A framework for the simulation of regional decadal variability for agricultural and other applications

Arthur M. Greene, Lisa Goddard and James W. Hansen
International Research Institute for Climate and Society
The Earth Institute, Columbia University
New York, NY

Abstract

Climate prediction on decadal time scales is currently an active area of research, and reliable model-based forecasts of regional "near-term" climate change have yet to be demonstrated. In the absence of such forecasts, synthetic data sequences that capture the statistical properties of observed near-term climate variability have potential value. Incorporation of a climate change component in such sequences can help define risk estimates for a range of climatic stresses, including those lying beyond what has been experienced in the past. Properly conditioned simulations can be used to drive agricultural, hydrological or other application models, enabling resilience testing of adaptation or decision systems. The use of statistically-based methods enables the efficient generation of large ensembles of synthetic sequences and consequently, the creation of well-defined probabilistic risk estimates.

In this report we examine some procedures for the generation of synthetic climate sequences that incorporate both the statistics of observed variability and expectations regarding future regional climate change. Model fitting and simulation are considered in the framework of classical time series analysis, with methodology conditioned by requirements particular to the decadal climate problem. A method of downscaling annualized simulations to the daily time step, while preserving subannual statistical properties, is presented and other possible methods discussed. Deployment in the applications setting, the details of which may vary considerably, depending on regional climate characteristics, available data and the design of follow-on models, is considered and elements of a case study presented.

Contents

List of Figures

List of Tables

1 Introduction

Decadal climate variability, sometimes referred to as "near-term climate change," has received increasing attention in recent years, and the potential for numerical climate models to forecast climate variations on time scales out to a few decades is currently an active area of research. At present, however, well-verified, reliable near-term climate forecasts for terrestrial regions, and particularly at local to regional scales, have not been demonstrated. Alternative methods for assessing near-term climate-related risks may thus have considerable value.

One technique that can useful in this regard involves stochastic simulation, the creation of synthetic climate sequences having statistical properties representative of a region or locality of interest. Such sequences, while not forecasts *per se,* can nonetheless help to quantify ranges of uncertainty associated with near-term climate variability. Simulations may be structured so as to incorporate the long-term climate change trends associated with anthropogenic (greenhouse) forcing. These trends then provide a slowly-changing background state on which decadal, and by extension, higher-frequency fluctuations are superimposed. Acting in concert, these influences can provide a better description of the expected range of near-term climate variations, and their potential impacts on statistics of interest for agriculture or other applications, than either considered alone. It is the generation of such sequences that constitutes the focus of the present report.

The discussion presented here can be considered an exploration of the practical considerations involved in generating such simulations. At the heart of the simulation process, as conceived herein, lies a statistical model. As will be seen, the structure of such a model can only be prespecified up to a point: Its detailed form can be expected to depend in a significant way on several factors, including characteristics of the local or regional climate that constitutes the simulation target, availability of suitable data with which to train the model (possibly including paleodata), the structure and requirements of follow-on applications models and the particular impacts that are of concern. Examples are provided that illustrate these contingencies.

The remainder of this presentation is organized as follows: In Sec. 2 we provide some theoretical background and describe the conceptual decomposition by time scale that underlies the proposed simulation methodology. Section 3 considers the various issues encountered in model design and specification. Section 4 presents a detailed rubric for the construction of a simulation model, in light of the information presented in earlier sections. In Sec. 5 elements of a case study that illustrate one possible realization of the simulation methodology is considered. A discussion and summary follow in Secs. 6 and 7, respectively.

2 Decomposition by time scale

Climate variability is often parsed according to time scale, the various canonical scales corresponding approximately to different classes of climate process. The simulation strategy to be discussed depends on such a decomposition, with each of several scales, or equivalently in this view, process types, receiving quasi-independent treatment, the results thereafter being combined. This treatment strategy has a number of implications, which are discussed in the relevant sections.

2.1 Climate change

On the longest scales to be considered here are the slow, secular climate shifts engendered by humankind's activities, which play out over the course of a century or longer. These are the scales at issue in recent reports of the Intergovernmental Panel on Climate Change [*Solomon et al.*, 2007], and involve anthropogenic "forcing" of the climate system through changes in the radiative properties of the Earth's atmosphere, with attendant adjustments in many aspects of the climate system. One such adjustment is reflected in temperature at the Earth's surface, which is expected to rise as the atmosphere becomes more and more opaque to the planet's thermal radiation, owing to increasing atmospheric burdens of carbon dioxide (CO_2) and other greenhouse gases. This is the source of the often-heard designation "global warming." We thus identify a "climate change" time scale, and associate with it slow, trend-like components in the signals to be analyzed.

2.2 Subannual variations

At the opposite end of the spectrum we find subannual variability, including the seasonal cycle and daily weather fluctuations. The former is of course a consequence the inclination of Earth's axis and its orbital motion around the sun, while day-to-day variations that we think of as "weather" are an expression of very short time scale processes acting to resolve atmospheric instabilities of various kinds. The chaotic nature of the atmosphere implies limited predictability for such fluctuations, and for application purposes such daily variability is often treated statistically through the use of "weather generators," that produce realistic sequences of synthetic data for use in agricultural or other models requiring inputs that are highly resolved in the time domain [*Wilks and Wilby*, 1999; *Wilks*, 1999]. Nonparametric resampling schemes are also used for weather generation; a scheme of this type is utilized in the case study to be discussed.

2.3 The decadal scale: "Near-term" climate change

2.3.1 Modes

The time scale on which we focus in this report, ranging from years to decades, with an emphasis on the longer-period end of this range, occupies a middle ground between the climate change and weather scales. Indeed, stochastic decadal simulation may be thought of as an analog of weather generation, with the year replacing the day as the simulation time step. Unlike climate change, which represents the response of the land-ocean-atmosphere system to external forcing (anthropogenic changes to the radiative properties of the atmosphere) the processes governing variability on decadal scales are believed to be intrinsic, or internal, to the climate system. Decadal variability is often described in terms of "modes," an implicit reference to the "normal modes" characterizing the natural behavior of many physical systems, examples being the tones of a bell or the natural oscillations of a swinging pendulum. Such modes can be "excited" by random disturbances — a pebble tossed at a bell will elicit the bell's natural frequencies — and are intrinsic to the physical system. Similarly, large-scale climate modes such as the El-Niño-Southern Oscillation (ENSO), the Atlantic Multidecadal Oscillation (AMO) or Pacific Decadal Oscillation (PDO) may be thought of as intrinsic to the climate system, excited by random perturbations within the chaotically-varying climate system, but evolving along trajectories that are determined by intrinsic properties of the system itself. In the language suggested earlier, decadal-scale variability is the expression of a process class that is distinct from climate change.

2.3.2 Noise, white and red

It is important to note that both the rapid fluctuations of daily weather and the slower variations of decadal variability may be, in part or whole, expressions of *randomness* within the climate system. Such variations are not driven by any particular mechanism and are essentially unpredictable, beyond the decay time associated with processes having long "memory.". The spectra of such random processes differ between atmosphere and ocean, as illustrated in Fig. 1, which shows realizations of white and red noise processes and Fig. 2, which shows the corresponding power spectra.

In the sequence shown in Fig. 1a, which may be taken as a surrogate for the atmosphere, fluctuations are uncorrelated: The level, or value of the process at any time t is unrelated to its value at any other time. Conditional on its mean and variance, such a process is completely unpredictable from a knowledge of its own history.

By contrast, the process illustrated in Fig. 1b, which we may take as a surrogate for variability associated with the oceans, exhibits a degree of serial autocorrelation.

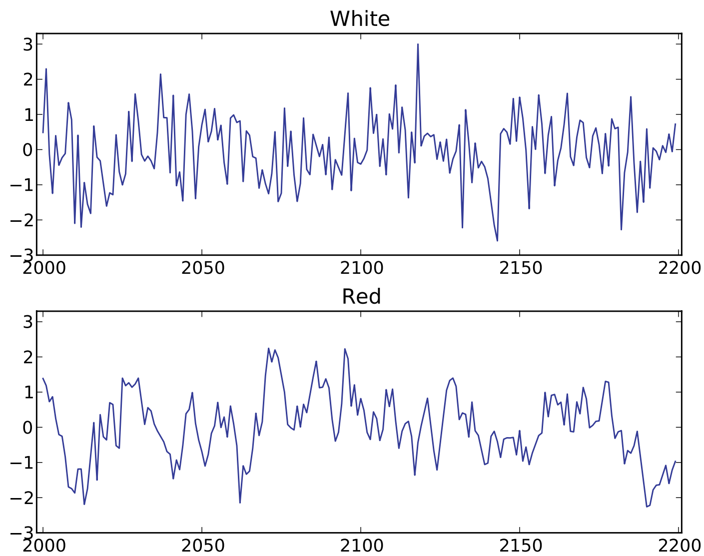

Figure 1: Realizations of a white (a) and red (b) noise processes. These are synthetic sequences; units are arbitrary.

It has been generated by the relation

$$x_t = \gamma x_{t-1} + \alpha \epsilon_t, \tag{1}$$

where x_t is the process level at time t, x_{t-1} its value at the preceding time step, γ and α are constants and ϵ_t is a white noise process with unit variance. In the sequence shown in Fig. 1b γ has been taken as 0.8 and α as 0.4. The lag-1 dependence introduced by the first term on the rhs of (1) gives the process a degree of *memory,* and as a consequence its shifts in level are more sluggish than those of the sequence in Fig. 1a. The memory of a red-noise process is equated with the decay time of its autocovariance function, which can be expressed as $-1/log(\gamma)$. For the process shown in Fig. 1b this would be about 4.5 yr, not very long compared with the extent of the series but enough to produce a distinct difference in the qualitative appearance of the plots. The two series of Fig. 1 have been adjusted so as to have equal variance.

Figure 2 shows power spectra for the sequences of Fig. 1. The white-noise spectrum is essentially flat, with equal power at all frequencies. In the case of red noise, the low-frequency end of the spectrum is enhanced, with spectral power dropping off in accordance with the decay time of the underlying process.

