1. Introduction

The global coupled atmosphere-ocean-land-cryosphere system exhibits a wide range of physical and dynamical phenomena with associated physical, biological and chemical feedbacks that collectively result in a continuum of temporal and spatial variability. The traditional boundaries between weather and climate are, therefore, somewhat artificial. The large scale climate, for instance, determines the environment for microscale and mesoscale processes that govern weather and local climate, and these small scale processes likely have significant impacts on the evolution of the large scale circulation (Fig. 1, from Meehl et al. 2001).

Figure 1. Schematic illustrating interactions between various time and space scales in the climate system. Space scales are indicated on the left, and possible forecasts are shown on the right. Though “synoptic” is the smallest time scale, these interactions could continue to infinitely short time scales and small space scales. Also note that the annual cycle and seasonal time scales are forced by the Sun, while the other timescales are internal phenomena.
The accurate representation of this continuum of variability in numerical models is, consequently, a challenging and desirable goal. Fundamental barriers to advancing weather and climate prediction on time scales from days to years, as well as long-standing systematic errors in weather and climate models, are partly attributable to our limited understanding and capability to simulate the complex, multiscale interactions intrinsic to atmospheric and oceanic fluid motions.

The purpose of this paper is to identify some of the research questions and challenges that are raised by this seamless prediction paradigm. For instance, shorter-term climate predictions may require the initialization of coupled models with best estimates of the observed state of the climate system in order to resolve decadal time scale processes. But what is the best method of initialization, and what effect does initialization have on the climate predictions? Moreover, what predictions should be attempted, and how would they be verified? To give a particular example, consider the case of hurricane prediction which has traditionally been regarded as a short-term weather prediction. However, hurricanes generate a cold wake as they churn up the ocean and extract considerable amounts of heat through evaporative cooling. Feedback from the cold wake is now thought to be important in getting the intensity and track right. Hence hurricane forecasting is a coupled problem requiring initialization of a model of the ocean and its heat content, as well as the atmospheric model. These and other issues are discussed in this paper.

2. The seamless prediction paradigm

Scale interactions, both spatial and temporal, are the dominant feature of all aspects of atmospheric and oceanic prediction. This dominance of scale interactions has driven our approach to weather and climate forecasting. We have long recognized that global atmospheric and oceanic models are essential to both weather and climate predictions and that the more scales that they can explicitly resolve, the better the predictions. As a result, weather and climate prediction has been a major driver for advanced numerical and physical techniques and for sophisticated computing systems.

State-of-the-art weather forecasting is carried out using Atmospheric General Circulation Models (AGCMs) that have traditionally been forced with sea surface temperature (SST) anomalies observed at initial time, but are then projected and damped toward climatological conditions as the integrations proceed out to typically 10-14 days. In these integrations, dynamical interactions of the atmosphere with other climate system components do not come into play, and are therefore not included. For longer term predictions, the radiative forcings and coupled interactions and feedbacks among the climate system components are critical, but usually the AGCMs are initialized from some arbitrary state.

There are, of course, good and well documented reasons for these different approaches. First and foremost is that the two time scales address two distinct scientific problems. For a weather forecast on the scale of days, deterministic time evolution of individual synoptic systems must be forecast as an initial value problem. The success of a short-term weather forecast, moreover, is not dependent on the correct simulation of the
meridional overturning circulation (MOC) in the ocean or other longer-term climate processes such as those involving changes in vegetation, sea ice or land ice.

For climate time scales of seasonal to interannual and beyond, statistics of the collections of weather systems are of interest, not the individual time evolution of the systems themselves. For seasonal predictions, coupled air-sea interactions become especially important, but the prediction of an El Niño event still does not depend critically on the state of the MOC. However, for longer time scales, the response of the climate system is dictated by interactions of the atmosphere with not only the ocean, but also the sea ice, land, snow cover, land ice, and fresh water reservoirs, and considerations of biogeochemistry become important as well. Moreover, external effects, such as changes in solar activity, volcanic eruptions, and human influences, all influence the evolution of the climate system. Yet, even on time scales of centuries, other components of the climate system can be considered as fixed: for example, the distribution of the continents and mountains. Variations in major ice sheets, such as Greenland and Antarctica become important over millennia, but can be assumed fixed for many purposes.

