. For example, for the ensemble forecast starting on June 1, 1986, the GSWP-derived soil moisture states for 1/6/86, 1/6/87, 1/6/88, ..., 1/6/95 could be used to initialize the ten ensemble members. The same ten sets would be used to initialize the forecasts starting on June 1, 1987, and so on. This approach would give the Series 1 and Series 2 forecasts the same "land initialization climatology" for the 10-year period.

We require, of course, that the approach used for initialization be reported, so that we can better interpret the results. If a modeling group has difficulty generating initial land conditions for either Series 1 or Series 2, they should contact the GLACE Data Center for assistance or advice.

c. *Meteorological datasets for land initialization*

GSWP2 (http://www.iges.org/gswp2/sensitivity.html) offers a number of 10-year forcing datasets. GLACE-2 will rely on the baseline (B0) forcing dataset, mostly because some groups have already performed the baseline runs and may be depending on using those data for GLACE-2. *However*, we recognize that the B0 dataset has been criticized for having an overcorrection for precipitation gauge undercatch and for having poor winds. Thus, if groups prefer to use another set of GSWP (or GSWP-type) forcing data because they feel that would better suit the needs of their institutions, that's fine, *as long as the precipitation and radiation data are suitably scaled to GPCP and SRB observations*, as in B0. By having all groups rely on the GPCP and SRB observations, we should maintain the necessary consistency between participants. Again, participants need to state the alternative forcing dataset they use, if they don't use the default.

d. *Atmospheric initialization*

If possible, the atmosphere should be initialized realistically, i.e., with fields representing the actual state of the atmosphere on the forecast start date. Appropriate atmospheric conditions will be extracted from existing reanalyses and will be provided to GLACE-2 participants at the resolution of the reanalysis. Participants will then aggregate/disaggregate/interpolate these atmospheric conditions onto their own model grids. They will generate 10 different sets of atmospheric initial conditions for each ensemble, using their choice of ensemble generation technique (e.g., sampling every few hours from a one-day simulation initialized with the reanalysis fields). Most centers already have standard ensemble generation techniques in place; if, however, a group needs help with this aspect, they can contact the GLACE-2 organizers.

Again, GLACE-2 will provide reanalysis fields. Participants, however, are free to extract atmospheric initial conditions from a different reanalysis if they wish.

If, for whatever reason, the atmospheric model cannot be initialized through the application of reanalysis fields, the different ensemble members can be initialized with atmospheric states derived from AMIP-style or coupled ocean-atmosphere simulations of the year and season in question. Participants following this course must make their choice clear to the organizers to ensure a proper interpretation of the forecast results.

e. Use of prescribed SSTs vs. the use of coupled ocean-atmosphere models

GLACE-2 is designed to isolate the impacts of land initialization on subseasonal predictability and forecast skill. Thus, model-to-model variations in predictability associated with ocean processes should be avoided if possible. Participants are thus encouraged to run their forecast simulations with prescribed SSTs. The SSTs to be prescribed during each forecast period will be provided by the GLACE-2 organizers; the time series of SST fields will be constructed by applying a simple persistence measure to the SST anomalies present on the forecast start date. The SST prescription is deemed acceptable because of the short length of the forecast simulations (2 months) relative to the long timescales of ocean variability.

Some groups (e.g., those involved in operational seasonal forecasting), however, may find it logistically easier to run their forecasts in coupled mode, i.e., with the full interactive ocean running together with the atmospheric model. In the coupled approach, the SSTs (and subsurface ocean states) are only initialized at the start of the forecast; SSTs during the forecast period are predicted rather than prescribed. The coupled approach, while not optimal, is acceptable for GLACE-2. In the analyses described in section 5, land impacts on predictability are isolated by subtracting "control" predictability diagnostics (Series 2) from those obtained when the land is initialized accurately (Series 1). The subtraction should allow most of the complications associated with intermodel variations in ocean state predictability to "cancel out" in our analyses, again isolating the land initialization effects.