The sequences and spectra of Figs. 1 and 2 are emblematic of a now-classical view of ocean-atmosphere interaction that likens atmosphere and ocean to the molecules in a liquid and and (much larger) particles suspended in that liquid, that are undergoing "sluggish" Brownian motion as a result of the rapidly fluctuating forces of molecule-

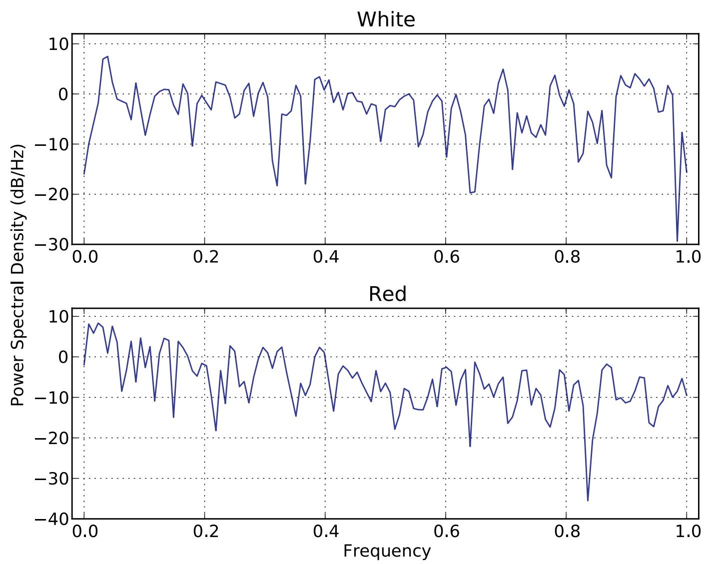

Figure 2: Power spectra for the sequences shown in Fig. 1. (Letters indicate the correspondences.) The x-axis is scaled so that the highest resolvable frequency (the Nyquist frequency) is unity.

particle collisions *Hasselmann* [1976]. The time series and spectra of Figs. 1a and 2a on the one hand, and Figs. 1b and 2b on the other, would then be associated with atmospheric (white noise) and oceanic (red noise) processes, respectively.

2.4 Simulations grounded in physical understanding

To bring this discussion back to the practical problem of decadal simulation, we recognize that the oceans are a likely source, through their "teleconnections" with terrestrial climates, of the variations that we wish to emulate in our stochastic sequences. We should not be surprised, therefore, if we find that this variability has the character of red (or "reddened") noise, or if such noise is present as a background against which more "deterministic" processes play out.

The stochastic sequences to be discussed are generated on an annual time step. To the extent that deterministic (e.g., quasi-periodic) signals are present in the observational records to be mimicked, these would be extracted, modeled and simulated as such, using one of a number of possible methods. The residual "noise" would then be treated as a random process, possibly (but not necessarily) having the form of (1). We regard this procedure as consistent with the Brownian-motion view of ocean-atmosphere interaction first articulated by *Hasselmann* [1976] and echoed in many subsequent studies [e.g., *Saravanan and McWilliams*, 1998; *Kleeman*, 2008].

2.5 A focus on low-frequency variability

Depending on the requirements of follow-on applications models, the inclusion of realistic subannual (i.e., high-frequency) variability in a final simulation product may be desirable, or possibly required. There are various ways of accomplishing this, including resort to the weather generators mentioned earlier, and we touch on this subject in Sec. 4. However, the focus of this work remains the annual and superannual scales; the generation of daily sequences, while afforded an appropriate degree of attention (particularly in the case study), necessarily assumes secondary importance in the overall plan.

2.6 Section summary

In sum, the simulation rubric to be presented assumes three broad classes of climate variability: A forced component, expressed as long-term trend and presumably owing to anthropogenic influence, a low-frequency, or decadal component, reflecting unforced, "natural" variability intrinsic to the climate system and ultimately deriving from the slow variations characteristic of oceanic processes, and subannual high-frequency variations representing short time scale weather fluctuations as well as the seasonal cycle. The latter two classes may have significant random components, although the precise character of this randomness is likely to differ between them, owing to the reddening of atmospheric noise through interaction with the ocean. Although a final simulation product may require the realistic representation of daily variability, it is the low-frequency component that constitutes the principal subject and focus of this report.

3 Modeling contingencies

The scheme we present here is based on the fit of a statistical model to observational data, regarding which several issues immediately arise. These include the availability of suitable data to which to fit the model, characteristics of the regional climate itself and requirements of the follow-on applications model or models. Each of these factors plays a role in simulation design; they are discussed in turn.

3.1 Role of observational Data

3.1.1 Observations, climate models, satellites

If the simulations are to have realistic properties, there must be available sufficient data of reasonably good quality to which to fit the statistical model that is ultimately used to generate them. The existence of GCMs, which provide data of unlimited length, complete spatial coverage and without missing values, would seem to obviate the need for station-based observational records, but in fact the situation is quite the opposite: observational data are *required* for the validation of GCM simulations. Representation by GCMs of the "internal" variability that is of interest here is often of questionable realism, with the models typically exhibiting biases in signal amplitudes, time scales, spatial patterns and very possibly, mechanisms of variability. Such models cannot in general substitute for good-quality observational records, particularly on the small spatial scales relevant for impacts modeling.

Satellite data is widely available but this too offers only a limited alternative to observational records. For one thing, these data share the characteristic of models that "ground truth" is required for their validation. In addition, satellite records commence only around 1979, thus are too short for the confident characterization of most decadal signals. Recall that we are interested in the *statistics* of low-frequency variability, meaning that a number of realizations are required for estimation. The lower the frequency at issue the longer the record required. The Atlantic Multidecadal Oscillation, (AMO) exhibits what appears to be oscillatory behavior (Fig. 3), but the period of these "oscillations," 65-70 yr, is so long compared to the length of the observational record (here 155 yr) that it cannot be said with confidence that the AMO is truly periodic. (Thus the more general designation, Atlantic Multidecadal *Variability,* that is sometimes applied.) Similarly, if one takes the satellite perspective and considers the AMO only beginning in 1979, the series has the appearance of an upward trend, giving little suggestion of oscillatory behavior. The difficulties imposed by limited record length on the confident characterization of decadal variability should be clear.

3.1.2 Paleodata

Given the limited length of many observational records, one can easily imagine a role for paleoclimate data, which may extend hundreds of years or more into the past. Tree-ring reconstructions have been used to good effect by *Prairie et al.* [2008], for stochastic simulations of Colorado river streamflow, for example. The paleorecord in that case exhibits "megadroughts," dry epochs whose lengths greatly exceed those of historical droughts in the American southwest. Clearly, such information has the potential to enhance our understanding of the range of "plausible" behaviors that might be expected in the future. Beyond the obvious requirement that suitable paleorecords, applicable to the region under study, must be available, the introduction of such data raises calibration and other technical issues that lie beyond the scope of the present study, and we do not pursue this option further in the present report. We note in passing that *Prairie et al.* [2008] utilize a preexisting streamflow reconstruction,

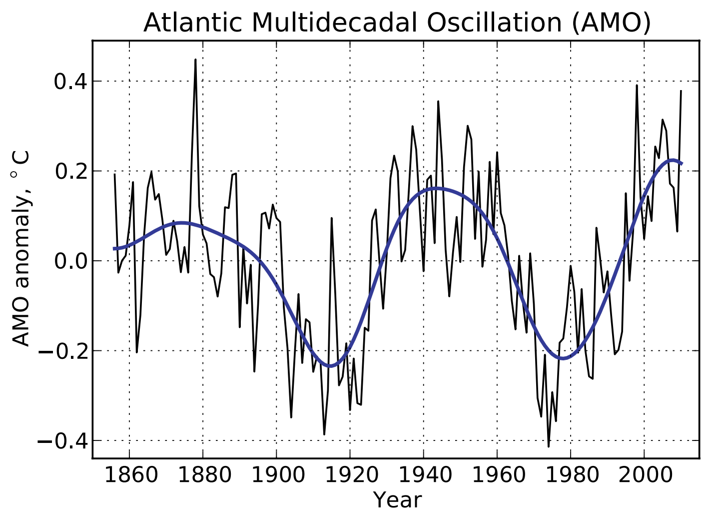

Figure 3: Atlantic Multidecadal Oscillation, and a version that has been lowpassed using a Butterworth filter of order 5 having a half-power point at period of 30 yr. Data from the Kaplan SST reconstruction [*Kaplan et al.*, 1997, 1998] was used.

thus avoiding the necessity of dealing with calibration issues.

3.1.3 Importance of spatial coverage

Complementary to coverage in the time domain is spatial coverage of the available records. Obviously, training data for the statistical model should be representative of the locality or region for which simulations are to be generated. However, these data serve an additional function that is critical to our modeling enterprise: the characterization of regional low-frequency variability. If the "region" under consideration is too small, any low-frequency component that is present may be masked by the relatively larger local variability (in this case to be considered "noise") in the record. Decadal-scale processes exhibit climate footprints, on land or ocean, that tend to be relatively large-scale. Data sources should thus be sufficiently extended in the spatial domain to capture such regional-scale fluctuations. At the same time they should be representative of the study area for which simulations are to be generated.

3.1.4 A possible role for "intermediate" data

The simultaneous requirements that the observational data be of sufficient spatial extent to capture regional decadal signals, while at the same time being representative of a (possibly smaller) study area, suggests a potential role for an "intermediate" data set that, in some circumstances, might be deployed to bridge the gap between regional and local scales. Thus, for example, we might be confronted with a small set of highly localized observational records specific to a particular watershed or subregion that represents a modeling objective. At the same time it is perhaps believed, from independent analysis of climate variability over a larger, enclosing region, that a significant decadal footprint is present, perhaps related to one of the recognized large-scale modes. This decadal signature, however, is detectable at only a relatively low level in the localized observational data. In such a situation one can imagine a hierarchical modeling regime, in which a modicum of the inferred regional decadal signal is introduced into simulations that are primarily based on the more localized observational data. Formalization of this type of model lies beyond the scope of the present report.

3.2 Regional climate characteristics

The goal of our enterprise is the simulation of regional decadal variability, respecting aspects of this variability that are important for the characterization of climate impacts, and in particular for estimating these impacts through the use of follow-on models. We have seen that the characterization of regional variability requires suitable observational data. Having such data in hand, then, what are the salient characteristics that simulations must respect?

3.2.1 Trends and trend-like behavior

Trends, which we associate with anthropogenically-forced climate change, were discussed in Sec. 2 from the perspective of time scale decomposition. Here we consider trend-like behavior from the modeling perspective. Trends represent nonstationarity in the mean: an average value, changing with time, around which decadal and high-frequency, interannual signals fluctuate. Even if the character of these fluctuations changes little over time, the addition of a significant trend will eventually bring about the occurrence of climate anomalies that lie outside the range of past experience. Natural systems, and by extension crops, may have thresholds that do not adjust with changing climate, beyond which critical stresses occur. A shifting mean state may imply that a particular threshold will be crossed more and more often as climate warms, or, alternatively, that certain detrimental thresholds, such as cold spell length or number of frost days, may be exceeded less and less often.