In essence, the validity of the assumptions made in designing and conducting numerical experiments must ultimately be evaluated in the context of the problem being studied. Yet, the seamless prediction paradigm explicitly recognizes the importance and potential benefit of greater convergence of methods used in weather and climate forecasting, in particular with regards to initialization of the climate system. As discussed in Section 4, phenomena with lifetimes ranging from hourly to decadal can all potentially benefit from accurate initial conditions in the full climate system.

The prediction of the El Niño/Southern Oscillation (ENSO) phenomenon is a good case in point. While ENSO can now be predicted with some skill with an initialized state of the atmosphere and at least an upper ocean model in the tropical Pacific, profound gaps in our prediction abilities remain, in no small part because of the presence of large systematic errors in the coupled models. Of relevance is that: (i) the coupled model mean state does not agree with the observed mean state with sufficient fidelity; and (ii) the space-time evolution of the simulated climate anomalies is not sufficiently realistic.

Historically, these two problems have been addressed from semi-empirical perspectives. The first approach has been to try to improve the individual physical parameterizations in the component models, which would then hopefully lead to an improved coupled simulation and prediction. The second approach has centered on how best to use imperfect models to make predictions – e.g., a multi-model ensemble. The seamless prediction paradigm points to a third consideration: namely, that current climate models poorly represent the statistics of internal atmospheric (e.g., synoptic weather systems and tropical waves) and oceanic (e.g., poorly resolved tropical instability waves) dynamics and, thus, the interactions of these intrinsic motions with climate. Moreover, the specification of accurate initial conditions in the full climate system may be critical to accurately capture the high frequency phenomena of relevance (e.g., the dependence of the Madden-Julian Oscillation (MJO) on the upper ocean state). This third approach, then, postulates that the two errors noted above are at least due in part to the fact that the
models do not accurately capture the weather-climate link. The issue then becomes what are the important missing elements of the statistics of internal dynamics and what is the best strategy for incorporating them in the coupled models.

In some sense the problem of improving the sub-grid scale physical parameterizations can be viewed as a procedure for including some aspects of the weather and climate link in the models. The typical assumption for sub-grid scale parameterization is to assume that the statistics of sub-grid scale processes can be parameterized in terms of the grid scale variables. However, it is noted here that in many cases this assumption may be seriously flawed and may be unable to capture the weather-climate connection. Hence an alternative strategy has been to reduce the grid size of the model and resolve more of the motions explicitly. While some improvements have resulted from this approach, it is inherently limited by the available computing power.

3. Improving climate models

a. Upscaling research

A basic requirement is that the research community needs to gain considerable experience running models in climate mode with mesoscale processes resolved. This is essential to improve our understanding of the multiscale interactions in the coupled system, to identify those of greatest importance, and to document their effects on climate. Ultimately, such basic research will help us determine the best methods of including upscaling processes in climate models, and it will help us differentiate between those processes that can be better or newly parameterized versus those that cannot.

There is a wide range of upscale interactions to be considered. One of the most critical is the manner in which moist convection and its associate mesoscale organization drives larger circulations. Current parameterization schemes do not adequately handle the mesoscale organization of convection, which is considered to be a critical missing link in the scale interaction process. The poor representation of cloud processes is likely a major factor in the inadequate simulation of tropical oscillations by today’s climate models (Fig. 2), from easterly waves to ENSO, and in the well known double ITCZ bias issue.

Such problems with the tropical modes directly affect the capacity to simulate important climatic features such as ENSO and tropical cyclones. They also effect simulations at higher latitudes, as the export of wave energy from the tropics is an important driver of mid-latitude circulations. As noted above, hurricanes likely play a key role in cooling and mixing the ocean, but as such processes are not represented in any way, how can the result be correct?