4. Required output diagnostics

a. Forecast output

GLACE-2 participants will provide, on their chosen model grid, the following fields for each of the 4 15-day periods of each 2-month (60-day) forecast simulation:

a. 15-day average total precipitation.
b. 15-day average near-surface (2-meter) air temperature.
c. 15-day total evaporation.
d. 15-day average net radiation (the sum of the downwelling shortwave and longwave average radiation fluxes at the Earth's surface minus the sum of the upwelling shortwave and longwave average radiation fluxes.
e. 15-day average vertically-integrated soil moisture content.
f. 15-day average near-surface (2-meter) relative humidity

The total number of required global fields is thus 48000, calculated with N_y x N_s x E x D x V x S, where

N_y = Number of years covered by the forecasts = 10
N_s = Number of start dates per year = 10
E = Number of ensemble members per forecast = 10
D = Number of 15-day periods per forecast simulation = 4
V = Number of variables required = 6
S = Number of forecast series required = 2

Each submitted file corresponds to a specific start date in Series 1 or Series 2. The file should contain 10 sets of 6 groups of 4 consecutive fields, one set for each ensemble member, one group for each output variable, and one field for each 15-day period of the forecast. The 240 data fields in a given file should thus be stored in the following order:

Ensemble member 1:
- Precipitation (4 fields)
- 2-meter air temperature (4 fields)
- Evaporation (4 fields)
- Net radiation (4 fields)
- Soil moisture content (4 fields)
- Relative humidity (4 fields)

Ensemble member 2:

Precipitation (4 fields)
2-meter air temperature (4 fields)
Evaporation (4 fields)
Net radiation (4 fields)
Soil moisture content (4 fields)
Relative humidity (4 fields)

.
.
.

Ensemble member 10:

Precipitation (4 fields)
2-meter air temperature (4 fields)
Evaporation (4 fields)
Net radiation (4 fields)
Soil moisture content (4 fields)
Relative humidity (4 fields)

Output filenames should have the form *model_YYmonXX_ serS*, where

model = 5-letter model identifier
YY = last two digits of forecast year's start date
mon = month of forecast start date ("apr", "may", "jun", "jul", or "aug")
XX = day of forecast start (01 or 15)
S = forecast series (1 or 2)

For example, the fields produced by the ensemble of forecasts beginning on May 15, 1990, for the experiment in which the land is not initialized, should be in a file called GSFC0_ 90may15 _ser2", if the model producing them was the GSFC model.

For the proper interpretation of the gridded data, participants should provide all relevant information regarding their model grid (e.g., resolution, where on the globe the first data point lies, land mask, latitudinal grid coordinates if not uniform, etc.)

b. Soil moisture memory output

In order to analyze the decay with lead time of the information provided by soil moisture initialization, an aspect of critical underlying importance to the forecast problem addressed by GLACE-2, an additional set of output files is required – output for soil moisture content on a daily timescale. To keep the size of the required output at a manageable level, only the first 30 days of soil moisture data are requested (for the longer timescales, we can look at the 15-day averages above), and participants are free to provide this information at a subset of the start dates.

Required data: GLACE-2 participants will provide, on their chosen model grid, the global field of vertically-integrated soil moisture for each day of the first 30 days of the forecast experiments (both Series 1 and Series 2) with the following start dates:

April 1, May 1, June 1, July 1, and August 1 of 1986
" " of 1988
" " of 1990

“ “ of 1992
“ “ of 1994
(Total # of fields: 15000)

Optional data: Same, but for all start dates.
(Total # of fields: 60000)

The daily soil moisture file names should have the form *model_YYmonXX_ serS_daily*, where

model = 5-letter model identifier
YY = last two digits of forecast year's start date
mon = month of forecast start date (“apr”, “may”, “jun”, “jul”, or “aug”)
XX = day of forecast start (01 or 15)
S = forecast series (1 or 2)

5. Proposed analyses

a. *Idealized analysis*

GLACE-2 will repeat the idealized analysis of K04, as outlined in section 1c (see Figure 3), for each participating model. The idea is to determine the degree to which each model can “predict itself” – to determine the extent to which simulated atmospheric chaos in the model can foil a forecast – both with and without a specific land initialization. The analysis will reveal the geographical variation of the model's potential predictability at subseasonal timescales as produced by both the ocean boundary condition (from the Series 2 experiment) and the land boundary condition (from a comparison of the Series 1 and Series 2 experiments). The fall-off of potential predictability in the model with increasing lead time (and how this varies over the course of the boreal summer) will be precisely quantified. In addition, the analysis will uncover the degree to which today's models differ in their intrinsic levels of potential predictability.