Thus, the inclusion of trends in simulated sequences (at least to the extent that such trends are believed to be significant) may be considered essential from the estimation-of-risk perspective. However, the question of how such trends are to be estimated then arises, since they could well differ, going forward, from trends recorded during the observational period. This issue is discussed further in Sec. 4.

A further issue with trend, and its association with anthropogenically-forced climate change, was suggested in the discussion of Sec. 3.1.1. If we were to consider, say the time period from 1970-2010, the AMO (Fig. 3) would appear very much like an upward-trending signal. Attempting to extract the signature of "climate change" by detrending such a series would yield the potentially misleading impression that a strong anthropogenic signal is present, when in fact we would be radically undersampling a largely trend-free signal that instead is exhibiting slow oscillatory behavior. This points up the importance of drawing inferences from the longest records that are available, or at least attempting to verify through independent means the presence or absence of low-frequency variability in the signal or signals to be deconstructed.

3.2.2 Deterministic vs. random variations

As we have suggested in Sec. 2, decadal variability as characterized simply by time scale is in fact not all of a piece. Broadly speaking, there are two types of processes that may engender low-frequency variations, which we characterize here as "deterministic" and "random." In the classical time series approach, for example, both trend and seasonality are considered deterministic, meaning that they are in some sense predictable. Thus, trend may be extrapolated or may be modeled as dependent on a process that can be projected forward in time, while the seasonal cycle is periodic and may be modeled as a sinusoid. When the effects of such processes have been accounted for, what remains is the random component of variability. The analysis then attempts to paramaterize this component as belonging to one or another family of known processes, and both elements — deterministic plus random — are projected forward in time, in order to generate "realizations" of the data being simulated.

We follow a similar procedure here, with the difference that since our decadal-scale modeling is based on annual values, seasonality ceases to be relevant, at least for the annual-superannual time scale of primary interest. That there may still exist significant trends should be evident. But this is not the end of the decomposition, since other deterministic components may still be present. The question then becomes one of discrimination, between (non-trend) components that are deterministic and those that are random, given that we are starting with annually-resolved data values.

As we have discussed, the very large thermal and physical inertia of the oceans, as compared with the atmosphere, results in variability that is essentially random, but having a degree of *memory,* as described by Equation (1). It is this type of process (i.e., red noise) that we will use as a *test,* to differentiate between deterministic and random behavior, as we examine the data to be simulated. That is, the data is tested, in a statistically rigorous manner, against the null hypothesis that its character is not different from that of red noise, and thus that it plausibly results from random atmosphere-ocean interaction. Rejection of this red noise null hypothesis is taken as evidence of the existence of a deterministic data component, which will then require its own model, just as seasonality requires its own model when analyzing data resolved on subannual time scales. Residuals from the deterministically-modeled series are then represented by a random time series model (which must encompass red-noise processes, since these are, by construction, not classified as deterministic).

What sorts of "deterministic" processes are likely to be distinguishable from red noise? Periodic, or quasi-periodic processes, if they represent a large enough fraction of variance of the original data, may meet this criterion. The use of a null hypothesis based on red noise, sometimes referred to as an autoregressive process of order unity, or AR(1), suggests an additional possibility, namely higher-order random processes, in which x_t may be dependent on two or more past values:

$$x_t = \gamma_1 x_{t-1} + \gamma_2 x_{t-2} + \ldots + \ \gamma_n x_{t-n} + \epsilon_t,\ n > 1. \tag{2}$$

Thus, a random AR(2) or higher-order process, under the red-noise null hypothesis, may be classified as deterministic. Processes also can be defined for which the current level x_t depends on m previous values of x and n previous values of ϵ; such processes, denoted as autoregressive-moving average of order (m,n), or ARMA(m,n), may also be classified as deterministic. An example of such a process, in the seasonal to interannual domain, is the Southern Oscillation (SO), which was modeled as ARMA(3,1) by *Trenberth and Hoar* [1996]. Similarly, Fig. 4 shows that for ENSO, as measured by the NINO3 sea-surface temperature (SST) index, the red noise null hypothesis can be rejected at the 10% significance level (see Sec. 4.4. Thus, if we were attempting to simulate the NINO3 time series, an AR(1) (red noise) process would not suffice: Such a series would require a higher-order model.

3.2.3 Subannual variability

Seasonality is a significant factor in many, if not most, regions. In the simulation context it may useful to model only the rainy season, for example, since this is likely also the growing season and outside this window little significant precipitation may accrue. On the daily time scale there may be significant covariation among climate parameters; if simulations are to be generated at this time step, such covariation will need to be accounted for. Other daily (or more generally, high-frequency) statistics may also be at issue in simulation. These include extremes as well as spell lengths, for both temperature and precipitation. The daily component of a complete simulation scheme should attempt to account for those aspects of high-frequency variability deemed to be significant for the intended application.

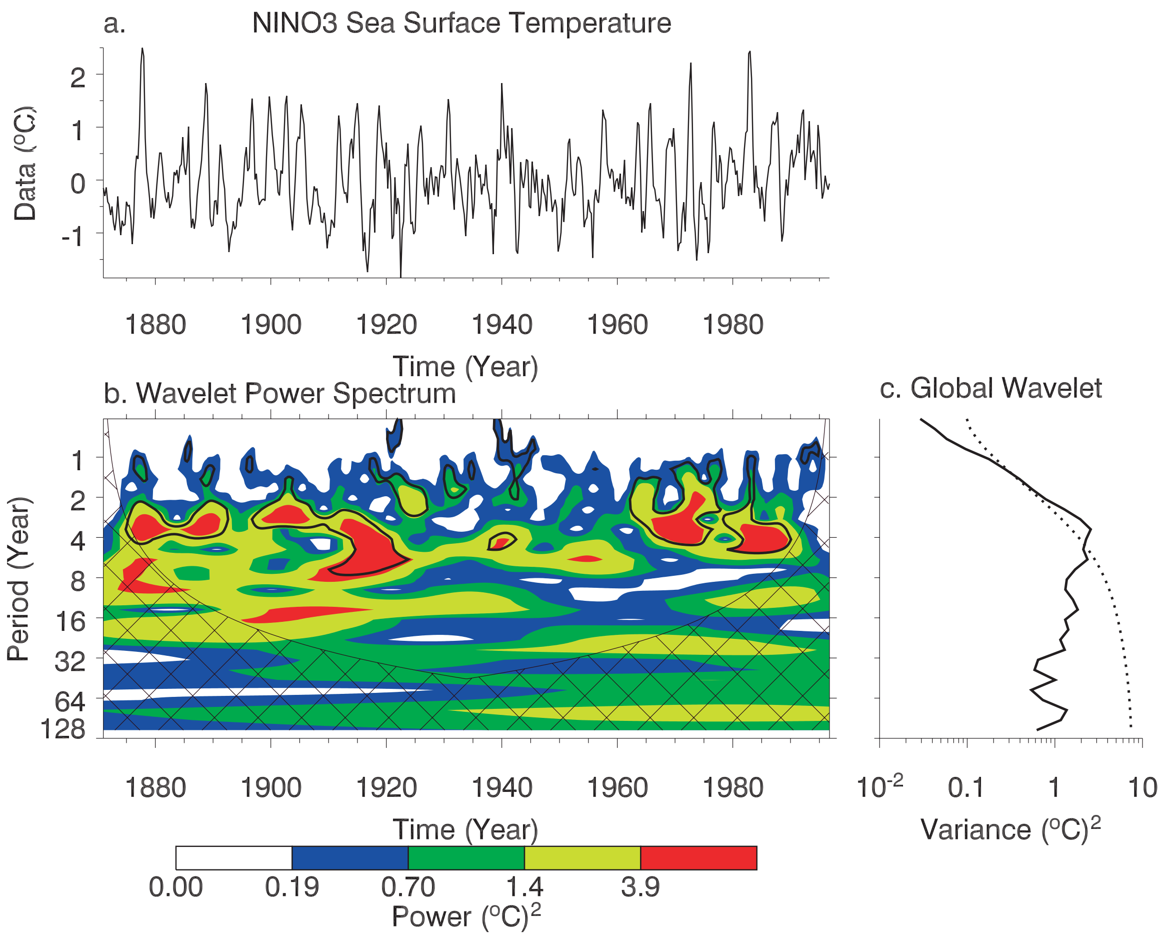

Figure 4: Time series plot of the NINO3 index (a), wavelet spectrum (b) and global wavelet spectrum (c), showing the 10% significance level for a red noise process. ENSO variability, as defined by this metric, differs from red noise at the 10% significance level. Analysis courtesy of http://paos.colorado.edu/research/wavelets.

3.3 Follow-on modeling requirements

As noted, some applications models require data on more highly-resolved time scales than annual. A crop model might utilize daily values of rainfall, temperature and insolation, for example. It is these requirements that will determine whether simulation must extend to the daily time step, and in fact which data must be simulated. A problem may arise if required variables (insolation comes to mind) have not been recorded, or if the available observational records are of short duration, in which case empirical rules, perhaps based on data from similar sites, may have to be devised in order to obtain reasonable simulation values. An example would be the creation of two insolation distributions, for wet and dry days. Insolation values could then be sampled from these distributions conditional on the occurrence of rain, as specified by the core simulation model.

In some cases, such as the Colorado River streamflow simulations mentioned earlier, univariate simulations are sufficient: The Lee's Ferry streamflow value encapsulates sufficient information for inference concerning relevant "downstream" impacts. In others, such as the case study to be described in Sec. 5, multivariate simulations are required, and intervariable correlation on annual-superannual time scales becomes an important simulation target. The selection of a univariate or multivariate model will depend in this manner on the exigencies of follow-on models to be driven with the simulation outputs.