Another scale interaction problem is the challenge in modeling the Subtropical Eastern Boundary (STEB) regime off the coasts of Southwest Africa, Peru-Ecuador-Chile, and Baja-Southern California. This regime is marked by marine stratus, equatorward alongshore winds, and ocean upwelling not well simulated by most, if not all, coupled models. With relatively coarse vertical and horizontal resolution, climate models do not adequately represent the small scale ocean processes that play an important role in
maintaining the cold SSTs along eastern boundaries. Nor do they adequately resolve boundary layer processes of the marine stratus and the offshore winds that are influenced by the narrow coastal mountains in the west coasts of North and South America, and the associated oceanic response. They also do not represent the effects of coastal fog that influence near-shore ecosystems and hydrology. Recent evidence suggests that better resolution of these features produces effects that propagate and strongly influence the large-scale climate system (Large and Danabasoglu 2005), reducing rainfall biases across the tropical oceans.

Figure 2: Space-time spectrum of the 15N-15S symmetric component of precipitation, divided by the background spectrum; Observations are in the left panel and from a coupled model simulation in the right panel. Note the poor simulation of the Kelvin and MJO modes and the nearly non-existence of the Rossby (ER) and the Inertial Gravity (WIG) modes.

Other examples of regions or “hot spots” with significant upscaled effects include the monsoon regions of India and Tibet and Central and South America where steep topographical gradients and mesoscale processes such as low-level jet and mesoscale convective complexes play an important role in the water and energy budgets locally and remotely. Over the Maritime Continent, Lorenz and Jacob (2005) presented a study of two-way coupling using global and regional models, and they demonstrated large and positive impacts on the tropospheric temperature and large scale circulation in the global climate simulation.

Clearly, addressing these errors is critical to climate prediction from seasonal to multi-decadal timescales. Therefore, there is a strong need to develop pilot projects to demonstrate the methodologies and impacts of multiscale interactions on the regional and global climate. While numerical models and techniques will be central to this effort, a successful attack on the problem of prediction across scales will also involve sophisticated theoretical and physical research to both understand and specify the critical interactions. It will also require adequate computing resources.
b. Value of testing models on all time scales

A paradigm has long been that it is not essential to get all of the details of weather correct as long as their statistically averaged effects on the climate system are adequately captured. A key question is whether the rectification effects of small scale and high frequency weather events are indeed adequately captured. Water resources are a case in point as they rely on good predictions of precipitation. This means not only precipitation amount but also precipitation intensity, frequency, duration and type (snow versus rain). The character of precipitation affects runoff and flooding, and thus soil moisture and stream flow. To test whether models have biases that then mean that rectification effects are not correct to the detriment of subsequent model evolution, it is important to utilize tests wherever possible.

The diurnal cycle, for instance, provides a systematic climate forcing that is strongest in summer over land, and it affects the timing, location and intensity of precipitation events. Models tested against observations typically have onset of precipitation that is too soon and with insufficient intensity compared with observations, demonstrating the need to improve boundary layer and convective processes in models (e.g., Trenberth et al. 2003). The annual cycle is an obvious strong test for measuring the response of a model to a major climate change, albeit one that affects only those parts of the climate system capable of responding on such a short time scale. Interannual variability, such as how well models simulate ENSO, provides another necessary but not sufficient test of models. These tests highlight the shortcomings and help identify the confidence that can be placed in models.

4. Prediction across scales

a. Effect of initial conditions

For weather prediction, detailed analyses of the observed state of the atmosphere are required but uncertainties in the initial state grow rapidly over several days. Other components of the climate system are typically fixed as observed. For climate predictions, the initial state of the atmosphere is less critical; states separated by a day or so can be substituted. However, the initial states of other climate system components, some of which may not be critical to day-to-day weather prediction, become vital. For predictions of a season to a year or so, the sea surface temperatures, sea ice extent and upper ocean heat content, soil moisture, snow cover, and state of surface vegetation over land are all important. Such initial value predictions are already operational for forecasting El Niño, and extensions to the global oceans are under way. On longer timescales, increased information throughout the ocean is essential. Initial conditions for the global ocean could conceivably be provided by ocean data assimilation exercises currently underway, but salinity reconstructions remain a significant problem. If model simulations are started prior to the availability of the ocean initial conditions, the model ocean would have to be nudged toward the observed values. How strongly this should be done, and what it implies about energy conservations are research issues that need to be explored.
The mass, extent, thickness, and state of sea ice and snow cover are vital at high latitudes. The states of soil moisture and surface vegetation are especially important in understanding and predicting warm season precipitation and temperature anomalies along with other aspects of the land surface. But at this time there is no direct way to provide soil moisture or ground conditions. Any information on systematic changes to the atmosphere (especially its composition and influences from volcanic eruptions) as well as external forcings, such as from changes in the sun, is also needed; otherwise these are specified as fixed at observed values. The potential errors induced by incorrect initial conditions should become less apparent as the simulations evolve, but could still be evident through the course of the simulations.