The daily soil moisture data will allow the analysis of soil moisture predictability at a much finer timescale. The goal is to quantify how the information content in a soil moisture initialization is lost with lead time over the course of a month.

b. *Evaluation of forecast skill*

The analysis that evaluates forecast skill will also parallel that performed by K04, as discussed in section 1c. The GLACE-2 analyses will consider separately the four 15-day periods within a forecast, so that the fall-off of forecast skill with lead time can be quantified. For each forecast period, the actual rainfall (and air temperature) will be regressed against the forecasted value, and the resulting r^2 value will be taken as the measure of forecast skill. Comparing the r^2 values obtained from the Series 1 and Series 2 forecasts will isolate the contribution of land initialization to the skill. Analyses will identify potential seasonal variations in skill and will quantify skill for the April-

September period as a whole. The result will be the first-ever comprehensive survey of forecast skill levels associated with land state initialization in today's state-of-the-art atmospheric GCMs

6. Optional Supplemental Analyses

a. Extension of base forecast years

The results of K04 suggest that the contribution of land initialization to forecast skill in regions of high potential predictability, though significant, may be modest. Teasing out the skill levels in the presence of large natural chaotic variability may turn out to be a challenge in GLACE-2. GLACE-2 participants are thus encouraged (though not required) to perform additional forecast simulations that span a much longer period of time.

Two options for extending the forecast period will be supported:

> (i) The forecasts can be extended to cover the 50-year period spanned by the global meteorological forcing dataset of Sheffield et al. (2006). The approach used to acquire land initial conditions with this dataset should parallel exactly that used to process the GSWP forcing data into land initial conditions. The forecast start dates for the multi-decade period should be the same as those used in the base experiment (April 1, April 15, May 1, May 15, June 1, June 15, July 1, July 15, August 1, and August 15). It is not necessary, however, to repeat the ten years that make up the base experiment. The optional additional forecasts should thus cover several decades not covered by GSWP.
>
> The longer period of time covered by this dataset allows for a much greater number of forecasts (500 rather than 100), thereby improving our ability to quantify land initialization impacts on skill. Two groups (GSFC and COLA) are already committed to performing the more extended forecasts.
>
> (ii) Participants can extend their forecasts and improve their statistics without running every possible start date in the supplemental decades by running only a select subset of start dates – dates for which large soil moisture anomalies (e.g., beyond a standard deviation) are seen in key areas. The more limited set of supplemental forecasts will allow the modeling group to focus their efforts on those periods with the greatest potential for a positive impact of land initialization, assuming that local influences prevail. While this option allows the participant to avoid performing a great many of the additional forecasts, the full multi-decade offline land forcing exercise would still need to be completed, in order to generate the forecasts' initial conditions.

Early in the project, a multi-decadal offline simulation with the recommended dataset will be performed at GSFC, for two purposes: (1) it will provide time series of land moisture states for other groups to use if for some reason they are unable to run the multidecadal offline simulations themselves (provided they follow the scaling protocols outlined in section 3b above), and (2) it will be used to identify the subset of forecast start dates that represent the less computationally-intensive option (ii) above.

b. "Pure" predictability study

The idealized analysis described in section 5a is suboptimal in the sense that the land surface initial conditions used, though scaled to be appropriate to the land surface model's parameterizations and to the climatology of the atmospheric model, may nevertheless lie outside the set of states that the model can produce naturally (e.g., in terms of spatial pattern). Forcing the initial conditions to look realistic necessarily "muddies" a pure idealized analysis. A cleaner, purer version of the idealized analysis would involve a series of "forecasts" that use as initial conditions the land states produced by the model at selected times during a pre-existing free-running AMIP-style simulation, one that is unencumbered by observational inputs.

Such a cleaner analysis is not part of the main GLACE-2 experiment because with the slightly "muddy" experiment outlined in section 2, we get the added benefit of being able to compare the model forecasts with observations. GLACE-2 participants, however, are welcome to perform a supplementary experiment that, except in the assignment of the land surface states, is identical to that described in section 2. In the supplementary experiment, the land initial conditions are taken from archived model restart files generated during a free running simulation with the atmospheric model rather than from an offline forcing/scaling exercise. Outputs generated in the supplementary experiment will be processed by the GLACE-2 Data Center along with the basic submissions. A comparison of the supplementary experiment results with those of the base experiment will (i) illustrate the degree to which the use of the observations-based initial conditions hampers the estimation of the model's true underlying potential predictability, and (ii) provide an optimal estimate of the model's true potential predictability to the participating modeling group.

c. Austral summer forecasts

The forecast experiments outlined in section 2 focus on boreal summer for two reasons: (1) land-atmosphere interaction tends to be larger in summer, when evaporation is higher, and (2) the northern hemisphere has a much greater land mass. GLACE-2 participants who are interested in forecast skill and predictability in the southern hemisphere summer are encouraged to repeat the Series 1 and 2 forecasts, shifting the noted start dates by six months. Submitted data will be analyzed by the GLACE-2 organizers.