3.4 A note on modeling philosophy

As with many statistical models, the machinery of simulation provides many "knobs" that the experimenter can turn at will, to generate scenarios having climate stresses of different types. It is our considered opinion that the best policy in the use of such controls is *discretion*: Simply because a parameter is modeled as adjustable does not mean that the user should take the first opportunity to modify it, simply in order "to see what happens." It is more sensible, in our view, to focus on climatic shifts that external evidence informs us are likely to occur, or at least whose likelihood is supported by theoretical arguments. Otherwise the risk exists of generating scenarios that have little likelihood of actually occurring, and possibly becoming lost in a maze of simulations whose relation to reality is tenuous. In effect, this amounts to nothing more than a recommendation for the principle of parsimony.

4 The simulation model

4.1 Overview

Model development is keyed to the decomposition by time scale (or equivalently, as we have argued, process class) discussed in Sec. 2. Beyond this it is conditioned by the contingencies considered in Sec. 3, including the availability of observational "training" data, requirements of agricultural or other follow-on models and, most importantly, characteristics of regional climate variability. Below we go step by step through these elements of the simulation process, noting at each stage the ways in which these contingencies affect the development of a suitable model.

The above description hints at an emergent theme, one that will become more evident as the discussion proceeds. This is the idea that a single, particularized model structure will not likely be equally applicable to all situations. Climates, observational data and application model requirements may all be expected to differ from region to region; models utilized for simulation will doubtless need to adapt to these differing requirements. This situation is common, perhaps even universal, in statistical modeling. Even though there is but one climate "system," that system is immensely complex, incorporates many subsystems and exhibits a correspondingly wide range of behaviors. What we attempt to do here is draw together those threads that are likely to be common to the simulation enterprise in many, perhaps most settings. Section 5, which follows, will provide an idea of how these common elements are made specific in one regional setting.

4.2 Dissasembly

As suggested above, three broad classes of climate process are recognized, each being treated separately. It is not necessarily the case that these classes are completely independent, in the sense that changes in one will not affect changes in (the statistics of) the others, but this is taken as a starting assumption. Consistent with the philosophy laid out in Sec. 3.4, we prefer a parsimonious approach: Should convincing evidence become available suggesting that that significant cross-scale coupling or dependence is likely to play a role in the future evolution of variability, then such interaction can be implemented in the simulation model. Absent such evidence we do not presume to tamper. We recognize that this approach may not be shared by all workers.

4.3 Treatment of trend

Although our focus is on the decadal scale, secular shifts in the mean that we identify here with anthropogenically forced climate change may also be important during the simulation target period. We therefore propose a mechanism for taking account of such trends.

There are two aspects to trend that concern us: First, we require a *detrending* procedure, to remove the (deterministic) forced response from the observational data on which the decadal component of the simulation model is to be trained. In other words, we expect to fit the simulation model to a detrended series. Second, in the simulation step we require some estimate of how the mean process level will evolve in the future, i.e., we need a plausible trend to project forward in time. These two trends, past and future, need not be the same.

There are many options available for fitting both linear and nonlinear trends to time series, the simplest perhaps being the straight line fit. Such a line can be extrapolated, providing in addition a trend for the future. However, past and future trends may differ, rendering such an approach questionable. Fitting nonlinear trends, using exponential or other parametric forms, does not address this problem. Further, trends fitted in the time domain have no *physical* basis: they are based purely on numerics, rather than any insight into the behavior of the system.

We propose instead to parameterize trend in terms of *climate sensitivity.* Since we associate trend with anthropogenic climate change, which takes the form of global warming, we model it in terms of regional sensitivity to global temperature change: In the absence of anthropogenic forcing the globe does not warm, and future trends are null. The globe is not expected to warm uniformly; some regions will warm more rapidly, others less so, than the global mean. Fitting local trends based on global climate sensitivity takes this variation into account.

For temperature the operative method used to date is to model both past and future trends based on the assumption that it is the spatial pattern of trends, rather than the rate of warming, that is stationary. Thus, regional or local temperature trends are regressed on a global mean temperature signal; the sensitivities so derived are used to both detrend the observed temperature record and project local temperature forward in time, based on the evolution with time of global mean temperature. For consistency, the global mean signal taken as regressand is derived from the multi-model mean of an ensemble of GCMs participating in the Fourth Assessment Report (AR4) of the Intergovernmental Panel on Climate Change (IPCC) [*Solomon et al.*, 2007], as listed in Table 1. These GCMs project climate changes through the end of the 21st century and beyond, making it a straightforward matter to derive simulation trends for the future.

The use of a single sensitivity for both past and future local temperature seems consistent, in view of the fact that both regressor and regressand variables are the same, thus are more directly coupled than might be the case for local precipitation and global temperature. This assumption can be tested, however, and future temperature trends conditioned in a different way if desired. The discussion of precipitation trends,

Group	Model
BCCR (Norway)	BCCR-BCM2.0
CCCMA (Canada)	CGCM-3.1
CCCMA	CGCM-3.1(T63)
CNRM (France)	CM3
CSIRO (Australia)	CSIRO-Mk3.0
CSIRO	CSIRO-Mk3.5
GFDL (USA)	GFDL-CM2.0
GFDL	GFDL-CM2.1
GISS (USA)	GISS-AOM
GISS	GISS-EH
GISS	GISS-ER
IAP (China)	FGOALS-g1.0
INGV (Italy)	INGV-ECHAM4
INM (Russia)	INM-CM3.0
IPSL (France)	IPSL-SXG
CCSR/NIES/JAMSTEC (Japan)	MIROC3.2(hires)
CCSR/NIES/JAMSTEC	MIROC3.2(medres)
MPI (Germany)	ECHAM5/MPI-OM
MRI (Japan)	MRI-CGCM2.3.2
NCAR (USA)	CCSM3
NCAR	PCM1
Hadley Centre (UK)	HadCM3
Hadley Centre	HadGEM1

Table 1: The 23 GCMs utilized from the IPCC AR4, from which the global mean multimodel mean temperature signal is extracted.

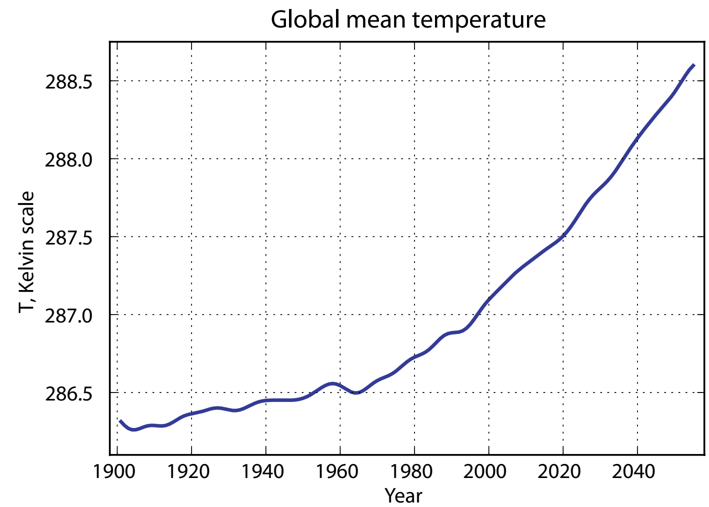

Figure 5: The smoothed global mean, multimodel mean temperature time series used as regressand in the detrending of observations and possibly for the projection of future trends.

below, offers one alternative.

The global temperature signal used as regressand is shown in Fig. 5. To obtain this series the global mean temperature record from each of the GCMs listed in Table 1 was first smoothed, using a Butterworth filter of order five and having a half-power point, or "cutoff," at a period of 10 yr. The plot shows the multimodel average of the 23 smoothed series thus obtained, extending from 1901 through 2049.

The critical operation, in forming the multimodel mean signal shown in Fig. 5, is not the filtering but the *averaging:* Since internal variability is intrinsic to each of the models, it is largely incoherent among them. Averaging thus acts to suppress this variability, while enhancing that part of the signal that the GCMs have in common. The one thing the GCMs share is a set of boundary conditions, which include the anthropogenic forcings that give rise to global warming. The resulting warming tendency is clearly evident in Fig. 5. Thus, model averaging enhances the climate change signal while attenuating internal variability noise, effectively increasing the signal-to-noise ratio in the resultant series. The filtering further smooths this signal, removing short-lived transients such as the effects of volcanic eruptions. Volcanos are treated here as unpredictable external forcing, unrelated to climate; no attempt is made to simulate their effects.

When local or regional signals are regressed onto the series of Fig. 5, the fitted values, now representing the local or regional climate change trend, appear as a scaled and shifted version of that series. This process and its result are shown in Fig. 6,

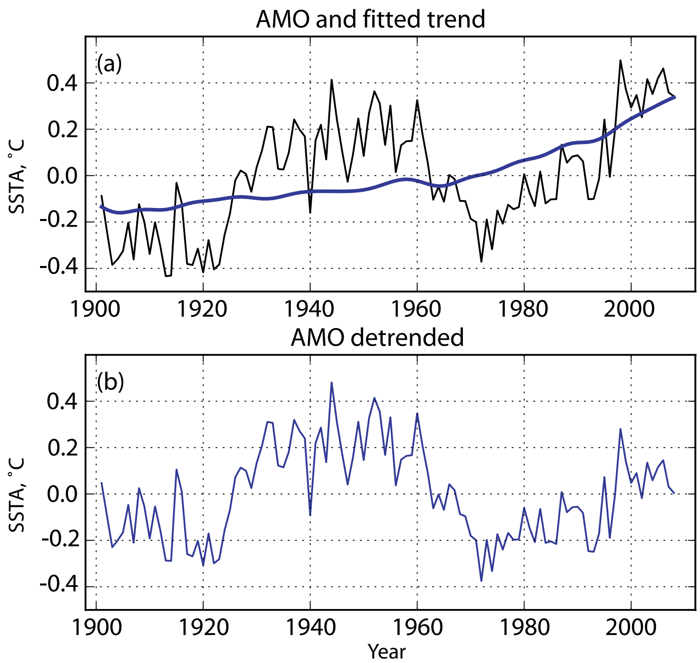

Figure 6: (a) The un-detrended AMO time series (Kaplan SST reconstruction) and a fitted trend computed by regression against the global mean temperature signal shown in Fig. 5. (b) Residuals from this regression, representing the "natural" component of variability.

where the detrending operation has been applied to the AMO signal beginning in 1901. The original signal, which has an upward trend, is shown in panel (a), along with the fitted trend, a scaled and shifted version of the the curve shown in Fig. 5. Note that the trend, although linearly dependent on the global mean temperature signal, is not linear with respect to time. In particular, because the globe has warmed more rapidly toward the end of the century, the trend accelerates during this time. The effect is that a greater portion of the AMO signal is assigned to anthropogenic causes than would be the case if the AMO were linearly detrended. The ability to follow changes in the planetary response to anthropogenic forcing, is what gives this detrending method its appeal.