A good rule of thumb for prediction is that it is possible to predict one lifecycle of a phenomenon. Using such a rule of thumb one could hope to predict a single convective element, cyclone wave, the MJO, an ENSO warm event, or a fluctuation of the Atlantic MOC over its lifecycle, but not the second generation event. This rule of thumb is consistent with the climate system being a chaotic dynamical system with limited predictability.

All of the above phenomena have different lifetimes and the predictions of all can potentially benefit from accurate initial conditions in the full climate system. High frequency atmospheric phenomena can usefully incorporate accurate specifications of the surface initial conditions of SST and soil moisture. However, since the information in these degrees of freedom is lost in a short period of time, the evolution of the land and SST is not important for the prediction of the atmospheric high frequency.

Nonetheless, the pathways leading rectified high frequencies into low frequencies may involve progressively more aspects of the climate system; e.g., the rectified convection of the MJO needs the ocean mixed layer to be accurately specified in the initial state, and the MJO influence on ENSO needs an accurate depiction of the initial state of the Southern Oscillation and the thermocline slope across the equatorial Pacific.

Seamless prediction, in principle, lets all of these interactions occur as they do in nature. If the models fall short, in principle one can track how and learn why. This is a good motivation for seamless prediction studies – letting models evolve on a hierarchy of time scales.

\textit{b. Effect of systematic errors}

Another significant obstacle is the large systematic errors present in today’s couple models, so that the simulated climate is, in some cases, far from the observed state at any given point in time. All models have systematic errors, some of which are common across all models. There are at least two approaches to reduce errors. In one approach, the parameters in various physical parameterizations are modified – usually within the range of uncertainty based on observations – in an effort to reduce the known biases. This ad hoc tuning can produce a simulation with smaller biases, which is an important goal, but offers only limited benefit in the long term, since such tunings are dependent on the
details of the model, and will likely not be portable to differing model resolutions and/or physics. A second approach is to improve the models so that they more accurately simulate the phenomena in question. This can occur through improved knowledge of the relevant physics, improvements in the parameterizations of unresolved physics, and numerical experimentation to better understand existing parameterizations.

Efforts to reduce the systematic biases are crucial, since biases in the mean state could affect both a model’s climate sensitivity (response to altered radiative forcing) and its utility as a predictive tool. For predictions, the use of anomaly models or flux-adjusted models can be viewed as a viable complement to non-flux adjusted models. However, with limited resources (human and computational) it is not clear whether – for the purposes of predictions – it is better to focus one’s efforts on the improvement of non-flux adjusted models to reduce bias, or to diversify efforts to complement existing models with flux-adjusted/anomaly models.

To understand the implications of the errors in the mean state on forecast skill we need to note how the coupled forecasts are initialized. In broad terms, the separate atmosphere and ocean initial states are model-based estimates that are intended to be best estimates of the observed climate. While they are typically generated acknowledging the errors in the uncoupled component models, they are constructed without recognition of the errors in the coupled model. In fact, the sub-surface ocean thermal climate associated with ocean initial conditions is significantly different than the climate of the free running coupled model. As a consequence, at forecast initialization, the coupled model rapidly adjusts away from the observed climate estimates towards the coupled model climate. This adjustment is primarily accomplished via Kelvin waves (in the tropics), which ultimately lead to an erroneous SST response 2-4 months into the forecast evolution. This is often referred to as an initialization shock.