7. Timetable / Other Issues

Funding of GLACE-2 (through the CPPA program of NOAA) begins in August, 2007. Shortly thereafter, GLACE-2 will provide participants with the needed land surface model forcing data, the reanalysis-based initial conditions for the atmosphere, and the persistence-based SST boundary conditions. Participants will be given to the end of 2008 to perform the simulations and to submit their data. The supplemental (and optional) simulations discussed in section 6 will be due by the end of 2009.

All data produced in this experiment, from all groups, will be made publicly available for analysis. If anyone has a concern about this, they should contact the GLACE-2 organizers.

Individual modeling groups will undoubtedly have model-specific questions regarding the set-up of the experiment. The GLACE-2 organizers will be available throughout the experiment period to provide guidance on such issues. (Contact randal.d.koster@nasa.gov) We foresee allowing a certain amount of flexibility in the experimental design to accommodate groups that would otherwise be unable to participate. All deviations from the basic experimental design should be cleared, however, with the organizers and will be documented in all write-ups.

GLACE-2 data storage and analysis will be based at the NASA Goddard Space Flight Center. Processing of data will begin immediately upon submission.

8. **References**

Dirmeyer, P. A., and L. Tan, 2001: A multi-decadal global land-surface data set of state variables and fluxes. COLA Technical Report 102 [Available from the Center for Ocean-Land-Atmosphere Studies, 4041 Powder Mill Road, Suite 302, Calverton, MD 20705 USA], 43 pp.

Dirmeyer, P. A., M. J. Fennessy, and L. Marx, 2003: Low skill in dynamical prediction of boreal summer climate: grounds for looking beyond sea surface temperature. *J. Climate*, **16**, 995-1002.

Entin, J. K, A. Robock, K. Y. Vinnikov, S. E. Hollinger, S. Liu, and A. Namkhai, 2000: Temporal and spatial scales of observed soil moisture variations in the extratropics, *J. Geophys. Res*., **105**, 11865-11877.

Koster, R. D., and P. C. D. Milly, 1997: The interplay between transpiration and runoff formulations in land surface schemes used with atmospheric models. *J. Climate*, **10**, 1578-1591.

Koster, R. D., M. J. Suarez, and M. Heiser, 2000a: Variance and predictability of precipitation at seasonal-to-interannual timescales. *J. Hydrometeorology*, **1**, 26-46.

Koster, R. D., M. J. Suarez, P. Liu, U. Jambor, A. Berg, M. Kistler, R. Reichle, M. Rodell, and J. Famiglietti, 2004a: Realistic initialization of land surface states: Impacts on subseasonal forecast skill. **J. Hydrometeorology**, *5*, 1049-1063.

Koster, R. D., P. A. Dirmeyer, Z. Guo, G. Bonan, E. Chan, P. Cox, H. Davies, T. Gordon, S. Kanae, E. Kowalczyk, D. Lawrence, P. Liu, S. Lu, S. Malyshev, B. McAvaney, K. Mitchell, T. Oki, K. Oleson, A. Pitman, Y. Sud, C. Taylor, D. Verseghy, R. Vasic, Y. Xue, and T. Yamada, 2004b: Regions of strong coupling between soil moisture and precipitation. *Science*, **305**, 1138-1140.

Koster, R. D., Z. Guo, P. Dirmeyer, and 23 others, 2005: GLACE, The global land-atmosphere coupling experiment, 1, Overview. **J. Hydrometeorology**, 7, 590-610. (Also available at http://glace.gsfc.nasa.gov/pdfs/GLACE_part1.pdf)

Kumar, A. and M. P. Hoerling, 1995: Prospects and limitations of seasonal atmospheric GCM predictions. *Bull. Amer. Met. Soc*., **76**, 335-345.

Seneviratne, S. I., and 10 others, 2006: Soil moisture memory in AGCM simulations: Analysis of Global Land-Atmosphere Coupling Experiment (GLACE) data. *J. Hydrometeorology*, in press.

Sheffield, J., G. Goteti, and E. F. Wood, 2006: Development of a 50-year high-resolution global dataset of meteorological forcings for land surface modeling. *J. Climate*, **19**, 3088-3111.

Shukla, J., 1998: Predictability in the midst of chaos: A scientific basis for climate forecasting. *Science*, **282**, 728-731.

Trenberth, K. E., G. W. Branstator, D. Karoly, A. Kumar, N. C. Lau, and C. Ropelewski, 1998: Progress during TOGA in understanding and modeling global teleconnections associated with tropical sea surface temperatures. *J. Geophys. Res*., **103**, 14291-14324.

Vinnikov, K. Y. and I. B. Yeserkepova, 1991: Soil moisture, empirical data and model results, *J. Climate*, **4**, 66-79.