Residuals from the regression fit of the AMO to global mean temperature are plotted against time in Fig. 6b. Two features of this curve stand out: First, there are large interdecadal swings, of peak-to-peak amplitude ~ 0.4°C. This slow "oscillation" identifies the AMO signal as one of the principal large-scale modes of *decadal* climate variability. Second, variations are not limited to the decadal/multidecadal band: There is considerable year-to-year variability as well. Treatment of such a residual signal, representing natural or unforced variability on multiple time scales, is discussed in following sections.

The case of precipitation differs somewhat, since we are now dealing with a variable that is qualitatively distinct from the global mean *temperature* against which regional temperatures were regressed. The response of precipitation to changes in global mean temperature, although it may be systematic in certain ways [*Held and Soden*, 2006], has a significant indirect component, in that it depends not just on shifts in temperature but also on changes in atmospheric circulation. Thus, projecting forward the results of a 20^{th}-century regression is a less certain enterprise when the regional variable is precipitation than with temperature.

Our recommendation in this case is to examine the GCMs, to see what they have to tell us about future precipitation trends. Although imperfect in many ways, GCMs are still the most sophisticated tools yet devised for the investigation of future climate. If there is robust agreement among models, perhaps supported by theoretical arguments, estimating regional precipitation sensitivity from them may be preferable to assuming that 20^{th}-century sensitivities will remain unchanged as climate warms. The case study, discussed in Sec. 5, provides an excellent illustration of this issue.

Precipitation trends are best parameterized in terms of fractional (or equivalently, percentage) change per degree of global temperature increase, rather than absolute change. This is in part because precipitation amounts cannot fall below zero, and in part because the modeling framework might call for an inferred regional sensitivity to be "propagated" to a network of stations having differing precipitation means. A sensitivity parameterized in terms of absolute amounts would not be appropriate in such a situation.

4.4 Treatment of "deterministic" components

The objective now is to fit some sort of statistical model to time series that are the analog of that shown in Fig. 6b, a detrended sequence comprising a possibly wide spectrum of variability on periods of one year and longer. The statistical model will be used to generate the "low-frequency" (i.e., annual-to-superannual) component of the simulated sequences. Continuing with a treatment analogous to the classic approach, we first attempt to determine the extent to which this "natural" component of variability represents the expression of deterministic, as opposed to random, processes. We do this following a procedure discussed in Sec. 3.2.2, by testing our series against a red-noise null hypothesis.

4.4.1 Wavelets

Figure 4 illustrates a wavelet decomposition of the NINO3 SST index, with Figs. 4a, 4b and 4c showing the NINO3 time series, the wavelet spectrum and the global wavelet spectrum, respectively. In the last of these three plots, the dotted line indicates the 10% red noise significance level. That the curve of spectral power exceeds this level in the ENSO band, corresponding to periods of roughly 2-8 yr, implies that the probability is less than 10% that ENSO is the expression of a red noise process. Thus, if we wished to generate simulations that remain faithful to the properties of the NINO3 signal shown in Fig. 4a, a red-noise model alone would not suffice. The analysis of *Trenberth and Hoar* [1996] suggests the use of an ARMA(3,1) model for this purpose. The residuals from such a model would likely have the character of red noise, and would be assigned to the "random" component of low-frequency variability.

The "spectral summary" of Fig. 4c is in effect a byproduct of the wavelet analysis, whose principal insights are provided via the full wavelet power spectrum, as shown in Fig. 4b. This spectrum shows not only that the NINO3 SST index has significant power in the 2-8 yr band, but that the expression of variability in this frequency band has not been constant over time, experiencing a period of relative weakness between about 1920 and 1960. On the one hand this shows that ENSO has been a robust feature of 20$^{\text{th}}$-century climate, and thus may be expected to continue into the future. On the other, it suggests the use of a second-order model, in which ENSO-band activity is amplitude-modulated on multidecadal time scales. In the latter case, if one were to generate, say, sequences of length 50 yr, some would exhibit strong ENSO variability while others would not. In the absence of additional guidance regarding the future behavior of ENSO, it may be necessary to take this secondary level of uncertainty into account in the design of decision-level modeling strategies.

Wavelet analysis might also reveal that a process associated with a significant spectral peak was active only during the early part of the century, having since apparently gone dormant. In this case judgment would be required, but arguably the choice not to model such a component explicitly might be a reasonable one. In general, modeling choices must be guided not only by the cold numerical information in the wavelet spectrum but, as in most enterprises, the exercise of good judgment.

4.4.2 Wavelets elaborations and alternatives

A systematic method for decomposing signals using wavelets, identifying components that differ significantly from the red noise background, modeling those components as higher-order stochastic processes and generating corresponding simulations is wavelet autoregressive modeling (WARM), described by *Kwon et al.* [2007, 2009]. Such a method could possibly also serve here as the low-frequency component of the simulation model. The exposition by those authors involves only univariate series; elaboration of the technique would be required for extension to the multivariate case.

An additional option for the low-frequency simulation component, also utilizing a sophisticated red-noise significance test (and also basically univariate), is Monte Carlo singular spectrum analysis (MCSSA) [*Allen and Smith*, 1996]. The resulting spectral decomposition resolves the target signal into deterministic and random (red noise) components; the former can be projected forward using a technique called linear predictive coding [*Press et al.*, 1986-1992]. Space limitations preclude more detailed discussions of the methods discussed in this section, but they should be kept in mind as options as a simulation strategy is developed.

4.4.3 Periodic phenomena; large-scale decadal "modes"

Besides stochastic processes of order > 1, the basic building blocks of the WARM method, a deterministic signal might be *periodic.* The 11-year solar cycle, for example, has been identified in many climate records. Periodic signals can of course easily be simulated, as with the seasonal cycle in subannually resolved time series. However, we recommend caution here, since red noise, as well as higher-order stochastic processes, can easily be mistaken for periodic phenomena [*Wunsch*, 1999]. Many of the solar-cycle findings reported with enthusiasm have ultimately proven elusive [*Pittock*, 1983]. Unless a plausible physical explanation or linkage can be offered to explain such periodic signals, accumulated experience would appear to recommend skepticism. Still, if persuasive evidence does exist, sinusoidal functions can always be deployed to represent periodic phenomena.

A dependence on one or more of the large-scale decadal modes provides one further example of possible "determinism" that might be present in the target signal. Studies suggest that at least some predictability may be associated with these modes [*Knight et al.*, 2005; *Newman*, 2007], raising the possibility of simulations that contain an actual predictive component. The modeling strategy in this case would begin with the predictions, which would then be propagated to the study region in a manner determined by observational analysis.

4.4.4 A range of models

It should be clear that "deterministic" elements can assume many forms, and this is one reason that a definitive formula for decadal simulation cannot be specified in complete detail *a priori.* For the deterministic component of natural variability the form of the target signal will dictate the choice of statistical model. The range of such models can thus be large, at least in principle. If, however, as suggested by WARM modeling, most deterministic elements can be treated as fairly low-order stochastic processes, the minimally-sufficient class of models would be limited to a reasonably small set. This may well be the case in practice, but application in a variety of simulation settings will be required in order to better delimit such a class.

4.5 Treatment of random components

That part of the target signal not deemed deterministic is treated via a purely stochastic model. It may be that much, or most terrestrial decadal variability falls into this category, since the recognized modes are principally oceanic, or ocean-atmosphere coupled. Their variations are communicated to land regions via so-called teleconnections, perturbations to atmospheric flow regimes giving rise to climate anomalies in locations remote from the oceanic sources of variability. As such, low-frequency signals may be "diluted" by the noise of atmospheric variability. True periodic phenomena, other than the diurnal and annual cycles, seem to be rare. So the red noise null hypothesis may prove a difficult barrier to surmount. The possible presence of a deterministic component notwithstanding, a first-order autoregressive model is thus taken as the basic building block for the random simulation component. Since the AR(1) process has *memory* (Figs. 1b, 2b), it can generate fluctuations on decadal time scales, including potentially long spells above or below the mean. However, although AR(1) processes may meander up and down, they cannot generate true periodic behavior: As Fig. 2b suggests, the spectra of such processes exhibit no peaks.

As has been mentioned, agricultural or other applications models may require multivariate input, in which case a natural generalization of the red noise process is the first-order vector autoregressive (VAR) model. In this case additional considerations come into play, since the variables, typically precipitation and temperature, are likely to be correlated. Thus, it should be verified that the selected model can reproduce not only the persistence characteristics of the target observations but also the intervariable correlation structure. In either case, suitable model checking is a sensible way to verify that the structure selected is appropriate for the target simulation data.

4.6 A nonparametric alternative

Nonparametric resampling techniques offer a possible alternative to the fitting of a (parametric) stochastic model, and, depending on data availability, might even obviate the necessity of separating deterministic and random components of the target signal. Methods such as K-nearest neighbor (K-NN) [*Lall and Sharma*, 1996; *Rajagopalan and Lall*, 1999] bypass the fitting procedure completely, simulating by resampling the data, conditional on a metric for the "distance," in parameter space between variables at the present time step and those at a predetermined number (K) of similar points in the past or future.

In pure form, K-NN admits only data points that belong to the original series, limiting the range of simulated variability somewhat. Modified schemes permit the addition of random elements to be added to the resampled values, adding a degree of variation beyond the simple reordering of data points afforded by the original method. Such a "randomness deficit" may become a handicap in the case of short observational records, since these may severely constrain simulation variability from one realization to the next. But given ample data from which to sample, the K-NN method can generate simulations having distributional or spectral properties that would otherwise be difficult or impossible to mimic.

4.7 Reassembly / downscaling

We have described the simulation process in terms of the decomposition of a target signal into trend, low- and high-frequency components, and the decomposition of the low-frequency component (periods of years to decades) into deterministic and random elements. In generating the simulated sequences the process is reversed: The components are simulated individually and the results combined.

Deterministic elements from the low-frequency signal component are simulated first. To these a random component, based on the selected stochastic model (AR(1) or VAR in this discussion), is added. Alternatively, the entire low-frequency component may be simulated via a nonparametric scheme such as K-NN. The result of this step is a simulation of regional low-frequency variability, annually resolved and without trend, of a specified length, typically several decades. If simulations are to be generated over a network, this low-frequency component, suitably scaled, shifted and mixed with uncorrelated noise, will, in a later step, be propagated to all stations in the network. The regional simulation thus serves as a sort of "master copy," a template to which stations across the network will be mapped. Ultimately, many such simulations will be generated, in order to define the uncertainty ranges of interest to applications modelers.