How is this mean state problem related to seamless weather and climate modeling? Some of the well known, but unsolved coupled model errors in the tropical Pacific provide the best answers to this question. For instance, all coupled models produce a cold tongue that is too narrowly confined to the equator and extends too far to the west. One hypothesis for the excessively narrow cold tongue in the eastern Pacific is that the coupled models do not accurately simulate how ocean currents, eddies and waves distribute the cold water offshore and how the cold water interacts with the atmospheric boundary and clouds. This is not merely a resolution problem nor can it be easily corrected with physical parameterizations. Similarly, the excessive westward extent of the cold tongue is most likely due to the fact that the models fail to capture the atmospheric wind variability associated with the MJO, westerly wind busts and, perhaps, even the diurnal cycle. Again, internal atmospheric dynamics has a rectification on the mean state that cannot be easily parameterized. Other recent research has also implicated biochemistry in this error, because changes in ocean color, which relate to small organisms, affect how solar radiation penetrates into the ocean and where it is deposited.

The implications are difficult to quantify. The basic idea is that the coupled “modes” (think ENSO) of the coupled model are different from the coupled modes of nature.
Often we think of these coupled modes as a slow manifold. One solution to this problem is to improve the simulation of the coupled modes and, in fact, there has been some progress in this regard. For example, preliminary results with the NOAA coupled forecast system (CFS) indicate that a poor simulation of the weather-climate link is responsible for the excessive variability and regularity in the operational version of CFS. Figure 3 shows the Nino3.4 sea surface temperature in the CFS at T62 atmospheric horizontal resolution (the operational version) and at T126 horizontal resolution. The higher horizontal resolution model has more irregularity and the amplitude is in better agreement with observed estimates. The working hypothesis is that the atmospheric internal dynamics are more accurately represented at higher resolution, which feeds back on the fidelity of the sea surface temperature variability on seasonal to interannual time scales. Atmospheric model resolution experiments conducted with the Italian SINTEX coupled model also indicate significant improvements in simulated ENSO periodicity with increasing atmospheric model resolution. However, it should be noted that improvements in model fidelity with increasing resolution are still not well understood, and they are not universal among all models, raising questions regarding what aspects of the weather and climate connection need to be improved in order to improve the simulation and prediction.

Despite these limited improvements in the ENSO variability, fundamental mismatches between the model and nature’s coupled modes remain. There are active research efforts that examine how to initialize the coupled modes of the coupled models given that they do not agree with those of nature. The basic idea is to recognize that the best state
estimate for the individual component models may not be the best initial condition for coupled forecasts. Much of the research focuses on how to identify the slow manifold described by the observed estimates and the coupled model and how a mapping between them can be derived. Much of the difficulty in identifying the slow manifold comes down to separating the signal associated with the coupled modes from the variability associated with internal dynamics.

c. Predictability

Climate predictability arises from at least two sources – initial conditions, and changing external forcing, such as increasing greenhouse gases and aerosols. In response to initial conditions, deterministic atmospheric predictability is limited to approximately two weeks. On longer time scales, a substantial source of predictability appears to be associated with ENSO and coupled ocean-atmosphere interactions, generating seasonal to possibly interannual predictability. In addition, seasonal predictability may arise from stratospheric influences on the state of the troposphere, from Atlantic SST influences on the North Atlantic Oscillation, and potentially from other sources. Decadal scale predictability in the ocean may arise from fluctuations in gyre and overturning circulations, particularly in the Atlantic; this could lead to atmospheric predictability. It is not clear how large such a signal would be, but this could potentially be very important. For example, any predictability of decadal scale Atlantic temperature fluctuations that appear to modulate hurricanes would be extremely important. In addition to the potential sources of predictability that arise from the initial values of the system, predictability may also be derived from changing radiative forcing. In this regard, projections of future radiative forcing changes are crucial, and present a number of problems. These include uncertain estimates of future emissions of radiatively important pollutants, uncertainties in our understanding and ability to predict changes in the carbon cycle given emission pathways, and uncertainties in the response of the climate system to a given forcing. Despite these uncertainties, it is likely that projected changes in radiative forcing are an important source of climate predictability, whose importance may grow with the time scale considered.

d. What predictions should be attempted?