Separately, trend is extended into the future. If the 21$^{\text{st}}$-century sensitivity is expected to be the same as that of the 20$^{\text{th}}$century, the trend is projected forward using the coefficients from the 20$^{\text{th}}$-century regression, applied to the 21$^{\text{st}}$-century global temperature signal. Alternatively, if the 21$^{\text{st}}$-century sensitivity is GCM-derived, this sensitivity is used instead as the basis for trend projection. For temperature the trends are applied additively; for precipitation, multiplicatively. For a network, the temperature trend computations are performed station by station, since differing subregions may warm at different rates. If precipitation trends are to be estimated from the 20$^{\text{th}}$-century sensitivity this is done for precipitation as well. GCM-derived regional precipitation trends based on 21$^{\text{st}}$-century data alone are applied uniformly over the entire network; these cannot be inferred at the station level.

Trend and low-frequency components together represent all simulation variability on time scales of one year or longer. What remains is the spatial/temporal downscaling step. As mentioned, trend computations are based either on station-level regressions or GCM-based estimates. Thus, these are already implicitly downscaled, or ready to be downscaled. For the low-frequency component the procedure differs somewhat: Over a region, local stations will not be perfectly correlated with the regional decadal signal. For this reason the regional simulation is propagated to the local level using linear regression, a residual being supplied in the form of uncorrelated noise to insure that station variance is conserved. The downscaling of simulated regional low-frequency variations, as distinct from trend, is performed in the same manner for both temperature and precipitation variables.

For the temporal downscaling a weather generator or resampling method can be used. Here, we have chosen a simpler alternative, in order to maintain the focus on the low-frequency simulation component: Entire years from the observational record are resampled at random, and the daily values rescaled to have the annual mean value of the regional simulation. This method preserves the seasonal cycle, daily variability and covariances, not only across variables but across the entire network being simulated, while stripping out the interannual component and substituting for it the simulation value. The result is a set of hybrid station-level simulations, each having the subannual characteristics of the station but the annual-superannual properties of the simulation.

For coding efficiency, the injection of low-frequency variations and trend occurs simultaneously with the spatial downscaling step. The result is a fully-resolved simulation, for a single station or an entire network, having realistic variability on all time and space scales. For a multivariate simulation (and assuming the low-frequency model is well chosen) intervariable correlations will be respected, as will serial autocorrelation in the individual variables, on annual and longer time scales, while the resampling-by-year procedure preserves these relationships on subannual time scales. The simulations will also have secular trends conditioned by best estimates of future behavior (conditioned, as we have noted, by the best judgment of the modeler). Such simulations should be suitable for driving agricultural, hydrological or other applications models, in explorations of the ranges of potential impacts engendered by decadal-scale climate variability.

4.8 Additional considerations

Although a simulation model can be fit to annual-mean values, this is not necessarily an optimal procedure. The rainy season, for example, may cross the calendar boundary from one year to the next, so it might be more sensible to define a hydrological year that differs from the calendar year. Modeling annual values for a specified season may also constitute a useful strategy, assuming that simulated output in this form is compatible with follow-on models. In the case study to be discussed the rainy season is Austral Winter, which does not cross the calendar boundary. However, the follow-on hydrological model requires full years of simulated data in order to close the soil moisture budget. Other modeling frameworks are likely to have differing requirements, and these will constrain simulation design accordingly.

Once the simulation model is complete, many sequences can be generated; as an ensemble these sequences can be used to delineate confidence intervals (CI) for climate change in the near term. It should be emphasized that relevant CI may be constructed, not just for climate uncertainty in a given year (or decade), but for temporal characteristics, such as dry and wet epoch lengths, as well. Using present climatology as a baseline (and assuming fixed thresholds), such characteristics would be expected to shift under any long-term trend. In the multivariate setting temperature trends may play a significant role in the hydrological cycle even when precipitation exhibits no significant trend, for example. In any event, a large ensemble of simulations, each representing a "plausible" realization of future climate, may be generated; these simulations then become the inputs to crop or other application models, and can be used, for example, to "stress test" adaptive systems for resilience in the near-term climate change context.

5 Elements of a case study: The Berg River watershed

The ideas and methods presented in this report represent a generalization of lessons learned in implementing a highly particularized simulation scheme, which is described in detail in *Greene et al.* [2011]. We abstract here certain elements of that implementation, in an attempt to show how the more generalized simulation rubric described in the foregoing sections is actually realized in a specific simulation context.

5.1 Setting

The Berg watershed lies to the north of Cape town, in the Western Cape province of South Africa (Fig. 7). It is about 300 km in length, with an area of 7715 km^2, about two thirds of which is agricultural land. It provides water for industrial uses and is the principal water source for the city of Cape Town. Because of its economic importance it has been the subject of considerable study, and a suite of models has been developed for the estimation of economic impacts owing to possible changes in climate [*Callaway et al.*, 2008]. A dam has recently been constructed on the river (http://en.wikipedia.org/wiki/Berg_River_Dam), partly in anticipation of a climatic

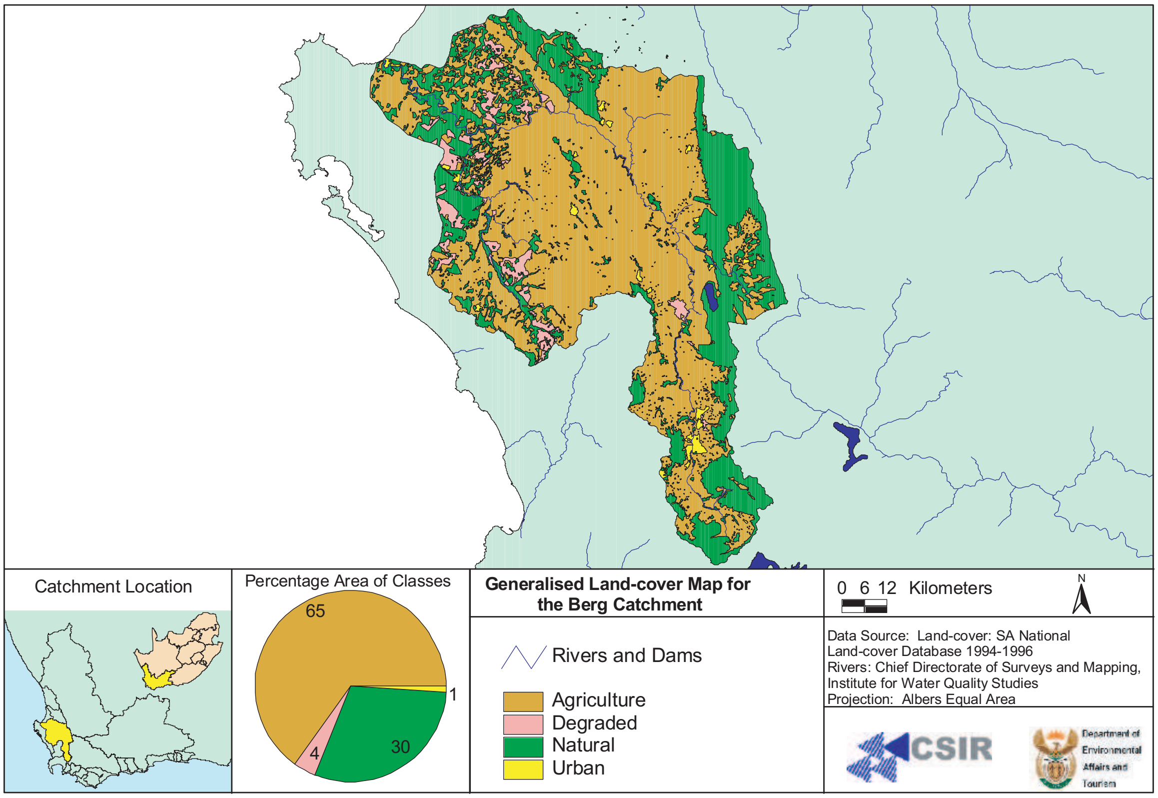

Figure 7: Map of the Berg River watershed. Cape Town lies off the large map to the southwest. Courtesy South Africa Dept. of Environmental Affairs and Tourism.

drying trend forecast by the IPCC suite of GCMs [see Ch. 11 in *Solomon et al.*, 2007]. A hydrology model, the Agricultural Catchments Research Unit (ACRU) model [*Schulze*, 1995; *Smithers and Schulze*, 2004] has been adapted for this watershed, and simulations were designed specifically with this model in mind. Considerations extended from the required suite of input variables down to the precise file formatting ACRU requires. The project itself is part of a broader modeling effort undertaken in collaboration with scientists at the University of Cape Town, University of the Free State and University of KwaZulu-Natal, in which the International Research Institute for Climate and Society (IRI) group was charged with the elucidation of "near-term" climate variability and its potential effects on hydrology in the watershed of the Berg.

An initial 20-yr target period, 2011-2030, was specified for investigation by the hydrology modeling group. The object of the simulation exercise was to investigate the degree to which decadal-scale variations might affect 20-yr mean runoff during this period. Follow-on modeling by the economic team will attempt to infer the potential

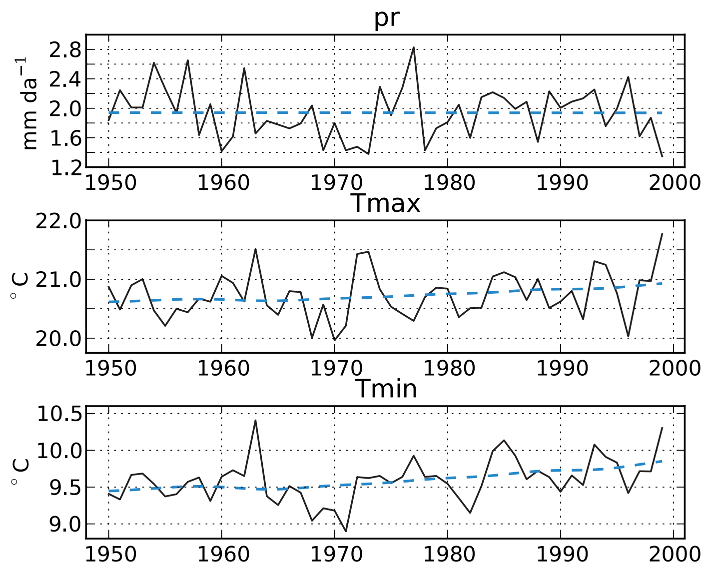

Figure 8: Regional (trivariate) signal for the Berg catchment.

economic effects owing to defined percentile climate deviations for the target period, as identified via the simulation process.