Predictions for ENSO are already well established. On time scales longer than ENSO, at least two types of predictions may be possible. The first is a prediction of the internal variability of the climate system based on initial conditions and the underlying dynamics of the system. One of the most promising physical phenomena for such predictability is the Atlantic MOC. Observational analyses (Folland et al., 1984; Kushnir, 1994; Delworth and Mann 2000) have revealed decadal and longer time scale fluctuations of Atlantic temperature, apparently with strong climatic relevance. Modeling studies have suggested that such decadal scale variability may result from fluctuations of the Atlantic MOC. Furthermore, a number of modeling studies have suggested the MOC is predictable on time scales of a decade or so, given accurate initial conditions, thus offering a potentially important source of decadal scale predictability (Fig. 4). It is also possible that decadal
scale predictability exists in the Pacific Ocean, associated with the Pacific Decadal Oscillation or other phenomena.

Figure 4. One example of decadal scale predictability of the Atlantic MOC as computed in the GFDL CM2.1 global coupled climate model. A 5-member ensemble of predictability experiments is shown, in which each ensemble member used identical initial conditions for the ocean, land, and sea ice. These are taken from 1 January 1101 in a long control integration. The ensemble members differed in their atmospheric initial conditions, which come from January 6, 11, 16, 21, and 26 from the same year in the control integration. The quantity plotted is an index of the MOC, defined as the maximum stream function value in the North Atlantic each year, indicating the northward mass flow in the upper layers of the North Atlantic (1 Sverdrup = 10^6 m^3 s^-1). The relatively low spread among ensemble members in the first 10 years suggests substantial decadal predictability. Additional ensembles were calculated, some of which had similar predictability, and others of which had very little predictability.

In addition, predictability may arise from changing radiative forcings. The altered radiative forcings change the background climate state. Decadal scale predictions need accurate representations of anticipated radiative forcing changes. This necessitates the best possible estimates of future emissions of radiatively important pollutants, as well as modeling capabilities to accurately simulate both how these pollutants affect the global energy, carbon and sulfur cycles, and how the climate system subsequently responds to that altered forcing. In this regard, unpredictable volcanic eruptions can be a significant “wild card” to such predictions, although techniques to handle this aspect (e.g., stochastic occurrences, based on the last 100 years) are under consideration.

e. Techniques

The purpose of ensemble prediction is to quantify the uncertainty in the forecast. This uncertainty may be due to errors in the initial conditions and/or errors in the model (or models) used to make the forecast. Both of these issues are encapsulated in Fig. 5. Here we show rainfall variability simulated by several state-of-the-art atmospheric models.
forced by observed sea surface temperatures. Each model simulation includes an ensemble of nine initial conditions. The differences in the initial conditions are designed to mimic potential observational errors. The first column shows the rainfall variance of the ensemble mean of each model. This is the signal variance. The second column shows the variance about the ensemble mean or the variance due to atmospheric internal dynamics. The last column is the ratio of the ensemble mean variance divided by the internal dynamics variance, i.e., a signal to noise ratio.

The models predict very different signal variance (first column) despite the fact that the same sea surface temperatures force all the models. There is a great deal of uncertainty in the signal due to the difference in the model formulation, and apparently the uncertainty due to model formulation is larger than the uncertainty due to initial conditions. The multi-model approach is specifically designed to quantify this uncertainty and the differences in the signal variance highlight the necessity for the multi-model approach. There are a number of different strategies currently employed to combine the models. The simplest and most common approach is to have the various modeling centers make ensemble predictions and then devise statistical strategies (i.e., Bayesian, linear regression) for combining the models. It is also possible to take a specific model and systematically probe the uncertainty in the model formulation by varying the parameters in the model. Both approaches have strengths and weaknesses, but neither strategy is completely satisfactory in terms of adequately resolving the uncertainty.