5.2 Observational data

The watershed area and surroundings have been mapped into higher-level subcatchments; modeling was carried out using a 50-yr daily data set, including precipitation and minimum and maximum daily temperatures, (pr, Tmin, Tmax) defined on a network of 171 quinary catchments, most lying within the mapped catchment (Fig. 7), the remainder in proximal regions. There is not a unique weather station located in each quinary catchment; rather, data from a somewhat smaller network of stations (one for each group of 3-5 catchments) is interpolated to the catchment locations using empirical relationships.

The "regional" signal, on which the low-frequency simulation model is based, is the 171-catchment mean climate record, reduced to annual resolution. Since the catchment records include pr, Tmax and Tmin, this regional signal consists of a 50-yr trivariate time series (1950-1999), the components of which are plotted in Fig. 8. Trend lines in this figure are based not on time, but on climate sensitivity, as discussed in Sec. 4.3.

5.3 Hydrology model requirements

The ACRU model requires, at a minimum, daily values for all three variables, implying the use of a multivariate simulation model. The Western Cape, unlike most of southern Africa, experiences a winter rainfall maximum, and modeling just this season was initially given consideration. However, ACRU, because of its inclusion of soil moisture, has year-to-year memory, and requires input data for the full year. Model fitting and simulation were therefore carried out on annually-averaged data. The Jan-Dec year was utilized; this does not have the undesired effect of dividing the wet season across years.

ACRU has exacting specifications for the formatting of input variable arrays, and requires as well that leap years be respected. This required the configuring of simulation code so as to take these requirements into account.

5.4 Implementation

To detrend the regional series, each component was regressed on the global temperature signal shown in Fig. 5. The fitted values are overplotted on the regional series in Fig. 8. The fitted trends were subtracted from the series to obtain the residual "natural" component of variability. This natural residual, like the regional series itself, has annual time resolution.

The wavelet spectrum of each of the three natural components (pr, Tmin, Tmax) was computed; these spectra gave no indication that the component series differed from red noise. In other words, deterministic components were not detected in the regional record. Because of this, a modeling step needed to account for such components would have been superfluous. Since the regional record is multivariate, a VAR model of order 1 was therefore selected for simulation. Experiments with two other model structures — a two-state hidden Markov model (an alternative not discussed in Sec. 4.5) and a K-NN resampling scheme applied to the annual values of the regional series — indicated that neither of these models did as good a job at capturing both intervariable correlation and serial autocorrelation in the individual variables as the VAR model, confirming its selection as preferable.

The VAR model was thus fit to the regional data series. A single long simulation (500 kyr) was then generated, using the fitted values. Such a simulation is equivalent to an ensemble of 10000 50-yr simulations; the single file was utilized for convenience in data manipulation and for bookkeeping purposes. Statistics computed on this sequence indicated that it preserved well both the intervariable correlations and individual serial autocorrelations (thus persistence, or low-frequency variability) of the individual regional series.

Requests from the hydrology modeling team included the 20^{th} and 80^{th} percentiles for 20-yr means of each of the variables, for the target simulation period 2011-2030. That is, simulations were requested for which the 20-yr mean precipitation (and eventually Tmin and Tmax) for the 2011-2030 period lay at these two percentiles. An additional simulation with all 2011-2030 means at their 50th percentiles completed the request, for a total of seven sequences. The simulated sequence length was set at 66 yr, extending from 2000 through 2065. This will eventually enable comparisons with downscaled climate change simulations that rely on daily data from the IPCC models, which is available for the 2046-2065 period.

The very long simulation sequence permitted fairly precise screening, to identify 20-year sequences for which (a) means of the specified variables fell at the designated percentile levels and (b) means of the "secondary" variables (Tmin and Tmax if the percentiles of pr were prespecified, and so on) lay close to their conditional mean values, given the selected percentile of the primary variable. Condition (b) was imposed so that hydrology driven by the simulations would not be biased by atypical values of the secondary variables.

Each of the sequences selected via the screening process was downscaled to all 171 catchments, for a total of 1197 output files, each containing 66 yr of daily data for pr, Tmin and Tmax. Running ACRU in distributed mode (i.e., over the entire watershed) provides a more granular description of the climate impacts on local hydrology and is deemed desirable by the hydrology team. Given the selected simulation, downscaling was performed one quinary catchment at a time, as described in Sec. 4.7.

For the temperature trend components, each catchment's record was regressed on the global mean signal; the derived coefficients were then used to project 21st-century trends. For precipitation the 21st-century trend was computed using the GCM ensemble alone, without reference to the observations. This was done because 20th- and 21st-century precipitation sensitivities were found to differ considerably: There is no significant precipitation trend during 1950-1999 in either the GCMs or observations, even though temperatures increased markedly during this period. Thus the sensitivity derived using data from this period would be null, and simulated precipitation would remain constant regardless of global temperature change. By contrast, a strong drying tendency is forecast for southwestern South Africa during the 21st century (in the GCMs), particularly for winter. This drying trend is robust in the GCM ensemble and consistent with theoretical expectations having to do with expansion of the dry subtropics, with a particularly robust response occurring toward the poleward margins of the subtropical dry zones. Southwestern South Africa lies in just such a zone [See Fig. 11.2 in *Solomon et al.*, 2007].

Again, the judgment of the modeler is required, to resolve these two conflicting inferences regarding future precipitation sensitivity: In this case it was decided to heed the message of the GCM simulations, in part because the temporal pattern of drying seems consistent with the poleward advance of a dry subtropical regime that reaches, and eventually overtakes, the region of the Western Cape. The 21^{st}-century regional precipitation sensitivity to global temperature change, as computed within the GCM ensemble, was thus utilized in the forward projection of simulated rainfall changes. Further investigation of the dynamics involved would certainly be desirable, but is beyond the purview of this report.

For the natural component of variability, individual catchment records were first regressed on the 171-catchment mean, to estimate the degree to which each catchment's record expresses the "regional" low-frequency signal. The derived regression coefficients were then used to propagate the simulations to the catchment level, uncorrelated noise being added in order to replace variance lost to the regression. The result of this step is a simulated local (trivariate) sequence with annual time resolution, having temperature trend components local to the catchment and a precipitation trend component derived from the 21^{st}-century GCM ensemble. In addition, this now-localized signal expresses the simulated low-frequency regional variability, to the same degree that that the catchment-level observations were found to express the signal of regional variability during the 20^{th} century. All of the catchment records, for all of the variables, are positively correlated with the 171-catchment mean signals, often strongly so, indicating that the Berg watershed is driven, on annual-superannual time scales, by a coherent large-scale signal.

To downscale to daily resolution, entire data years (365 or 366 values at a time, of all three variables) were randomly resampled from the catchment record for each simulation year. The individual variables are stripped of their intrinsic annual means, for which the simulated mean values coming from the previous step are now substituted. The resulting daily sequences are "hybrids," having the subannual variations of the randomly selected year, but on annual and longer time scales (excluding trend) carrying the signature of the simulated variability. These sequences, appropriately formatted, constitute the data passed to ACRU.

A simulated catchment-level sequence is shown in Fig. 9. The years 1950-1999 consist of the original catchment-level observations; in the present scheme these years are not simulated. The 21^{st}-century multimodel-based drying trend, as well as warming trends in both Tmin and Tmax, can be seen clearly. There is little or no precipitation trend during the 20^{th} century.

Less obvious is the negative precipitation anomaly for years 2011-2030, representing the 20^{th} percentile for 20-yr means, as computed in the "raw" simulation sequence, i.e., before propagation to the catchment level. (Recall that the propagation step involves some "dilution" of the decadal signal.) Statistics of 10- and 20-year means at the catchment level suggest that on decadal time scales, natural variability is large enough to completely offset (or potentially double), with one chance in five, the long-range drying tendency owing to anthropogenic climate change, although details will depend on the degree to which individual catchments "subscribe" to the

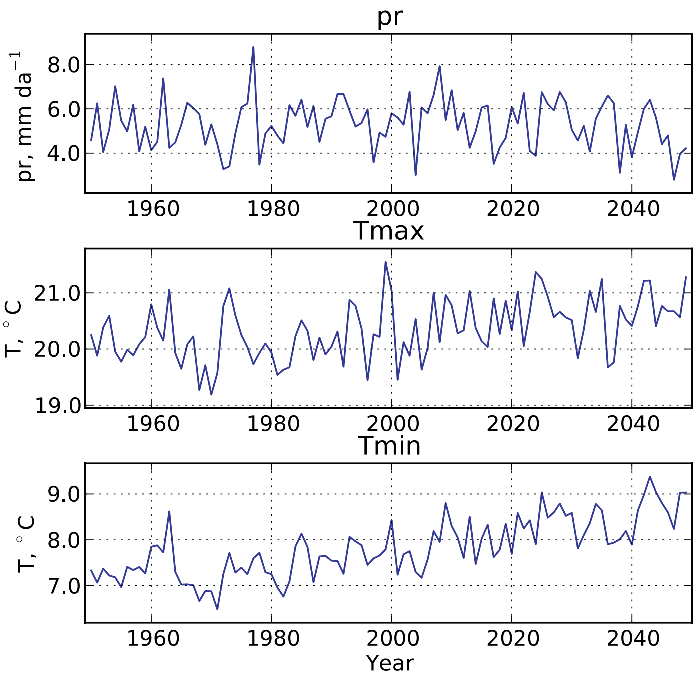

Figure 9: A typical simulated catchment-level sequence, reduced to annual time resolution. The declining trend in precipitation is derived from the IPCC multimodel ensemble; increasing temperature trends are based on 20$^{\text{th}}$-century sensitivities.

regional decadal signal. This is a significant conclusion for adaptation planning. Another observation is that both decadal and anthropogenic signals appear against a background of significant interannual noise. Thus, planning for climate stresses owing to year-to-year variability remains an important consideration for the overall risk assessment portfolio.