**Seasonal Model Inter-comparison Project (SMIP): JJA Rainfall Variance**

![Image showing seasonal model inter-comparison project (SMIP): JJA rainfall variance](image)

Figure 5. Note a large signal to noise in this case may not be the correct result or desirable result but may merely indicate deficiencies in natural variability.
One possible interpretation of the large differences in the signal variance is that we still do not know how to parameterize the atmospheric physics. Another possibility is that the parameterization problem is not well posed and the models are failing to capture the interaction between the internal dynamics and the forced signal. Indeed, the surprisingly large difference in the variability due to internal atmospheric dynamics suggests that the models have entirely different estimates of tropical weather statistics and how this weather “noise” interacts with the signal. The implications of the differences become even more important and dramatic once the atmospheric models are coupled to ocean models – the differences in signal and noise variance impact the evolution of the ocean component. The seamless prediction paradigm is relevant to these aspects.

f. Verification

Metrics

A quick scan through a year of *J. Climate* will reveal a dizzying array of different climate metrics both interesting and important. They differ in variable, time scale, space scale, or functional representation. Given a 1000 different climate metrics pulled out of the last decade of climate research papers, how would one go about determining which ones constrain global temperature climate sensitivity to greenhouse gas forcing such as doubled CO2? And to what accuracy? The ENSO index is one of the best studied, yet no clear link has been found between a climate model's ability to predict past ENSO cycles and its ability to predict climate sensitivity. This situation is basically true of all of our climate metrics.

Note that the same is not true in weather prediction, where some estimates of both prediction limits, and the impact of different weather prediction metrics can be determined. But long-term climate change is a “boundary condition” problem and is fundamentally unlike the initial value weather prediction problem. Even if we could test long term climate models with all possible climate metrics proposed in the last decade of journal papers, we have no current method to prioritize or weight their impact in measuring uncertainty in predicting future climate change for temperature, precipitation, soil moisture and other variables of critical interest to society.

There has been some recent progress in this direction recently using perturbed physics ensembles (PPEs; Stainforth et al., 2005). PPEs are climate models that perturb uncertain physical parameterizations instead of initial conditions. An example of comparing climate metrics from this new approach can be found in Murphy et al. (2004) and is shown in Fig. 6. This figure shows the performance of different climate variable metrics in climate models that showed climate sensitivity changes to doubled CO2 of a factor of 6. The non-dimensional error shown in Fig. 6 is defined as the ratio of climate model rms error versus observations, and the interannual natural variability of the same climate variable metric: in essence a signal to noise measure. A large range of a given non-dimensional climate metric shows that the metric varies greatly when climate sensitivity varies. A small range indicates insensitivity. The whisker plots show minimum, 25/50/75th percentiles, and maximum for 58 climate model runs. The plot does not yet
show, however, how to weight the variables, or how to predict uncertainty from the climate metrics. It does confirm the intuition that climate variables associated with energetics (cloud, radiation, sea-ice) appear more sensitive than classical weather dynamical variables (e.g. 500hPa streamfunction). Further work along these lines is critically needed to discover methodologies to define rigorous climate metrics capable of determining climate prediction uncertainty. The essential question is this: what climate metrics for hindcast climate prediction accuracy can be used to determine the uncertainty bounds on future climate prediction accuracy? If this question can be answered, a second benefit will be the ability to more rigorously define climate observation requirements. While weather prediction can perform Observing System Simulation Experiments to define physically based observing requirements, no similar capability exists for long term climate.

Use of probabilistic information
As for weather, climate metrics will inevitably be determined in terms of probabilistic information. Figure 6 shows one example of such information. The probabilistic metrics must include both measures of intrinsic natural variability of the climate system (i.e. unpredictable background noise), as well as the fact that prediction uncertainties or errors will also be determined in terms of probabilities. Also like weather systems such as low or high pressure centers, some climate metrics may be critical to capture in terms probability distributions of systems or objects. An example is cloud systems, where it may be very important to have separate climate metrics by cloud type. This becomes especially important on long time scales where the normal seasonal average for a regional grid box may obscure nonlinear physical relationships that do not survive the time and space averaging of variables needed to reduce weather and short term climate noise. Compositing climate probabilities by cloud type, or atmosphere/ocean dynamic state can
potentially retain climate physics lost in grid based averaging. This could be especially
critical in climate metrics attempting to separate aerosol effects on cloudiness (the highly
uncertain aerosol indirect effect) from dynamics effects on cloudiness.

5. Concluding remarks
There are currently several seamless prediction activities, although all are still in their
infancy. The efforts typically fall into one of the following categories: (i) using the IPCC
class models for days to decadal prediction; (ii) using NWP class models for seasonal to
decal prediction; (iii) developing very high resolution models with mesoscale processes
explicitly resolved, either globally or by nesting high resolution regional models within
global climate models. There are other emerging approaches as well, such as the concept
of beginning integrations with higher resolution to satisfy weather forecast requirements,
then cascading down to lower resolution versions of the model with consistent physical
parameterization schemes. All of these approaches have various different scientific issues,
as well as advantages and disadvantages, but they are seamless approaches in the sense
that they all attempt to blur the distinction between weather and climate.

Moreover, there are other potential benefits of a focus on the seamless paradigm
(reference THORPEX). Among them are: (i) skill improvement in both weather and
climate forecasts; (2) stronger collaboration and shared knowledge among those in the
weather and climate “communities” working on physical parameterization schemes, data
assimilation schemes and initialization methods; (3) and shared infrastructure and
technical capabilities.

Beyond broadly noting these emerging efforts and potential benefits, we have not
attempted to prescribe specifically how the seamless prediction paradigm should be
implemented by the modeling community over the next few years. There is no single way
forward, and each of the above approaches has merit. Most important is that the research
community gain considerable experience over the next few years running models in
climate mode with mesoscale processes resolved, so that some of the fundamental
scientific questions concerning upscaling be addressed. The use of such models could
point to resolutions that may be more generally possible and deemed necessary in a few
years, could isolate phenomena on which more detailed data are required, or they could
help in understanding how to parameterize processes in lower resolution models.
Moreover, the benefit of subjecting climate models to data assimilation and integrating
those models in numerical weather prediction mode could be substantial.

Over the longer term (10 years and beyond) with increased computational resources,
higher resolution climate models, perhaps with non-hydrostatic atmospheric dynamical
cores and eddy-resolving ocean components with global coupled initialization schemes to
run predictions for several decades, will be developed and used more routinely. Yet, even
these models will likely not be run at resolutions that can explicitly resolve, for example,
tropical convective processes, at least for global climate studies over decades to millennia.

On the other hand, if it is found that it is essential to include phenomena and processes
that are currently not included, alternative strategies must be developed. A case in point
could be the need to represent hurricanes in climate models in order to adequately depict
their effects on the ocean, and the energy and water cycles, something that is not parameterized at present.

As the community pursues a seamless prediction capability, it must be understood that there are a variety of prediction problems that are not necessarily seamless. Lorenz first pointed this out in his conceptual model of two kinds of climate prediction. As knowledge is gained and understanding increased, the gaps will be filled in and the range of problems will appear more seamless. For the foreseeable future, however, there are still important, pressing problems, particularly the decadal to centennial scale climatic effects of altering the Earth’s atmospheric composition that will benefit from a focused effort on the time and special scales of interest. The focus is required because the evaluation of model skill, and therefore the usefulness of model predictions, is highly dependent on the predictive requirements imposed on the simulation.

Additionally, current efforts to develop Earth System Models (ESMs) will allow for more complete assessments of the physics of climate change by including components (e.g., carbon and nitrogen cycles, dynamic vegetation, computed aerosols of all types, and chemistry) that are not essential to the shorter time scales. Therefore, there must be different classes of models that are related, but focused on the specific scientific problems of interest. Moreover, the computational burden of the ESMs will likely not allow for explicit resolution of multiscale interactions and more regional discrimination of climate change impacts. Given the large systematic errors, the additional feedbacks from more interactive components of an earth system-type model clearly increase the uncertainty in the magnitude and nature of the climate changes projected in future scenario simulations. Additionally, the time-evolving ingredients required for such future scenario integrations with earth system-type models are just starting to be addressed. These, along with logistical issues related to coupling strategies, coupled initialization (particularly with regards to salinity), and the scientific issues related to the myriad of unconstrained and poorly understood feedbacks, are significant aspects of these emerging earth system-type models that will continue to stretch both computational and human resources.

Acknowledgements

The authors wish to thank Kevin Trenberth, Joe Tribbia and Greg Holland for their constructive comments and suggestions. Their voluntary efforts helped improve the white paper considerably.
References