6 Discussion

The methodologies we have discussed above comprise at least two levels of generality: The simulation "rubric" itself is broadly sketched out, while realization at the level of the case study is considerably more particularized. Thus, the former should be considered a framework, within which simulation details must be elaborated according to the available evidence and particularities of the simulation setting. Evidence to be considered should include the observational record, information from GCMs, theoretical expectations and possibly paleorecords, if the latter are available. These sources must be weighed with respect to both content and reliability and a coherent narrative woven from the various threads that they present. A "one-size-fits-all" simulation scheme is thus unlikely to fit any; in the end, as in most enterprises, skilled judgment will be required in order to achieve a meaningful result.

In addition to decade-length swings in mean precipitation or temperature, changes in features such as dry-spell length and extreme events merit attention. Where long paleorecords are available, such as in the American southwest, such records may reveal wet or dry epochs, such as the well-known "megadroughts," that do not have parallels in the observational record [*Stahle et al.*, 2010]. Simulations based on such evidence, although perhaps less certain than those based solely on observational records, have the potential to expand our conception of the possible, in ways that the latter do not. With respect to simulations that are based even on observed variability alone, previously unexperienced climatic stresses, perhaps in the form of threshold exceedance frequencies, may arise as a result of interactions between low-frequency variations and climatic trends of anthropogenic origin. In either case, study of temporal behavior through the use of decadal simulations would seem warranted.

A large ensemble of simulations, exemplified here by the 500-kyr sequence utilized in the case study, presents an opportunity to compute simulation statistics to high precision. This is a direct consequence of the statistical nature of the simulation methodology, which is computationally inexpensive. This advantage must be weighed against the inability of any statistical method to anticipate changes owing to processes or interactions that are not included *a priori* in the model.

An illustration of the way such issues might be addressed can be provided for the case of temporal precipitation variability. As a result of anthropogenic warming it is widely believed that such variability will increase, owing to the rapid increase of water saturation vapor pressure with temperature: A warmer atmosphere can transport more water vapor. Because of this it can rain more but also become drier, since exchanges of water between land and atmosphere are bidirectional. Thus the variability of precipitation could increase as climate warms. There is no mechanism in the statistical model we have described that would act to bring this about, however. Independently (outside the model) precipitation in western South Africa in the individual GCM runs was inspected, and it was found that, at least on interannual time scales, this variability *decreased* very slightly as climate warmed during the first half of the 21$^{\mathrm{st}}$ century. (The decrease was not statistically significant.) It was concluded that there was no basis for including a cross-scale mechanism linking precipitation variability in the subject region to global temperature change. Such a procedure, while not a definitive determination of what might occur in the future, can at least serve as an indicator for processes under consideration for inclusion in a simulation model.

Simulation covariance among pr, Tmin and Tmax is found to vary from realization to realization, a not-unexpected result, given the stochastic nature of the simulation process. It thus becomes natural to think of covariance in the observational record itself as a *sample,* rather than as a standard to which all realizations must be rigorously held. In this way the simulation exercise can also throw light backward, onto aspects of the observed climate that are perhaps deserving of contemplation.

7 Summary

We have described a rubric, or framework, for the generation of stochastic simulations, with the end in mind of driving agricultural or other applications models that require detailed climate information having a realistic representation of decadal variability. The incorporation of such variability into impacts studies represents an advance over the simple comparison of mean states that has typically been performed in climate change impact studies.

The approach presented is based loosely on classical time series analysis, in that an observational record, which is taken to represent regional climate variability, is decomposed into trend, deterministic and random components, each of these being treated independently. An association is made between trend — a secular shift in the mean — and anthropogenic forcing. Accordingly, this component of variability is modeled by regression on a global mean temperature signal, meaning that it is modeled as a sensitivity to global temperature change, rather than as time-dependent. Detrending, as refracted through this procedural prism, then amounts to separating climatic changes due to anthropogenic effects, and natural variability intrinsic to the climate system itself. Possible problems that arise in attempting to effect such a separation using short time series were discussed.

Trend having been removed, the residual variability is examined for evidence of deterministic processes, in the sense that the residual variations differ significantly from red noise. If such processes are identified, they would be modeled as separate independent components, with the residual from this step modeled as a stochastic process. It is this component of the analysis — the deterministic and stochastic elements of low-frequency non-trend variability — that offers perhaps the widest latitude for the modeler, depending on the types and characteristics of the available records and the nature of the information sought, as well as the characteristics of the data itself. There is a trade-off in this richness, between flexibility in "customizing" the modeling approach and a lack of specificity with regard to how to accomplish such a customization. But every climate setting being unique in some way, flexibility in modeling (with the attendant role for creativity) would seem on balance to constitute a net positive.

It is hoped that the methodology outlined here will prove useful in delineating uncertainties owing to natural internal variability, in the context of a background climatic state undergoing secular, forced shifts. For better or worse, this is the situation in which we are likely to find ourselves in coming decades. The investigation and characterization of such uncertainties will play an important role in anticipating potential climate risks in the near term, and the more confidently such risks can be defined, the better prepared we will be to deal with them in coming years.

Acknowledgments

This work has been supported by donors to the CGIAR through the CCAFS secretariat hosted by the Department of Agriculture and Ecology at the Faculty of Life Sciences at University of Copenhagen. This document has been produced with the financial assistance of the European Union, Canadian International Development Agency, World Bank, New Zealand Ministry of Foreign Affairs and Trade and Danida and with the technical support of IFAD. The views expressed herein can in no way be taken to reflect the official opinion of these agencies.

References

Allen, M. R., and L. A. Smith, Monte Carlo SSA: Detecting irregular oscillations in the presence of colored noise, *J. Clim.*, *9*, 3373–3404, 1996.

Callaway, J. M., D. B. Louw, J. C. Nkomo, M. E. Hellmuth, and D. A. Sparks, Benefits and costs of adapting water planning and management to climate change and water demand growth in the western cape of south africa, in *Climate Change and Adaptation*, edited by N. Leary, J. Adejuwon, V. Barros, I. Burton, J. Kulkarni, and R. Lasco, chap. 3, pp. 53–70, Earthscape, 2008.

Greene, A. M., M. Hellmuth, R. Schulze, and T. Lumsden, Stochastic decadal simulations for the berg river watershed, western cape province, south africa, *Water Resourc. Res.*, 2011, in preparation.

Hasselmann, K., Stochastic climate models part 1. Theory, *Tellus*, *28*, 473–485, 1976.

Held, I. M., and B. J. Soden, Robust responses of the hydrological cycle to global warming, *J. Climate*, *19*, 5686–5699, 2006.

Kaplan, A., Y. Kushnir, M. Cane, and M. Blumenthal, Reduced space optimal analysis for historical datasets: 136 years of Atlantic sea surface temperatures, *J. Geophys. Res.*, *102*, 27,835–27,860, 1997.

Kaplan, A., M. Cane, Y. Kushnir, A. Clement, M. Blumenthal, and B. Rajagopalan, Analyses of global sea surface temperature 1856-1991, *J. Geophys. Res.*, *103*, 18,567–18,589, 1998.

Kleeman, R., Stochastic thoeries for the irregularity of enso, *Philos. T. R. Soc.*, *366A*, 2509–2524, 2008.

Knight, J. R., R. J. Allan, C. K. Folland, M. Vellinga, and M. E. Mann, A signature of persistent natural thermohaline circulation cycles in observed climate, *Geophys. Res. Lett.*, *32*, 2005, l20708.

Kwon, H.-H., U. Lall, and A. F. Khalil1, Stochastic simulation model for nonstationary time series using an autoregressive wavelet decomposition: Applications to rainfall and temperature, *Water Resour. Res.*, *43*, 2007.

Kwon, H.-H., U. Lall, and J. Obeysekera, Simulation of daily rainfall scenarios with interannual and multidecadal climate cycles for South Florida, *Stoch. Environ. Res. Risk. Assess.*, *23*, 879–896, 2009.

Lall, U., and A. Sharma, A nearest neighbor bootstrap for time series resampling, *Water Resour. Res.*, *32*, 679–693, 1996.

Newman, M., Interannual to decadal predictability of tropical and north pacific sea surface temperatures, *J. Climate*, *20*, 2333–2356, 2007.

Pittock, A. B., Solar variability, weather and climate: an update, *Q. J. R. Meteorol. Soc.*, *109*, 23–55, 1983.

Prairie, J., K. Nowak, B. Rajagopalan, U. Lall, and T. Fulp, A stochastic nonparametric approach for streamflow generation combining observational and paleoreconstructed data, *Water Resour. Res.*, *44*, 2008.

Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in Fortran*, 2nd ed., Cambridge University Press, 1986-1992.

Rajagopalan, B., and U. Lall, A knearest-neighbor simulator for daily precipitation and other weather variables, *Water Resourc. Res.*, *35*, 3089–3101, 1999.

Saravanan, R., and J. McWilliams, Advective ocean-atmosphere interaction: An analytical stochastic model with implications for decadal variability, *J. Climate*, *11*, 165–188, 1998.

Schulze, R. E., Hydrology and Agrohydrology: A Text to Accompany the ACRU 3.00 Agrohydrological Modelling System, *Report TT 69/95*, Water Research Commission, Pretoria, RSA, 1995.

Smithers, J. C., and R. E. Schulze, ACRU Agrohydrological Modelling System: User Manual Version 4.00, *Tech. rep.*, School of Bioresources Engineering and Environmental Hydrology, University of KwaZulu-Natal, Natal, South Africa, 2004.

Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K. Averyt, M. Tignor, and H. Miller (Eds.), *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2007.

Stahle, D. W., F. K. Fye, E. R. Cook, and R. D. Griffin, Tree-ring reconstructed megadroughts over North America since a.d. 1300, *Clim. Change*, *83*, 133–149, 2010.

Trenberth, K. E., and T. J. Hoar, The 1990-1995 El Niño-Southern Oscillation event: Longest on record, *Geophys. Res. Lett.*, *23*, 57–60, 1996.

Wilks, D. S., Multisite downscaling of daily precipitation with a stochastic weather generator, *Clim. Res.*, *11*, 125–136, 1999.

Wilks, D. S., and R. L. Wilby, The weather generation game: a review of stochastic weather models, *Prog. Phys. Geog.*, *23*, 329–357, 1999.

Wunsch, C., The interpretation of short climate records, with comments on the North Atlantic and Southern Oscillations, *Bull. Am. Met. Soc*, *80*, 245–255, 1999.